

TESIS DOCTORAL

uc3m | Universidad **Carlos III** de Madrid

ESCUELA POLITÉCNICA SUPERIOR

DEPARTAMENTO DE INFORMÁTICA

HACIA UNA REPRESENTACIÓN DEL CONOCIMIENTO BASADA EN LA PERCEPCIÓN

Autor:

D. Lisardo Prieto González

Directores:

Dr. D. Antonio de Amescua Seco

Dr. D. Ricardo Colomo Palacios

Dr. D. Vladimir Stantchev

Tutor:

Dr. D. Antonio de Amescua Seco

Leganés, febrero de 2017

TESIS DOCTORAL

**HACIA UNA REPRESENTACIÓN DEL CONOCIMIENTO BASADA
EN LA PERCEPCIÓN**

Autor:

D. Lisardo Prieto González

Directores:

Dr. D. Antonio de Amescua Seco

Dr. D. Ricardo Colomo Palacios

Dr. D. Vladimir Stantchev

Firma del Tribunal Calificador:

Firma

Presidente: (Nombre y apellidos)

Vocal: (Nombre y apellidos)

Secretario: (Nombre y apellidos)

Calificación:

Leganés, ____ de _____ de _____

AGRADECIMIENTOS

Por un lado, me gustaría agradecerle a usted, querido lector, por haber decidido invertir parte de su recurso más preciado, **el tiempo**, al embarcarse en la revisión de la presente tesis doctoral.

Por otro lado, tienen mi más sincero agradecimiento **todos** aquellos que **colaboraron**, **colaboran**, y **colaborarán** de algún u otro modo en el desarrollo del trabajo aquí expuesto. Esperemos que la consecución y el buen uso del marco de referencia propuesto traigan muchas dichas, tanto al género humano como al entorno que lo rodea, y que amplíe de forma notable nuestro conocimiento sobre el Universo en que vivimos.

RESUMEN

Esta Tesis define un marco de referencia para crear un modelo de representación del conocimiento inspirándose en la componente biológica cognitiva. Con ello se pretende conseguir un meta-modelo que se comporte de manera análoga a cómo funciona el sistema de adquisición, representación y almacenamiento de la información en los seres humanos, considerando los principales componentes cognitivos asociados: la percepción, el conocimiento y la memoria. Se espera de dicho meta-modelo el poder extender las capacidades del ser humano (limitación perceptual, tanto en capacidad como en variedad), reduciendo sus inconvenientes, de manera que sea posible integrar más información semántica en cada una de las entidades almacenadas en la Base de Conocimiento Global. Dicho conocimiento global podrá ser usado posteriormente en una gran variedad de Sistemas de Información, tales como motores de inferencia avanzados, planificadores y sistemas de recomendación, entre otros.

El marco de referencia está compuesto por tres partes bien diferenciadas:

- Una serie de Principios Fundamentales identificados y obtenidos mediante un análisis exhaustivo de la literatura, en el área de la Psicología y la Neurociencia Cognitiva.
- Una serie de Características Tecnológicas inferidas por dichos principios y soportadas por un estudio del estado actual en el área de las Ciencias de la Computación.
- Una implementación modular, distribuida, escalable, eficiente y segura del modelo de representación basada en las características tecnológicas y en los principios fundamentales expuestos.

El marco de referencia ha sido validado en sus componentes teóricas mediante expertos en la materia para ambos campos, y en su componente práctica mediante una implementación física y su correspondiente experimentación.

PALABRAS CLAVE: *modelado del conocimiento, cognición, percepción, conocimiento, memoria, marco de referencia, modelo, aprendizaje automático, inteligencia artificial, arquitectura de computadores, computación en la nube, Internet de las cosas, sensores, seguridad de la información, expansión perceptual.*

ABSTRACT

This Thesis defines a framework to create a knowledge representation model inspired by the biological cognitive component. With this, it is sought to achieve a meta-model that behaves in an analogous way as the acquisition, representation and storage systems in humans, considering the main associated cognitive components: perception, knowledge and memory. This meta-model is expected to extend the capabilities of the mankind (perceptual limitation, both in capacity and variety), reducing its drawbacks, so that it is possible to integrate more semantic information in each one of the entities stored in the Global Knowledge Base. Such global knowledge can later be used in a wide variety of Information Systems, such as advanced inference engines, planners and recommendation systems, among others.

The framework is composed by three distinct parts:

- A set of Fundamental Principles identified and obtained through an exhaustive analysis of the literature, in the area of Psychology and Cognitive Neuroscience.
- A series of Technical Characteristics inferred by such Fundamental Principles and supported by a study of the current state in the area of Computer Science.
- A modular, distributed, scalable, efficient and secure implementation of the representation model based on the technological characteristics and fundamental principles exposed.

The framework has been validated in its theoretical components by experts in both fields, and in its practical component through a physical implementation and its corresponding experimentation.

KEYWORDS: *knowledge modelling, cognition, perception, knowledge, framework, model, machine learning, artificial intelligence, computer architectures, cloud computing, Internet of things, sensors, information security, perceptual expansion.*

Índice general

RESUMEN.....	vi
ABSTRACT.....	vii
Índice de ilustraciones.....	xi
Índice de tablas.....	xii
Capítulo 1. Introducción.....	1
1.1. Contexto.....	1
1.2. Relevancia.....	3
1.3. Objetivo e hipótesis.....	4
1.4. Plan de acción de la investigación.....	5
1.5. Del plan de acción a la metodología de la investigación.....	6
1.6. Estructura del documento.....	8
Capítulo 2. Cognición humana.....	9
2.1. Introducción.....	9
2.2. Percepción.....	10
2.2.1. Características de la percepción.....	13
2.2.1.1. Constancia.....	14
2.2.1.2. Agrupamiento.....	14
2.2.1.3. Efectos de contraste.....	15
2.2.2. Sentidos. La visión.....	15
2.2.2.1. Relevancia del contexto y la experiencia. Procesamiento “ <i>top-down</i> ”.....	18
2.2.2.2. Combinación del procesamiento “ <i>bottom-up</i> ” y “ <i>top-down</i> ”. Reconocimiento visual humano.....	20
2.2.2.3. Modelos de coincidencia de características.....	21
2.2.2.4. Modelos de coincidencia con plantillas.....	22
2.2.2.5. Modelos de reconocimiento basados en componentes.....	23
2.2.2.6. Modelos de configuración.....	24
2.2.3. Sentidos. El Oído.....	24
2.2.3.1. Identificación de los objetos. Percepción del timbre.....	27
2.2.3.2. Información utilizada para la identificación de distintos objetos audibles.....	29
2.2.3.3. Frecuencia fundamental y regularidad espectral.....	30
2.2.3.4. Las disparidades de inicio.....	32
2.2.3.5. Contraste con sonidos anteriores.....	33
2.2.3.6. Cambios correlacionados en la amplitud o en la frecuencia.....	34
2.2.3.7. Posicionamiento del sonido.....	35
2.2.3.8. Segmentación de flujos.....	36
2.2.4. Principios en la organización perceptual.....	37
2.2.4.1. Principio de similitud.....	37
2.2.4.2. Principio de buena continuación.....	38
2.2.4.3. Principio del destino común.....	39

2.2.4.4. Principio de la asignación disjunta	39
2.2.4.5. Principio de clausura	40
2.2.4.6. El principio de la atención y de figura – fondo.....	40
2.2.5. Otros sentidos e interacciones sensoriales	41
2.3. Conocimiento	48
2.3.1. Representaciones.....	49
2.3.1.1. Representaciones de modalidad específica: imágenes	50
2.3.1.2. Representaciones de modalidad específica: registro de características	51
2.3.1.3. Representaciones Amodales	52
2.3.1.4. Representaciones amodales estadísticas. Redes neurales	53
2.3.2. Formatos de representación para la percepción y para la composición	54
2.3.3. Categorización.....	55
2.3.4. Estructuras de categorías asociadas al conocimiento.....	56
2.3.4.1. Ejemplares y reglas.....	56
2.3.4.2. Prototipos	57
2.3.4.3. Conocimiento de base y esquemas	57
2.3.4.4. Representación dinámica.....	58
2.4. Memoria.....	58
2.4.1. Memoria a largo plazo	58
2.4.2. Codificación	60
2.4.3. Consolidación	62
2.4.4. Recuperación	62
2.4.4.1 Problemas asociados a la recuperación.....	64
2.5. Conclusiones del capítulo	65
Capítulo 3. Entorno computacional	69
3.1. Introducción	69
3.2. Equipos y líneas de investigación paralelas	69
3.2.1. Algunos sistemas cognitivos que hacen uso de la percepción computacional	70
3.2.2. Algunos servicios (software) que hacen uso de la percepción computacional	77
3.2.3. Tendencias de mercado, e innovación de las empresas más importantes del sector	80
3.3. Cloud Computing.....	87
3.3.1 Virtualización	90
3.3.2 Ejecución mediante contenedores.....	92
3.3.3 Gobierno de la nube, o <i>cloud governance</i>	94
3.4. IoT. <i>Internet of Things</i>	98
3.5. Seguridad de la información y de las comunicaciones	102
3.6. Modelos de representación del conocimiento.....	106
3.6.1. Redes Is-As	107
3.6.2. Marcos.....	107
3.6.3. Redes semánticas como modelo de visualización y presentación	109
3.6.4. Redes bayesianas como modelo estocástico.....	110

3.6.5. Ontologías como modelo de representación	110
3.6.6. SGML como modelo tipo clave-valor y marcado	113
3.6.7. JSON y JSON-LD	114
3.7. Conclusiones del capítulo	115
Capítulo 4. Propuesta (marco teórico)	119
4.1. Introducción	119
4.2. Planteamiento del problema.....	120
4.3. Características y requisitos generales de la cognición. Principios Fundamentales	121
4.4. Características técnicas de diseño para un modelo computacional de representación del conocimiento bio-inspirado.....	147
4.5. Aproximación del modelo biológico de representación del conocimiento a la computación	157
4.5.1. Patrón de análisis sensorial.....	158
4.5.2. Arquitectura lógica.....	159
4.5.3. Entradas perceptuales - La visión	162
4.5.4. Otras entradas perceptuales – Beneficios del IoT	167
4.6. Conclusiones del capítulo	171
Capítulo 5. Validación	173
5.1. Introducción	173
5.2. Validación de la solución.....	173
5.2.1. Validación teórica. Principios Fundamentales y Características Técnicas de diseño.....	173
5.2.2. Validación empírica, o experimental. Arquitectura del modelo	174
5.3. Análisis de resultados. Validación de hipótesis y objetivo	187
5.4. Conclusiones del capítulo	189
Capítulo 6. Epílogo.....	193
6.1. Recapitulación.....	193
6.2. Aportaciones a la Ciencia.....	195
6.3. Líneas futuras de investigación.....	196
Chapter 6. Epilogue (English version).....	199
6.1. Recapitulation.....	199
6.2. Contributions to Science	201
6.3. Future research lines.....	202
Referencias bibliográficas	205

Índice de ilustraciones

Ilustración 1. Metodología de la investigación	7
Ilustración 2. Áreas del Conocimiento comprendidas en la Tesis	8
Ilustración 3. Dos ilusiones ópticas (izquierda: la ilusión Zöllner; derecha: un blivet)	11
Ilustración 4. Sentidos en el ser humano	12
Ilustración 5. Procesamientos “ <i>top-down</i> ” y “ <i>bottom-up</i> ”	16
Ilustración 6. Ilusión de Ebbinghaus.....	19
Ilustración 7. Estructura anatómica del oído humano	25
Ilustración 8. Representaciones amodales para elementos de una misma escena	53
Ilustración 9. Modalidades de memoria a largo plazo	59
Ilustración 10. Esquema conceptual simplificado de la cognición humana	65
Ilustración 11. Esquema conceptual simplificado de la percepción	66
Ilustración 12. Esquema conceptual simplificado del conocimiento	67
Ilustración 13. Esquema conceptual simplificado de la memoria	68
Ilustración 14. Ciclo de “ <i>hype</i> ” para tecnologías emergentes (2016) según Gartner [316]	81
Ilustración 15. Definiciones de NIST para la computación en la nube.....	87
Ilustración 16. Arquitectura de Hipervisor (Máquinas Virtuales) frente a Arquitectura de Contenedores.....	92
Ilustración 17. Relaciones entre estándares y marcos de referencia para el gobierno de las TI.....	96
Ilustración 18. Ecosistema del Internet de las Cosas	98
Ilustración 19. Microcontrolador Arduino UNO con escala de medida	100
Ilustración 20. <i>System-on-chip</i> Intel Edison con escala de medida.....	101
Ilustración 21. Ejemplo de red semántica	109
Ilustración 22. Ejemplo de red bayesiana	110
Ilustración 23. Ejemplo básico de una ontología.....	111
Ilustración 24. Ejemplo de entidad codificada mediante JSON-LD	115
Ilustración 25. Posición del marco de referencia propuesto en la Tesis.....	116
Ilustración 26. Patrón de análisis sensorial	158
Ilustración 27. Propuesta de arquitectura lógica	159
Ilustración 28. Varios escenarios asociados a la visión. Analogía hombre-máquina	163
Ilustración 29. Arquitectura física del modelo para el módulo de visión	165
Ilustración 30. Intel Edison montado en placa de desarrollo, y conectado a sensores específicos	168
Ilustración 31. Algunos sensores utilizables en plataformas de desarrollo	169
Ilustración 32. Arquitectura física del modelo para percepción multi-sensorial	170
Ilustración 33. Esquema de flujo del desarrollo del marco de referencia	172
Ilustración 34. Validación de los conceptos teóricos.....	174
Ilustración 35. Parte implementada del modelo con fines de prueba	175
Ilustración 36. Ejemplo de imagen con una distribución uniforme de cuadrados	177
Ilustración 37. Elementos prototípicos para cada categoría del escenario II	179
Ilustración 38. Ejemplo de red de neuronas.....	180
Ilustración 39. Parte del conjunto de entrenamiento para cada BPNN	181
Ilustración 40. <i>Raspberry Pi</i> conectada a dos cámaras iguales (visión estéreo).....	183
Ilustración 41. <i>Pipeline</i> de reconocimiento utilizando el modelo propuesto	185
Ilustración 42. Ejemplos de representación de información en el sistema mediante JSON	191

Índice de tablas

Tabla 1. Receptores del tacto en las distintas capas de la piel y sus características	45
Tabla 2. Sectores afectados en base al número de incidentes	103
Tabla 3. PF-1.....	121
Tabla 4. PF-2.....	121
Tabla 5. PF-3.....	122
Tabla 6. PF-4.....	122
Tabla 7. PF-5.....	122
Tabla 8. PF-6.....	123
Tabla 9. PF-7.....	123
Tabla 10. PF-8.....	123
Tabla 11. PF-9.....	124
Tabla 12. PF-10.....	124
Tabla 13. PF-11.....	125
Tabla 14. PF-12.....	125
Tabla 15. PF-13.....	125
Tabla 16. PF-14.....	126
Tabla 17. PF-15.....	126
Tabla 18. PF-16.....	126
Tabla 19. PF-17.....	127
Tabla 20. PF-18.....	127
Tabla 21. PF-19.....	127
Tabla 22. PF-20.....	128
Tabla 23. PF-21.....	128
Tabla 24. PF-22.....	129
Tabla 25. PF-23.....	129
Tabla 26. PF-24.....	129
Tabla 27. PF-25.....	130
Tabla 28. PF-26.....	130
Tabla 29. PF-27.....	130
Tabla 30. PF-28.....	131
Tabla 31. PF-29.....	131
Tabla 32. PF-30.....	131
Tabla 33. PF-31.....	131
Tabla 34. PF-32.....	132
Tabla 35. PF-33.....	132
Tabla 36. PF-34.....	132
Tabla 37. PF-35.....	133
Tabla 38. PF-36.....	133
Tabla 39. PF-37.....	133
Tabla 40. PF-38.....	134
Tabla 41. PF-39.....	134
Tabla 42. PF-40.....	134
Tabla 43. PF-41.....	135
Tabla 44. PF-42.....	135
Tabla 45. PF-43.....	135
Tabla 46. PF-44.....	136
Tabla 47. PF-45.....	136
Tabla 48. PF-46.....	136
Tabla 49. PF-47.....	137
Tabla 50. PF-48.....	137

Tabla 51. PF-49.....	137
Tabla 52. PF-50.....	138
Tabla 53. PF-51.....	138
Tabla 54. PF-52.....	138
Tabla 55. PF-53.....	139
Tabla 56. PF-54.....	139
Tabla 57. PF-55.....	139
Tabla 58. PF-56.....	140
Tabla 59. PF-57.....	140
Tabla 60. PF-58.....	140
Tabla 61. PF-59.....	141
Tabla 62. PF-60.....	141
Tabla 63. PF-61.....	141
Tabla 64. PF-62.....	142
Tabla 65. PF-63.....	142
Tabla 66. PF-64.....	142
Tabla 67. PF-65.....	143
Tabla 68. PF-66.....	143
Tabla 69. PF-67.....	143
Tabla 70. PF-68.....	144
Tabla 71. PF-69.....	144
Tabla 72. PF-70.....	144
Tabla 73. PF-71.....	145
Tabla 74. PF-72.....	145
Tabla 75. PF-73.....	145
Tabla 76. PF-74.....	146
Tabla 77. PF-75.....	146
Tabla 78. PF-76.....	146
Tabla 79. PF-77.....	147
Tabla 80. PF-78.....	147
Tabla 81. CT-1.....	148
Tabla 82. CT-2.....	148
Tabla 83. CT-3.....	148
Tabla 84. CT-4.....	149
Tabla 85. CT-5.....	149
Tabla 86. CT-6.....	149
Tabla 87. CT-7.....	149
Tabla 88. CT-8.....	150
Tabla 89. CT-9.....	150
Tabla 90. CT-10.....	150
Tabla 91. CT-11.....	151
Tabla 92. CT-12.....	151
Tabla 93. CT-13.....	151
Tabla 94. CT-14.....	152
Tabla 95. CT-15.....	152
Tabla 96. CT-16.....	152
Tabla 97. CT-17.....	152
Tabla 98. CT-18.....	153
Tabla 99. CT-19.....	153
Tabla 100. CT-20.....	153
Tabla 101. CT-21.....	154
Tabla 102. CT-22.....	154
Tabla 103. CT-23.....	154

Tabla 104. Matriz de trazabilidad - Características Técnicas de Diseño a Principios Fundamentales (1 a 39)	155
Tabla 105. Matriz de trazabilidad - Características Técnicas de Diseño a Principios Fundamentales (40 a 78)	156
Tabla 106. Resultados promedio de la ejecución en el procesamiento de imágenes individuales	178
Tabla 107. Resultados de reconocimiento al aplicar la segunda implementación del modelo	182
Tabla 108. Rangos de acidez promedio para categorías cítricas (fuente: FDA [498])	184
Tabla 109. Resultados de reconocimiento al aplicar la implementación final del modelo	186
Tabla 110. Validación de la sub-hipótesis H_1	187
Tabla 111. Validación de la sub-hipótesis H_2	187
Tabla 112. Validación de la sub-hipótesis H_3	188
Tabla 113. Validación de la sub-hipótesis H_4	188
Tabla 114. Validación de la sub-hipótesis H_5	189

Capítulo 1. Introducción

1.1. Contexto

Hoy en día Internet cuenta con más de dos mil cuatrocientos millones de usuarios alrededor del globo [1] y el aumento de conexiones y dispositivos a la red de redes sigue una tendencia lineal de crecimiento. Tal cantidad de usuarios interconectados, crean y comparten diariamente a través de los distintos servicios que brinda Internet cantidades ingentes de datos, de información, de conocimiento. Esta descomunal cantidad de datos, en muchas ocasiones redundantes, ha derivado en el problema conocido como *information overload*, o sobrecarga de información [2].

Según la Real Academia Española en su vigésima tercera edición del diccionario de la lengua española [3], el término conocimiento tiene múltiples acepciones, entre las que destacan:

- Entendimiento, inteligencia, razón natural.
- Noción, saber o noticia elemental de algo.
- (Psicol.) Estado de vigilia en que una persona es consciente de lo que le rodea.

Debido a la profusión de definiciones es necesario recurrir a la rama de la ciencia dedicada a estudiar el conocimiento. Dicha rama es la epistemología, que, atendiendo al citado diccionario, es la doctrina de los fundamentos y métodos del conocimiento científico. La epistemología proporciona múltiples definiciones y diversas teorías acerca del conocimiento. Desde un punto de vista simplificado, la epistemología relaciona las acciones con la mente y con el conocimiento, atendiendo este último al conjunto de hechos e información que adquiere un ser vivo a través de la experiencia o educación, y su comprensión práctica y teórica de elementos referentes a la realidad. Entre los tipos de conocimiento acerca de dichos elementos se destaca el “saber qué”, el “saber cómo”, el “saber cuándo” y el “saber dónde” [4].

Los sistemas de representación del conocimiento han evolucionado en las últimas décadas. Cabe destacar la evolución adoptada por las tecnologías web, comenzando con páginas en las que la información estaba estructurada en etiquetas simples con bajo contenido semántico hasta alcanzar actuales tecnologías como la web semántica, en las que la información de las etiquetas es mucho más rica y completa, permitiendo a los motores de búsqueda especializados obtener resultados más acordes con las búsquedas realizadas en los mismos, o la web social, que básicamente se basa en un ecosistema de participación entre usuarios, donde el contenido se crea dinámicamente mediante la agregación de contribuciones individuales de múltiples usuarios que pueden estar relacionados directa o indirectamente [5]

Como se ha comentado anteriormente en este punto, la red de redes permite intercambiar diariamente ingentes cantidades de información entre sus usuarios. Información que puede estar representada de diversas formas (fotografías e

imágenes, vídeo, audio, texto) y tener diversas naturalezas (contenido estático, como en una web, o bien generado dinámicamente, como en una videoconferencia). Tal grado de heterogeneidad en la naturaleza de los contenidos hace que el poder clasificarlos, indexarlos y acceder de forma precisa a los mismos sea una tarea extremadamente costosa. Suficientemente complicada como para ser inabarcable por el ser humano, a pesar de las diferentes tecnologías semánticas surgidas para ayudar a tal propósito.

A fin de poder clasificar parte de la información accesible en Internet, se han desarrollado diferentes elementos basados en el área de las ciencias de la computación, y de la inteligencia artificial. Tecnologías como son los buscadores (desde texto plano hasta similitud entre imágenes) y arañas de búsqueda, las cuales pueden considerarse agentes inteligentes que recorren e indexan automáticamente el contenido. Sin embargo, la labor de clasificación e indexación dista mucho del proceso de interpretación humano, provocando en ocasiones imprecisión a la hora de recuperar los contenidos a pesar de la semántica asociada a los datos indexados.

Por otra parte, cada vez existen más sistemas dotados de conectividad a la red, y un amplio rango de sensores diferentes, capaces de percibir diferentes características del entorno. Dichos dispositivos, considerados parte de la corriente denominada *Internet of Things* (IoT) [6]–[8], se emplean en un variado número de entornos. Desde sistemas domóticos [9] para hogares hasta en la denominada Industria 4.0 [10], [11]. Sin embargo, aunque dichos sistemas tienen un enorme potencial a la hora de enriquecer la información relativa al entorno en el que actúan gracias a sus sensores y sus capacidades de comunicación, por lo general, suelen aplicarse únicamente en tareas específicas, en las cuales los datos de los sensores suelen ser empleados únicamente para determinar acciones a realizar en base a eventos detectados (p. ej. levantar una persiana, encender y apagar luces, activar climatización, activar una alarma, comenzar o detener un proceso de producción, etcétera). Si bien, es cierto que existen servicios de Internet que aprovechan dicha potencia “sensorial” para proporcionar sus funcionalidades clave, como es el caso de la predicción del tráfico por Google [12] utilizando la información en tiempo real facilitada por los *smartphones* dotados del servicio, o la navegación y geo-posicionamiento en interiores, como por ejemplo, en la red de metro, mediante el uso de los sensores inerciales y del barómetro (en dispositivos dotados de los mismos) [13], es cierto que todavía no se está aprovechando al máximo el rendimiento que podría darse a dicha información para enriquecer las características semánticas del entorno en el que operan.

1.2. Relevancia

Tras realizar un estudio del estado del arte en el ámbito tecnológico, y tras consultar a distintos expertos en la materia, no se han encontrado evidencias sobre la existencia de un sistema singular capaz de aprender conocimiento de cualquier disciplina y realizar un proceso de razonamiento combinando el conocimiento de diferentes áreas. Igualmente, la capacidad de aprender de distintas fuentes, al igual que hacen los seres humanos, podría mejorar drásticamente la reducción de redundancia en la base de conocimiento de sistemas existentes, como puede intuirse de la evolución tecnológica de los sistemas de clasificación y recomendación presentada en el estudio de [5]. Nótese que la redundancia es negativa en cuanto al almacenamiento de los prototipos representativos en la base de conocimiento, pues puede aumentar significativamente el tamaño de la misma, y es dicha redundancia la que pretende minimizarse, agrupando múltiples representaciones de un mismo elemento clasificadas bajo el mismo concepto en base a sus características esenciales, y extendiendo el prototipo mediante características adicionales generando sub-clases o sub-categorías, en lugar de almacenar una nueva instancia completa (tarea intuitiva que ocurre de manera natural en los seres humanos, como se enuncia en la teoría de prototipos formulada por Rosch [14]). Sin embargo, es cierto que la redundancia puede ayudar en el entrenamiento de sistemas de aprendizaje automático, y al igual que con la teoría de prototipos, se verá más adelante la manera en la que se tratará el conjunto de entrenamiento. Dicho sistema también mejoraría la precisión en la recuperación de información, ya que disponer de mayor representatividad podría ayudar a crear motores de búsqueda semánticos muy precisos con capacidad de autoaprendizaje, útiles en el Internet actual, donde la tasa de ruido en los principales buscadores es relativamente alta. Asimismo, podría ayudar a la investigación en nuevas áreas del conocimiento, desarrollar sistemas expertos en un amplio rango de dominios, incluyendo, por ejemplo, diagnósticos médicos, asistencia a la toma de decisiones empresariales o paliar el efecto negativo de las noticias falsas [15], [16], con un alto índice de repercusión social.

Desde un punto de vista a largo plazo, un sistema cuya base de conocimiento extienda un modelo de representación de conocimiento basado en la percepción, proporcionaría ciertas características similares a las del ser humano en cuanto a aprendizaje y procesamiento asociativo y relacional de la información, lo cual ayudaría a incrementar la calidad de vida de las personas a nivel global, pues la aplicabilidad sería prácticamente universal.

1.3. Objetivo e hipótesis

Se puede considerar como principal objetivo de la tesis la creación de un nuevo modelo de representación de conocimiento que tenga en cuenta la percepción, de una forma similar a como lo hacen los humanos (mediante sus sentidos), a fin ser utilizado posteriormente en sistemas complejos, capaces de emplear módulos de razonamiento automático, o módulos de aprendizaje y adquisición de datos desde el entorno.

Teniendo en cuenta el objetivo expuesto, la hipótesis de investigación que se trata de validar, puede formularse de la siguiente manera:

LA EXISTENCIA DE UN MODELO DE REPRESENTACIÓN DEL CONOCIMIENTO BASADO EN LA PERCEPCIÓN FACILITARÁ, AL SER APLICADO:

- UNA MEJORA EN LA HOMOGENIZACIÓN E INTEGRACIÓN DEL CONOCIMIENTO DE DIFERENTES ÁREAS BAJO UN MARCO COMÚN DE REPRESENTACIÓN.
- LA DETECCIÓN Y PALIACIÓN DE REDUNDANCIA ASOCIADA A LOS CONCEPTOS Y TÉRMINOS CLASIFICADOS Y ALMACENADOS EN UNA BASE DE CONOCIMIENTO GLOBAL.
- EL DESARROLLO DE NUEVAS TÉCNICAS DE APRENDIZAJE AUTOMÁTICO.
- UNA MEJORA EN LA CAPACIDAD DE RECONOCIMIENTO SEMÁNTICO POR PARTE DE LOS SISTEMAS QUE IMPLEMENTEN EL MODELO PROPUESTO.

A su vez, esta hipótesis puede ser subdividida en cinco sub-hipótesis tal y como se expresa a continuación:

Sub-Hipótesis H₁: La integración de espacios de conocimiento correspondientes a diferentes áreas se verá facilitada mediante la adopción de un modelo de representación del conocimiento basado en la percepción.

Sub-Hipótesis H₂: Es posible mejorar la calidad del aprendizaje automático al reducir la redundancia de términos y conceptos almacenados en una base de conocimiento global, mediante la aplicación de un modelo de representación del conocimiento basado en la percepción.

Sub-Hipótesis H₃: El tratamiento semántico de la información resultará más preciso al aplicar de un modelo de representación del conocimiento basado en la percepción.

Sub-Hipótesis H₄: Será posible alcanzar mayor precisión en la recuperación de información al contener ésta mayor diversidad en la naturaleza de los datos que forman su definición (más descriptiva). El tratamiento de la misma será más completo y la generación de nuevos conceptos o relaciones entre los mismos podrá adecuarse a modelos de razonamiento similares a los humanos.

Todo gracias a la aplicación de un modelo de representación del conocimiento basado en la percepción.

Sub-Hipótesis H₅: Es muy posible que la complejidad computacional aumente considerablemente al ser un modelo complejo basado en diferentes módulos interrelacionados.

1.4. Plan de acción de la investigación

Para llevar a cabo el objetivo definido en la Tesis, se realizarán las siguientes actividades:

- **Actividad 1.** Establecer en base a un estudio del estado del arte una aproximación al funcionamiento del sistema de creación de recuerdos y almacenamiento de los mismos en la memoria humana desde un punto de vista psicológico y neurocientífico, considerando el funcionamiento de cada uno de los sentidos (percepción) en el proceso de recogida de información.
- **Actividad 2.** Estudiar los principales modelos computacionales actuales de representación del conocimiento, evaluando la capacidad de los mismos para almacenar información de origen perceptual de una naturaleza similar a la estudiada en la actividad 1.
- **Actividad 3.** Definir una serie de principios fundamentales derivados del estudio realizado en la actividad 1, con el objetivo de asentar las bases, características y requisitos que deberá presentar el modelo computacional de representación del conocimiento basado en la percepción.
- **Actividad 4.** Basándose en los resultados de las actividades 2 y 3, definir una serie de características técnicas de diseño sobre las que desarrollar las arquitecturas lógica y física del modelo computacional, escalable y extensible, que permita representar el conocimiento obtenido perceptualmente, de forma estructurada.
- **Actividad 5.** Evaluar los principios fundamentales y las características técnicas de diseño obtenidas en las actividades 3 y 4. Evaluar la factibilidad y funcionamiento de la arquitectura definida en la actividad 4, analizando el impacto y los resultados del modelo propuesto respecto a los modelos analizados en el segundo objetivo.

El proceso de evaluación será ejecutado mediante validación teórica sobre los principios fundamentales y características técnicas de diseño (actividades 3 y 4), contando con la ayuda de expertos en la materia, y mediante validación empírica de

las arquitecturas propuestas (actividad 4), a fin de verificar que el modelo propuesto en la tesis doctoral mejora las capacidades de representación de conocimiento de otros modelos actuales, proporcionando ventajas sobre los mismos.

Según los resultados a validar, se propone un plan de ejecución en cuatro fases bien diferenciadas:

- **Fase 0:** Evaluación del modelo en base al estudio sobre las áreas relativas a la psicología, neurociencia y ciencias de la información. Establecimiento de un ciclo de refinamiento y mejora del modelo en base a nuevos estudios en las diversas materias que fundamentan el modelo.
- **Fase 1:** Conceptualización de la arquitectura correspondiente al modelo propuesto. Evaluación de su factibilidad y funcionalidad.
- **Fase 2:** Comparación entre la eficiencia en el reconocimiento aplicando conceptualización mediante modelos actuales frente a conceptualización mediante el modelo propuesto en la tesis.
- **Fase 3:** Análisis de resultados de las fases mencionadas, verificando o rechazando las hipótesis planteadas.

1.5. Del plan de acción a la metodología de la investigación

La consecución del objetivo establecido, dota a la tesis de un carácter teórico-práctica, requiriendo tanto investigación documental como investigación de campo.

A fin de abordar las actividades citadas en el apartado anterior, se han clasificado y asociado a un enfoque metodológico de investigación, dependiendo del tipo de actividad a desarrollar dentro del marco de la tesis. La secuencia del método científico asociado a dichas actividades es la siguiente:

1. **Observación – estudio (estado de la cuestión).** Consistente en un estudio mediante una revisión sistemática de la literatura, a fin de concretar las diferentes partes sobre las que se asentará el modelo de representación del conocimiento. Esta actividad se divide principalmente en tres partes:
 - a. Estudio referente a la componente psicológica del modelo (funcionamiento de la percepción y la memoria en los seres humanos).
 - b. Estudio de modelos de representación computacionales sobre los que podría basarse el modelo objetivo de la tesis.
 - c. Búsqueda de soluciones existentes híbridas, que combinen algún tipo de percepción con un modelo de representación (por ejemplo, modelos que almacenen información visual, o auditiva).

Para la realización de estas actividades se ha contado con el soporte de expertos en las distintas materias.

2. **Inducción e hipótesis (formalización del modelo).** A partir de las observaciones sobre teorías existentes y experiencias concluidas de la primera actividad se extraen los principios particulares y conceptos que conforman la solución al problema. Dicha solución es la formalización del modelo de representación del conocimiento basado en la percepción. Al igual que en un proyecto de software, es necesario cubrir una serie de sub-actividades para finalizar la presente actividad. En primer lugar, es necesario realizar un estudio de requisitos, seguido de un análisis del modelo y sistemas asociados al mismo. A partir de este punto, se puede realizar un diseño inicial del modelo y seguidamente una implementación de prueba para poder realizar una evaluación del mismo. Siguiendo con el ciclo de vida del software, será necesario revisar algunas sub-actividades en un futuro, a fin de ser ajustadas, teniendo que propagar los cambios realizados en el resto de sub-actividades.

3. **Prueba y demostración de la hipótesis (evaluación del modelo).** A fin de comprobar la validez del modelo generado, es necesario realizar una serie de pruebas teóricas y experimentales. Dichas pruebas son definidas en base a las características desarrolladas e inferidas desde la actividad 1.

4. **Tesis (documentación y conclusiones).** A lo largo del desarrollo de la tesis, es conveniente y necesario documentar el proceso de diseño del modelo. Esto, además de justificar su validez, hace que la base de la que partir se encuentre bien documentada en caso de requerir cambios para ser adaptado a comportamientos futuros más precisos o complejos, facilitando el proceso de cambio o adaptación. Igualmente, en la documentación se proponen futuras líneas de investigación en base a las conclusiones obtenidas del trabajo de investigación.

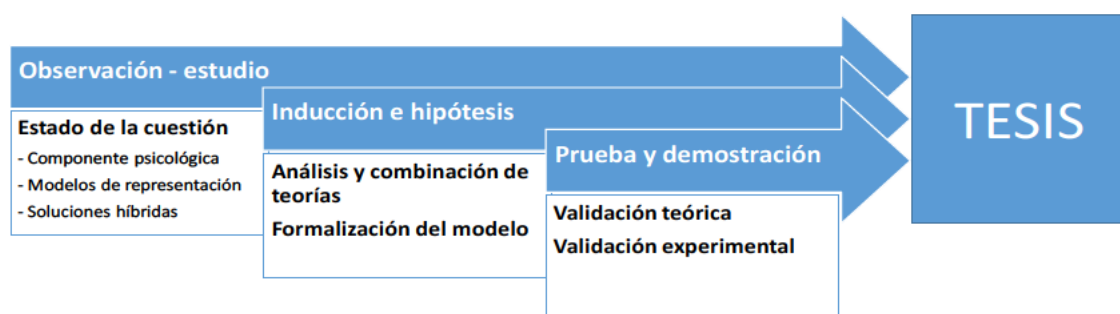


Ilustración 1. Metodología de la investigación

1.6. Estructura del documento

La estructura seguida por el presente documento de tesis es la siguiente: en el capítulo 2, se presentan un estudio sobre las características fundamentales de la cognición humana, desde un punto de vista psicológico y neurocientífico, del estado de la cuestión referente a la investigación planteada, los objetivos a conseguir junto con las metodologías y actividades necesarias para alcanzarlos, y el estado actual de dichas actividades; en el capítulo 3 se realiza la evaluación del estado de la cuestión sobre el marco tecnológico que sustentará el modelo. Seguidamente, el capítulo 4 presenta las diferentes aportaciones a la ciencia desarrolladas en la tesis. Comenzando por una serie de principios fundamentales con los que debe contar un modelo como el propuesto, inspirados en el estudio presentado en el capítulo 2, siguiendo con una serie de características técnicas de diseño sobre las que sustentará una arquitectura que modele el sistema, y la arquitectura en sí junto con una serie de recomendaciones. El capítulo 5 muestra el plan de validación empleado, y la correspondencia entre los resultados y las hipótesis a comprobar. Finalmente, los capítulos sexto y séptimo esbozan una serie de conclusiones y líneas de trabajo futuras hacia las que orientar la investigación. Nótese que el capítulo 7 se corresponde a la versión traducida al idioma inglés del capítulo 6.

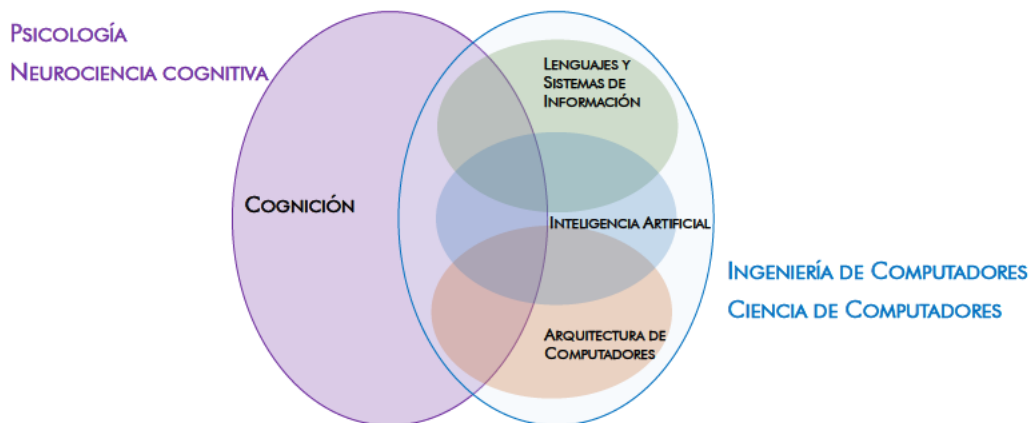


Ilustración 2. Áreas del Conocimiento comprendidas en la Tesis

Capítulo 2. Cognición humana

2.1. Introducción

El aprendizaje humano (adquisición de conocimiento) es una tarea muy compleja que se lleva a cabo a lo largo de toda la vida [17], y el proceso de la percepción juega un papel fundamental en él al recibir estímulos del exterior, hasta el punto de que "*sin percepción no hay conocimiento*" [18]. Aquí entra en juego el término cognición, definido como la *habilidad de procesar información desde la percepción, el conocimiento adquirido (experiencia) y las características subjetivas que nos permiten evaluar la información*, lo cual puede ser visto como un enlace entre el aprendizaje humano y la percepción humana.

El diccionario Oxford define la percepción como la habilidad de ver, escuchar, o ser consciente de algo a través de los sentidos. Acorde a la Psicología y la Zoología: *el proceso neuropsicológico, incluida la memoria por la que un organismo toma conciencia de e interpreta los estímulos externos*. Como se muestra en la definición anterior, la memoria juega un papel importante en la percepción. La forma en la que los estímulos son procesados y almacenados por el cerebro humano podría ser la clave para enriquecer el proceso de aprendizaje automático no supervisado.

En las siguientes sub secciones se analizarán los aspectos correspondientes a la percepción y almacenamiento de la información en el ser humano, pero no entrando en detalle sobre el razonamiento ni en la inteligencia y los distintos procesos que derivan en ella, al ser un tema demasiado extenso y que queda fuera del objetivo de la presente tesis doctoral. Cabe destacar que, como líneas futuras de investigación, a fin de aplicar o mejorar el modelo que se propone en el presente trabajo, resultaría útil evaluar dichos procesos.

Así pues, este capítulo pretende definir las bases psicológicas y neurocientíficas acerca de la cognición en el ser humano, centrándose en las tres características principales de la misma: la percepción, el conocimiento y la memoria. El contenido del capítulo recopila y analiza las teorías fundamentales y diferentes estudios significativos conducidos hasta la fecha sobre estos tres pilares principales. Finalmente se exponen una serie de conclusiones e ideas clave que facilitarán el desarrollo del modelo, mediante la creación de los elementos necesarios en sinergia con la adaptación de la tecnología existente.

2.2. Percepción

Para el ser humano, la percepción puede ser considerada como la identificación y la interpretación de información sensorial a fin de representar y comprender su entorno [19]. Cada sistema perceptual implica la generación y envío de señales a través del sistema nervioso central, las cuales son producto de la estimulación física de los órganos sensoriales [20]. Por ejemplo, escuchar comienza con la estimulación de una onda mecánica (variación de presión) en los tímpanos, la visión implica la estimulación de la retina en el interior de los ojos, y el olfato depende directamente de la estimulación producida por determinadas sustancias químicas en partes específicas de la nariz. Sin embargo, la percepción no consiste en la recepción pasiva de este tipo de señales. Es ajustada y afinada mediante el aprendizaje, la memoria y la anticipación [21], [22]. La percepción requiere procesamiento “*top-down*” y “*bottom-up*” al tratar la información originada por los diferentes órganos sensitivos [22]. El procesamiento tipo “*bottom-up*” (de abajo hacia arriba) comienza con la información perceptual de bajo nivel recogida del entorno (p. ej. Reconocimiento visual de patrones para identificar objetos). El enfoque “*top-down*” (de arriba hacia abajo) se ocupa de la recuperación activa de información, y está basado en el conocimiento, la experiencia, las creencias, las expectativas y los objetivos propios. Así pues, la percepción depende de muchas y complejas funciones del sistema neural, pero desde un punto de vista subjetivo, parece ser fácilmente realizado ya que todo el procesamiento se realiza subconscientemente [20]. Ambos tipos de procesamiento, “*bottom-up*” y “*top-down*”, al igual que sus implicaciones en la percepción serán discutidas posteriormente en el capítulo.

Desde el inicio de la psicología experimental hacia el siglo XIX, la comprensión de la psicología de la percepción ha progresado mediante la combinación de una variedad de técnicas [21]. La psicofísica mide el efecto en la percepción de la variación de las cualidades físicas de la entrada. La neurociencia sensorial estudia los mecanismos cerebrales subyacentes a la percepción. Los sistemas de percepción pueden también estudiarse computacionalmente, en términos de la información que procesan. En la filosofía se tienen en cuenta cuestiones de percepción como en qué medida existen cualidades sensoriales (olores, colores) en la realidad objetiva en lugar de en la mente del perceptor [21].

Aunque los sentidos hayan sido considerados tradicionalmente como receptores pasivos, el estudio de las ilusiones [23] y las imágenes ambiguas (ver Ilustración 3) ha demostrado que son en realidad los sistemas perceptuales del cerebro los que intentan proporcionar sentido a los estímulos recibidos de manera una forma pasiva y pre-consciente [20]. Todavía existe un debate activo sobre si la percepción es un proceso de pruebas de hipótesis similar a la ciencia o la información sensorial realista es lo suficientemente completa como para hacer este proceso innecesario.

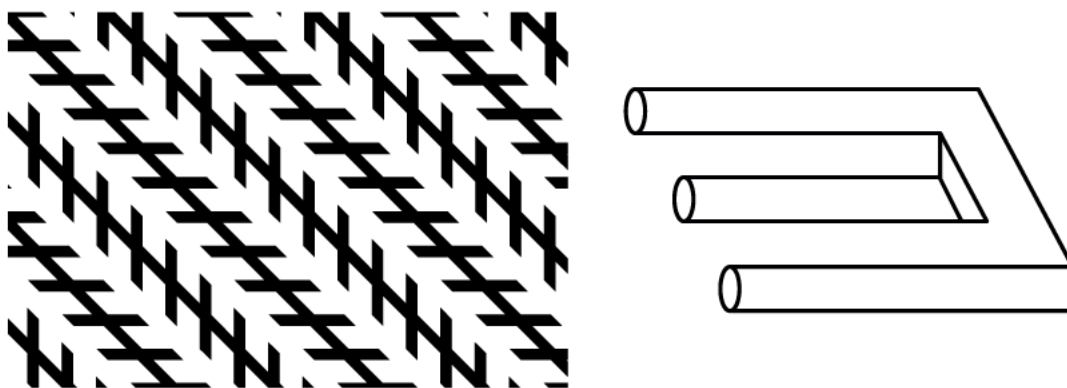


Ilustración 3. Dos ilusiones ópticas (izquierda: la ilusión Zöllner; derecha: un blivet)

Desde Aristóteles [24], muchos filósofos, psicólogos e investigadores han estudiado acerca de la percepción y las interacciones entre los sentidos. Irvin Rock, centrándose en la percepción visual, considera que el conocimiento en forma de representaciones almacenadas afecta a la percepción, habilitando capacidades en los seres humanos tales como el reconocimiento y la interpretación, la discriminación perceptual entre miembros similares de una categoría, el acceso a soluciones previamente aprendidas para aplicar en casos donde ocurre la resolución perceptual del problema, entre otros [25]. Muchos autores centraron su investigación en visión argumentando que es el sentido que proporciona la más rica y detallada información sobre el entorno, pero el sistema perceptual es muy complejo y se basa en variados sistemas interconectados. Por ejemplo, Harry McGurk y John MacDonald publicaron un estudio en 1976 analizando la influencia de la visión sobre la percepción del discurso [26]. Igualmente existen interacciones bien conocidas, como entre el olfato y el gusto, y estudios sobre cómo estas relaciones afectan la percepción del sabor [27].

Acorde con la clasificación perceptual clásica propuesta por Aristóteles [28], existen cinco sentidos principales en el ser humano: la vista, el oído, el olfato, el gusto y el tacto. Cada sentido está asociado con un sistema sensorial: ojos, orejas, nariz, lengua y piel. Esos sentidos permiten a las personas percibir el entorno que les rodea. Sin embargo, muchos otros sentidos han sido definidos e incluidos en esta lista. Sentidos que añaden información sobre el entorno o sobre el ser en sí mismo. La *Propriocepción*, o sentido kinético, proporciona a los seres consciencia sobre la posición de las partes del cuerpo, consciencia sobre el movimiento corporal y cuánta fuerza es necesaria para mover cada parte. Este sentido está basado en receptores dedicados situados en los músculos, tendones y articulaciones [29]. La *Equilibriocepción* (sentido vestibular, o sentido del equilibrio) permite la detección del movimiento del cuerpo, la dirección y la aceleración al sentir el campo gravitatorio gracias a canales semicirculares llenos de líquido y órganos otolíticos en los oídos [29]. La *Cronocepción*, o sentido del tiempo, es responsable de la consciencia y experiencia del paso del tiempo [30], [31]. La *Termocepción* [32] y la *Nociocepción* [33] determinan la percepción de temperatura y dolor respectivamente, y tienen una relación muy cercana ya que ambos sentidos utilizan canales TRP (*transient receptor potential*, canal iónico receptor de potencial transiente) para el proceso de la detección. Como se puede observar, algunos de dichos sentidos detectan objetos y propiedades del mundo externo al cuerpo (como la vista y el oído). Dichos sentidos se

denominan *exteroceptivos*. Por otra parte, algunos otros sentidos detectan cambios en el cuerpo. Estos se denominan *interoceptivos*. La combinación de sentidos *exteroceptivos* e *interoceptivos* le proporciona al cerebro múltiples fuentes de información para componer la percepción del entorno. Es intuitivo pensar que cuanto mayor información sensorial sea recibida, más compleja y precisa será la definición del entorno percibido. Por ejemplo, algunos animales tienen sentidos especiales adaptados al medio en el que viven, como el caso de los delfines, equipados con *Ecolocalización*, habilidad de emitir ultrasonidos a fin de determinar obstáculos o bancos de peces que componen su comida; o el caso de las palomas, las cuales utilizan *Magnetocepción* (habilidad de sentir el campo magnético) a fin de volver a casa; o los tiburones, que utilizan la *Electrocepción* para detectar movimientos musculares o espasmos en animales vivos como los peces (similar a las máquinas de electrocardiograma para monitorizar el latido cardíaco).

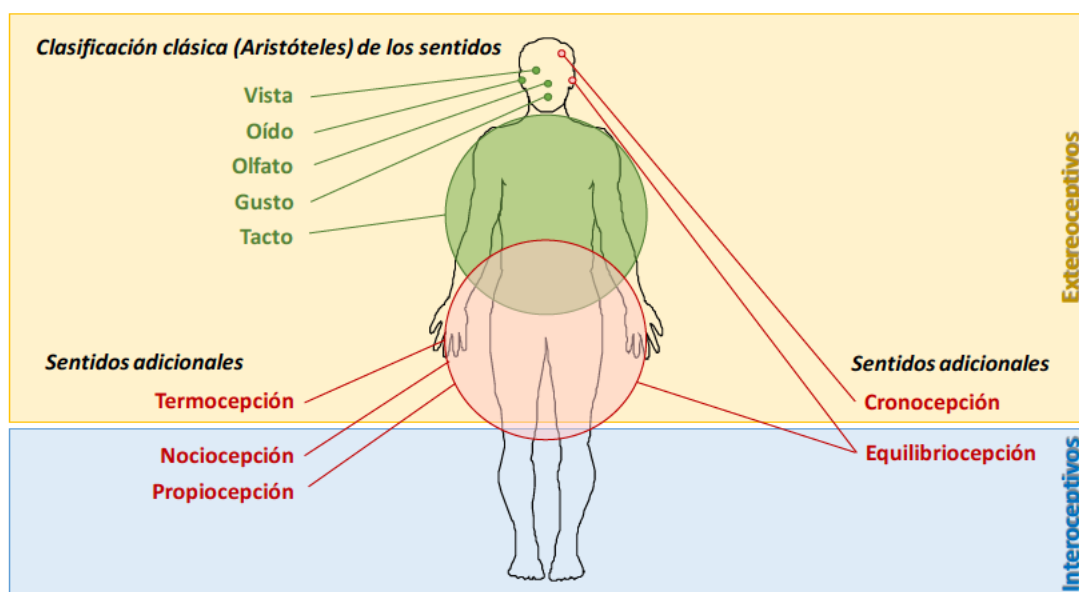


Ilustración 4. Sentidos en el ser humano

Los sistemas perceptuales del cerebro permiten a la gente comprender el entorno como un sistema estable y coherente, aunque la información sensorial pueda ser incompleta o variar rápidamente [34]. Tanto los cerebros humanos como los animales están estructurados de manera modular [34], presentando áreas de procesamiento por defecto diferentes para cada tipo de estímulo sensorial. Algunos de dichos módulos asignan características del entorno a lo largo de la superficie cerebral y se encuentran interconectados entre sí, interactuando mutuamente. Por ejemplo, el sentido del gusto está fuertemente relacionado con el olfato [35]. Respecto a esto último y en el contexto de la visión, una de las interacciones más importantes entre los sentidos que afecta directamente a la visión es la interacción entre la *Propiocepción*, la *Equilibriocepción* y la *Cronocepción*. La falta de interacción y/o el comportamiento alterado de dichos sentidos puede derivar en consecuencias catastróficas [36] debido a una percepción errónea del entorno.

La adición de nuevos sentidos al ser humano podría resultar extremadamente complicado a pesar de los importantes avances en la ciencia (principalmente en bioingeniería y medicina), sin embargo, resulta una tarea trivial en entornos computacionales, ya que básicamente requiere la adición de un nuevo periférico específico, como por ejemplo una cámara de infrarrojos, y su configuración para funcionar en el sistema. De esta manera es posible extender las capacidades sensoriales y la precisión del proceso sensorial mediante la adición de receptores sensoriales más exactos, y/o aumentando el número de los mismos. Por ejemplo, añadiendo diferentes cámaras calibradas, y orientadas hacia a un mismo elemento en un escenario se puede computar en mayor o menor grado información tridimensional de dicho elemento. Pero más aún, es posible utilizar la combinación de diferentes tecnologías para mejorar la precisión de la percepción. En este sentido, por ejemplo, la *Termocepción* [32] de los seres humanos es percibida a través de receptores específicos situados en la piel, por lo que para sentir la temperatura, por lo general requiere un nivel determinado de proximidad a los objetos, mientras que en un Sistema de Información, la medida de temperatura podría realizarse a través de un amplio rango de sensores, basados en diferentes principios físicos, tales como termómetros electrónicos basados en resistencia eléctrica, fibras ópticas (en DTS, *Distributed Temperature Sensing Systems* [37]) y cámaras infrarrojas, haciendo posible la detección de temperatura desde corta hasta larga distancia, y a través de diferentes interacciones, tales como el contacto directo o la visión.

Procesar y combinar todos esos estímulos entrantes, y almacenar el conocimiento inferido de los mismos (información semántica extendida) sobre el entorno, conforme a un modelo de representación unificado, puede ayudar a procesar y relacionar la información de forma similar a como los seres humanos realizan dicha tarea. Este hecho representa un avance hacia el razonamiento avanzado y el proceso de aprendizaje no supervisado, proporcionando capacidades y beneficios como la generalización de elementos redundantes en el almacén de conocimiento conforme a la teoría de prototipos [14], la cual es una de las teorías fundamentales relacionadas con la percepción humana y las bases de la memoria. Esta generalización de la información redundante, con la consiguiente reducción del espacio de almacenamiento requerido para almacenar los datos resultantes puede ser de gran relevancia y utilidad, pudiendo aportar entre sus múltiples beneficios la mitigación del fenómeno denominado "*information overload*", comentado al principio de la presente tesis doctoral.

2.2.1. Características de la percepción

A lo largo del tiempo se han determinado diferentes características en la percepción. A continuación, se recogen las principales, atendiendo a diversos estudios.

2.2.1.1. Constancia

La constancia perceptual es la capacidad de los sistemas de percepción para reconocer el mismo objeto desde distintas entradas sensoriales [22] [38]. Por ejemplo, una persona concreta puede ser reconocida a partir de puntos de vista diferentes (de frente, de perfil) que forman imágenes muy distintas en la retina. Una moneda puede ser vista de frente, formando una circunferencia, o ligeramente ladeada, formando una elipse [39]. En la percepción normal, dichos objetos son reconocidos como objetos tridimensionales. Sin este proceso de corrección, un objeto acercándose desde la distancia parece ganar tamaño [39], [40].

Un tipo de constancia perceptiva es la constancia del color. Por ejemplo, un trozo de papel blanco puede ser reconocido como tal en diferentes colores e intensidades de luz [39].

Otro ejemplo es la constancia de la rugosidad: cuando se pasa la mano por una superficie rugosa, los nervios del tacto se estimulan más intensamente. El cerebro compensa esta, por lo que la velocidad de contacto no afecta a la rugosidad percibida [39].

Los sistemas de percepción del cerebro logran la constancia perceptual en una variedad de formas, cada una especializada para el tipo de información a procesar [41].

2.2.1.2 Agrupamiento

Los principios de agrupamiento (o *leyes de la Gestalt de agrupación*) son un conjunto de principios de la psicología, expuestos inicialmente por los psicólogos de la Gestalt para explicar cómo los seres humanos perciben de una forma natural los objetos como patrones organizados y objetos. Los psicólogos de la Gestalt argumentaron que existen estos principios debido a que la mente tiene una disposición innata para percibir patrones basados en estímulos que siguen ciertas reglas. Estos principios están organizados en seis categorías [42][43][44][20]:

- El **principio de proximidad** sostiene que, en igualdad de condiciones, la percepción tiende agrupar estímulos cercanos como parte de un mismo objeto, y los estímulos separados como objetos separados.
- El **principio de similitud** dicta que, en igualdad de condiciones, la percepción se presta a ver a estímulos que se asemejan físicamente entre sí como parte del mismo objeto, y los estímulos que son diferentes como parte de un objeto diferente. Esto permite a las personas a distinguir entre objetos adyacentes y superpuestos en función de su textura visual y semejanza.
- El **principio de cierre** se refiere a la tendencia de la mente para ver figuras completas o formas, incluso si una imagen está incompleta, parcialmente oculta por otros objetos, o si parte de la información necesaria para identificar una imagen completa en nuestra mente no se encuentra presente.

- El **principio de la buena continuación** da sentido a los estímulos que se superponen, como cuando hay una intersección entre dos o más objetos. Las personas tienden a percibir esto como un único objeto ininterrumpido.
- El **principio del destino común de los grupos de estímulos en base a su movimiento**. Cuando los elementos visuales se aprecian moviéndose en la misma dirección y a la misma velocidad, se asocia la percepción del movimiento, como parte de un mismo estímulo.
- El **principio de la buena forma** se refiere a la tendencia a agrupar las formas en base a características similares, como modelo, color, etc.

Aunque posteriores investigaciones han identificado principios de agrupación adicionales [45][20] relacionados con los anteriores. Los principios de agrupamiento serán analizados en mayor detalle posteriormente, en la sección correspondiente a la percepción auditiva.

2.2.1.3 Efectos de contraste

Una característica común en diferentes tipos de percepción es que las cualidades percibidas de un objeto pueden ser afectadas por las cualidades del contexto [46], [47]. Esto provoca que, si un objeto es mucho más grande en alguna dimensión, los objetos vecinos se perciben como más lejos de dicho objeto.

Atendiendo al orden temporal de la percepción, el “*efecto de contraste simultáneo*” hace referencia a cuando dos estímulos se perciben al mismo tiempo, mientras que “*contraste sucesivo*” se aplica cuando los estímulos se presentan espaciados en el tiempo [48].

El efecto de contraste fue modelado inicialmente por el filósofo del siglo XVII John Locke, quien observó que el agua tibia puede sentirse fría o caliente, dependiendo de si la mano que toca el agua fue introducida previamente en agua caliente o fría [49]. En el siglo XX, Wilhem Wundt, identificó el contraste como un principio fundamental de la percepción, y desde entonces, el efecto se ha confirmado en muchas áreas diferentes [49]. Estos efectos determinan no sólo cualidades visuales como color o brillo, sino otros tipos de percepción, como el peso que aparenta tener un objeto [50]. O que una pieza musical sea percibida como buena o mala (puede depender de si la música escuchada anteriormente era desagradable o agradable) [51].

2.2.2. Sentidos. La visión

La percepción humana basada en la visión permite comprender el significado esencial de una escena tras haber sido observado mediante los ojos durante un breve lapso de tiempo. A pesar de que el todo el sistema visual es extremadamente complejo y potente, está limitado, por lo que demasiada información sensorial lo puede colapsar, con la consecuente pérdida de coherencia en los elementos percibidos. Desde hace

En el lado derecho de la imagen, los elementos difieren en múltiples características, como el color, la orientación o la forma. A fin de encontrar el objetivo es necesaria información adicional, como, por ejemplo: “el objetivo es el rectángulo negro vertical”. Esto puede ser utilizado para llevar a cabo procesamiento tipo “*top-down*” a fin de encontrar el elemento que encaja con la descripción-objetivo. Este ejemplo demuestra cómo la percepción humana resulta en interpretaciones de lo que es percibido, más aún, en representaciones mentales producidas por los procesamientos conjuntos “*bottom-up*” y “*top-down*”.

La interpretación del entorno está determinada principalmente por dos hechos: la estructura biológica del entorno, y la experiencia, la cual es enriquecida y modifica la estructura cerebral con los nuevos estímulos perceptuales recibidos. En los recién nacidos, el sistema visual está casi desarrollado por completo, y es bien sabido que muchos de los principales y finales cambios estructurales se llevan a cabo durante el primer año de vida [55], [56]. Tras abrir los ojos, los bebés miran alrededor buscando elementos relevantes (en el sentido subjetivo de la palabra). Cada mirada dura aproximadamente medio segundo en promedio, y en este primer año de vida, procesan una gigantesca cantidad de información (alrededor de diez millones de miradas), entre las que se incluyen caras, objetos desde diferentes perspectivas y en diferentes contextos, y con diversas condiciones de iluminación. Mientras los estímulos visuales son procesados, interpretados y almacenados como recuerdos, los canales neuronales responsables de almacenar dicha información se vuelven más fuertes al ser reforzados con estímulos similares, y la nueva información recogida se combina con recuerdos existentes, proporcionando refuerzo para posteriores tareas de reconocimiento. De forma análoga, estudios conducidos sobre el desarrollo perceptual en animales recién nacidos han concluido que las características del entorno durante ciertas fases del desarrollo tienen consecuencias determinantes en etapas posteriores. En las primeras fases de la vida, el animal tiene que desarrollar determinadas habilidades en momentos específicos o no será capaz de desarrollar las citadas habilidades posteriormente. Incluso aunque sea expuesto a los estímulos ambientales requeridos. Una de dichas habilidades es la percepción de la profundidad, la cual depende de la integración de información percibida por ambos ojos [57], [58]. Además de habilidades específicas como la citada percepción de la profundidad, distintos procesos sensitivos compiten al representar y procesar la entrada en el córtex cerebral [59]. Si la señal percibida por un ojo es más intensa que la señal percibida por el otro, esto es, si es más activa, mediante la plasticidad neural, los recursos del córtex son redistribuidos en dicha dirección a fin de procesar la cantidad adicional de información, pero debilitando con ello las capacidades del otro ojo. Más casos de competición por los recursos perceptuales han sido identificados y estudiados. Por ejemplo, la competición táctil entre diferentes dedos [60] o la competición sensorial entre la vista y el oído [61], [62].

2.2.2.1. Relevancia del contexto y la experiencia. Procesamiento “*top-down*”

El reconocimiento a través del procesamiento “*top-down*” es asistido por la experiencia. Esto es, gracias a los elementos previamente reconocidos. La percepción puede ser conseguida mediante la combinación del procesamiento “*bottom-up*” (la información es recogida por los diferentes órganos sensoriales y cruza una jerarquía de análisis), y procesamiento “*top-down*”, que compara la información recibida con el conocimiento memorizado, las creencias, inferencias y expectativas, afectando a dicho procesamiento “*bottom-up*”. De esta forma, los seres humanos pueden completar información ausente en la percepción sensorial utilizando su experiencia anterior basada en el contexto.

La percepción humana de objetos y sus propiedades, como el color, es poco precisa como ha sido demostrado a través de un amplio conjunto de experimentos en las últimas décadas [42], [44], [46], [63]. A pesar de ello, los seres humanos desempeñan de una forma extraordinariamente eficiente las tareas de reconocimiento e inferencia a través del uso de heurísticas alimentadas por las entradas sensoriales. Dicho procesamiento heurístico que puede acelerar el reconocimiento y mejorar su precisión tiene algunas desventajas, como la reconstrucción errónea de información basada en hechos adquiridos parcialmente. Es posible demostrar este problema a través de las ilusiones ópticas, como las mostradas en la Ilustración 3. Existen muchos tipos de ilusiones causadas por el sistema visual, a fin de producir una percepción coherente del mundo basada en experiencias pasadas. De esta manera, contornos ilusorios pueden producir figuras completas, diferencias en el brillo pueden ser homogeneizadas causando la ilusión de un color uniforme (o el efecto opuesto), y objetos de diferente tamaño pueden ser percibidos de manera diferente dependiendo del contexto en el que se encuentren localizados. En la Ilustración 6 los círculos centrales en los grupos izquierdo y derecho tienen exactamente el mismo tamaño, pero el círculo rodeado por el grupo de círculos mayores parece más pequeño [46], [64], [65]. De la misma manera, cuando existe un grupo que comparte las mismas características, tales como el mismo movimiento, es más sencillo reconocer el elemento que se mueve en una dirección diferente, ya que es percibido como fuera del contexto global.

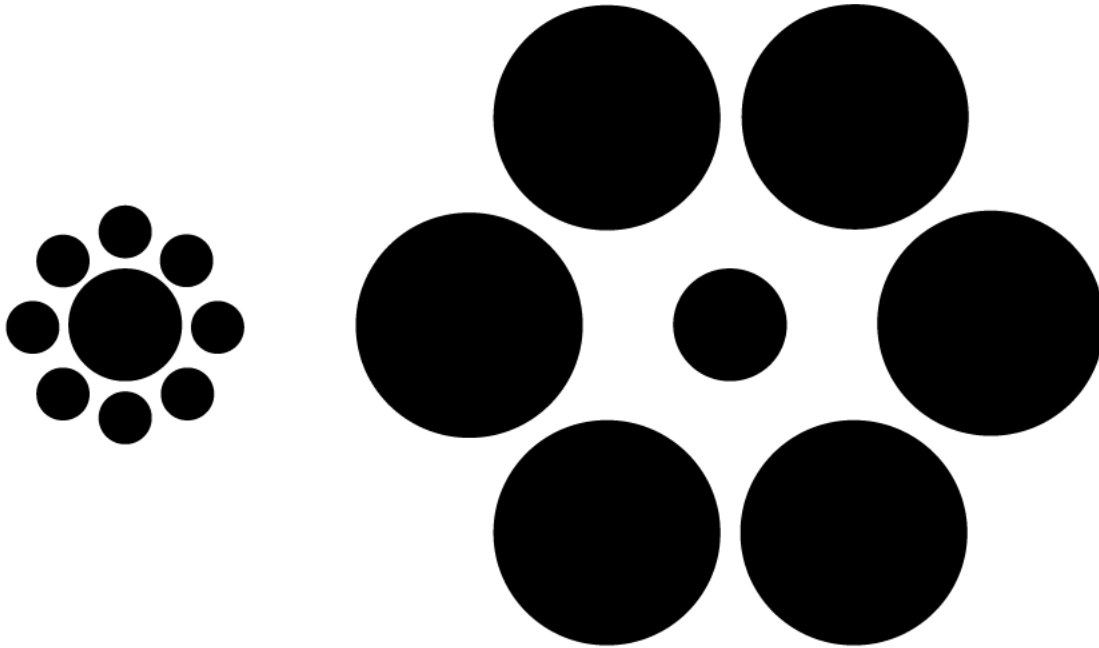


Ilustración 6. Ilusión de Ebbinghaus

Por otra parte, un objeto puede ser fácilmente reconocible si se espera en el contexto en el que está localizado (por ejemplo, un compañero de trabajo en la oficina), o si se encuentra en un contexto frecuente común. Sin embargo, el mismo objeto será más difícilmente reconocible si aparece fuera de los contextos esperados o en una nueva experiencia (por ejemplo, un compañero de trabajo durante unas vacaciones en las Antípodas). Ha sido probado experimentalmente que el contexto tiene una notable influencia en el reconocimiento de objetos mediante estrategias mnemónicas [66], y mediante la atención humana [67].

El procesamiento “*top-down*” afecta a cómo los objetos son percibidos, por ejemplo, en el contexto de la escritura. Si falta una letra en una palabra impresa, o se encuentra en un estado corrupto, el cerebro puede reconstruir la palabra ajustando la palabra que encaja de manera más coherente en el contexto de la oración (significado) [68], [69]. Este efecto se denomina “supremacía de la palabra” [70].

En un escenario general, tener conocimiento previo sobre los objetos a ser identificados es la base del reconocimiento, así como para seleccionar la interpretación más probable del entorno. Este conocimiento previamente aprendido facilita la tarea de completar información ausente en los canales sensitivos de entrada, extendiendo bordes, líneas o normalizando colores a fin de hacer los elementos percibidos como coherentes respecto a los prototipos mentales. Esta interacción existente entre la entrada real percibida y el conocimiento almacenado es la base de la percepción. Además, el contexto que rodea a un objeto o grupo puede influenciar positivamente la eficiencia del reconocimiento si el objeto se comporta de una forma diferente (por ejemplo, un pez nadando fuera de un banco de peces), o puede alterar los objetos percibidos (por ejemplo, los “monstruos” que perciben algunos niños en base a las sombras en oscuridad relativa). Igualmente juega un papel fundamental al

reconocer objetos similares en un contexto común uniforme (por ejemplo, en los libros “¿Dónde está Wally?”).

2.2.2.2. Combinación del procesamiento “bottom-up” y “top-down”. Reconocimiento visual humano

El reconocimiento visual es el proceso de emparejar elementos reconocidos anteriormente (memorizados) con representaciones de entradas sensoriales organizadas [71]. En la evolución esto es determinante, como es el caso de la capacidad de reacción, que es fuertemente dependiente de la capacidad de encontrar eficientemente el emparejamiento entre la entrada sensorial percibida por los ojos y entradas sensoriales previamente percibidas y memorizadas, así como las experiencias asociadas a las mismas. Por lo general, si una persona entra en una habitación y es familiar con los contenidos de la misma, puede reconocer fácilmente sillas, mesas, cuadros y demás. Las capacidades de reconocimiento dependen fuertemente de la combinación de los sentidos que el individuo puede utilizar, pero no están ligadas a ningún órgano sensorial específico. Por ejemplo, una persona ciega reconocerá los objetos usando, por lo general, una combinación del tacto, el olfato y el oído. Sin embargo, la capacidad de reconocimiento parece no estar relacionada con los órganos sensoriales (ojos, nariz, orejas, etcétera), como fue estudiado en el caso de una persona afectada por agnosia visual [72]. Esta persona era capaz de reconocer características y grupos de objetos, tales como formas, y era capaz de copiar con fluidez dibujos o reproducir imágenes desde su memoria, pero era totalmente incapaz de obtener una representación visual de los objetos en su memoria y emparejarla con la información visual organizada a la que tenía acceso. Este es un hecho importante, ya que demuestra que existen algunos procesos visuales de reconocimiento que no son utilizados ni para recuperar ni para organizar las características visuales del entorno.

El reconocimiento visual parece una tarea trivial, pero está muy lejos de ser algo sencillo. En el área de la investigación parece existir una sinergia entre los modelos de reconocimiento por ordenado y el conocimiento sobre el funcionamiento de la visión humana. Esto es así gracias a los avances surgidos en las últimas tres décadas. Afrontar el problema desde una perspectiva matemática y computacional ha revelado algunos retos inherentes al reconocimiento. Uno de dichos retos es la “dependencia del punto de vista”. Esto es, un objeto tridimensional puede ser observado desde un conjunto ilimitado de perspectivas variando la distancia hacia el objeto y el ángulo de visión. Cada perspectiva está representada por una proyección planar bidimensional, la cual varía en tamaño, posición y apariencia. No es lo mismo observar una mesa desde arriba que desde la parte frontal. Ni que decir tiene que para seres humanos adultos sanos (sin ningún trastorno perceptual o mental) es muy sencillo, o casi instantáneo, reconocer el objeto como una mesa, incluso si la entrada sensorial no proporciona suficiente información. Otro reto es el de la “variación del elemento”. Esto es, para cada categoría de objetos, existe un enorme conjunto de casos particulares. Utilizando el ejemplo anterior, no hay un único tipo de mesa. Existen mesas de té, mesillas de noche, altares, mesas modernistas con diseños poco convencionales,

etcétera. Y como ocurría anteriormente, los seres humanos con experiencia previa en sentir e interactuar con mesas, pueden identificarlas de forma sencilla incluso siendo significativamente diferentes, bajo la categoría común de “mesa”. Este reto implica el procesamiento de una vasta cantidad de información desde la entrada sensorial. A fin de abordar los distintos retos inherentes al reconocimiento visual, varias categorías principales de modelos han surgido con el tiempo:

- Modelos de coincidencia de características
- Modelos de coincidencia con plantillas
- Modelos de reconocimiento basados en componentes
- Modelos de configuración

Cada uno de los modelos tiene una serie de ventajas e inconvenientes que los hacen aplicables para el reconocimiento de tipos específicos de objetos. Algo similar ocurre con el ser humano, el cual puede utilizar múltiples conjuntos de representaciones y procesos, los cuales son más o menos efectivos dependiendo de los diferentes tipos de objetos y entornos a reconocer [71].

2.2.2.3. Modelos de coincidencia de características

Los modelos de coincidencia de características buscan por características simples pero únicas en los objetos, y posibles emparejamientos entre ellos. Como ejemplo, es posible identificar un pájaro considerando que tiene dos alas, un pico, un tamaño determinado que puede variar en un rango, dos patas y está recubierto de plumas, incluso aunque no se sepa con exactitud dónde están situadas dichas plumas, ni las dimensiones exactas de cada parte del cuerpo. Las características a ser encontradas dependen del tipo de objeto. Durante una primera fase de análisis visual, los bordes y colores son detectados. Esas son características por sí mismas, y son usadas en algunos modelos, como el reconocimiento óptico de caracteres, donde las letras impresas pueden ser consideradas como un conjunto limitado de segmentos con diferentes orientaciones y grados de curvatura [73].

Este tipo de modelos funciona bien siempre que el objeto a reconocer presente las características identificables, o si tiene más características que puedan ser consideradas en otra configuración. Por otra parte, el repositorio para almacenar este tipo de modelos puede ser relativamente pequeño, ya que un sucinto conjunto de características distintivas pueden ser utilizadas para reconocer un enorme número de objetos distintos pertenecientes a la misma clase, incluso si no son idénticos (considérese el ejemplo anterior acerca del reconocimiento de aves).

Los modelos de coincidencia de características, al igual que otros modelos, pueden ser paralelizados y distribuidos, de forma similar al comportamiento del cerebro y cómo la información es procesada en el mismo.

Existen diversos estudios que demuestran que las redes neuronales humanas proporcionan respuestas selectivas ante elementos de entrada, de manera similar al funcionamiento del modelo de coincidencia de características. En [74], [75] se estudia

la respuesta de las neuronas del córtex visual a segmentos con determinada inclinación y curvatura. En [76], [77] se estudia la respuesta selectiva a otras características, tales como color, tamaño, textura y forma. En [78] se demuestra que existen neuronas que desarrollan respuestas ante partes específicas del objeto a ser reconocido, como los ojos en una cara, y pueden ser más selectivas ante ciertas características basadas en la experiencia.. Más aún, las especies que pueden clasificar objetos como miembros de distintas categorías tienen grupos neuronales que aumentan su selectividad a fin de activarse ante la detección de determinadas características, lo cual ayuda a identificar mejor dichas categorías [79]. Esta sensibilidad (respuesta) selectiva de grupos neurales sugiere que las características fundamentales que hacen posible el reconocimiento podrían variar dependiendo del nivel de detalle requerido en cada momento.

Los modelos de coincidencia de características parecen funcionar mejor al reconocer categorías que entidades individuales [80]. Los modelos de este tipo propuestos inicialmente no eran capaces de reconocer objetos compuestos por las mismas características distribuidas en diferentes relaciones espaciales, pero evolucionaron hacia modelos que tienen en cuenta las relaciones espaciales (intercesiones, uniones y demás), haciéndolos más precisos y flexibles [81].

2.2.2.4. Modelos de coincidencia con plantillas

En los modelos de coincidencia con plantillas, una plantilla se compara con la entrada sensorial a fin de comprobar su correspondencia. Esto puede ser utilizado para comparar elementos individuales con un prototipo. En la versión inicial de dichos modelos, la correspondencia con la plantilla era directa, por lo que los objetos a ser comparados tenían que ser casi idénticos y diferentes a otros. La falta de flexibilidad en la plantilla hace casi imposible reconocer el mismo objeto si ha experimentado alguna variación, como, por ejemplo, en tamaño o posición. Esto puede ser contraproducente, ya que tal tipo de variaciones ocurren continuamente en la experiencia sensorial. En los modelos actuales de coincidencia con platilla, la imagen de entrada es transformada acorde a distintos filtros tales como rotación y escala, y comparan los resultados con la plantilla. Estos modelos son más flexibles que los originales, y pueden ser útiles en distintas áreas, tales como en seguridad (para reconocer huellas dactilares [82], [83]), en mercado (para leer códigos de barras [84]). Cuando el objeto a ser reconocido está bien definido o es único, el modelo de coincidencia con plantilla puede ser rápido y muy preciso.

El problema principal de dicho tipo de modelos consiste en obtener suficiente flexibilidad como para tener buen rendimiento y precisión. Considerando el caso de reconocer texto escrito, existen una inmensa cantidad de formatos de fuente, o tipografías, que en el caso del ámbito digital pueden ser utilizadas fácilmente como plantillas para reconocer letras y palabras, pero en el caso de la escritura humana, los textos manuscritos varían de persona a persona, o incluso en la misma persona dependiendo de sus circunstancias. En este caso, para conseguir suficiente flexibilidad, es posible generar plantillas desde una jerarquía de componentes de la

plantilla, de forma que cada una detecte una parte relevante del modelo. Algunos sistemas de reconocimiento biométrico [85] se aprovechan de este nuevo tipo de modelos más flexibles.

2.2.2.5. Modelos de reconocimiento basados en componentes

Los modelos de reconocimiento basados en componentes abordan el desafío del reconocimiento de objetos tridimensionales, observados desde diferentes puntos de vista. Este problema es realmente complejo para otros modelos (características y coincidencia de plantillas), ya que podría requerir el representar y almacenar un gigantesco número de características y relaciones, o plantillas ajustadas para cada variación de ángulo y posición.

Los estudios sobre el tema sugieren que el cerebro puede descomponer objetos en sus partes funcionales y en relaciones espaciales entre dichas partes [86], [87]. Los modelos de reconocimiento basados en componentes se basan en dicha premisa [88], utilizando un conjunto geométrico básico de elementos tridimensionales, tales como cubos o cilindros, para aproximar casi cada objeto mediante la combinación de elementos del conjunto. Estas formas básicas se denominan *geones* [88]. Estos modelos también consideran las relaciones geométricas, de pertinencia y de adyacencia, tales como “pegado a”, “encima de”, “dentro de” y demás.

Los modelos de reconocimiento basados en componentes tratan de detectar los geones y las relaciones espaciales, y las comparan con representaciones tridimensionales de objetos previamente analizados o conocidos [89].

La utilidad básica de los geones deriva de que sus propiedades son invariantes a la perspectiva. Esto es, se mantienen en la imagen independientemente del punto de vista de la escena. Las propiedades invariantes son líneas rectas, esquinas y vértices. Por ejemplo, los bordes de un cuadrado proyectarán una línea recta en cada plano bidimensional, independientemente de la perspectiva. Cada geón está asociado a un conjunto de propiedades invariantes a la perspectiva, lo cual los definen de forma inequívoca y los hacen únicos respecto a otros geones. De este modo, cada definición estructural de un objeto puede ser independiente de la perspectiva incluso cuando la forma percibida del objeto cambia dependiendo de las circunstancias de observación.

Los sujetos humanos de prueba probaron que pueden reconocer de manera sencilla los geones que componen objetos manufacturados, más aún cuando experimentan el fenómeno de *priming visual* [90]. Esto es, un reconocimiento más rápido del objeto tras haber visto el mismo objeto (o similar) por segunda vez. El *priming* es un efecto de la memoria implícita que ocurre cuando un estímulo o tarea facilita el procesamiento de una tarea o estímulo posterior. Irving Bierdman [91] desarrolló varias pruebas creando conjuntos de imágenes que representaban objetos con partes faltantes y los mismos objetos dibujados desde un punto de vista diferente, y mostrando pares de imágenes del mismo objeto a los sujetos de pruebas. Los experimentos concluyeron que el reconocimiento de los objetos era más rápido cuando la segunda imagen estaba compuesta con el mismo conjunto de geones que la primera imagen. De forma

análoga, otros experimentos han probado que algunos objetos sólo pueden ser reconocidos desde un número limitado de perspectivas, como las cabezas, cuya representación frontal difiere mucho de la representación lateral [92], [93] en términos de composición y distribución de geones. Esto es por lo que los modelos de reconocimiento basados en componentes funcionan mucho mejor con elementos manufacturados que con objetos naturales, o incluso con elementos particulares de una clase.

2.2.2.6. Modelos de configuración

Los modelos de reconocimiento basados en la configuración abordan las limitaciones de los modelos basados en componentes al representar los objetos como una composición de sub-objetos, y una estructura común que almacena las diferentes relaciones entre los sub-objetos y la desviación espacial respecto a un objeto prototipo previamente definido. Este tipo de modelos ayuda a reconocer individuos específicos pertenecientes a la misma clase de objetos, tales como caras [94], [95]. En un modelo de configuración, una cara está descrita dependiendo en su desviación de una cara prototipo estándar, definida por la cara media cuantificada en una población. Todas las caras tendrán los mismos atributos localizados en una distribución espacial similar, pero las proporciones y la posición específica de los sub-objetos (ojos, nariz, boca y demás) hacen que cada cara sea única. Los estudios en neurociencia apoyan el reconocimiento de caras mediante modelos de configuración. Existen registros de neuronas individuales en el lóbulo temporal de los primates que actúan selectivamente ante la visión de diferentes partes de caras [96]. En los seres humanos, el daño del giro fusiforme, una parte específica del lóbulo frontal, produce prosopagnosia, esto es, la incapacidad de reconocer diferentes tipos de caras. Los seres humanos con prosopagnosia pueden reconocer fácilmente la categoría "cara", pero no pueden reconocer caras específicas de la categoría global. Una variante de esta teoría es la hipótesis de la experiencia, que afirma que existe un sistema neural experto desarrollado para producir una discriminación visual selectiva, y así identificar diferencias sutiles dentro de una categoría visual específica [97]. Ambos enfoques pueden ser válidos, ya que algunas de las funciones identificadas del giro fusiforme son el reconocimiento de la cara y el cuerpo, el reconocimiento de pertenencia a categorías, el procesamiento de la información de color y el reconocimiento de palabras.

2.2.3. Sentidos. El Oído

El sistema auditivo en el ser humano, al igual que otros sentidos, es un sistema complejo e interrelacionado con ellos. El sistema auditivo está basado en los estímulos percibidos por el oído. El oído humano convierte las ondas sonoras en señales eléctricas que se transmiten por el nervio acústico hasta el cerebro, donde dichas señales son interpretadas.

El oído puede dividirse en tres partes bien diferenciadas: el oído externo, el medio y el interno (ver Ilustración 7). El oído externo está compuesto por el pabellón auditivo, o, dicho de otra forma, la oreja, el conducto auditivo y el tímpano. Las ondas sonoras se recogen por el pabellón, que las propaga a través del conducto auditivo hasta la membrana del tímpano. A efectos comparativos, el oído externo es el equivalente a un micrófono de alta precisión, capaz de detectar variaciones de presión en el aire con una sensibilidad de menos de una mil millonésima parte de la presión atmosférica, y en un rango de frecuencias comprendido aproximadamente entre los 20 Hz y los 20 kHz [98]. Igualmente, el rango dinámico de la audición, o rango de estimulación es el más amplio de los sentidos. Dicho rango va comprendido entre el umbral de audición (cero decibelios) hasta el umbral del dolor (130 decibelios) [99]. Dicho rango es mejorado por la estructura de amplificación efectiva que amplía el extremo inferior del rango y por el mecanismo protector de la membrana timpánica, que amplía su extremo superior, reduciendo la transferencia de fuerza a la ventana oval del oído interno.

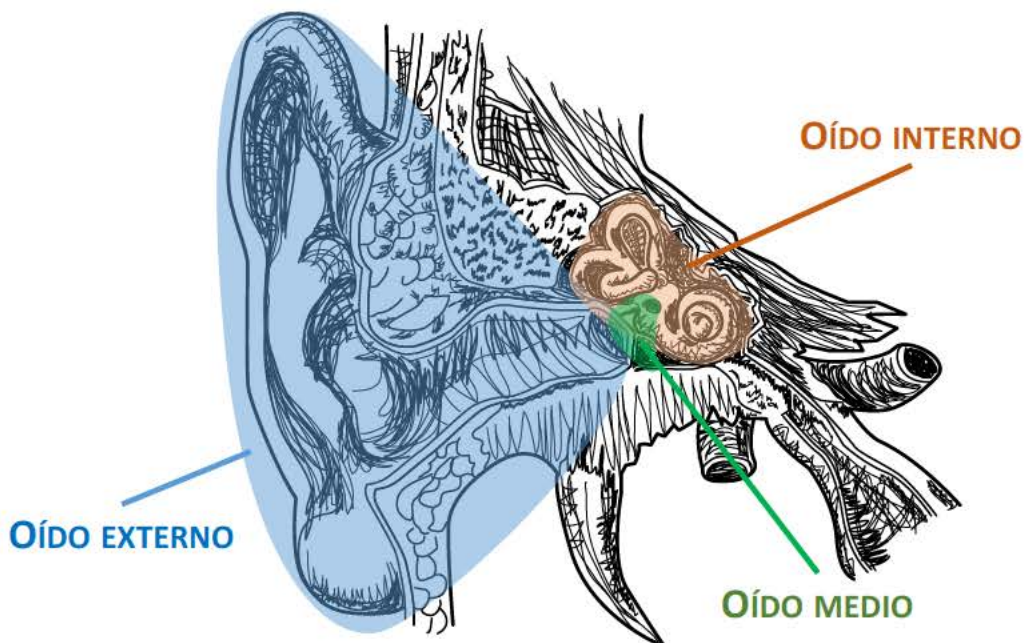


Ilustración 7. Estructura anatómica del oído humano

El oído medio es una cavidad comprendida entre el tímpano y la base de la cóclea. En su interior hay tres pequeños huesos, denominados martillo, yunque y estribo. La cabeza del martillo se apoya sobre el tímpano, transmitiendo las vibraciones a través del yunque, al estribo. Este último se apoya en una de las dos membranas que cierran la cóclea, denominada ventana oval. Por así decirlo, el oído medio se comporta como un pre-amplificador y procesador del sonido. Gracias a él, el sentido auditivo tiene la inmensa sensibilidad y rango dinámico comentados anteriormente.

Finalmente, el oído interno es una cavidad hermética anegada por un líquido denominado linfa. Está compuesta por los canales semicirculares, el vestíbulo y la cóclea. Los canales semicirculares no tienen relación directa con la audición, pero sí

con el equilibrio [100]. Las vibraciones de la ventana oval del vestíbulo son transformadas en la cóclea. Las señales de la cóclea son codificadas y transformadas en impulsos electroquímicos que se propagan por el nervio acústico hasta llegar al cerebro. Entrando en detalles, el vestíbulo contiene la ventana oval y redonda (membranas). La ventana oval está unida al estribo, y recibe del mismo las vibraciones. La cóclea está dividida longitudinalmente por la membrana basilar, sobre la que se asientan los filamentos terminales del nervio auditivo. Cuando la ventana oval es empujada por el estribo, una sobrepresión es producida en la parte superior de la cóclea, que obliga a circular el fluido linfático hasta la cavidad inferior a través del helicotrema. En el proceso, la membrana basilar se deforma hacia abajo, y la membrana elástica que cierra la ventana redonda cede hacia afuera. Cuando el estribo se mueve hacia la izquierda y derecha, aumentando y disminuyendo la presión del líquido contenido encima de la membrana basilar, aparece una onda que se desplaza lateralmente (de izquierda a derecha) a lo largo de la membrana. Dicha onda puede visualizarse como un movimiento de traslación vertical (de arriba hacia abajo) en la membrana. Su velocidad de avance depende de la frecuencia y de las características de la membrana basilar. En algún punto de la cóclea, la velocidad de avance es cero. Cerca de ese punto, la oscilación de la membrana es más acentuada, absorbiendo la energía de la onda. Cada punto de la membrana basilar es estimulado y responde así por una determinada frecuencia. Cuando el oído recibe un sonido compuesto por diferentes frecuencias, cada una de ellas excita un punto diferente de la membrana basilar, de modo que el cerebro puede interpretar además de la altura del sonido su timbre, requiriendo únicamente conocer qué terminaciones nerviosas fueron excitadas con qué nivel de intensidad. A efectos comparativos, el oído interno se comportaría como un analizador de frecuencias.

Los estímulos auditivos son enviados al cerebro, donde son interpretados, generando la sensación auditiva. Dicha sensación se compone de una serie de atributos, como son el tono, la localización subjetiva y la intensidad (o volumen percibido), los cuales pueden ser relacionados de una forma directa con las propiedades físicas del estímulo. Existe un complejo sistema neural subyacente encargado de identificar dichos atributos, aunque cabe destacar que habitualmente dichos estímulos no son percibidos en un medio aislado. Más bien, el entorno auditivo es analizado como fuentes discretas de sonido u objetos audibles, cada uno de los cuales tiene su propio tono, timbre, posición e intensidad. A veces las fuentes de sonido pueden ser reconocidas como elementos familiares, como es el caso de una persona específica hablando. A menudo el objeto percibido puede ser identificado, como, por ejemplo, una palabra hablada o un instrumento musical reproduciendo una pieza específica. En este apartado se discutirán cuatro aspectos relevantes de los objetos audibles y de la percepción de patrones teniendo en cuenta los estudios analizados en el estado del arte. Los aspectos a tratar son: los factores involucrados en la identificación de un objeto particular entre un conjunto de objetos, las reglas que gobiernan la organización perceptual del entorno auditivo, la percepción de secuencias de sonidos y los elementos utilizados para analizar un conjunto complejo de sonidos como fuentes discretas.

2.2.3.1. Identificación de los objetos. Percepción del timbre.

Muchos atributos de la sensación acústica, como el tono y la intensidad pueden ser considerados unidimensionales. Si analizamos una amplia variedad de sonidos con diferentes tonos, es posible ordenarlos en una escala de comprendida entre graves y agudos [101]. De manera similar, sonidos con diferente intensidad pueden ser ordenados con otra escala, comprendida entre más débil y más intenso. La capacidad humana de identificar un objeto entre un gran conjunto de objetos depende de que existan distintas dimensiones entre las que los objetos a identificar presenten diferentes características. Cuando un conjunto de estímulos varía en una única dimensión sólo es posible identificar, o asociar una etiqueta verbal, un máximo de cinco a seis elementos diferentes [102]. Esto es cierto siempre que los tonos sean separados en un amplio rango de frecuencias (distintas octavas) o si los tonos están concentrados en un rango relativamente estrecho, como por ejemplo en la misma octava.

A fin de identificar más estímulos, se requieren dimensiones adicionales. En el oído, las dimensiones adicionales surgen principalmente por dos vías. En primer lugar, para el caso de estímulos complejos, es importante establecer patrones de energía en base a la frecuencia. En segundo lugar, los estímulos auditivos suelen variar con el tiempo, con lo que el establecimiento de patrones temporales resulta crucial para la percepción.

En caso de percibir una onda sinusoidal estacionaria, entonces el patrón de sonido puede ser representado únicamente por dos números, especificando su frecuencia e intensidad. Sin embargo, casi todos los sonidos que el ser humano percibe en su vida cotidiana son mucho más complejos que una simple onda sinusoidal, conteniendo una multitud de frecuencias con distintos niveles y fases relativas. La distribución de energía sobre la frecuencia es uno de los mayores determinantes de la calidad del sonido, o dicho de otra forma, su timbre. El timbre se puede definir como “*el atributo de la sensación auditiva que permite a un oyente juzgar que dos sonidos no idénticos, presentados de manera similar y teniendo la misma intensidad y tono, son diferentes*” [103]. Las diferencias en el timbre permiten, por ejemplo, distinguir la misma nota tocada por un violín, un piano o una flauta.

El timbre depende de más factores que de la magnitud del espectro del sonido. Las fluctuaciones a lo largo del tiempo juegan un papel importante. Al contrario que el tono o la intensidad, consideradas variables unidimensionales, el timbre es multidimensional. No existe una escala única sobre las que se puedan comparar u ordenar los timbres de sonidos diferentes. Por ello, es necesario definir un método para describir el espectro de un sonido, teniendo en cuenta este aspecto multidimensional relacionado con el timbre subjetivo. Una primera aproximación es considerar la distribución espectral de la energía. Por ejemplo, los tonos complejos con armónicos inferiores fuertes (por debajo de la sexta) suenan melosos, mientras que los tonos con armónicos fuertes por encima de la sexta o séptima suenan agudos y penetrantes. Sin embargo, un enfoque cuantitativo mucho más completo fue descrito en [104]–[107], donde se identifica que las diferencias perceptuales entre sonidos distintos, tales como las vocales o tonos estacionarios producidos por instrumentos

musicales, están estrechamente relacionadas con las diferencias espectrales de los sonidos, cuando los espectros son especificados como los niveles en bandas de frecuencia de 18 octavas y un tercio. Un ancho de banda de un tercio de octava es sensiblemente mayor que el ancho de banda crítico, esto es, por encima del rango de frecuencia audible. Así, el timbre está relacionado con el nivel relativo producido por un sonido en cada banda crítica. Dicho de otra forma, el timbre de un sonido está relacionado con el patrón de excitación de dicho sonido. Aparentemente, el número de dimensiones necesarias para caracterizar el timbre está limitado por el número de bandas críticas requeridas para cubrir el rango de frecuencias audibles. Obviamente en un sistema computacional, dicha limitación estará impuesta por el rango de frecuencia a detectar por el dispositivo de captura. En el ser humano, el número es aproximadamente de 37 dimensiones. Para cada clase o categoría restringida de sonidos, sin embargo, se requiere un número muy inferior de dimensiones. Por lo general, tanto para sonidos relacionados con el habla como para los que no lo están, los timbres de tonos fijos son determinados principalmente por la magnitud espectral, aunque las fases relativas de sus componentes también tienen un papel relevante en el proceso [108]–[110].

Aunque las diferencias en timbres estáticos permitan distinguir dos sonidos que son presentados sucesivamente, no son siempre suficientes como para permitir la identificación absoluta de un objeto audible, como por ejemplo un instrumento musical específico. Una razón para esto es que la magnitud y fase (espectro) del sonido podría ser alterado considerablemente por el camino de propagación y por las reflexiones del entorno. En la práctica, el reconocimiento de un timbre específico, y por ende, el reconocimiento de un objeto audible podría depender de diversos factores. En [111] se sugieren algunos de ellos, entre los que cabe destacar:

1. *Periodicidad*, cuando el sonido presenta una calidad tonal para tasas de representación entre 20 y 20000/s, o *irregularidad*, cuando presenta un comportamiento de carácter ruidoso.
2. Si la forma de la onda es *constante* o *fluctuante* en función del tiempo, y en el segundo caso, cómo son las fluctuaciones.
3. Si cualquier otro aspecto del sonido *cambia con el tiempo*, como por ejemplo el espectro o la periodicidad.
4. Cómo son los sonidos *antecedente* y *precedente*.

El reconocimiento de instrumentos musicales, por ejemplo, tiene una fuerte dependencia con las transiciones iniciales y con la estructura temporal del sonido [112]. El tono característico de un piano depende del hecho de que las notas tienen un inicio rápido y un deterioro gradual. Si a una grabación de un piano se le aplica una inversión temporal, el timbre es completamente diferente. De hecho, asemejándose al armónico de un acordeón, a pesar del hecho de que el espectro de magnitud a largo plazo permanezca invariante a la inversión temporal. La percepción de los sonidos con envolvente asimétrica ha sido estudiada en [113], [114] por Patterson utilizando portadoras sinusoidales moduladas en amplitud mediante la repetición de una función exponencial. En ellas, la envolvente se aumentaba abruptamente y se decaía gradualmente (sonido amortiguado), o se incrementaba gradualmente y se decaía abruptamente (sonido aumentado). Los sonidos aumentados eran versiones invertidas

en el tiempo de los sonidos amortiguados y tenían el mismo espectro de magnitud a largo plazo. Los sonidos estaban caracterizados por el periodo de repetición de la envolvente, y por su vida media. Para sinusoidales amortiguadas, la vida media es el tiempo para el que la amplitud decrece en un factor de dos. Patterson identificó que, para la vida media en el rango de los 2 a los 32 ms, los sonidos amortiguados y aumentados presentaban diferentes cualidades. Para una vida media de 4 ms, los sonidos amortiguados eran percibidos como una única fuente, similar a un redoble de tambor en una superficie hueca resonante. Los sonidos aumentados eran percibidos como dos sonidos distintos: un tono continuo correspondiente a la frecuencia de la portadora, y un redoble de tambor en una superficie no resonante. En [115] Akeroyd y Patterson utilizaron sonidos con envolventes similares, pero en este caso, la portadora era ruido de banda ancha en lugar de una sinusoidal. Informaron de que los sonidos atenuados eran percibidos como un tambor golpeado por cepillos de alambre. No mostraba ninguna cualidad similar a un silbido. En contraste, los sonidos aumentados se escuchaban como un ruido, con características similares a un silbido bruscamente cortado con el tiempo. Dichos experimentos mostraron la importancia de la envolvente temporal en la percepción del timbre.

Muchos instrumentos presentan características similares al ruido que influyen fuertemente su calidad subjetiva. Una flauta, por ejemplo, tiene una estructura harmónica relativamente simple, pero los tonos sintetizados con la misma estructura harmónica no suenan como una flauta a menos que cada nota sea precedida por un pequeño de ruido similar a un soplo. En general, los tonos de los instrumentos musicales estándar son simulados de una forma muy pobre mediante la suma de frecuencias estables, ya que tales síntesis no pueden producir la variación dinámica con las características temporales de dichos instrumentos. Los sintetizadores actuales modelan las envolventes de los sonidos que producen y por tanto son capaces de generar imitaciones acústicas más precisas y convincentes de los instrumentos que simulan. Para que una simulación resulte completamente convincente, a veces es necesario dar diferentes envolventes a diferentes armónicos, formando un sonido complejo [112], [116].

2.2.3.2. Información utilizada para la identificación de distintos objetos audibles

Es importante hacer una distinción entre dos conceptos, fuente y flujo [117], [118]. Una fuente es una entidad física que da lugar a ondas de presión acústicas, por ejemplo, un piano siendo tocado. Por otra parte, un flujo es la percepción de un grupo sucesivo y/o simultáneo de elementos sónicos formando un todo coherente, y pareciendo emanar de una misma fuente. Por ejemplo, la percepción de la audición de un piano siendo tocado.

Por lo general los sonidos percibidos por el ser humano son originados por diversas fuentes. Si se comparan la forma de onda de las muestras de sonidos percibidos individualmente con la forma de onda de la composición de las mismas (sonido escuchado), es muy difícil apreciar alguna de las muestras individuales en la onda

resultante. Aunque cuando se escuchan los sonidos conjuntamente en el entorno, es fácil percibir cada uno como un flujo separado.

El sistema auditivo periférico actúa como un analizador de frecuencias, separando los componentes con distinta frecuencia en sonidos complejos. En el cerebro, las representaciones de dichas componentes de frecuencia deben ser asignadas a sus fuentes correspondientes. Si la entrada proviene de dos fuentes distintas A y B, entonces los componentes de frecuencia deben ser divididos en dos grupos: los componentes originados en la fuente A deberán ser asignados a un flujo y los componentes emanados de la fuente B deberán ser asignados a otro. El proceso encargado de dicha discriminación se llama “agrupación perceptual”. También recibe el nombre “*parsing*” [118] o “análisis auditivo de la escena” [119]. El proceso de separar los elementos derivados de dos fuentes diferentes se denomina “segregación”. Muchos tipos distintos de características físicas pueden utilizarse para derivar flujos perceptuales separados, correspondientes a fuentes individuales, lo que da lugar a una entrada acústica compleja. Existen dos aspectos relacionados con este proceso [118]: el agrupamiento conjunto de todas las componentes de frecuencia que emanan simultáneamente de una fuente única en un momento determinado, y la conexión temporal de las frecuencias que cambian en una fuente concreta entre un instante y el siguiente. Dichos aspectos se denominan respectivamente como “agrupamiento simultáneo” y “agrupamiento secuencial”. Cabe destacar que el proceso de agrupamiento perceptual auditivo [120] no es un proceso absoluto, o de todo o nada. Muchos experimentos al respecto se basan en el efecto de un atributo específico de los sonidos, como su tono, su timbre, su posición subjetiva, en el agrupamiento. Dichos experimentos han mostrado que una característica que puede ser efectiva para un atributo podría ser menos efectiva o totalmente inefectiva para otro. Igualmente, la efectividad de las características puede variar para el agrupamiento secuencial y simultáneo.

2.2.3.3. Frecuencia fundamental y regularidad espectral

Al escuchar dos tonos complejos fijos de manera conjunta (como por ejemplo el sonido de dos vocales o el sonido de dos instrumentos musicales), los seres humanos en general no confunden los armónicos pertenecientes a cada tono. Por el contrario, suelen escuchar cada tono como una fuente separada, incluso aunque los armónicos estén intercalados e incluso puedan coincidir. Esto sólo puede realizarse si los dos tonos tienen frecuencias fundamentales (F_0) diferentes. En un experimento llevado a cabo por Broadbent y Ladefoged [121] se utilizaron sonidos sintetizados correspondientes a diferentes vocales, observando que sólo se percibían como una voz fusionada cuando todos los armónicos contenían la misma F_0 . Si los armónicos se dividen en dos grupos con distintas F_0 , entonces la percepción identifica dos sonidos diferentes.

La frecuencia fundamental es importante por diferentes motivos. Los componentes de un sonido periódico tienen frecuencias que forman series armónicas simples. Las frecuencias son múltiplos de la frecuencia fundamental. Esta propiedad se denomina

“armonicidad”. Los armónicos inferiores se determinan en el sistema auditivo periférico. El espaciado regular de los armónicos inferiores puede derivar en su fusión perceptual, haciendo que sean percibidos como un único sonido. Si un componente sinusoidal no forma parte de esta serie de armónicos, tiende a ser escuchado como un sonido diferente.

En [122], [123], Culling y Darwin sugieren que hay otro mecanismo aparte del basado en la regularidad espectral o en la armonicidad, que juega un papel importante en la segregación perceptual de los armónicos inferiores. Mostraron que, para pequeñas diferencias de la frecuencia fundamental entre dos vocales, sólo las frecuencias más bajas contribuyen a aumentar la precisión de la identificación de la vocal con un incremento de la diferencia en la frecuencia fundamental. Sugirieron que la mejora ocurría por las interacciones temporales entre los armónicos inferiores en las dos vocales. Dichas interacciones rítmicas, tienen un efecto en la respuesta neuronal, siendo dominante en primer lugar una vocal y alternativamente la otra. Parece ser que el sistema auditivo es capaz de escuchar selectivamente en el tiempo para extraer una representación de cada vocal. Para los armónicos superiores, sin embargo, la frecuencia fundamental juega un papel diferente. Los armónicos superiores de un sonido complejo periódico no se resuelven en la membrana basilar, sino que dan lugar a una onda compleja con una periodicidad igual a la de la frecuencia fundamental. Cuando dos sonidos complejos con diferentes frecuencias fundamentales son percibidos a la vez, cada uno formará ondas en la membrana basilar con una periodicidad igual a la de su frecuencia fundamental. Si los dos sonidos tienen diferente espectro, entonces cada uno generará la respuesta en ciertos puntos de la membrana basilar. El sistema auditivo podría agrupar las regiones de forma conjunta con una frecuencia fundamental común y segregarlas de las regiones con una frecuencia fundamental diferente [124]. También podría darse el caso de que tanto las componentes resueltas como las no resueltas sean agrupadas en base al patrón detallado temporal de los picos neuronales [125]. Este proceso se puede explicar de manera cualitativa asumiendo que el tono de un sonido complejo es el resultado de la correlación o comparación de los intervalos de tiempo entre picos de neuronas ante diferentes frecuencias características. Sólo aquellos canales que muestren una alta correlación podrían ser clasificados como pertenecientes al mismo sonido. Tal mecanismo podría agrupar de forma conjunta componentes con la misma frecuencia fundamental. Sin embargo, en [126] se muestra una evidencia en contra de dicho mecanismo. En el estudio se muestra que la identificación de una vocal específica en presencia de otra vocal sonando de fondo era mejor identificada cuando la vocal de fondo era armónica que cuando no lo era. En contraste, la identificación de la vocal objetivo no depende de si es o no armónica. En el estudio completo, los autores proponen un mecanismo basándose en la idea de que un sonido de fondo armónico puede ser cancelado por el sistema auditivo, mejorando con ello la representación de la vocal objetivo. En cualquier caso, debería recalarse que la habilidad del sistema auditivo para segregar un sonido dado, basándose en la información de su frecuencia fundamental, en una región espectral específica, es limitada, como demuestra el fenómeno de la interferencia de la discriminación de tono [127].

2.2.3.4. Las disparidades de inicio

El ser humano tiene la habilidad de escuchar un tono complejo en presencia de otro [128]. Rasch llevó a cabo un experimento en el que uno de los tonos era tratado como una máscara y el nivel de la señal del tono (con mayor frecuencia fundamental) fue ajustado para encontrar el punto en el que era detectable. Cuando los dos tonos eran iniciados a la vez y presentaban la misma envolvente temporal, el umbral de la señal se localizaba entre los cero y los -20dB relativos al nivel de la máscara.

Aunque los sujetos de prueba percibieron que la señal continuaba a través de la máscara, Rasch mostró que esta percepción no estaba basada en la información sensorial recibida durante el tiempo de reproducción de la máscara. Encontró que umbrales idénticos eran obtenidos si la señal se apagaba justamente después de iniciar la máscara. Aparentemente el sistema perceptual asume que la señal continúa, ya que no hay evidencia de lo contrario. La parte de la señal que ocurre simultáneamente con la máscara podría estar completamente enmascarada. Este efecto está relacionada con el fenómeno de la continuidad [129].

Rasch mostró que, si dos tonos son iniciados a la vez, pero tienen distintos tiempos de incremento, esto puede resultar en umbrales muy bajos para la señal que tiene el tiempo más corto de incremento (la que antes se incrementa). Bajo dichas condiciones y con asincronías de inicio de hasta 30 ms, las notas suenan como si comenzaran síncronamente. Por tanto, el ser humano no necesita ser consciente de las diferencias de inicio para que el sistema auditivo pueda realizar correctamente, y a nivel perceptual, la separación de tonos complejos. Rasch también indicó que, en la música de orquesta, los distintos músicos no tocan exactamente en sincronía incluso si la partitura lo indica así. Las diferencias de inicio utilizadas en sus experimentos se corresponden aproximadamente con las asincronías percibidas como simultáneas en las notas percibidas por la música de orquesta. Esto soporta la visión de que las asincronías componen un factor importante en la percepción de las distintas partes o voces de la música polifónica.

Finalmente, las asincronías de inicio juegan un papel importante a la hora de determinar el timbre de los sonidos complejos. En un estudio conducido por Darwin y Sutherland [130] se concluyó que un tono que comienza o finaliza en un momento distinto del de una vocal es menos propenso a ser oído como parte de la vocal que si es simultáneo con ella. Por ejemplo, incrementando el nivel de un único armónico puede producir un cambio significativo en la calidad, o timbre, de una vocal. Sin embargo, si el armónico incrementado comienza antes de la vocal, el cambio en la calidad de la vocal es significativamente menor.

2.2.3.5. Contraste con sonidos anteriores

El sistema auditivo puede analizar de forma extremadamente eficiente los cambios en la entrada sensorial, y particularmente los cambios en el espectro a lo largo del tiempo [131], [132]. El aspecto que ha cambiado destaca perceptualmente del resto, como se vio anteriormente en el caso de la visión. Es posible que existan mecanismos especializados para llevar a cabo la detección de cambios en el espectro. Adicionalmente, la estimulación con un sonido estacionario puede conducir a una adaptación al mismo. Es fácil comprobar este efecto, por ejemplo, en un entorno de trabajo en el que hay sonidos constantes, como el zumbido de los condensadores de aparatos electrónicos. Cuando algún aspecto de un estímulo varía, ese aspecto es liberado de los efectos de la adaptación y por su detección será incrementada. En el ejemplo anterior esto ocurre cuando se desconectan los dispositivos electrónicos que producen el sonido de fondo, con la consiguiente liberación de los efectos de la adaptación a dicho sonido y la notable sensación de “silencio” (había algo sonando que ya no suena). Aunque el mecanismo subyacente de adaptación es un tema de debate, su efecto perceptual sin duda no lo es.

Aparte del ejemplo citado anteriormente, una demostración muy potente del efecto de adaptación se puede obtener al escuchar un estímulo con una estructura espectral específica, y entonces cambiar rápidamente a un estímulo con un espectro plano, como por ejemplo el de un ruido blanco. El ruido blanco percibido en aislamiento puede ser descrito como no coloreado, esto es, no tiene tono y presenta un tipo neutral de timbre. Sin embargo, cuando un ruido blanco sigue un estímulo con estructura espectral, el ruido suena “coloreado”. La tonalidad del ruido se corresponde con la inversión del espectro correspondiente al sonido precedente. Por ejemplo, si el sonido precedente contiene un corte o una parada, el ruido blanco presentará una calidad tonal correspondiente a la frecuencia central del corte o la parada [133], percibiéndose como un ruido con un pequeño pico correspondiente al corte. También se ha observado que un tono armónico complejo con un espectro plano puede producir una calidad similar a la vocal cuando es precedida por un tono armónico complejo con un espectro correspondiente a la inversa de un sonido de voz, como podría ser una vocal [134].

Otra demostración de los efectos del cambio en un estímulo puede ser obtenido al escuchar un tono complejo fijo, compuesto de diferentes armónicos. Por lo general, ese tipo de tono es percibido con un tono fijo correspondiente a su frecuencia fundamental, y los armónicos individuales no son percibidos por separado. Sin embargo, si uno de los armónicos se modifica de alguna manera, ya sea mediante la alteración de su fase relativa o de su nivel, entonces el armónico permanece fuera de la percepción completa como un todo. Durante un corto periodo de tiempo tras el cambio, se puede percibir una calidad tonal pura. La percepción del armónico se disipa entonces de manera gradual, hasta que se vuelve a mezclar con el complejo de nuevo.

La detección del cambio es obviamente algo importante al asignar componentes de sonido con sus respectivas fuentes. Habitualmente los seres humanos escuchan con ruido de fondo, tales como diálogos de otras personas, el sonido del tráfico, del viento,

etc. Dicho ruido de fondo suele ser relativamente uniforme. Un cambio súbito en el mismo por lo general indica que una nueva fuente ha sido activada, y los mecanismos de detección del cambio permiten aislar los efectos de dicho cambio e interpretarlos adecuadamente.

2.2.3.6. Cambios correlacionados en la amplitud o en la frecuencia

Como se ha comentado anteriormente, la separación perceptual de dos tonos complejos puede ser mejorada al introducir asincronía inicial. El umbral de detección del tono con la frecuencia fundamental más alta puede ser notablemente reducido si el tono comienza 30 milisegundos antes que el otro tono. Incluso cuando dos tonos comienzan de manera síncrona, es posible mejorar la detección del tono con mayor frecuencia fundamental mediante la modulación de su frecuencia. Esta modulación es similar al *vibrato* producido habitualmente en los tonos musicales, y también se aplica a los componentes cuyo mayor tono se amplía o reduce en sincronía. Rasch encontró que la modulación podría reducir el umbral para detectar el mayor tono en 17 dB, así cuando dos tonos complejos son presentados al mismo tiempo, la detección de uno de los tonos puede ser mejorada si sus componentes son modulados en frecuencia, dejando el otro tono estacionario. Un efecto similar puede aparecer por la modulación en amplitud de uno de los tonos. La modulación mejora la prominencia del sonido modulado, destacándolo sobre el sonido no modulado.

Parece claro que la modulación de los componentes de uno de los sonidos, ya sea en amplitud o en frecuencia pueden ayudar con la segregación perceptual del sonido respecto a un sonido de fondo uniforme [135]–[139], sin embargo, no está tan claro cuando la segregación perceptual de sonidos simultáneos está afectada por la coherencia de los cambios en la amplitud o frecuencia cuando ambos sonidos están modulados. En este caso, la coherencia se refiere a cuando los cambios en amplitud o frecuencia de los dos sonidos tienen el mismo patrón o distintos patrones a lo largo del tiempo. Diversos experimentos han reportado que la coherencia en los cambios de la amplitud juega desempeña cierto papel. Los sonidos con cambios coherentes tienden a fusionarse perceptualmente, mientras que los sonidos con cambios incoherentes tienden a segregarse [140]–[142]. Sin embargo, en el estudio desarrollado por Summerfield y Culling [137] se observa que la coherencia de la modulación en amplitud no afecta la identificación de pares de vocales simultáneas cuando las vocales están compuestas por componentes colocados aleatoriamente en la frecuencia, a fin de evitar los efectos de la armonía. No existe una evidencia clara de que la coherencia de la modulación en frecuencia influya el agrupamiento perceptual cuando ambos sonidos están modulados.

La fusión de componentes que varían en amplitud de manera coherente puede afectar la percepción de un tono complejo con un único armónico. Considerando un tono complejo conteniendo el tercer, cuarto y quinto armónico de una frecuencia fundamental de 200 Hercios, esto es, 600 Hz, 800 Hz y 1000 Hz, el tono resultante puede ser percibido de dos maneras. Puede escucharse de forma analítica, a través de la audición de uno o más componentes individuales, o puede escucharse de forma

sintética, como un único tono correspondiente a la frecuencia fundamental ausente. La forma analítica de audición es más propensa a ser percibida si el tono complejo se presenta de manera continua o si dicho tono complejo es construido añadiendo cada componente en distintos momentos. Sin embargo, el modo sintético es más propenso a ser percibido en caso de que todos los componentes del tono compuesto se modulen en amplitud de la misma manera. El comportamiento coherente de los componentes origina su fundición en una percepción única y uniforme.

2.2.3.7. Posicionamiento del sonido

Las señales utilizadas en la localización pueden ayudar en el análisis de entradas auditivas complejas. Un fenómeno relacionado con esto es el de enmascaramiento de diferencia de nivel biaural (MLD, *masking level difference*) [143]. Este fenómeno ocurre siempre que la diferencia de fase o nivel de una señal en los dos oídos no sea el mismo que el de la máscara, pues la habilidad de detectar la señal mejora dependiendo de si la señal y la máscara tienen la misma relación de fase y nivel en los dos oídos. La implicación práctica es que se puede detectar mejor la señal cuando se sitúa en una posición diferente en el espacio que la máscara.

Un ejemplo del uso de señales biaurales en la separación de un objeto audible del sonido de fondo puede observarse en el experimento conducido por Kubovy et al. [144]. En dicho experimento se emplearon ocho señales sinusoidales continuas a cada oído a través de unos auriculares. Las sinusoidales tenían frecuencias correspondientes a las notas de una escala musical, siendo la más baja de 300 Hz. La entrada al oído izquierdo se producía con un retardo de 1 milisegundo respecto a la entrada del oído derecho, por lo que las sinusoidales se escuchaban a través del lado derecho de la cabeza. Entonces, la fase de cada una de las sinusoidales se incrementaba en el oído izquierdo, mientras que la fase en el oído derecho se retrasaba, hasta que la entrada del oído izquierdo precedía a la entrada del oído derecho por un milisegundo. Este proceso de cambio de fase ocurría en una ventana temporal de 45 milisegundos. La fase permanecía en el valor modificado durante una cantidad específica de tiempo, y posteriormente se revertía de manera gradual a su estado inicial, nuevamente en una ventana de 45 ms. Durante el tiempo que se producía el cambio de fase, la sinusoidal con la fase cambiada se percibía desde el lado opuesto de la cabeza (lado izquierdo), destacándola perceptualmente. Una secuencia de cambios de fase en las diferentes componentes se percibía de forma clara, como una melodía. Esta melodía era totalmente indetectable al escuchar la entrada auditiva de un oído al otro. Los autores interpretaron sus resultados indicando que las diferencias en la fase relativa a los dos oídos podrían permitir aislar a un objeto audible ante la ausencia de otras señales.

Cuando dos sonidos simultáneos presentan diferencias en sus parámetros interaurales, se mejora su segregación perceptual y se mejora su detección y discriminación. Sin embargo, el procesamiento biaural no es siempre efectivo, y en algunos casos puede desarrollar un papel poco significativo.

2.2.3.8. Segmentación de flujos

Al escuchar secuencias rápidas de sonidos, los sonidos pueden agruparse conjuntamente, esto es, pueden parecer como percibidos de una misma fuente, fenómeno denominado fusión o coherencia. También puede ocurrir que se perciban como flujos diferentes, esto es, percibidos como si fuesen originados por más de una fuente. Este fenómeno es denominado fisión o segregación de flujos. [119], [145]–[147].

Diversos factores pueden producir la segmentación de flujos. Entre ellos, cabe destacar la diferencia entre tonos creados por interacciones biaurales (como en el caso de los tonos de Huggin [148]), la percepción de distintos salientes entre tonos sucesivos, la diferencia de intensidad entre tonos puros, o su diferencia de localización espacial [149].

El proceso de segregación de flujos está directamente relacionado con el proceso del reconocimiento del habla. Los sonidos del habla se suceden rápidamente formando secuencias, y las ligaduras y ligaduras parciales entre dichos sonidos, observadas en los componentes acústicos del discurso son un factor importante a la hora de percibir el habla como un flujo unificado [150]. El agrupamiento secuencial también es clave para la función del reconocimiento del habla. En un estudio conducido por Darwin y Hukin [151] se demuestra que el agrupamiento secuencial está fuertemente influenciado por la diferencia temporal interaural. En su estudio condujeron un experimento en el cual variaron los tiempos interaurales de dos frases habladas en el rango comprendido entre 0 y 181 μ s. Cada una de las frases era reproducida en un oído diferente, aplicando sobre una de ellas un desfase de 45 μ s. Las frases, basadas en el lenguaje natural hablado, habían sido pre procesadas a fin de ser reproducidas mono tónicamente, o lo que es lo mismo, manteniendo su frecuencia fundamental constante. La diferencia de frecuencia fundamental entre las dos frases se varió entre cero y cuatro semitonos. Los sujetos de prueba debían prestar atención a una de las dos frases. En un momento específico de la frase, dos palabras diferentes eran reproducidas coincidiendo en tiempo de inicio y duración (perro y pájaro). Igualmente, las frecuencias fundamentales y las interacciones biaurales de ambas palabras objetivo fueron modificadas independientemente respecto al resto de las frases. Los sujetos de prueba debían indicar cuál de las dos palabras habían escuchado en la frase sobre la que debían poner su atención. En el estudio se observó que la palabra objetivo que tenía las mismas interacciones biaurales que la frase a escuchar era percibida mucho más frecuentemente que la palabra que tenía diferentes interacciones biaurales. Esto es, la palabra y frase con iguales interacciones biaurales eran agrupadas de manera conjunta, incluso aunque tuviesen distinta frecuencia fundamental. Así, los sujetos agrupaban las palabras a lo largo del tiempo respecto a su localización subjetiva percibida, independientemente de las diferencias en su frecuencia fundamental. La conclusión del estudio es que los oyentes intentan seguir una fuente de sonido específica a lo largo del tiempo, dirigiendo su atención hacia objetos audibles localizados en una posición subjetiva específica. Igualmente, dichos objetos audibles podrían estar definidos por otras características aparte de las interacciones biaurales, como, por ejemplo, la armonicidad y la asincronía entre los tiempos de inicio y final.

Un gran número de compositores de música han explotado el fenómeno de la aparición de segregación de flujo cuando los tonos se separan ampliamente en el rango de frecuencias. Mediante la reproducción de una secuencia de tonos en la cual notas de diferentes rangos de frecuencias son reproducidas de manera alternativa, un instrumento como la flauta, en la que sólo es posible reproducir una nota cada vez, es posible percibir que dos notas son emitidas al mismo tiempo. Es posible encontrar ejemplos de este fenómeno en trabajos de grandes compositores clásicos como Bach, Telemann y Vivaldi, entre otros. Existen también ocasiones en la que los compositores intercalan notas de diferentes instrumentos para ser percibidas como una única melodía. Por ejemplo, entre los primeros y segundos violines de una orquesta. Este fenómeno ocurre cuando los distintos instrumentos presentan timbres y localizaciones similares, y son tocados en rangos tonales similares.

2.2.4. Principios en la organización perceptual

Como se comentó anteriormente en el estado del arte correspondiente a la percepción en términos generales, los psicólogos de la Gestalt [152] describieron diversos factores que rigen la organización perceptual. Tanto sus descripciones como principios también se ajustan a la forma en la que las señales físicas son procesadas para conseguir el agrupamiento perceptual de las entradas acústicas. El conjunto de reglas definido permite agrupar las partes de las entradas acústicas que surgen de la misma fuente, así como la segregación de las que no lo hacen. Las reglas en sí mismas no tienen por qué ser suficientes para llevar a cabo dicho proceso, sino más bien, su aplicación conjunta y coordinada (lo cual puede resultar complejo). De esta manera se consigue la interpretación correcta de la entrada sonora. Las reglas presentadas a continuación se pueden aplicar tanto a la visión como a la audición.

2.2.4.1. Principio de similitud

Este principio define que los elementos son agrupados si son parecidos entre sí. En la audición, la similitud por lo general implica características como el tono, el timbre, la intensidad, la duración o la posición (localización) subjetiva. Por ejemplo, si se escucha una secuencia rápida formada por tonos puros, a una tasa de unos diez tonos por segundo, entonces los tonos que con una frecuencia cercana (parecidos), son interpretados como originados por un único flujo perceptual, mientras que los tonos que tienen frecuencias muy separadas son percibidos como flujos distintos.

Para los tonos puros, la frecuencia es uno de los factores más importantes en cuanto a la similitud se refiere, aunque diferencias en la intensidad y en la posición subjetiva entre distintos tonos sucesivos también puede conducir a la separación de flujos. Para los tonos complejos, las diferencias en el timbre producidas por las diferencias espectrales parecen ser el factor más importante. Nuevamente, no es el único factor a tener en cuenta. Otros factores relevantes son las diferencias en la frecuencia

fundamental, las diferencias en la posición subjetiva percibida y las diferencias en el timbre producidas por las distintas envolventes temporales.

2.2.4.2. Principio de buena continuación

Dicho principio se basa en el comportamiento general de las fuentes de sonido. Habitualmente varían sus propiedades físicas, como frecuencia, intensidad, espectro o posición, de forma continua y suave, en lugar de hacerlo de forma súbita. Por tanto, una pequeña variación en alguno de dichos aspectos indica un cambio en una única fuente, mientras que un cambio súbito indica la activación de una nueva fuente de sonido. En un ejemplo descrito anteriormente [150], Bregman y Dannenbring mostraron que la tendencia de una secuencia de tonos agudos y graves a ser separada se reducía cuando los tonos sucesivos eran interconectados mediante la ligadura de sus frecuencias.

Otro ejemplo proviene de la experimentación con la síntesis de voz. En el habla sintética, grandes fluctuaciones de un tipo inesperado de frecuencia fundamental, y por tanto en el tono, generan la impresión de que un nuevo hablante ha intervenido para decir algunas sílabas del hablante principal. Sin embargo, parece que, en este caso, la suavidad del contorno de la frecuencia fundamental no es el único factor que determina cuándo se percibe una intrusión. La frecuencia fundamental para un único hablante cubre por lo general el rango de una octava, pero el efecto de intrusión puede ser observado para rangos mucho menores que éste, siempre y cuando sean inconsistentes con los cambios de frecuencia fundamental requeridos por el contexto fonético y lingüístico. Por tanto, parece que la asignación de un patrón espectral a un hablante específico sólo se basándose en la frecuencia fundamental de manera parcial. Aparentemente se requiere un conocimiento de las reglas de entonación (buena continuación). Sólo las desviaciones que no forman parte de dichas reglas son interpretadas como un nuevo hablante. Los resultados del estudio sobre dicho efecto [153] fueron obtenidos mediante un experimento llevado a cabo por Darwin y Bethell-Fox. En el experimento se sintetizaron patrones espectrales que cambiaban suave y repetidamente entre los sonidos de dos vocales. Cuando la frecuencia fundamental de los patrones de sonido era constante, se percibía como originada por una única fuente, y los sonidos del habla incluían ligaduras y semivocales. Sin embargo, cuando se aplicaba un contorno discontinuo con cambios en la frecuencia fundamental a los patrones anteriores, estos se percibían como dos flujos de habla diferentes, y en la conversación se percibían consonantes de parada. A menudo, un grupo específico de componentes se percibe únicamente como parte de un solo flujo, así, la segregación perceptual produce silencios ilusorios en cada flujo durante las porciones de la señal atribuidas al otro flujo, y dichos silencios son interpretados conjuntamente con los patrones espectrales correspondientes a las ligaduras con las vocales, tal y como indica la presencia de las consonantes de parada. La percepción del habla está fuertemente influenciada por la organización del flujo sonoro (continuidad).

2.2.4.3. Principio del destino común

Los distintos componentes de frecuencia surgidos de una única fuente suelen variar de una manera muy coherente. Tienden a comenzar y terminar juntos, cambiando en intensidad y frecuencia de manera conjunta. El sistema perceptual utiliza dicho fenómeno, agrupando sonidos complejos si dos o más componentes de los mismos presentan similitudes o cambios al mismo tiempo, y percibiéndolos como parte de la misma fuente de origen.

Ejemplos de dicho principio pueden encontrarse en la música. Si dos tonos complejos comienzan de manera asíncrona con 30 ms de diferencia o más, son percibidos como dos tonos diferentes. Las asincronías observadas en la música de orquesta por lo general son mayores a dicho valor, por eso al escuchar música polifónica es sencillo identificar las líneas melódicas de cada instrumento prestando la atención suficiente.

2.2.4.4. Principio de la asignación disjunta

También conocido como principio de pertenencia, hace referencia a que un único componente de sonido puede ser asignado únicamente a una fuente cada vez. Una vez que un componente ha sido asignado para la generación de un flujo, no puede ser utilizado en la formación de un segundo flujo. Para algunos tipos de estímulo, la organización perceptual podría resultar ambigua al haber más de una manera de interpretar la entrada sensorial. Cuando un componente específico puede ser interpretado como perteneciente a distintos flujos, la percepción puede variar dependiendo del flujo en el que dicho componente está incluido.

Bregman y Rudnický [154], idearon un experimento en el que se reproducían cuatro tonos muy breves sucesiva y rápidamente. Dos de los tonos tenían la misma frecuencia, y los dos tonos del medio tenían frecuencias diferentes. La secuencia de cuatro tonos podía variar la posición de los tonos del medio. Los oyentes debían identificar el orden de los tonos centrales. Se observó que dicha tarea les resultaba compleja cuando los tonos centrales se reproducían de forma única (sin los tonos inicial y final, con misma frecuencia), ya que los tonos centrales eran percibidos como parte de un patrón formado por cuatro tonos, incluyendo dos tonos de distracción (inicial y final). Por otra parte, integraron la secuencia de cuatro tonos en una secuencia más larga, con tonos denominados captosres. Cuando los tonos captosres tenían frecuencias similares a las de los tonos de distracción, producían el efecto de percibir los tonos captosres como parte de un flujo diferente (los capturaban), dejando a los tonos centrales en su propio flujo, y por ende, haciendo sencillo identificar el orden de dichos tonos centrales. Al parecer, los tonos distractores no podían ser percibidos como parte de ambos flujos. Cuando sólo un flujo es sujeto de análisis, y, por tanto, de atención, el otro flujo puede facilitar la eliminación de distractores del dominio de atención.

Cabe destacar que el principio de la asignación disjunta no siempre funciona, sobre todo en situaciones donde existen dos o más organizaciones perceptuales válidas [155], donde un elemento sonoro podría ser percibido como parte de más de un flujo.

2.2.4.5. Principio de clausura

En los entornos cotidianos, por lo general los sonidos de fuentes concretas pueden ser temporalmente enmascarados por otros sonidos. Durante el tiempo que el sonido enmascarador está presente, no hay evidencia sensorial de si el sonido enmascarado se sigue reproduciendo o ha terminado. Bajo dichas condiciones, los sonidos enmascarados tienen a percibirse de manera continua. Este proceso es denominado clausura.

Cuando un sonido X es alternado con un sonido Y, y el sonido Y tiene mayor intensidad que el X, entonces X podría percibirse como continuo, incluso tras haber sido interrumpido. Los sonidos no tienen por qué ser continuos. Por ejemplo, si Y es ruido y X es un tono que aumenta progresivamente su frecuencia, dicho aumento de frecuencia es percibido como continuo incluso aunque ciertas partes de dicho aumento no estén presentes en la señal [156]. De forma similar, cuando un tono modulado en amplitud o en frecuencia es interrumpido por ruido, la modulación se percibe como continua a través del ruido [157]. En este caso, los huecos en el tono deben ser rellenados por ruido y el ruido debe ser un enmascarador potencial del tono (si fueron reproducidos a la vez). En ausencia del ruido, la continuidad no es percibida. El efecto de continuidad también sirve en el caso del habla, cuando el flujo vocal se alterna con ruido. Si en ausencia de ruido el flujo vocal presenta huecos, se percibe como estridente. Cuando los huecos son completados por ruido, el diálogo se percibe como más natural y continuo [158]. Para tasas de interrupción del diálogo moderadas, esto puede conducir a una mejora de la inteligibilidad, ya que los cambios abruptos del habla producen señales engañosas (donde se encuentran los huecos), que el ruido puede enmascarar.

2.2.4.6. El principio de la atención y de figura – fondo

En general el ser humano es incapaz de prestar atención a cada aspecto individual de la entrada auditiva (ni siquiera en otras entradas sensoriales). Lo que ocurre es que únicamente un subconjunto de información sensorial es seleccionado para ser analizado conscientemente. En un principio parece que podría ser posible prestar atención y comparar cualquier pequeño grupo arbitrario de elementos pertenecientes a una señal acústica, sin embargo, no es posible. Por el contrario, el sonido complejo se descompone y analiza en distintos flujos, y sólo se presta atención a un flujo cada vez. Dicho flujo destaca perceptualmente, mientras que el resto del sonido es menos predominante. Los psicólogos de la Gestalt denominaron a la separación de flujos de atención y no atención el “fenómeno de figura - fondo”.

En un evento saturado de estímulos auditivos, como por ejemplo en una multitud, los seres humanos pueden prestar atención únicamente a una conversación cada vez, mientras que el resto de conversaciones forman un sonido de fondo. De manera similar, al escuchar una composición musical polifónica, la atención se centra en una línea melódica cada vez. Por supuesto es posible intercambiar la atención de una conversación a otra o de una línea melódica a otra, y aunque sea posible ser consciente de las otras voces, sólo es posible realizar un análisis consciente de un flujo cada vez. Neisser sugiere que la atención es ejercida tras el análisis inicial de los flujos [159]. Así pues, los procesos previos a la atención establecen restricciones sobre la misma, haciendo difícil atender simultáneamente o realizar inferencias sobre elementos que forman parte de dos flujos perceptuales diferentes.

Aunque la determinación de flujos auditivos impone limitaciones sobre la atención, el proceso es de dos vías. La atención también puede influir en la formación de flujos [160]. Al escuchar una secuencia de tonos alternantes, con un intervalo de aproximadamente siete semitonos entre los tonos, podría darse el caso de percibir fisión o fusión dependiendo de la atención prestada [161]. Es posible realizar una audición selectiva, intentando captar cada uno de los tonos inferiores o superiores, o por el contrario, se puede intentar hacer una audición exhaustiva tratando de captar todos los tonos juntos. Al escuchar sin un propósito específico, se puede comenzar con un tipo de audición y variar al otro. El cambio ocurre de forma espontánea y a intervalos irregulares [162], produciéndose un efecto de inversión perceptual que conduce a la ambigüedad en la interpretación de la entrada sensorial.

La importancia de los cambios en los estímulos puede interpretarse en el contexto de figura - fondo. Siempre que cambia algún componente del sonido mientras que el resto de componentes permanecen relativamente inalterados, entonces dicho aspecto es promocionado automáticamente a la atención del oyente (figura), pasando el resto del flujo al fondo. Este hecho revela que los cambios significativos en el entorno están asociados a cambios de algún tipo. La sensibilidad al cambio proporciona una manera de focalizar y dirigir la atención hacia eventos nuevos y potencialmente importantes del entorno.

2.2.5. Otros sentidos e interacciones sensoriales

Hasta ahora se han analizado características de sistemas perceptuales telerreceptores, como la vista y el oído. Como se ha comprobado, las bases fundamentales de funcionamiento en ambos son radicalmente diferentes. El primer caso está basado en la recepción de los fotones reflejados por superficies iluminadas, o los fotones emitidos directamente por una fuente lumínica. El segundo caso está basado en la detección de ondas acústicas de presión en un medio físico (por lo general, el aire). Podría decirse que en ambos casos la estimulación es producida por formas físicas de energía. Sin embargo, existen otro tipo de sentidos cuyo funcionamiento está basado en la percepción de sustancias químicas en el ambiente. Es el ejemplo del gusto y del olfato (éste último, también telerreceptor).

En líneas generales, el sentido del olfato funciona de la siguiente manera [163]: en primer lugar, las moléculas de olor presentes en el aire penetran en las fosas nasales y son disueltas en las mucosidades situadas en la parte superior de las fosas nasales. Debajo de dichas mucosidades, en el epitelio olfatorio, las células receptoras especializadas (neuronas receptoras el olfato) son capaces de detectar entre mil y cuatro mil olores diferentes [164] y su intensidad. Estas neuronas transmiten la información recibidas a los bulbos olfatorios, situadas en la parte de atrás de la nariz. Los bulbos olfatorios a su vez presentan receptores sensoriales, que forman parte del cerebro y se encargan de enviar directamente la información percibida tanto a los centros más primitivos del cerebro (estructuras del sistema límbico), lugar donde se estimulan la memoria y las emociones, como al neocórtex, lugar en el que se modifican los pensamientos conscientes. Ambos centros cerebrales perciben olores y tienen acceso a recuerdos que pueden evocar rápidamente memorias sobre personas, situaciones o lugares relacionados con dichas sensaciones olfativas. El sentido del olfato es aproximadamente diez mil veces más sensible que cualquier otro de los sentidos [163], y el reconocimiento del olor es inmediato. Mientras que, en el tacto y el gusto, los estímulos deben viajar por el cuerpo a través de vías neuronales y la espina dorsal antes de llegar al cerebro, el olfato está directamente conectado con el cerebro, haciendo que la respuesta sensorial sea inmediata.

Los olores percibidos por el sistema olfativo pueden ser clasificados en seis grandes grupos [164]: florales, afrutados, picantes, resínicos, quemados y pútridos.

La mucosa olfativa consiste en una capa de epitelios en forma de columna, rodeados por millones de neuronas olfativas cubiertas de mucosidad. Dichas neuronas son las únicas sometidas directamente a estimulación por parte del entorno externo, y se encuentran sometidas a un reemplazo constante. Las células basales cercanas se ocupan de la diferenciación y evolucionan en las neuronas olfativas cada 5-8 semanas. Aunque la densidad de receptores olfativos en las neuronas olfativas es muy elevada en los seres humanos, es aún mayor en un gran número de animales [165].

A nivel químico se pueden considerar dos tipos de moléculas olorosas, las hidrófilas y las hidrofóbicas. Las primeras son absorbidas directamente por la mucosa, mientras que las segundas necesitan ser combinadas con proteínas odorantes específicas para ser transportadas a los cilios de las neuronas olfativas. La interacción de las distintas moléculas olorosas con los cilios produce ionización, cuyo efecto es la generación de potencial eléctrico. Las corrientes iónicas generadas en respuesta a la tasa de flujo de moléculas odorantes y a su concentración, se combinan produciendo patrones espaciales específicos de actividad, que contribuyen a la codificación neuronal de los olores. Dichas respuestas a lo largo de la mucosa olfativa pueden ser monitorizados mediante el uso de electrodos superficiales (electroolfatograma) [166]. La acción de las moléculas odorantes es variable, puesto que su densidad e intensidad suele cambiar a lo largo del tiempo. Esto genera cambios en la tensión y por ende en los patrones detectados de manera proporcional a la concentración de las moléculas odorantes. La acción (o estimulación) producida por dichas corrientes es atenuada por la adaptación (o desensibilización) de los receptores [164]. Es por este fenómeno por el que el ser humano puede dejar de percibir ciertas fragancias u olores cuando lleva una cierta cantidad de tiempo expuesto a dicho olor, siempre que este último se

mantenga constante. El fenómeno de adaptación sensorial se ha observado (y comentado) anteriormente en otros sentidos, como la vista y el oído.

La adaptación rápida y la eliminación de odorantes, permite el reconocimiento y discriminación continua de los nuevos aromas inhalados en cada ciclo respiratorio. Las diferencias de potencial generadas en los axones de las neuronas activadas son propagadas al bulbo olfativo. Los bulbos olfativos tienen muchos tipos distintos de neuronas en una distribución laminar [164]. Millones de axones convergen en unos pocos miles de glomérulos en el interior de los bulbos olfativos, donde cada bulbo está interconectado con unas 75000 células mitrales. Dichas células mitrales son neuronas sensoriales de segundo orden, cuyos axones penetran en el tracto olfativo y ascienden hacia el córtex olfativo. Esta convergencia o divergencia entre los axones de las neuronas olfativas y las células especializadas de los bulbos olfativos generan potenciales excitativos post-sinápticos en las células mitrales y, por consiguiente, diferencias de potencial interpretadas como estímulos olfativos. Existe un complejo patrón de integración neuronal cuyo objetivo es la discriminación de diversas moléculas odorantes. Este patrón es generado por los mecanismos de convergencia o divergencia de éstas neuronas sensoriales de segundo orden. La complejidad de este patrón se relaciona con el hecho de que un único olor puede estimular diferentes receptores y neuronas [164].

Una característica del sentido del gusto es que requiere la interacción de distintos sistemas sensoriales. El gusto y el olfato son los sistemas principales encargados de distinguir sabores. Sin embargo, otros sistemas sensoriales como el tacto, la termocepción y la nocicepción de la mucosa oral, contribuyen a determinar la calidad sensorial de la comida [164], [167]. La saliva también tiene un papel importante al mantener la acuosidad de las células receptoras del sabor. Igualmente, contribuye a disolver los solutos polares, transportar dichos solutos a los receptores del gusto, atenuar las acciones de los alimentos ácidos y reparar el epitelio lingual.

Los bulbos gustativos están localizados en las papilas y distribuidos a lo largo de la superficie de la lengua. Los bulbos están presentes también en la mucosa oral del paladar y de la epiglotis. Dichas estructuras en forma de pera contienen unas ochenta células distribuidas alrededor de un poro gustativo [164]. Las células receptoras del gusto son células neuro-epiteliales que se extienden desde la base hasta la cima de los bulbos gustativos y que son reemplazadas cada nueve o diez días [164]. Las microvellosidades de cada célula receptora del gusto se proyectan hacia el interior de los poros gustativos, donde entran en contacto con los solutos disueltos en la superficie de la lengua. Los solutos que contienen el sabor son transportados hacia los poros gustativos y son distribuidos a través de la capa líquida para entrar en contacto con las proteínas receptoras de la membrana apical y de las microvellosidades. El sentido del gusto, al igual que el olfato, depende de la concentración de moléculas. En este caso, moléculas de sabor, y de su solubilidad en la saliva. Por ello, si la lengua está seca, los poros gustativos recibirán pocas señales químicas y por ende, experimentarán dificultad a la hora de identificar el sabor de los alimentos [168].

La sensación del gusto puede ser evocada por un gran número de diversos solutos que contienen sabor. La interacción de dichos solutos con los receptores produce patrones de cambio de potencial (despolarización, la despolarización seguida por

hiperpolarización, o únicamente hiperpolarización) en las respectivas membranas [164]. Hay cinco tipos de receptores para el gusto, respondiendo cada uno de ellos a un cierto tipo de componente químico. Los cinco grandes grupos de sabores asociados a dichos componentes son: dulce, ácido, salado, amargo y umami (nombre derivado de la palabra japonesa para “delicia”). Este último responde específicamente al glutamato y a ciertas proteínas [168]. En cualquier caso, la especificidad de un terminal gustativo no es absoluta. Una misma fibra puede responder a diferentes estímulos químicos, por tanto, el gusto percibido de cualquier estímulo es el resultado de un patrón de activación en los diferentes receptores gustativos afectados [169]. Aunque históricamente se establecieron diferentes regiones para la detección de distintos sabores a lo largo de la superficie de la lengua (dulzura en la parte delantera, acidez y salinidad en los laterales, amargura en la parte posterior), estudios sobre el gusto basándose en la respuesta de los diferentes nervios craneales demuestran que las comidas con gusto similar producen un patrón de actividad parecido. Dichos patrones de actividad indican que la formación de la codificación neural del gusto ocurre en diferentes neuronas y células gustativas respondiendo a un estímulo particular. Este hallazgo indica que no existe una única fibra que proporcione una única cualidad gustativa (dulce, salado...), sino que determinadas fibras pueden responder mejor o peor a determinadas cualidades gustativas [164]. Igualmente existen diferencias significativas en las preferencias del gusto, tanto dentro de la misma especie como entre especies. Por ejemplo, el ser humano y las ratas comparten el gusto por las soluciones de sacarosa y sacarina, mientras que, por ejemplo, los perros rechazan la sacarina, y los gatos son indiferentes a ambas. Por otra parte, dentro de la especie humana, existen diferencias significativas entre los umbrales gustativos. Las personas ancianas tienen por lo general umbrales más altos, en parte debido a la notable reducción de poros gustativos con el envejecimiento. Los niños tienen menor tolerancia a las especias y a los sabores amargos, pues su sentido del gusto es más intenso. Y dentro de los adultos existen diferencias absolutas: ciertas personas perciben algunos gustos como fuertes y desagradables mientras que otras personas son indiferentes a ellos (por ejemplo, con tipos de quesos) [168], [169]. Finalmente, se ha observado que el género sexual también puede influir en la detección de sabores, pues en promedio las mujeres tienen más papilas fungiformes y bulbos gustativos [170], lo que las hace más sensibles a los cambios y matices en las cualidades gustativas.

Además de las modalidades sensoriales exteroceptivas analizadas, existen cuatro sentidos adicionales, clasificados como sentidos somáticos, importantes a la hora de percibir características adicionales del entorno y características propias del ser humano o de su interacción con el medio. Dichos sentidos son el tacto, la termocepción, la propiocepción y la nocicepción. Los receptores para estos sentidos somáticos se encuentran tanto en la piel como en las vísceras. La compleja estructura neural, así como los canales que transmiten la información de estos receptores al cerebro pasan por el cerebelo, de forma que es posible utilizar la información sensorial para coordinar equilibrio y movimiento [171]. La información recogida en el cerebro se procesa en la corteza somatosensitiva. Gracias a experimentos con pacientes durante la cirugía cerebral para la epilepsia [172] se ha determinado que cuanto más sensible a la estimulación táctil es una región del cuerpo, mayor es su cantidad correspondiente de espacio asociado en la corteza somatosensitiva. El tamaño de dichas regiones, por

tanto, no es fijo. Si se utiliza más una parte particular del cuerpo, su región topográfica en la corteza aumenta. Por ejemplo, los discapacitados visuales que aprenden a leer Braille con los extremos de sus dedos tienen una región más amplia en la corteza somatosensitiva específica para esta parte de los dedos [171].

Los receptores del tacto se encuentran entre los más comunes del cuerpo. Dichos receptores responden a muchas formas de contacto físico, como el estiramiento muscular o de la piel, la presión constante, el movimiento de golpeteo o aleteo, la vibración o la textura. Existen tanto en la zona superficial de la piel como en regiones más profundas del cuerpo, como músculos órganos internos o articulaciones.

RECEPTOR	ESTÍMULO ACTIVADOR	LOCALIZACIÓN	ADAPTACIÓN
Corpúsculo de Pacini	Vibración	Capas profundas (piel)	Rápida
Corpúsculo de Ruffini	Estiramiento de la piel	Capas profundas (piel)	Lenta
Receptores de Merkel	Presión constante y temperatura	Capas superficiales (piel)	Lenta
Corpúsculo de Meissner	Golpeteo constante y toque delicado	Capas superficiales (piel)	Rápida
Terminación nerviosa libre	Estímulos nocivos y de presión	Bajo la superficie de la piel	Variable
Terminación nerviosa libre capilar	Movimiento del vello	Raíz del pelo	Variable

Tabla 1. Receptores del tacto en las distintas capas de la piel y sus características

Como puede observarse en la Tabla 1, los receptores del tacto se presentan en diversas formas. Algunas de ellas simples, como las terminaciones nerviosas libres, y otras mucho más complejas, como los corpúsculos de Meissner. Cada uno de ellos responde a distintos estímulos. Por ejemplo, los corpúsculos de Pacini responden bien a vibraciones de alta frecuencia, cuya energía es transferida a los terminales nerviosos. Los corpúsculos de Pacini son receptores de adaptación rápida, lo que les permite responder a un cambio en el tacto pero luego ignorarlo [171].

Los receptores de temperatura, o termorreceptores, están compuestos por terminaciones nerviosas libres que terminan en las capas subcutáneas de la piel. Los receptores de frío son sensibles principalmente a temperaturas inferiores a la corporal, estableciéndose su rango de activación entre los 8 y los 20 grados Celsius, mientras que los receptores de calor se activan entre los 25 y los 45 grados Celsius [173]. Fuera de dichos rangos, los nocioreceptores, o receptores del dolor, entran en acción para prevenir daños potenciales, con lo que la sensación térmica se transforma potencialmente en dolor [174]. El campo perceptivo para un termorreceptor tiene un tamaño aproximado de un milímetro de diámetro, y los receptores están distribuidos por todo el cuerpo. Existen muchos más receptores de frío que de calor. Los termorreceptores se adaptan de forma lenta entre los 20 y 40 grados Celsius. Los cambios iniciales informan en cambios de la temperatura, y una estimulación sostenida y constante informa sobre la temperatura del entorno. Fuera del rango comprendido entre los 20 y 40 grados Celsius la probabilidad de daño tisular es mayor, y los receptores no se adaptan. El papel desempeñado por los termorreceptores es de vital

importancia en la termorregulación corporal [171], al igual que en la interacción con otros sentidos como el olfato y el gusto, variando la sensación percibida ante un alimento o aroma dependiendo de su temperatura. En cualquier caso, las preferencias olfativas y gustativas se aprenden a lo largo del desarrollo, y son asociadas a recuerdos mediante la memoria [175].

Otro sentido relacionado con la detección de la situación o estado del ser humano en el entorno es el equilibrio. El equilibrio es el estado de estabilidad, o la posición del cuerpo en el espacio. En la ubicación del cuerpo, la información sensorial originada en el oído interno y en los propioceptores musculares y articulares le indica al encéfalo la localización de las diferentes partes del cuerpo en relación unas con otras y con el entorno [171]. Igualmente, la vista también desempeña un rol importante en el equilibrio [176], [177]. El sentido del equilibrio tiene su componente más importante en la sensación recogida por diversos componentes del oído interno. Principalmente, por las células ciliadas que reviste en aparato vestibular lleno de líquido (endolinfa), y los conductos semicirculares del oído interno. Dichos receptores no neurales responden a cambios en la aceleración rotacional, vertical y horizontal. Las células ciliadas funcionan igual que las situadas en la cóclea, pero en lugar de detectar ondas sonoras, detectan la interacción de los estereocilios con fuerzas como son la gravedad y la aceleración.

El sentido espacial del equilibrio puede dividirse en dos componentes principales, uno dinámico, encargado de registrar los movimientos a través del espacio y otro estático, que registra la posición de la cabeza. Los tres conductos semicirculares del oído interno detectan la aceleración rotacional en distintas direcciones. Los órganos otolíticos registran la aceleración lineal y la posición cefálica. Las células ciliadas de ambos funcionan como las células ciliadas del órgano de Corti, responsable de la audición. Cuando los cilios se inclinan en una dirección se despolarizan y al inclinarse en la dirección opuesta se hiperpolarizan. Al lado del haz ciliar, existe un componente denominado cinocilio, responsable de crear un punto de referencia para determinar la dirección de la inclinación [171]. En el caso de la rotación, el cráneo óseo y las paredes membranosas del laberinto se mueven, pero el líquido interno (endolinfa) tiende a permanecer en su sitio debido a la inercia. En las ampollas del aparato vestibular, la resistencia de la endolinfa inclina la cúpula y las células ciliadas en dirección opuesta a la dirección en la cual está girando la cabeza. Si la rotación perdura en el tiempo, la endolinfa seguirá el movimiento. Si la rotación de la cabeza se detuviera súbitamente, la inercia del líquido impediría su detención inmediata, con lo que la sensación sería de seguir rotando. Ante dicha sensación se activarían estímulos reflejos tratando de compensar la aparente pérdida de equilibrio [171]. En el caso de la aceleración lineal y la determinación de la posición del cráneo, son los órganos otolíticos del utrículo y el sáculo los dispuestos para sentir fuerzas lineales. Las máculas del utrículo permanecen horizontales cuando la cabeza está en posición erecta (posición normal). Si la cabeza se inclina, por ejemplo, para mirar hacia arriba, la gravedad hace que los otolitos (partículas minerales y proteicas) de la membrana otolítica gelatinosa sean deslizados hacia atrás, doblando las células ciliadas y disparando con ello una señal. Sin embargo, las máculas de los sáculos, al contrario que las del utrículo, están orientadas verticalmente cuando la cabeza está en posición erecta. Esto hace que sean sensibles a las fuerzas verticales, como las producidas por

caídas. El encéfalo analiza el patrón de células ciliadas hiperpolarizadas y despolarizadas para determinar la posición de la cabeza y la dirección del movimiento [171].

Como se ha podido observar en el estudio de la literatura, existen diversos sentidos, encargados cada uno de determinar características internas o externas al cuerpo humano, pero dichas características son percibidas con mayor precisión y detalle, o con un menor tiempo de respuesta cuando existe interacción entre dichos sentidos. Así pues, por citar algunas de estas interacciones, es posible identificar sustancias potencialmente dañinas mediante el gusto y el olfato, es posible detectar temperaturas potencialmente peligrosas para los tejidos mediante el tacto y la nociocepción, o posibles alteraciones en el movimiento o en el terreno mediante el equilibrio, que, aunque principalmente es responsabilidad de órganos en el oído interno, depende fuertemente de interacciones con la vista y del tacto. Sin embargo, una de las interacciones más importantes es la existente entre la vista y el oído a la hora de posicionar objetos en el espacio. Por lo general la posición de una fuente de sonido estacionaria es percibida como invariante, o fija, en el espacio. Si se mueve la cabeza, el sonido que incide en los dos oídos cambia, pero la imagen sonora permanece estacionaria en el espacio. De algún modo, la información sobre la posición de la cabeza es combinada con información sobre la posición relativa de la cabeza para generar una percepción constante. La visión puede desempeñar un papel importante en este proceso al ayudar a definir la posición de la cabeza en el espacio. Este hecho se describe en una serie de experimentos realizados por Wallach [178]. En ellos, los sujetos de prueba tenían sus cabezas fijas en el eje vertical de una pantalla cilíndrica que rotaba sobre ellos. La pantalla estaba cubierta de bandas verticales, y tras observar su movimiento durante unos momentos, los sujetos percibían que se encontraban en movimiento respecto a una pantalla fija. Entonces se activaba una fuente inmóvil de sonido justo delante de los sujetos de prueba. Este sonido era percibido como si tuviese su origen encima o debajo. Aparentemente el cerebro trata de determinar un origen de la fuente de sonido que sea consistente con todas las evidencias sensoriales disponibles. Los sujetos en esta situación se percibían en movimiento, pero el sonido incidente en sus oídos no tenía disparidad interaural, ni cambio alguno. Perceptualmente este fenómeno es coherente sólo si el sonido es originado por encima o por debajo de los sujetos. Si la fuente de sonido tiene un azimut de, por ejemplo, 24 grados hacia la derecha del sujeto, la fuente de sonido no puede ser interpretada como procedente de arriba o abajo, ya que en este caso aparecen diferencias interaurales en intensidad y tiempo. En este caso, la fuente es percibida en movimiento, rotando con el sujeto a una elevación aproximada correspondiente al complemento del azimut (en este caso, unos 66 grados). De esta forma, la interpretación del posicionamiento de las señales auditivas está fuertemente influenciado por la orientación visual, y, por consiguiente, para conseguir un mayor nivel de percepción espacial se requiere la integración de información de diferentes sentidos.

Por último, y no menos importante, cabe destacar un fenómeno que ocurre en los diferentes sistemas sensoriales y permite evitar la saturación de estímulos tanto a los sistemas sensoriales como al cerebro. Dicho fenómeno es el de la adaptación, o fatiga [179], [180]. La fatiga se produce en respuesta a una estimulación constante sobre

determinados sistemas sensoriales. Por ejemplo, el tacto se adapta a la estimulación producida por las ropas (si no, serían una distracción constante). Como se ha comentado, la fatiga se da también en otros sentidos, como la visión, el oído, el olfato, el gusto o la termocepción [181], [182], [174], [183], [184]. La adaptación implica mecanismos a nivel de las células receptoras, incluyendo la desactivación de los canales iónicos de la membrana que genera la respuesta eléctrica. De una forma simplificada, tras la estimulación, la célula produce una señal eléctrica. Al poco, la membrana celular impide que los iones sigan fluyendo, previniendo el envío de más señales. Este proceso puede ser más lento o más rápido, como por ejemplo en el caso de las células del dolor, cuya adaptación es variable dependiendo del tipo de estímulo [174], [171]. La eliminación del estímulo seguido por la re-estimulación vuelve a iniciar el proceso.

2.3. Conocimiento

El conocimiento para el ser humano está constituido por un conjunto particular de hechos, técnicas y procedimientos que desarrollan las culturas, y que tienen una influencia constante en el desarrollo de la vida cotidiana. De forma más extensa, el conocimiento puede ser definido como la información acerca del mundo que puede ser cierta, es coherente, y está justificado creerla [185], [186]. Así pues, el conocimiento posibilita el buen funcionamiento de muchos procesos mentales, no solo de la memoria, del lenguaje y del pensamiento, sino también de la percepción y de la atención. Sin el conocimiento, los procesos mentales resultarían ineficaces. Por ejemplo, sería imposible determinar el significado o la historia de los elementos percibidos en el entorno. Incluso la clasificación o categorización de dichos elementos sería inviable, pues se actuaría de forma totalmente pasiva ante los estímulos percibidos al no ser capaz de relacionarlos con experiencias anteriores. La categorización es la capacidad de establecer que una entidad percibida pertenece a un grupo específico de cosas que comparten características clave, como estructura o uso [71]. Por ello, sin conocimiento la categorización sería imposible de llevar a cabo.

Una vez que una entidad percibida es categorizada, es posible disponer de más conocimientos sobre la categoría para aplicarlos y extraer deducciones. Es decir, permitir obtener información que no se encuentra presente explícitamente en un único miembro de la categoría, pero de la que se puede disponer gracias al conocimiento de las características del grupo o grupos a los que pertenece. De manera similar, las acciones involucradas y la experiencia dada por el conocimiento ayudan a actuar dependiendo de la situación en la que el ser humano se encuentre. Sin conocimiento no existiría reflejo biológico alguno que permitiera interactuar de manera apropiada acorde a las diferentes situaciones cotidianas. Tampoco sería posible completar la percepción parcial (fenómenos comentados en la visión o en la audición, relativos a la reconstrucción mental de escenas o sonidos basándose en la experiencia previa), pero que sí puede ser realizado gracias al conocimiento, al igual que el reconocimiento e interpretación del lenguaje [71].

2.3.1. Representaciones

Un aspecto fundamental del conocimiento se basa en las representaciones. De acuerdo a Dietrich y Markman [187], una representación es un estado físico (como las conexiones neurales en un determinado cerebro) que representan un objeto, un suceso o un concepto. Las representaciones también pueden transmitir información de aquello que representan. Por ejemplo, un mapa de la red de transportes es una representación, puesto que representa los diferentes medios de transporte, sus líneas, paradas, conexiones y contiene información de todas ellas en mayor o menor grado (el orden de las paradas y en algunos casos los horarios). Pero las representaciones, en general implican más que eso.

Una propiedad básica del cerebro es que, hasta cierto punto, pero lejos de la perfección, es capaz de almacenar las experiencias percibidas, o recuerdos. Sin embargo, para almacenar un recuerdo deben cumplirse ciertos requisitos [188], [189]:

- **Criterio de intencionalidad:** una representación tiene que construirse deliberadamente para representar algo. Aunque el ser humano no pueda obligarse a sí mismo a recordar un objeto tras verlo por primera vez, por lo general el cerebro se ocupará de almacenar la información al respecto automáticamente, incluso si no se está tratando de fijar dicha información en la memoria de forma consciente. Es un fenómeno que evolutivamente se ha asignado a mecanismos automáticos e inconscientes.
- **Criterio de transmisión de información:** una representación debe contener información acerca de lo que representa. Los recuerdos almacenados en la memoria contienen detalles que permiten describir objetos y experiencias a posteriori. Por ejemplo, forma, color, olor, función, desarrollo en el tiempo, etc. Igualmente es posible establecer categorías basándose en los recuerdos. Por ejemplo, si un ser humano observa un coche, al observar otro diferente, en modelo o en color puede automáticamente determinar que el nuevo ejemplar pertenece a la misma categoría que el que observó inicialmente, puesto que la representación del ejemplar observado inicialmente conlleva información sobre sus características, que puede utilizar para reconocer elementos similares. De modo similar, si se ha observado algún comportamiento o característica interactiva en el coche inicial, como la apertura de las puertas, al observar el segundo será posible acceder a los recuerdos del primero para conjeturar que en éste nuevo coche también será posible abrir las puertas de alguna manera.

Así pues, las representaciones establecen las bases del conocimiento. Una vez que el cerebro establece los recuerdos que contienen información sobre el entorno, se pueden llevar a cabo muchas y sofisticadas capacidades cognitivas [189].

Un aspecto fundamental de la representación es su formato. El formato no sólo se refiere a los elementos que conforman una representación y cómo están dispuestos los elementos, sino que se basa en las características de los procesos que operan en ellos para extraer información [71]. A grandes rasgos, las representaciones pueden ser:

- **Modalidad específica:** pueden valerse de los sistemas perceptivos o motores
- **Amodales:** basándose en elementos externos a las modalidades perceptivas o motoras.

Igualmente, el contenido (o información expresada) por una representación es un aspecto importante a tener en cuenta.

2.3.1.1. Representaciones de modalidad específica: imágenes

Las imágenes, análogas a las que podría capturar una cámara, son un modelo de representación que describe información. En caso de una imagen registrada por una cámara digital, las unidades de información de la misma son los píxeles. Generalizando un poco más, se puede considerar que una imagen tiene tres elementos que en conjunto determinan su contenido: una ventana espacio-temporal, unidades de almacenamiento e información almacenada [71]. Considerando el ejemplo de una fotografía, la imagen mostrada en ella no capta toda la escena, sino sólo la parte que se encuentra en una ventana espacio-temporal. Existen un número ilimitado de fotografías que podría tomar una cámara para una misma escena, dependiendo de la posición relativa a dicha escena. Temporalmente, la escena tampoco se capta de manera continua en el tiempo, sino únicamente en el momento en el que se realiza la fotografía. Respecto a las unidades de almacenamiento, para la ventana espacio-temporal, una imagen contiene un conjunto de unidades de almacenamiento que conforman la imagen. Para una imagen digital, un conjunto de píxeles. Para una imagen analógica, un conjunto de gránulos sensibles a la luz. En cualquier caso, dichas unidades de información también se rigen por la ventana espacio-temporal, pues cada una capta la información existente dentro de una pequeña región espacial y temporal determinada, contenida en la ventana mayor que abarca el conjunto de la imagen. Finalmente, la información contenida en las unidades de almacenamiento para una fotografía son las intensidades de la luz en las longitudes de onda visibles que inciden en cada una de las unidades de almacenamiento. O, dicho de otro modo, los colores para cada pixel. El conjunto completo de información contenida en las unidades de almacenamiento define el contenido de la imagen.

Existe una gran cantidad de información importante adicional que reside en la imagen, de forma implícita. Por ejemplo, un conjunto de píxeles podría formar una figura geométrica, como un cuadrado o una elipse. Igualmente, a más alto nivel, es posible determinar distancias entre objetos, o sus tamaños relativos, pero dicho análisis requiere un procesamiento adicional. Procesamiento que el cerebro es capaz de llevar a cabo [190], [191]. Existen evidencias científicas que existen imágenes mentales tanto el cerebro humano como en el de otros animales [192]–[198]. Como ejemplo, es

posible citar las investigaciones anatómicas del cerebro de un mono ante la estimulación visual [190]. La activación en la corteza occipital del mono, registrada mediante un marcador neuronal mientras el mono observaba el estímulo, mostró una correspondencia entre el patrón de activación cerebral en la superficie y del cerebro y la forma del estímulo observado. La corteza de las áreas de procesamiento visual inicial está dispuesta de forma similar a los píxeles de una imagen digital, y responde de forma similar. Al dispararse las neuronas dispuestas de esta manera, el patrón de activación forma un mapa topográfico: su disposición espacial en el cerebro es análoga a la disposición del espacio en el entorno. Dicho fenómeno sugiere la presencia de representación basada en imágenes.

Las imágenes mentales distan de las imágenes que podrían capturarse mediante una cámara. Las imágenes cerebrales no son tan continuas ni completas. Igualmente no presentan un nivel uniforme de detalle (algunas áreas no están tan bien definidas y representadas como otras) [199]. La consecuencia de este fenómeno es la atención visual. Por lo general, las partes mejor representadas en la escena son aquellas donde se ha prestado más atención [200]. Al no centrar la atención en ciertas regiones de la escena, dichas regiones no se codifican igual de bien en la imagen mental. También puede ocurrir que dependiendo de dónde se focalice la atención, en una misma imagen se puedan reconocer diferentes elementos, como ocurre en algunas imágenes bi-estables o reversibles, consideradas ambiguas (p. ej. el pato-conejo de Wittgenstein [201]).

2.3.1.2. Representaciones de modalidad específica: registro de características

La base de una representación sofisticada reside en la categorización de entidades significativas. En comparación con las imágenes, una entidad significativa puede ser considerada como un suceso u objeto que juega un papel importante en la supervivencia de un organismo vivo y en su persecución de metas [71], mientras que un pixel apenas contiene significado en sí mismo. Sin embargo, un conjunto suficientemente grande de píxeles puede contener patrones relacionados con características del entorno, que pueden ayudar a construir representaciones más significativas.

Un ejemplo de representación sofisticada se puede apreciar al estudiar el sistema visual de la rana. Para la supervivencia de una rana, es necesario poder detectar insectos en su hábitat natural. En un estudio liderado por Lettvin [202] se observó que las neuronas del sistema visual de las ranas responden de manera diferente ante pequeños objetos que se mueven dentro de su campo visual. Tras acoplar electrodos en el cerebro de una rana y variar estímulos frente a sus ojos (objetos redondos estacionarios o en movimiento), se identificaron distintas respuestas en varias poblaciones de neuronas, concretamente ante el movimiento independientemente del objeto, y ante el objeto independientemente de su movimiento. En base a estas respuestas, se concluyó que la información detectada por estas neuronas es significativa para las ranas: redondo y pequeño como características del objeto y en movimiento como acción llevada a cabo por el objeto. Estas son características

comunes a los insectos voladores que abundan en los hábitats de las ranas, y no constituyen los fragmentos de una imagen, sino interpretaciones de regiones de imágenes en forma de características particulares. Al activarse dichas neuronas asociadas al reconocimiento, clasifican una región de una imagen como contenedora de una característica significativa de un objeto o suceso. Estas características, como se ha comentado, son detectadas por poblaciones de neuronas, y no por neuronas aisladas, de esta forma es posible obtener una respuesta gradual en lugar de una respuesta absoluta (sí o no), lo cual aumenta la fiabilidad [202]. Por otra parte, estos grupos neuronales suelen responder a más de una característica, y la información a la que responden puede variar con la experiencia y con el objetivo del organismo en un momento concreto [203].

En el sistema visual, las señales transmitidas desde la corteza visual primaria hasta los lóbulos temporales y parietales son analizadas, y en dicho proceso se extraen varios tipos de características, como son la forma, la orientación, el color (textura) o el movimiento de los objetos. Además, las neuronas asociativas de las vías de procesamiento se ocupan de integrar la información de las características extraídas, formando representaciones de objetos [203].

Este formato de representación, construida a partir de características, complementa una imagen del mismo objeto, almacenada en la corteza cerebral y organizada topográficamente. Cabe destacar que la detección de características no se realiza únicamente para la visión, sino para otras modalidades sensitivas, como la audición, el tacto, el olfato y el gusto [204].

2.3.1.3. Representaciones Amodales

Las representaciones amodales se basan en los llamados símbolos amodales, los cuales describen de manera abstracta las propiedades y relaciones entre las entidades significativas del entorno [71]. Los símbolos amodales, al residir fuera de las modalidades y poseer características ajenas a una modalidad específica, radican en un sistema de conocimiento que construye y maneja descripciones de estados perceptivos y motores. Su fundamento está fuera del sistema visual y son parte de un sistema más general que se utiliza en el lenguaje y en otras tareas que no implican específicamente una modalidad sensitiva concreta.

Aunque en las representaciones amodales se utilicen palabras, los símbolos amodales no son exactamente las palabras en sí [71]. Las palabras representan a los símbolos amodales subyacentes (palabra “computador” vs. concepto “computador”).

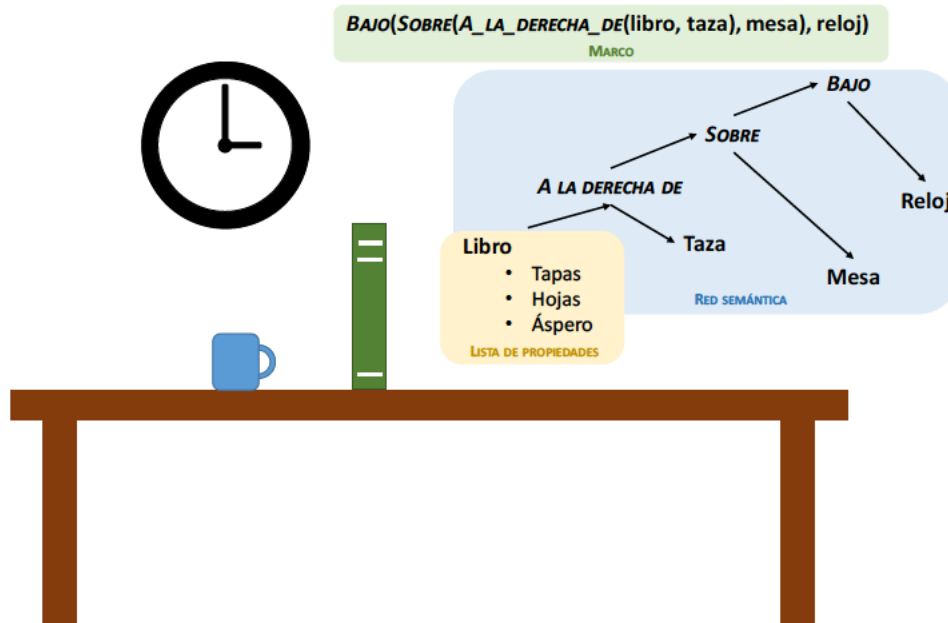


Ilustración 8. Representaciones amodales para elementos de una misma escena

En la Ilustración 8 se pueden observar tres tipos de representación amodal para los elementos mostrados en la escena: un marco, una red semántica y una lista de propiedades. Un marco es una estructura que especifica un conjunto de relaciones que unen objetos del entorno. La red semántica representa esencialmente las mismas relaciones y objetos en forma de diagrama. La lista de propiedades enumera las características de las entidades que pertenecen a una misma categoría. Por ejemplo, en la lista mostrada en la ilustración, algunas propiedades de los libros, como su tapa o las hojas. Este último tipo de representación, como puede apreciarse, omite las relaciones entre propiedades, mostrando únicamente aspectos abstractos de un objeto.

Los símbolos amodales también complementan las imágenes, pues clasifican las regiones de las mismas tras el proceso de detección de características [71].

2.3.1.4. Representaciones amodales estadísticas. Redes neurales

Este tipo de representación conlleva la asociación de un patrón estadístico, como una secuencia binaria, basada en la detección o no de características respecto a la activación neural, al elemento percibido. Las dos ventajas principales de este tipo de representación son las siguientes [205]: la primera es que mientras en un sistema modal, un único símbolo amodal por lo general sólo representa una categoría, en una red neural, múltiples modelos estadísticos pueden representar la misma categoría. La segunda, es su representación. Los elementos de un patrón estadístico se pueden modelar como neuronas o poblaciones de neuronas que están activadas o desactivadas. En dicha base binaria, cada 1 en el patrón representa una neurona o población que es disparada, y un 0 representa otra que no lo hace. Esta interpretación

neural natural hace que el patrón estadístico sea un candidato plausible para la representación biológica.

2.3.2. Formatos de representación para la percepción y para la composición

Todo conocimiento está asociado con, al menos, un formato de representación descriptivo [71]. El cerebro humano utiliza el conocimiento, o concretamente, sus representaciones, en la mayoría de procesos que componen la cognición [189]. Para ello, utiliza múltiples formatos de representación, como los analizados anteriormente.

Al percibir una escena, el cerebro construye una imagen visual fragmentada. Principalmente en la corteza occipital. Con la elaboración de la imagen, diversos sistemas neurales detectores de características extraen las partes significativas. Finalmente se activa un modelo estadístico en los lóbulos temporales para representar la información de la imagen y sus características extraídas previamente, y para asociar toda esa información [71]. Debido a que las neuronas que representan el patrón estadístico son asociativas, o integradoras, tanto las neuronas que representan la imagen como las que extraen sus características son integradas con ellas. Esto permite que los elementos del modelo estadístico establezcan asociaciones retroactivas con la imagen y las características que se han activado.

Aunque se esté haciendo mención principalmente a las representaciones basadas en imágenes, por ser uno de los campos perceptuales más importantes, cabe destacar que los diferentes sistemas sensoriales también aportan información a la representación, que es asociada de forma análoga por dichas neuronas integradoras [206]. Esta secuencia de fases de procesamiento establece una representación en múltiples niveles.

El proceso de recordar, también denominado composición o simulación, conlleva reactivar la información de la imagen y de características incluso cuando la escena original ya no está presente. Cualquier estímulo de entrada, como el sonido asociado a una palabra o un patrón de olor específico pueden reactivar la información de la imagen y características incluso cuando la escena original ya no esté presente. Este proceso conlleva inicialmente la activación de un patrón estadístico correspondiente al estímulo de entrada. Desde ahí, en un proceso tipo “*top-down*”, el patrón estadístico reactivará parcialmente las características obtenidas de la situación correspondiente, como aspectos de la imagen que la representa y diversas características. La estructura asociativa que liga toda la información capturada por los diferentes sistemas cognitivos y perceptuales permite simular la experiencia original. Por otra parte, un procesamiento tipo “*bottom-up*” a través de un sistema perceptivo produce una representación estadística, un procesamiento “*top-down*” que vuelve por otra vía reconstruye de forma parcial el procesamiento perceptual original. Esta última capacidad de procesamiento “*top-down*” es la que permite crear imágenes mentales y recordar sucesos pasados [189].

2.3.3. Categorización

Como se ha observado, el conocimiento es el fin que permite comprender y analizar nuevos estímulos e información. El medio para alcanzar dicho conocimiento son las diferentes representaciones que pueden realizarse sobre los datos. Sin embargo, todavía no se ha mencionado cómo se crean los conjuntos de representaciones para proporcionar conocimientos sobre una categoría.

El conocimiento de categorías se elabora a partir de representaciones sobre miembros individuales de una misma categoría, y posteriormente a partir de la integración de dichas representaciones [14], [79], [206].

Al observar elementos que pueden pertenecer a una misma categoría, como una mesa, en el cerebro se crean modelos estadísticos similares, que comparten muchas de sus unidades básicas de información. Esto permite que la agregación de dichos elementos se pueda realizar de una forma natural, pues al compartir dichas unidades básicas, todos los elementos son asociados a un mismo *hub* [71]. El resultado de dicha asociación es la representación de una categoría (basándose en el grupo de unidades básicas comunes).

Las unidades básicas de información compartidas también ofrecen un modo de recuperar miembros de una categoría desde la memoria. Dado que los miembros de una misma categoría comparten un *hub* común, dicho *hub* actúa como activador para recordar los miembros de la categoría en ocasiones posteriores. Mediante un procesamiento tipo “*top-down*”, el *hub* recupera la imagen y características asociado a un miembro de la categoría específica, aunque a menudo se pueden mezclar recuerdos de múltiples miembros de categorías durante la recuperación y producir una combinación con las características de los mismos [207], [208]. Por consiguiente, el elemento recordado para la categoría suele ser más similar a un miembro típico de la categoría que a uno específico. Este proceso de simulación de miembros típicos de una categoría proporciona un mecanismo para generar prototipos [14].

Al reconocer a un miembro de una categoría, se activan los conocimientos asociados y relevantes para esa categoría general, lo que proporciona una enorme cantidad de información útil para actuar en consecuencia con dicha entidad. El incremento de la experiencia reconociendo e interactuando con elementos de la misma categoría refuerza y amplía el conocimiento de la misma. Debido a que el conocimiento de categorías se conforma por información de distintos orígenes (visual, auditivo, táctil, etc.), es posible obtener deducciones útiles y realizar diversas funciones inteligentes en distintos ámbitos cognitivos [206]. Obviamente, no todos los elementos categorizados contienen la misma cantidad de información procedente de las mismas modalidades sensoriales. Por ejemplo, considerar un plato de comida y un piano. Dependiendo de la categoría, el perfil de información puede ser diferente, en base a los sentidos involucrados en su percepción inicial [209]. La motivación y la emoción ofrecen otros modos de experiencia que forman parte de una representación de categoría. Así, existen elementos relacionados con emociones (o experiencias) positivas, otros relacionados con negativas, otros con sensaciones de relajación, de estrés, de hambre, etc. La etiqueta, o nombre asociado a una categoría, permite el

acceso directo al conocimiento de dicha categoría. La integración de la información entre el nombre de la categoría y la información pertinente a la misma en el cerebro es la clave. Existen diversas teorías para explicar el fenómeno. La más generalizada es la teoría de la zona de convergencia o área de asociación [210], [211] se basa en que un grupo de neuronas asociativas permiten la integración de información de una imagen y del análisis de características dentro de una modalidad dada, como por ejemplo la visión o el olfato, aunque la información de características podría estar integrada dentro de otras modalidades perceptivas o de acción. Investigaciones neurocientíficas revelan que las áreas de asociación almacenan información de modalidad específica [212], [213]. Igualmente, las zonas de convergencia de nivel superior, situadas en los lóbulos temporal, parietal y frontal, integran el conocimiento de categorías de distintas modalidades junto con el nombre (etiqueta) de dicha categoría [210]. Las áreas de asociación no se corresponden con una modalidad específica, lo que sugiere la importancia de los símbolos amodales en la representación del conocimiento. Por lo general, las zonas de convergencia de nivel superior integran la información de las neuronas asociativas localizadas en zonas de convergencia anterior, correspondiente a una modalidad específica. En todo el cerebro, las zonas de convergencia integran el conocimiento de varias maneras, un conocimiento de categorías así tiene un carácter multimodal de los miembros de la categoría. Por ende, todas las características pertinentes de las diversas modalidades relativas a una determinada categoría se integran juntas, dando lugar a la recuperación de todas ellas en caso de necesidad.

2.3.4. Estructuras de categorías asociadas al conocimiento

Las categorías asociadas al conocimiento no se almacenan de manera desestructurada y heterogénea. Por el contrario, están almacenadas en diferentes estructuras organizadas de diversos modos, como ejemplares, reglas, conocimiento de base, prototipos y esquemas. A continuación, se detallan algunos de ellos.

2.3.4.1. Ejemplares y reglas

Las estructuras más sencillas que contienen conocimiento para una determinada categoría son los recuerdos de miembros individuales pertenecientes a dicha categoría. A estos miembros se les denomina ejemplares [214], [215]. Por otra parte, las reglas son definiciones precisas de los criterios de una categoría. Diversos estudios realizados indican que diferentes sistemas cerebrales son activados para representar ejemplares y para representar reglas [216]–[218]. Por otra parte, los sistemas particulares que se activan refuerzan la idea de que el conocimiento de categorías se representa en áreas de modalidad específica: las áreas visuales representan el contenido de los ejemplares y las áreas motoras efectúan el proceso de ensayo de reglas [218].

2.3.4.2. Prototipos

Los prototipos especifican qué propiedades son más probables para la pertenencia a una determinada categoría. Para ello, mediante la identificación de las propiedades que aparecen más frecuentemente en los miembros de una misma categoría, y la exclusión de las propiedades poco frecuentes, es posible generalizar los prototipos. [219]–[221]. Si una categoría tiene asociada un prototipo, los miembros similares al prototipo se consideran miembros típicos de la categoría. Los miembros que difieren del mismo se consideran atípicos. Sin embargo, el tipismo no es una condición binaria o absoluta, ya que en general los miembros de una categoría varían continuamente en similitud respecto al prototipo. Estos gradientes de tipismo aparecen en las distintas categorías [222], [223], incluso en aquellas que contienen reglas precisas para su definición [224].

Los gradientes del tipismo influyen en el procesamiento de las categorías. Por ejemplo, está demostrado que los individuos típicos son clasificados mentalmente más rápido que los atípicos [225] y de manera más precisa [226], que los miembros típicos de una categoría se aprenden más fácilmente que los atípicos [227], y que las deducciones más sólidas respecto a los miembros de una categoría son aquellas que se basan en los miembros típicos [228].

2.3.4.3. Conocimiento de base y esquemas

Por lo general, para establecer un recuerdo concreto de un ejemplar perteneciente a una categoría, el objeto percibido simplemente se añade a un conjunto de ejemplares memorizados de dicho objeto. Para actualizar una regla o prototipo, las nuevas propiedades se integran en la información establecida previamente. En cualquier proceso de aprendizaje, las propiedades se acumulan en condiciones de aislamiento relativo, sin embargo, diversos investigadores consideran que las propiedades activan un conocimiento de base en la memoria que especifica cómo se originan las propiedades y cómo se relacionan unas con otras [229]–[231]. Al parecer, las propiedades se procesan dentro de un contexto de conocimientos asociados.

Como ejemplo de conocimiento de base, en un experimento Barsalou [232] proporcionó el nombre de varios objetos a una serie de sujetos de prueba, y se les pidió identificar una categoría común para los mismos (roca, silla, ladrillo y maceta). Partiendo de un conocimiento base como que el día es cálido y ventoso y se necesita mantener una puerta abierta, se puede asociar instantáneamente a los diferentes objetos como pertenecientes a la categoría “objetos que permiten tener la puerta abierta en un día ventoso”. En este caso, la categoría no se identifica hasta que se presenta el conocimiento de base apropiado.

Como estructura de representación el conocimiento de base, se puede considerar el esquema. Un esquema es una representación estructurada que capta la información que por lo general se aplica a una situación o un suceso [233]–[235]. Los esquemas se consideran estructurados puesto que no son listas de propiedades independientes,

sino que establecen conjuntos coherentes de relaciones que asocian propiedades. Los esquemas se asemejan a reglas y prototipos en el hecho de que resumen los miembros de una categoría, pero difieren de los mismos al contener mucha información no esencial para clasificar entidades, pero importante para entender los sucesos asociados a las mismas, de modo que se pueden obtener deducciones útiles acerca de por qué un objeto puede formar parte del entorno y cómo interactuar con él.

2.3.4.4. Representación dinámica

No toda la información posible de una categoría se activa al acceder a la misma, sino que se activa preferentemente la información oportuna acorde al contexto o entorno. La representación dinámica se refiere a la capacidad del sistema cognitivo para construir y recuperar muchas representaciones diferentes para una categoría, cada una de las cuales sintetiza el conocimiento de la categoría más relevante para el contexto dado [236], [237]. Por tanto, la información más activa varía en cada ocasión, siendo por lo general la de mayor interés acorde al entorno.

2.4. Memoria

El recuerdo de los elementos del entorno a lo largo de la vida es una forma fundamental de cognición que guía la conducta [71]. La memoria se basa en un conjunto de procesos mediante los cuales la información es codificada, consolidada y recuperada. La memoria es esencial para el funcionamiento y para la supervivencia de diferentes organismos vivos. Sin memoria resultaría imposible aprender de la experiencia, establecer comunicación a través del lenguaje o incluso desarrollar tareas motoras.

El tipo de memoria utilizada en estos casos es la memoria a largo plazo. Esto es, la información que se adquiere en el transcurso de una experiencia y que persiste, de modo que es posible recuperar dicha información mucho después de la ocurrencia de la experiencia [71].

2.4.1. Memoria a largo plazo

Los investigadores han determinado diferentes tipos de memoria a largo plazo, las cuales difieren en sus propiedades básicas de procesamiento de la información, y en las estructuras del cerebro que permiten dicho procesamiento. Estas formas de memoria pueden clasificarse en dos grandes categorías generales [238]: memoria declarativa y memoria explícita (ver Ilustración 9, adaptada de [239]).

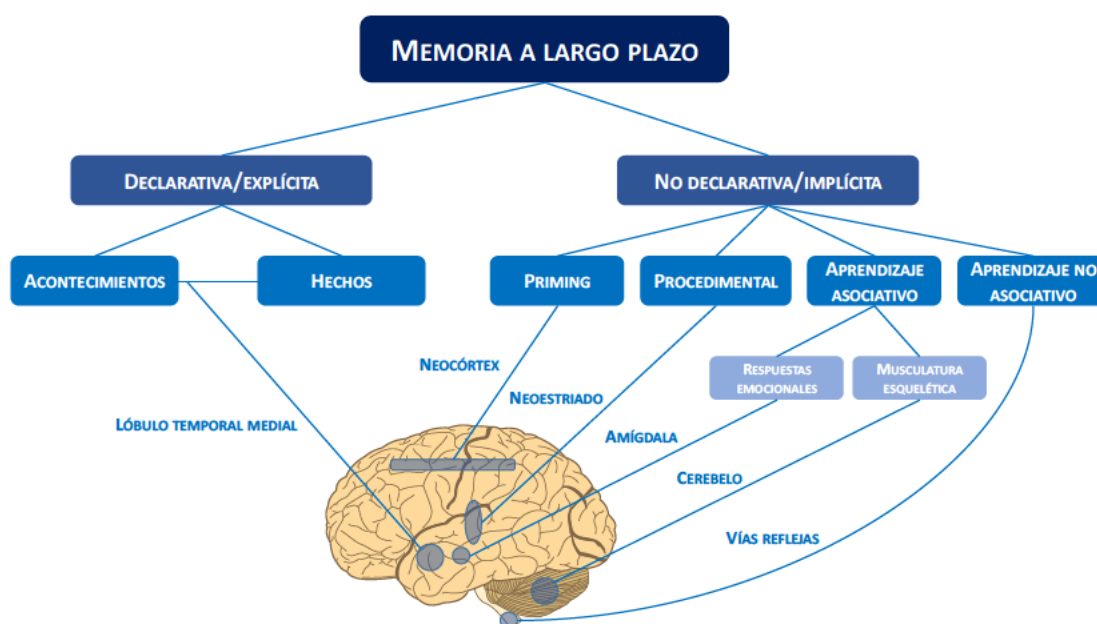


Ilustración 9. Modalidades de memoria a largo plazo

La memoria declarativa, o explícita, se refiere a las formas de memoria a largo plazo que por lo general pueden recordarse de forma consciente y describirse a otras personas, como por ejemplo la memoria de hechos, acontecimientos e ideas. La memoria declarativa incluye la memoria episódica: memoria en la que se almacena el conocimiento consciente de acontecimientos o episodios datados temporalmente, localizados espacialmente y experimentados personalmente; y la memoria semántica: memoria que almacena el conocimiento general relativo a los elementos del entorno, palabras y conceptos, sus propiedades e interrelaciones. [240]. La diferencia principal entre ambos tipos de memoria es el contexto o su ausencia.

La memoria episódica tiene un contexto y un estado temporal, de forma que, al recordar, se accede a la información relativa a la experiencia en particular. Sin embargo, la memoria semántica no está ligada al contexto específico en el que se adquirieron los conocimientos sobre el elemento en cuestión, ya que probablemente se acumuló en múltiples experiencias en diversos contextos. Las pruebas que evalúan la memoria declarativa se denominan pruebas de memoria explícita porque requieren la recuperación de una descripción explícita basada en la memoria. La memoria declarativa es muy flexible e implica asociar múltiples elementos de información formando una representación de memoria unificada [240].

La memoria no declarativa, o implícita, está relacionada con formas no conscientes de memoria a largo plazo manifestadas como un cambio de conducta sin que haya un recuerdo consciente. Las pruebas de memoria implícita no requieren una descripción del contenido de la memoria, sino que lo revelan implícitamente mediante los cambios observados en el rendimiento, como por ejemplo la adquisición gradual de una habilidad motora (p. ej. aprender a conducir, a montar en bicicleta o el desarrollo de hábitos). En comparación con la memoria declarativa, la memoria implícita suele ser más restrictiva en cuanto al modo en el que se puede recuperar el conocimiento [238].

2.4.2. Codificación

La codificación hace referencia los diferentes procesos mediante los cuales la información se transforma en una representación de memoria. Dichos procesos se ejecutan durante la experiencia, dando lugar a una representación mental que registra alguno o algunos aspectos de dicha experiencia [241].

Tanto la memoria declarativa como la no declarativa comienzan con la codificación de información, sin embargo, al ser la memoria episódica la asociada con momentos temporales en la vida de cada persona, es una base coherente para estudiar el funcionamiento de la codificación. Al analizar los factores que fortalecen el proceso de codificación, se han identificado como importantes el grado de atención y el nivel de elaboración, o análisis del significado [242]–[244]. La elaboración implica interpretar la información, relacionarla con otra información y ejecutar un proceso de reflexión. Igualmente, la codificación es influenciada por la recuperación consciente de la información y la ejecución o práctica espaciada en el tiempo [245].

La atención es un elemento muy importante a la hora de codificar la información obtenida del medio. Muchos casos de olvido, como nombres o lugares en los que se ha dejado un objeto (como las llaves), son simplemente la consecuencia natural de una codificación ineficaz de una experiencia, para transformarla en una memoria episódica en primer lugar.

Uno de los motivos principales de una codificación deficiente es el fallo en atender a un acontecimiento mientras está sucediendo. Cuando se presta menos atención a un acontecimiento, por ejemplo, al estar distraído, la codificación de la información es más débil y los intentos posteriores para recordarla serán ineficaces. En una serie de experimentos [243] se les pidió recordar quince palabras presentadas auditivamente a un grupo de sujetos de prueba, en dos condiciones: la condición de atención plena, donde la única tarea era recordar las palabras, y en la condición de atención dividida, donde los sujetos tenían que vigilar la posición de un asterisco en un monitor mientras escuchaban las palabras y a la vez tenían que presionar un botón en función de la localización del asterisco. En atención plena, los sujetos recordaron un promedio de nueve palabras, mientras que en atención dividida recordaron un promedio de sólo cinco palabras. Estudios de neuroimagen apoyan también la teoría que para realizar una codificación efectiva es necesaria la atención, ya que la pauta de activación neural durante la codificación en atención plena es diferente de la pauta en atención dividida [246], [247].

La codificación es una consecuencia inmediata de la atención a un estímulo y procesarlo [248]. Lo que influye en la eficacia de la codificación es el modo en el que se procesa el estímulo, esto es, las circunstancias en que se conoce inicialmente la información, y no la razón por la que se realizó el procesamiento. Hay diferencia si se presta o no atención y si se elabora o no la información. Existen otros factores que determinan la fortaleza de la representación codificada dependiendo del procedimiento empleado. Un procedimiento se vale del efecto de generación [249], por el cual el aprendizaje episódico es mejor si se puede generar la información objetivo a partir de la memoria en lugar de a partir de información compartida por otra persona. Otro

procedimiento utiliza el efecto de espaciamento [250], por el que la codificación a lo largo de múltiples ensayos de estudio de la misma información es óptima cuando se sigue una pauta determinada de secuencia temporal de los ensayos de estudio [65], [66], [69], [71]–[73], [75].

El efecto de generación describe el fenómeno observado [249] de que es más probable recordar la información que uno mismo ha recuperado o generado (durante el estudio) que la información que simplemente se recibe e intenta memorizar. Dicho efecto es una demostración experimental de la idea generalmente aceptada de que las cosas se aprenden mejor haciéndolas. Se cree que generar la información en base a la memoria es un fenómeno de codificación más potente que sólo procesar la información presentada externamente, porque para generar se requiere tanto elaboración como más atención. Esta conclusión proviene de estudios de neuroimagen [251]. Igualmente se sugiere que la generación es un modo particularmente eficaz de aprender debido a que es más probable que los procesos involucrados durante la generación inicial en la codificación sean solapados con los que se requieren para generar la información a partir de la memoria en la recuperación.

A la hora de memorizar elementos, existen diferentes técnicas. Dos de ellas consisten en: revisar repetidas veces el mismo elemento antes de pasar al siguiente elemento a analizar, y la otra, en revisar un conjunto completo de elementos y luego volverlo a repasar repetidas veces. La primera técnica para abordar el problema, en la cual se realizan varios intentos con el mismo estímulo sin interrupción se conoce como práctica masiva. La segunda, en la que los intentos con el mismo estímulo están separados por otros estímulos se denomina práctica distribuida. Hermann Ebbinghaus estudió los efectos de la práctica masiva frente a la distribuida, identificando un efecto de espaciamento [250]. Tal y como lo describió Ebbinghaus: “es indudablemente más ventajoso utilizar, con una serie considerable de repeticiones, una distribución conveniente de ellas en un espacio de tiempo, que amontonarlas en una única ocasión”. Es decir, para una codificación más efectiva es conveniente realizar una práctica distribuida. Se puede justificar este efecto de espaciamento mediante varias razones. Por ejemplo, cuando los intentos de estudio se agrupan todos juntos, es menos probable prestar atención completa a cada presentación. De hecho, a medida que se suceden las presentaciones es más probable la disminución de la atención sobre cada elemento del estudio al creer que ya está memorizado convenientemente. Por otra parte, al espaciar los intentos, el contexto en el cual se procesa el estímulo ha cambiado en mayor grado que la práctica masiva, con lo que se consigue una representación de memoria más rica y rutas de recuperación adicionales hacia la memoria. Finalmente, la práctica distribuida fomenta una mayor variabilidad de codificación, consistente en la codificación de diferentes aspectos de un estímulo al seleccionar diferentes características para su codificación en ocasiones posteriores. Es más probable recordar un estímulo al procesarlo de diferentes maneras en los sucesivos ensayos del estudio.

El rasgo característico distintivo de la codificación episódica es ligar las diversas características de un estímulo o acontecimiento, formando una representación de memoria integral [252]. Al enfrentarse a un elemento en un contexto determinado, el ser humano codifica diversas características del mismo de una forma más o menos

satisfactoria, como son los aspectos perceptivos de la apariencia visual, sonidos, contexto espacial y temporal, etc. Dichas características son procesadas por distintas redes neurales del cerebro. Así, la memoria integral de dicha experiencia requiere ligar los distintos elementos de información que la constituyen. En el ser humano se ha demostrado que dicho ligamento de información se realiza en el lóbulo temporal medial del cerebro [253].

2.4.3. Consolidación

Las memorias episódicas codificadas experimentan un fenómeno denominado consolidación, o fijación de la memoria. Dicho proceso permite que con el tiempo se vuelvan más estables e independientes de los lóbulos temporales mediales del cerebro. Al parecer, las interacciones entre el lóbulo temporal medial y algunas regiones corticales laterales permiten almacenar memorias fuera de los lóbulos temporales mediales formando lentamente conexiones directas entre las representaciones corticales de la experiencia. Según algunos investigadores, la consolidación de la memoria en la corteza cerebral ocurre mediante un proceso de repetición llevado a cabo durante el estado de sueño y durante el recuerdo, en los cuales el lóbulo temporal medial repite la pauta de activación existente durante el aprendizaje, reforzando las conexiones directas entre las regiones corticales laterales oportunas [254], [255]. De esta manera, una vez consolidadas, las memorias pueden ser directamente recuperadas de las regiones corticales [256].

2.4.4. Recuperación

El fenómeno de recuperación produce la experiencia subjetiva de recordar conscientemente episodios del pasado. La recuperación es un fenómeno complejo compuesto de diversos mecanismos. La recuperación episódica se trata de un fenómeno cognitivo que transforma el estado mental actual, de manera que el presente establece conexiones con el momento correspondiente en el que fueron identificados los elementos de información que componen la memoria recuperada. Elementos como la cara de una persona o el olor característico de un tipo de flor pueden actuar como desencadenante de una cascada de procesos que traen de vuelta a la mente multitud de detalles asociados con dichos elementos [252]. Estos elementos son denominados claves de recuperación.

Como se ha observado, las memorias episódicas se codifican ligando diversas características de un estímulo o acontecimiento en una representación integral, de forma que una memoria episódica consiste en una combinación de características relacionadas. Cualquiera de dichas características representa una posible vía a la memoria, lo cual aumenta el número de accesos al recuerdo, permitiendo su recuperación incluso con una cantidad limitada de información disponible. Cuando una clave de recuperación que corresponde a parte de la información codificada (como un aroma) se aloja en la representación almacenada, otras características ligadas a la

representación (como un lugar concreto) vuelven a activarse, ya que así se construye un todo a partir de la relación entre sus partes. Este proceso se conoce como conclusión de modelos [257].

Al igual que con la investigación de la codificación, al analizar los factores que se encuentran presentes cuando la recuperación se realiza con éxito se han podido identificar ciertos mecanismos que afectan al proceso. Una de las conclusiones fundamentales, es que la recuperación depende de la clave [258], o dicho de otra forma, la recuperación se estimula por pistas y claves procedentes del estado del entorno y del estado propio o interno. Cuando no están disponibles las claves o no son utilizadas, es menos probable que la recuperación de lugar a la conclusión de modelos. En muchas de las ocasiones, el olvido se produce porque las claves utilizadas para explorar la memoria no son eficaces, y no porque la información que se buscaba se haya borrado de la memoria. El contexto proporciona claves de recuperación muy sólidas, proporcionando recuerdos más fuertes y detallados en determinadas circunstancias que cuando simplemente se recuerda sin usar dichas claves. Este fenómeno determina que en la recuperación se produce un efecto dependiente del contexto. La recuperación es generalmente mejor cuando el entorno físico en el que se da coincide con el existente durante la codificación [258].

Los efectos dependientes del contexto apoyan la teoría de que la recuperación es dependiente de las claves y revelan una característica importante de la memoria episódica: al codificar un estímulo o acontecimiento, las características del entorno físico son ligadas habitualmente formando la representación de memoria episódica resultante. Esto genera otra vía para recordar dicho episodio. Si las mismas características están presentes en el entorno en el momento de la recuperación, sirven como claves adicionales para el recuerdo y aumentan la probabilidad de recuperar detalles adicionales. Igualmente, aspectos del estado interno, como por ejemplo el ánimo, alguna dolencia o la afección por algún tipo de droga, son codificados también en la memoria y proporcionan claves importantes para la recuperación. Se ha identificado que existen efectos dependientes del estado [259], [260]. Esto es, que la recuperación es mejor cuando el estado interno durante el proceso coincide con el que se dio durante la codificación de la información. Estos efectos son análogos a los efectos dependientes del contexto.

La memoria permite a un organismo la distinción entre un estímulo nuevo y otro anteriormente reconocido. La capacidad de reconocer objetos, lugares o personas se basa en dos procesos [261]: el recuerdo, o evocación consciente del contexto específico y sus detalles asociados en el momento en el que se reconoció el estímulo, y la familiaridad, o sensación subjetiva e inespecífica de haber reconocido antes un estímulo. Las teorías del doble proceso en el reconocimiento sostienen que éste puede basarse tanto en el recuerdo como en la familiaridad, identificando a ambos como procesos independientes. En el caso del recuerdo, se depende de los mismos mecanismos de conclusión de modelos que permiten evocar detalles episódicos asociados con una clave de recuperación, mientras que, en la familiaridad, en lugar del detalle se tiene en cuenta el parecido general. Un estímulo se coteja con la información en la memoria, y si el grado de parecido es suficiente, se puede recuperar la información sin detalles específicos. La investigación sobre el comportamiento soporta también esta distinción entre procesos de memoria [261]. Además se han

identificado algunas características, como que el recuerdo es un proceso más lento que la familiaridad, por lo que al tener que realizar una decisión de reconocimiento de forma rápida, se depende en mayor grado de la evaluación sobre la familiaridad del estímulo que de la evaluación del recuerdo, pues el conocimiento recordado implica más tiempo de recuperación, no siendo adecuado como para generar una decisión a tiempo [262], [263]. El recuerdo también depende de la atención en el momento de la codificación y la recuperación. Si la atención está dividida, la contribución del recuerdo a las decisiones del reconocimiento disminuye considerablemente [264], [265].

2.4.4.1 Problemas asociados a la recuperación

En algunas ocasiones los recuerdos están distorsionados acorde a las expectativas, y por tanto se recuerdan sucesos que en realidad nunca han ocurrido. Se han identificado diversos tipos de errores de memoria entre los que se incluyen el sesgo, la atribución errónea y la sugestión [266], [267].

2.5. Conclusiones del capítulo

A lo largo del capítulo se han expuesto las características fundamentales que conforman la cognición humana, considerando los principales estudios realizados en la materia. Cabe destacar la gran complejidad del sistema cognitivo, y de sus diferentes subsistemas, tanto a nivel estructural como funcional.

A fin de presentar de manera sintética los contenidos de los principales apartados del capítulo, se han desarrollado cuatro ilustraciones, mostradas a continuación

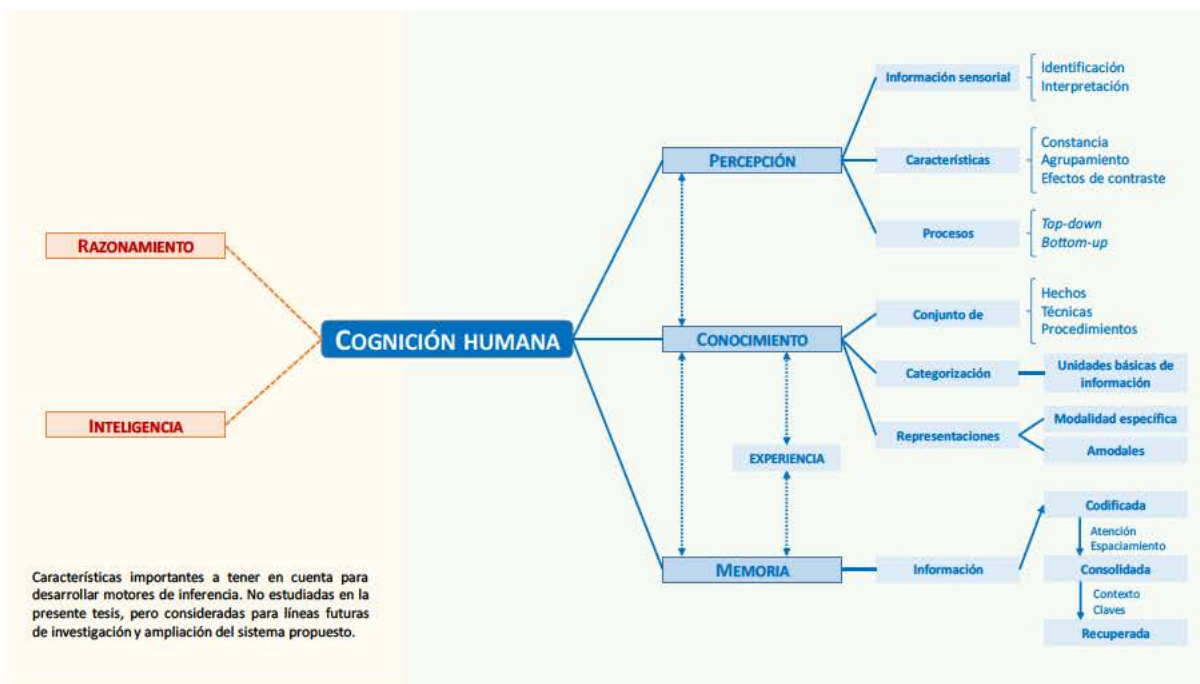


Ilustración 10. Esquema conceptual simplificado de la cognición humana

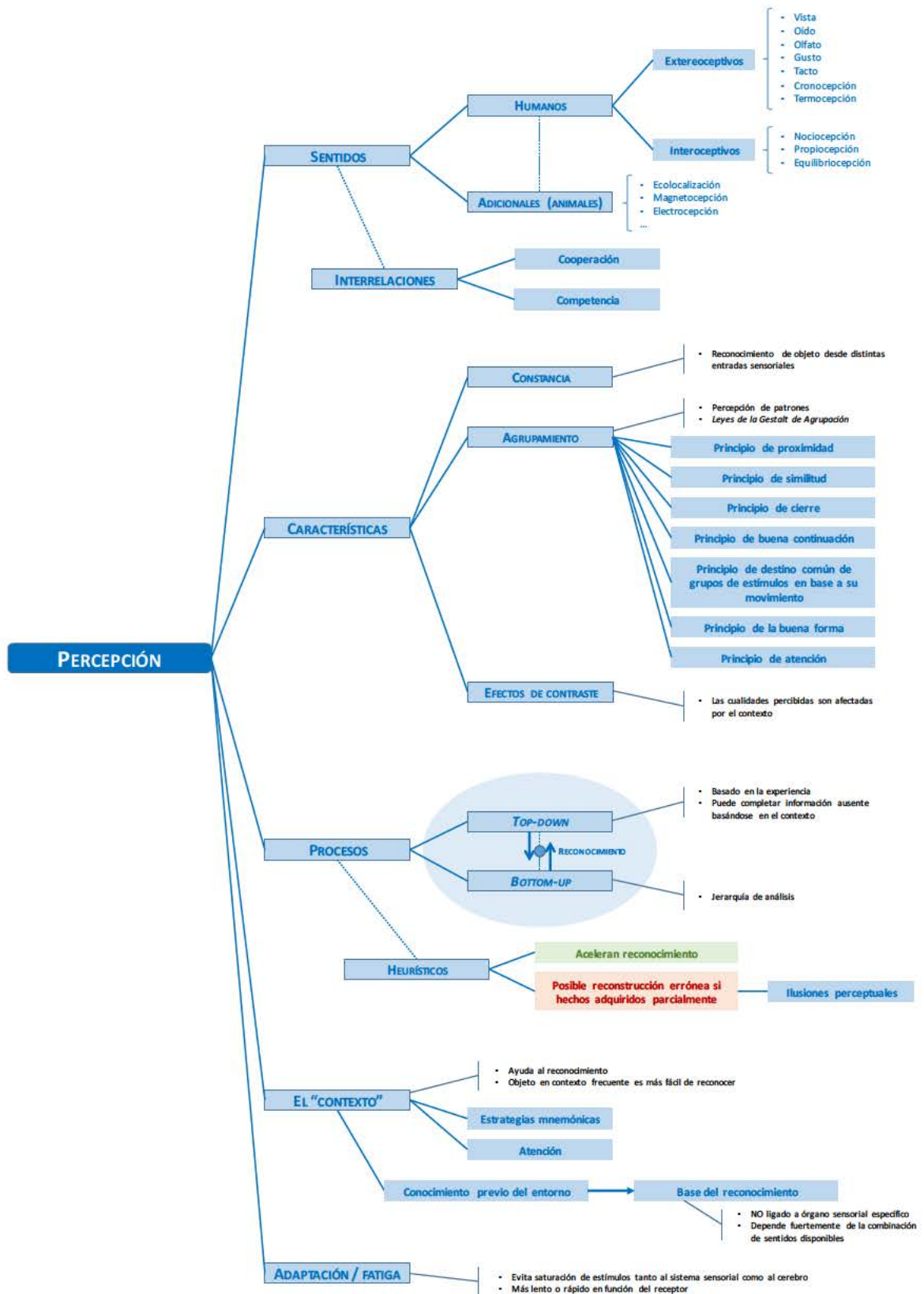


Ilustración 11. Esquema conceptual simplificado de la percepción

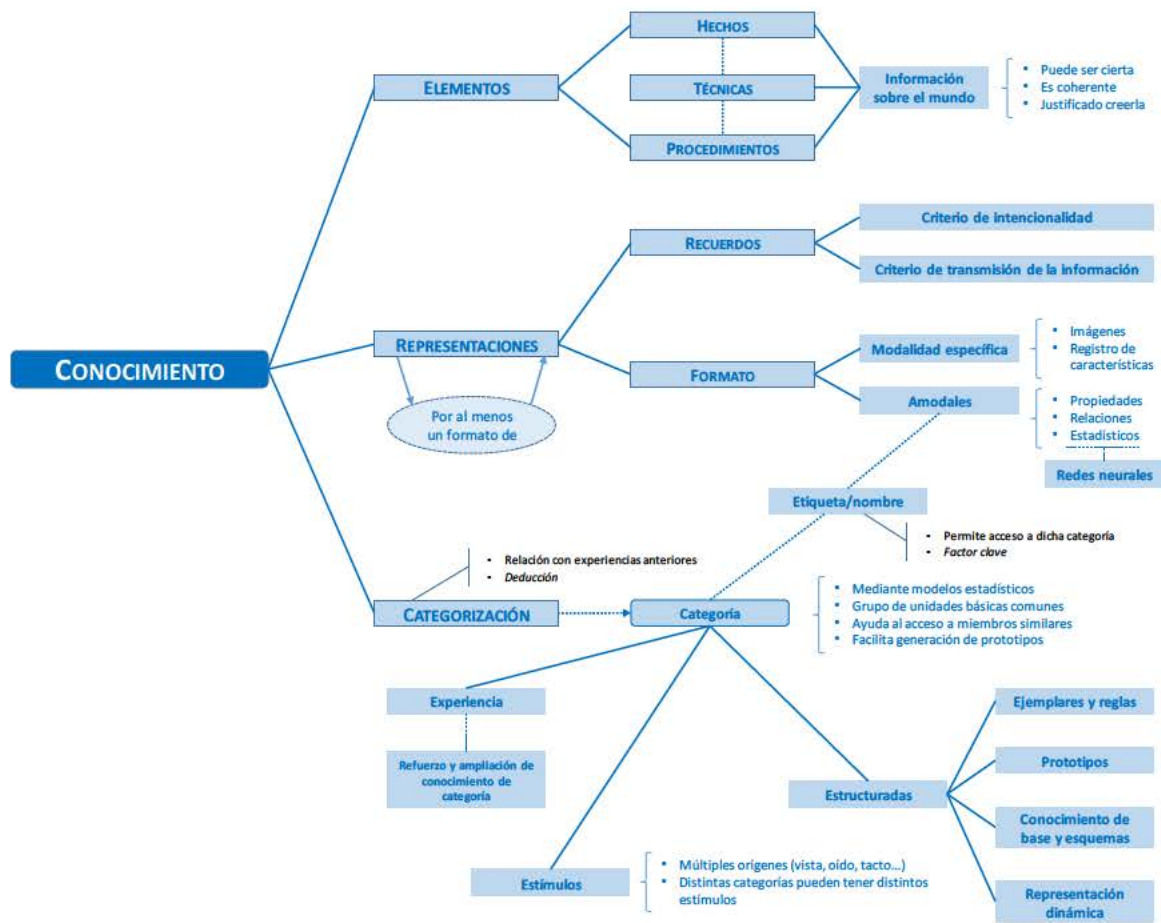


Ilustración 12. Esquema conceptual simplificado del conocimiento

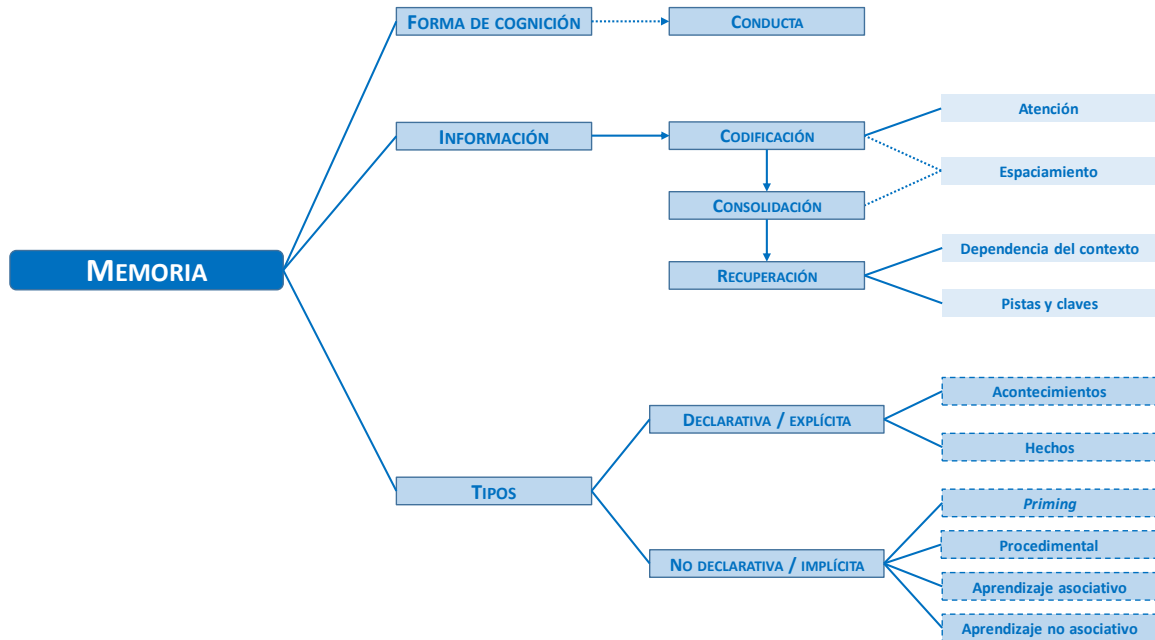


Ilustración 13. Esquema conceptual simplificado de la memoria

En el capítulo también se han analizado las características específicas relacionadas con los distintos sentidos perceptuales. Dichas características serán sintetizadas en la sección correspondiente al análisis de requisitos para desarrollar la solución, junto con el resto de características del sistema cognitivo.

Desde este punto, es más claro el siguiente paso hacia el desarrollo de un modelo que siga los principios elementales estudiados del sistema cognitivo y de sus diferentes y subsistemas. Mediante un análisis de las distintas soluciones tecnológicas y procedimientos existentes, relacionados con las características y los problemas derivados del sistema cognitivo, se determinarán las características y requisitos computacionales que deberá presentar el modelo a desarrollar. Por ello, el siguiente capítulo se centra en el estudio sobre el entorno computacional.

Capítulo 3. Entorno computacional

3.1. Introducción

Antes de abordar de lleno las diferentes tecnologías existentes que pueden ayudar en su conjunto a la resolución del problema propuesto, o al menos de parte del mismo, es conveniente realizar una reflexión y analizar las distintas características conocidas y aprendidas sobre la cognición. Mediante un asesoramiento inicial y un soporte posterior por parte de expertos en el área tecnológica, a los que se informa sobre las conclusiones derivadas del estudio anterior en el área de la cognición, se desarrolla una revisión sistemática de la literatura (guiada inicialmente por los expertos, como en el capítulo anterior), identificando distintos grupos que están trabajando en líneas similares de investigación, identificando dos grandes grupos principales, así como los enfoques seguidos por los mismos. Considerando estos resultados, sus características y los recursos de los que hacen uso, se procede a realizar un estudio en profundidad de las distintas tecnologías y tendencias relevantes para conseguir la formalización del modelo en el que se basa la presente tesis doctoral.

3.2. Equipos y líneas de investigación paralelas

Una vez conocidas y analizadas las diferentes características fundamentales de la cognición, principalmente en seres humanos, es conveniente tener una idea global sobre el panorama científico y empresarial respecto a esta área de investigación. Estudiar las tendencias, aproximaciones a soluciones y estado de las mismas.

La percepción computacional es una disciplina científica que se centra en las tecnologías mediante las cuales un computador puede reconocer el entorno en el que se encuentran sus usuarios, cuáles son las necesidades de dichos usuarios, y qué acciones deben llevar a cabo sin que el usuario tenga que proporcionar dicha información o proporcionar instrucciones detalladas [268]. Esta disciplina tiene mucho que ver con la ciencia cognitiva, y con sus orígenes en la cibernética (1943 – 1953) [269]. En los sistemas cognitivistas, la percepción proporciona la interfaz entre el mundo externo y la representación simbólica de dicho mundo. La percepción proporciona una abstracción fiel de las características espacio-temporales del entorno a través de los datos sensoriales. En los sistemas cibernéticos, la percepción también ayuda a determinar cambios en el estado del sistema, y en las perturbaciones ambientales, con lo que pueden tomarse las medidas oportunas a fin de mantener la estabilidad.

Existen numerosas iniciativas y sistemas que se apoyan en modelos perceptuales a fin de proporcionar características específicas, resolver los problemas para los que han

sido diseñados y brindar funcionalidades avanzadas respecto a los sistemas clásicos. En cualquier caso, casi la totalidad de los sistemas existentes se centra en la percepción parcial, combinando, en los casos más completos y complejos, un pequeño subconjunto de áreas sensoriales de las que extraer conocimiento sobre el entorno y/o sobre el propio sistema. Estos sistemas pueden clasificarse en dos grandes grupos: sistemas físicos y sistemas lógicos. Los sistemas físicos son los que mayor relación guardan con los sistemas cibernéticos y la robótica. El uso de la percepción se suele emplear para recoger información del entorno a través de un diverso número de sensores, distribuidos en configuraciones específicas sobre un sistema electro-mecánico. Los segundos, o sistemas lógicos, por lo general se corresponden con sistemas software bastante complejos, que principalmente utilizan uno o varios tipos de sensores para facilitar la interacción con los usuarios. De esta manera, se pueden proporcionar servicios de búsqueda, manejo de herramientas y, en general, un amplio abanico de interacción y nuevas interfaces, basado en el conocimiento específico sobre el usuario y el entorno, inferido por medio del sistema computacional.

A continuación, se describen los principales sistemas e iniciativas correspondientes a ambos grupos (físicos y lógicos), junto con las tendencias de investigación en el área, llevadas a cabo por las compañías líderes en el sector.

3.2.1. Algunos sistemas cognitivos que hacen uso de la percepción computacional

Dentro de los sistemas físicos desarrollados en los últimos años, y basados en arquitecturas cognitivas que hacen uso de la percepción computacional para su correcto desempeño, cabe destacar una serie de iniciativas concretas.

EPIC [270] (*Executive Process-Interactive Control*) presenta una arquitectura cognitiva especialmente adecuada para modelar la realización de tareas humanas multimodales. La arquitectura EPIC incluye procesadores sensoriales motores periféricos que rodean a un procesador cognitivo de reglas de producción, y ha sido utilizado para desarrollar modelos computacionales precisos en una amplia variedad de situaciones interactivas entre humanos y computadores. EPIC requiere tanto el conocimiento encapsulado en las reglas de producción, como los parámetros perceptivo-motores. Sobre los parámetros, se definen dos tipos principales: los estándares (o de sistema), fijos para todas las tareas (p. ej. duración de un ciclo de producción en el procesador cognitivo, correspondiente a 50ms) y parámetros típicos, que tienen valores convencionales, pero pueden variar entre las tareas (p. ej. el tiempo necesario para realizar reconocimiento visual).

EPIC está compuesto por un procesador cognitivo (con un intérprete de reglas de producción y una memoria de trabajo), un procesador auditivo, un procesador visual, un procesador oculomotor, un procesador de síntesis de voz, un procesador táctil y un procesador motriz manual. Todos los procesadores funcionan en paralelo. Los procesadores perceptuales simplemente modelan los aspectos temporales de la percepción, pero no realizan ningún proceso perceptivo en sí mismos. Por ejemplo, el procesador visual no realiza reconocimiento de patrones. En cambio, sólo modela el

tiempo que tarda una representación de un estímulo dado en transferirse a la memoria declarativa. Un estímulo sensorial dado puede tener varias representaciones posibles (por ejemplo, color, tamaño, etc.) con cada representación entregada a la memoria de trabajo en momentos diferentes. De manera similar, los procesadores motrices no se ocupan de generar los pares necesarios para la producción de cierto movimiento; En cambio, sólo se ocupan de estimar el tiempo que se tarda en producir una salida del motor después de que el procesador cognitivo la haya solicitado. Finalmente cabe destacar que EPIC no tiene implementado ningún mecanismo de aprendizaje.

ACT-R [271], [272] (*Adaptive Character of Thought – Rational*), la cognición artificial surge de una interacción de conocimiento procesal y declarativo. El conocimiento procedimental está representado en unidades llamadas reglas de producción, y el conocimiento declarativo está representado en unidades llamadas *chunks* (piezas). Las unidades individuales se crean mediante codificaciones simples de objetos en el entorno (*chunks*) o simples codificaciones de transformaciones en el entorno (reglas de producción). Un gran número de estas unidades de conocimiento son análogas a las comprendidas por la cognición humana. A partir de esta gran base de datos, las unidades apropiadas son seleccionadas para un determinado contexto mediante procesos de activación, que se ajustan a la estructura estadística del entorno. Según la teoría de ACT-R, el poder de la cognición humana depende de la cantidad de conocimiento codificado y del despliegue efectivo de dicho conocimiento codificado.

ACT-R se centra en la descomposición modular de la cognición y ofrece una teoría sobre cómo dichos módulos se integran para producir cognición coherente. Respecto a la percepción, ACT-R únicamente utiliza un módulo visual (para determinar la identidad y posición de los objetos en el campo visual) y un módulo para controlar las manos del sistema.

ICARUS [273]–[275] es una arquitectura cognitiva que explota (al igual que EPIC y ACT-R) las representaciones simbólicas del conocimiento, el uso del reconocimiento de patrones para determinar elementos relevantes de conocimiento, la operación acorde con el ciclo definido de reconocimiento-acción, y un enfoque incremental al aprendizaje.

ICARUS distingue entre conceptos y habilidades, y define dos tipos diferentes de representación y memoria para ellos, con variantes a largo y corto plazo. La memoria conceptual codifica el conocimiento sobre clases generales de objetos, y relaciones entre ellos. La memoria de habilidades codifica el conocimiento sobre maneras de actuar y alcanzar metas. Además, ICARUS adopta una organización jerárquica fuerte para su memoria a largo plazo, con la memoria conceptual dirigiendo la inferencia “*bottom-up*” y la memoria de habilidades estructurando la selección descendente de acciones. En la arquitectura general de ICARUS se habla sobre la necesidad de la percepción, y se incluye un buffer perceptual relacionado con la memoria de creencias, sin embargo, no se detalla cómo se realiza dicha percepción ni cómo se aprende o estructura la información aprendida de manera perceptual. Sólo parece hacerse hincapié en la habilidad de generar inferencias en base al conocimiento previo adquirido.

La fundación de ADAPT [276], arquitectura cognitiva para la robótica, se basa en el argumento propuesto por sus autores, consistente en que las arquitecturas cognitivistas existentes tales como, ACT-R y EPIC, no soportan fácilmente ciertos paradigmas de robótica convencional como la adaptación dinámica y la percepción activa. Gran parte del software en robótica comprende varios comportamientos simultáneos de comunicación en tiempo real y distribuidos. Estas arquitecturas no son adecuadas, ya que su enfoque se centra principalmente en estrategias de "búsqueda secuencial y selección", sus mecanismos de aprendizaje se centran en componer acciones secuenciales en lugar de concurrentes y tienden a organizarse de forma concurrente en lugar de distribuida.

En ADAPT se presenta una arquitectura cognitiva con capacidades adaptativas dinámicas y percepción activa. También adopta características de ACT-R (como la memoria declarativa a largo plazo, en la que se almacenan esquemas sensomotores para controlar la percepción y la acción), y de EPIC (todos los procesos perceptuales disparan en paralelo), pero los datos sensoriales de bajo nivel se colocan en la memoria de trabajo a corto plazo donde son procesados por el mecanismo cognitivo. ADAPT tiene dos tipos de objetivos: de tareas (como "encontrar el objeto con determinadas características") y de arquitectura (como "iniciar una acción para reconocer la escena"). También presenta dos tipos análogos para las acciones: de tareas (como "recoger el objeto con determinadas características") y arquitectónicas (como "iniciar una acción para realizar un agarre"). Los objetivos arquitectónicos y las acciones están representados mediante producciones, mientras que los objetivos y acciones de la tarea están representados de forma declarativa, tanto en la memoria de trabajo como en el procedimiento.

ADAPT utiliza RS (*Robot Schemas*), para la representación perceptual y la planificación. Los esquemas son patrones generales de percepción y acción, y se almacenan permanentemente en la memoria de trabajo. Cada esquema tiene una parte sensorial y una parte motora, representando el objetivo de ejecutar un movimiento específico dependiendo de determinadas condiciones perceptuales. Si bien, tampoco se entra en detalle sobre cómo se trata y combina la información perceptual para generar conocimiento. Al igual que en ICARUS, parece un sistema adaptativo que aprende a actuar en base a lo que percibe, pero sin detenerse a generar conocimiento semántico conceptual sobre el entorno percibido (por lo que, aun siendo muy complejo, puede considerarse un sistema reactivo).

Global Workspace Cognitive Architecture [277]–[279] propone una arquitectura cognitiva neuronal, biológicamente plausible, inspirada en el cerebro. Permite funciones cognitivas como la anticipación y la planificación, a través de la simulación interna de la interacción con el entorno. La arquitectura se basa en un bucle sensomotor externo y un bucle sensomotor interno en el que la información pasa a través de múltiples áreas corticales que compiten en un espacio de trabajo global. En contraste con la manipulación de representaciones simbólicas declarativas, como lo hacen las arquitecturas cognitivistas, la función cognitiva se logra aquí a través de mapas neurales topográficamente organizados que pueden ser vistos como una forma de representación analógica, o icónica. En el modelo se observa que estas representaciones analógicas son particularmente apropiadas en la cognición espacial, que es una capacidad cognitiva crucial, pero que es notoriamente difícil con los

enfoques tradicionales basados en la lógica. Se defiende que la brecha semántica entre la entrada sensorial y las representaciones analógicas es mucho menor que con las representaciones simbólicas de lenguaje y, por lo tanto, se minimiza la dificultad del problema de la conexión de símbolos. La arquitectura cognitiva propuesta por Shanahan, se basa también en la importancia fundamental del paralelismo como componente constituyente en el proceso cognitivo en lugar de ser un mero problema de implementación. En la arquitectura, muchos procesos paralelos separados y especializados compiten y cooperan para el acceso a un espacio de trabajo global. Aquellos que ganan, pasan a controlar el acceso al espacio de trabajo global y luego puede retransmitir la información a los otros procesos especializados (competencia). La arquitectura cognitiva de Shanahan se compone de los siguientes componentes: un bucle sensoriomotor de primer orden, cerrado externamente a través del mundo, y un bucle sensitivo-motor de orden superior, cerrado internamente a través de memorias asociativas. El lazo de primer orden comprende la corteza sensitiva y los ganglios basales (control de la corteza motora), proporcionando conjuntamente un subsistema reactivo de selección de acciones. El bucle de segundo orden comprende dos elementos de corteza asociativa que llevan a cabo simulaciones sobre el comportamiento sensorial y el sistema motor. La primera corteza asociativa simula una salida del motor, mientras que la segunda simula el estímulo sensorial que se espera por una salida dada del motor. El bucle de orden superior modula la selección de la acción de los ganglios basales en el bucle de primer orden a través de un componente de amígdala, impulsado por el afecto. Así pues, esta arquitectura cognitiva permite anticiparse y planificar el comportamiento potencial, a través de la simulación sensoriomotora asociativa interna. El espacio de trabajo global no se corresponde con ningún área cortical específica, más bien es una red global de comunicaciones. La arquitectura se implementa como un sistema conexionista que utiliza memorias de acceso aleatorio generalizadas (G-RAM).

SDAL (*self-directed anticipative learning*) [280] describe un enfoque interactivista-constructivista (I-C) para modelar la inteligencia y el aprendizaje como una forma dinámica encarnada de adaptación y explora algunas aplicaciones de I-C para comprender la forma en que el aprendizaje cognitivo es llevado a cabo en el cerebro. En SDAL se proponen dos ideas clave para la conceptualización de la inteligencia dentro de este marco: la inteligencia se ocupa centralmente de la capacidad de una gestión coherente, sensitiva al contexto y auto dirigida de la interacción; y el modelo primario para el aprendizaje cognitivo es la construcción de habilidades anticipativas. La autodirección es una capacidad para la modulación del proceso integrador que permite que un sistema se "guíe" a sí mismo a través de su mundo, ajustando anticipadamente sus propios requisitos de viabilidad a la interacción con su entorno. Debido a que los procesos de interacción adaptativa requeridos de los sistemas inteligentes son demasiado complejos para que se pueda pre especificar la acción efectiva (por ejemplo, genéticamente), el aprendizaje es un componente importante de la inteligencia. Al presentar su enfoque, se utiliza el término "señales de normas explícitas" para las señales que un sistema utiliza para diferenciar un contexto apropiado que realiza una acción. Estas señales normalizadas reflejan las condiciones para el mantenimiento de la autonomía del sistema (p. ej. señales correspondientes al hambre agotan los niveles nutricionales). El conjunto completo de señales de normalización se denomina "matriz de norma". También se distinguen dos niveles de

gestión: de bajo y alto orden. La gestión de bajo orden emplea señales normalizadas, o dicho de otra forma, un pequeño conjunto de parámetros, para dirigir la acción (p. ej. un mosquito utiliza seguimiento de calor y seguimiento de gradiente de dióxido de carbono para buscar portadores de sangre). Gracias al uso de dicho escaso número de parámetros, el éxito depende en última instancia de la simple regularidad en el entorno. Estos parámetros también tienden a ser localizados en el tiempo y el espacio. Por otra parte, las estrategias de gestión de alto orden, también dependen de la regularidad del entorno, pero explotan parámetros que se extienden más en el tiempo y el espacio, y usan más aspectos del proceso interactivo, incluyendo la capacidad de anticipar y evaluar el rendimiento del sistema.

SASE (*self-aware self-aware*) [281]–[283] se centra específicamente en la cuestión del desarrollo, por lo que quiere decir que a priori, el procesamiento realizado por la arquitectura no se especifica o programa, sino que es el resultado de la interacción en tiempo real del sistema con el entorno, incluidos los seres humanos. Por lo tanto, la arquitectura no es específica de las tareas, desconocidas en el momento de creación de la arquitectura, sino que es capaz de adaptarse y desarrollarse para aprender tanto las tareas que se requieren como la manera de realizar las mismas.

La arquitectura implica una distinción importante entre los sensores y los actuadores que están asociados con el entorno (incluyendo el sistema físico, y por tanto la detección propioceptiva), y los sensores y actuadores que están asociados con el sistema cerebral o sistema nervioso central del sistema.

El “cerebro” de la arquitectura no es directamente accesible a los diseñadores una vez que se pone en funcionamiento. Así pues, la única manera de afectar a un agente de tipo SASE es a través de los sensores externos y actuadores. Por lo tanto, la arquitectura SASE es muy fiel a los paradigmas emergentes de la cognición: su origen y desarrollo evolutivo es fijo, de manera que el sistema únicamente puede aprender a operar en su entorno a través de la formación (aprendizaje).

La arquitectura cognitiva de Weng se basa en los procesos de decisión de Markov (MDPs), específicamente un proceso de decisión de Markov auto-consciente. Con ello se pretende proporcionar una serie de atributos adicionales, específicamente mayor abstracción, contextos auto-generados, y un mayor grado de integración sensorial. El ejemplo del sistema de visión descrito en [282] profundiza en la arquitectura cognitiva, detallando tres tipos de mapeo en el flujo de información dentro de la arquitectura: mapeo sensorial, mapeo cognitivo y mapeo motriz.

DARWIN [284]–[288] comprende un conjunto de dispositivos basados en el cerebro (BBDs, *brain-based devices*), los cuales explotan un sistema nervioso simulado que puede desarrollar memoria espacial y episódica, así como capacidades de reconocimiento a través del aprendizaje autónomo experiencial. El enfoque se diferencia de los demás proyectos conectivos en que la arquitectura está mucho más fuertemente modelada en torno a la estructura y organización del cerebro que las redes neuronales artificiales convencionales. Esto es, se centran en el sistema nervioso como un todo, sus partes constituyentes y su interacción, en lugar de basarse en una implementación neuronal de alguna función de memoria, control o reconocimiento individual. Los principales mecanismos neuronales de los dispositivos

son la plasticidad sináptica, un sistema de recompensa (o valor), la conectividad reentrante (*feedback*), la sincronización dinámica de la actividad neuronal y las unidades neuronales con propiedades de respuesta espaciotemporal. El comportamiento adaptativo se consigue mediante la interacción de estos mecanismos neuronales con correlaciones sensomotoras que han sido aprendidas de forma autónoma por la detección activa y el movimiento propio.

Darwin VIII es capaz de discriminar objetivos visuales razonablemente simples (formas geométricas coloreadas) asociándolo con una señal auditiva preferida innata. Su sistema nervioso simulado contiene 28 áreas neurales, aproximadamente 54.000 unidades neuronales y aproximadamente 1,7 millones de conexiones sinápticas. La arquitectura comprende regiones de visión, seguimiento, valor o saliencia y audición. El área de seguimiento, basada en proyecciones excitatorias de la región auditiva, determina la dirección de la mirada de la cámara de Darwin VIII. Esto hace que Darwin se oriente hacia una fuente de sonido. Darwin IX puede navegar y categorizar texturas basándose en la estructura artificial somatosensorial artificial simulada, inferida de las ratas, que comprende 17 áreas, 1101 unidades neuronales y aproximadamente 8400 conexiones sinápticas. Darwin X es capaz de desarrollar memoria espacial y episódica basada en un modelo del hipocampo y regiones circundantes. Su sistema nervioso simulado contiene 50 áreas neurales, 90.000 unidades neuronales y 1,4 millones de conexiones sinápticas. Incluye un sistema visual y un sistema de dirección de la cabeza entre otros. En Darwin X la visión se utiliza para reconocer objetos y luego calcular su posición, mientras que la odometría se utiliza para sentir la dirección de la cabeza.

HUMANOID [289] presenta una arquitectura cognitiva híbrida para la aplicación en un robot humanoide. Se fundamenta en la interacción de componentes paralelos basados en el comportamiento, que comprende un subsistema de percepción jerárquica de tres niveles, un sistema de manejo de tareas jerárquico de tres niveles, un subsistema de memoria a largo plazo basado en una base de datos de conocimiento global que utiliza diversos esquemas representacionales (incluyendo ontologías de objetos y modelos geométricos y modelos de Markov ocultos), un gestor de diálogo que media entre la percepción y la planificación de la tarea, un supervisor de ejecución y un subsistema de memoria de corto plazo de “modelos activos” al que tienen acceso todos los niveles de percepción y gestión de tareas. Estos modelos activos desempeñan un papel central en la arquitectura cognitiva: son inicializados por la base de conocimiento global y actualizados por el subsistema perceptual y pueden ser actualizados y reorganizados de forma autónoma. El subsistema de percepción comprende una jerarquía de tres niveles con módulos de percepción bajos, medios y altos. El módulo de percepción de bajo nivel proporciona interpretación de datos de sensores sin tener acceso a la base de datos de conocimiento del sistema central, típicamente para proporcionar un control de robot de bajo nivel reflejo. Se comunica tanto con el módulo de percepción de nivel medio como con el módulo de ejecución de tareas. El módulo de percepción de nivel medio proporciona una variedad de componentes de reconocimiento y se comunica tanto con la base de datos de conocimiento del sistema (memoria a largo plazo) como con los modelos activos (memoria a corto plazo). El módulo de percepción de alto nivel proporciona servicios de interpretación más

sofisticados como el reconocimiento de la situación, la interpretación de los gestos, la interpretación del movimiento y la predicción de la intención.

CEREBUS [290] se trata de una arquitectura cognitiva híbrida, que combina los principios de las arquitecturas basadas en el comportamiento, con algunas características de la IA simbólica, como la inferencia de encadenamiento hacia atrás y hacia delante mediante lógica de predicados. Representa un intento de mejorar robots basados en el comportamiento, sin recurrir a un sistema tradicional de planificación central. Combina un conjunto de sistemas sensoriales-motores basados en el comportamiento con una red semántica de paso de marcadores y una red de inferencia. La red semántica efectúa la memoria declarativa a largo plazo, proporcionando conocimiento reflexivo sobre sus propias capacidades, y la red de inferencia le permite razonar sobre su estado actual y procesos de control. Juntos, implementan la característica clave de la arquitectura de Cerebus: el uso del conocimiento reflexivo sobre sus sistemas perceptivo-motores, para realizar un razonamiento limitado sobre sus propias capacidades.

Cog [291] se trata de un torso humanoide, construido para explorar aspectos de la estructura del desarrollo, la incorporación física, la integración de múltiples sistemas sensoriales y motores, y la interacción social. El robot tiene veintiún grados de libertad y una variedad de sistemas sensoriales, incluyendo los sentidos visual, auditivo, vestibular, kinestésico y táctil. Presenta una serie de rutinas visual-motrices implementadas, como el seguimiento suave, reflejos vestíbulo-oculares y reflejos opto cinéticos. También presenta comportamientos de orientación, técnicas de control motor y comportamientos sociales.

La arquitectura propuesta para Cog se basa en la Teoría de la Mente [292], [293], descomponiendo el problema en conjuntos de habilidades precursoras y módulos de desarrollo. La Teoría de la mente de Leslie enfatiza módulos independientes de dominio específico para distinguir el comportamiento de los objetos inanimados, el comportamiento de los objetos animados y las creencias e intenciones de los objetos animados. La teoría de la mente de Baron-Cohen consta de tres módulos, uno de los cuales se ocupa de la interpretación de los estímulos perceptivos (visuales, auditivos y táctiles) asociados con el movimiento autopropulsado. Ambos alimentan un módulo de atención compartida, que a su vez recae sobre un módulo de Teoría de la Mente, el cual representa el conocimiento intencional o "estados mentales epistémicos" de otros agentes. El enfoque de la Teoría de la Mente para Cog, al menos inicialmente, se centra en la creación de las habilidades perceptivas y motoras precursoras sobre las cuales se puede construir una teoría más compleja de la capacidad mental: distinguir entre el movimiento inanimado y el movimiento animado e identificar la dirección de la mirada. Estos explotan varias capacidades visuales integradas tales como detección de color, detección de movimiento, detección de color de la piel y estimación de disparidad, un módulo visual de búsqueda y atención y control visual-motor, reflejo vestibular-ocular, así como movimiento de la cabeza y cuello. Los comportamientos visual-motores primitivos, por ejemplo, para encontrar caras y ojos, se resuelven mediante impulsos de motivación predefinidos y estrategias de búsqueda visual.

Kismet [294], [295] utiliza una cabeza articulada antropomórfica robótica con un total de 21 grados de libertad (orientación de la cabeza, dirección de la mirada, párpados,

cejas, labios y orejas). Kismet tiene un sistema de visión binocular y gran angular, consistente en cuatro cámaras en color, y dos micrófonos, uno montado en cada oreja. Kismet está diseñado para interactuar con las personas cara a cara, de forma natural y expresiva, percibiendo las señales sociales naturales y respondiendo a través de la dirección de la mirada, la expresión facial, la postura corporal y el balbuceo vocal. El objetivo final de Kismet es aprender de las personas a través del compromiso social, aunque Kismet no tiene ninguna capacidad adaptativa (es decir, de aprendizaje o de desarrollo) o anticipatoria. Kismet tiene dos tipos de motivaciones: impulsos y emociones. Los impulsos establecen los objetivos de nivel superior del robot, como puede ser involucrar a las personas (unidad social) o descansar (fatiga de unidad). El comportamiento del robot se centra en saciar sus impulsos básicos. Estas unidades tienen una constante de tiempo más larga en comparación con las emociones, y operan cíclicamente, aumentando en ausencia de interacción satisfactoria y disminuyendo con la habituación. El objetivo es mantener el nivel de la unidad en algún lugar de una región homeostática entre la estimulación y el exceso de estimulación. Así, el modelo tiende a hacer que el robot entre en contacto con cosas que promueven su "bienestar" y evita las que no lo hacen. Las emociones son desencadenadas por condiciones antecedentes pre-especificadas, que se basan en estímulos perceptivos, así como el estado actual de la unidad y el estado conductual.

Kismet tiene cinco módulos distintos en su arquitectura cognitiva: un sistema perceptivo, un sistema de emoción, un sistema de conducta, un sistema de accionamiento y un sistema motor. El sistema perceptivo comprende un conjunto de procesos de bajo nivel que detectan estímulos visuales y auditivos, realizan extracción de características (por ejemplo, color, movimiento, frecuencia), extraen descripciones afectivas del habla, orientan la atención visual y localizan rasgos relevantes como caras, ojos, objetos, etcétera. Estos son introducidos en un sistema perceptivo de alto nivel donde, junto con la entrada afectiva del sistema emocional, la entrada desde el sistema de conducción y el sistema de comportamiento, están unidos por procesos de liberación, que codifican el conjunto actual de creencias del robot sobre su estado y su relación con el entorno.

3.2.2. Algunos servicios (software) que hacen uso de la percepción computacional

La evolución tecnológica, y el aumento de las capacidades computacionales, como la memoria y la velocidad de los componentes, permiten que algoritmos diseñados tiempo atrás pero tecnológicamente inviables en su momento (debido a restricciones computacionales), sean aplicados y mejorados en la actualidad. Esto, combinado con la emergencia de nuevas tendencias tecnológicas (fundamentadas también por el masivo incremento en las comunicaciones y conectividad), como, por ejemplo, el uso de la computación en la nube (*cloud computing*), ha derivado en un gran número de iniciativas que prestan servicios avanzados, basados en estos enfoques algoritmos definidos hace años, pero que en su momento no pudieron ser llevados a la práctica. Un ejemplo, sobre un campo muy utilizado en la actualidad, es el correspondiente a determinar información semántica en imágenes. Este no es un problema nuevo. Se

han realizado numerosos estudios a lo largo de los años para intentar proporcionar soluciones más o menos específicas para diferentes contextos. Por ejemplo, en [296] se define el uso de una red de retropropagación para reconocer códigos postales escritos a mano en el correo postal de Estados Unidos. En [297], se propone una red de neuronas híbrida para el reconocimiento facial humano, y dicha red es comparada con otros mecanismos de reconocimiento. En [298] se describe un algoritmo basado en reglas para reconocimiento facial robusto, combinada con una red neuronal convolucional. En [299] se diseña un mecanismo objetivo de medida para validar automáticamente la calidad percibida en vídeos digitales. En este caso, el análisis trasciende de una imagen estática a una sucesión de imágenes a lo largo del tiempo (vídeo). Igualmente, existen avances recientes en el área de los modelos computacionales de la emoción humana [300].

Una de las últimas tendencias en los mecanismos de reconocimiento está basada en técnicas tipo “*deep learning*” [301], [302], ya que los resultados teóricos sugieren las “*deep architectures*” pueden resultar un buen método para aproximar el tipo de funciones complejas que pueden representar abstracciones de alto nivel (p. ej. En visión, lenguaje y otras tareas a desarrollar por una inteligencia artificial).

El aprendizaje profundo (*deep learning*) [303] permite a los modelos computacionales compuestos de múltiples capas de procesamiento, aprender representaciones de datos con múltiples niveles de abstracción. El aprendizaje profundo descubre la estructura de grandes conjuntos de datos mediante el uso del algoritmo de retropropagación (*backpropagation*) para indicar cómo una máquina debe cambiar sus parámetros internos, utilizados para calcular la representación en cada capa de la representación en la capa anterior. Las redes convolucionales profundas han traído consigo avances en el procesamiento de imágenes, vídeo, voz y audio, mientras que las redes recurrentes los han traído en datos secuenciales, como el texto y el habla.

Esto parece ser cierto acorde a recientes sistemas desarrollados que utilizan dicho enfoque, como la empresa emprendedora respaldada por Google, Clairifai [304]. Dicha empresa ha desarrollado un algoritmo que analiza cientos de imágenes por segundo para clientes como empresas de redes sociales, agencias inmobiliarias y empresas de comercio electrónico. La tecnología de Clairifai puede identificar una imagen, clasificarla y agrupar imágenes similares juntas. También permite clasificación automática de grandes bibliotecas de medios, búsqueda en imágenes y vídeos a través de etiquetas o similitud visual, o recomendaciones basadas en auto-categorización de imágenes y comportamiento del usuario. Otro sistema interesante es el propuesto por TripleLift [305], orientado a optimizar la integración de anuncios en el publicador mediante el reconocimiento y análisis de las imágenes. Las técnicas basadas en “*deep learning*” han sido evaluadas, consiguiendo unos resultados similares a los obtenidos por los cerebros de los monos en cuanto a reconocimiento de imágenes se refiere [306]. Por desgracia, hasta el momento no existe un sistema capaz de competir con el ser humano en términos de reconocimiento de imágenes, extracción de características y detección de inferencias.

Igualmente, la aplicación de diversas técnicas algorítmicas ha tenido mucho éxito en el campo auditivo, facilitando una serie de novedosos servicios. La recuperación de información musical (MIR, *music information retrieval*), área interdisciplinar de la

ciencia, basada entre otros campos, en la musicología, psicología, estudio académico musical, procesamiento de señales y el aprendizaje automático, ha hecho posible servicios como Shazam [307], SoundHound [308] o Echoprint [309], para el reconocimiento automático de canciones o sonidos a través de dispositivos móviles con conectividad a Internet, mediante el envío de un pequeño fragmento de muestra, capturado por el micrófono de dichos dispositivos. Este tipo de servicios suele hacer uso de la técnica conocida como “huella auditiva”, que se puede considerar como un resumen digital de la señal de audio, donde la presencia de picos en los espectrogramas correspondientes a una señal en un entorno ruidoso frente a una señal en un entorno silencioso permite determinar, con alta probabilidad, si una muestra de audio coincide con una grabación en la base de datos [307].

También se ha recurrido a la combinación de múltiples dominios para proporcionar soluciones a problemas muy costosos de resolver hace unos años. Un ejemplo es el caso del reconocimiento de voz audiovisual (AVSR). Este tipo de técnica es considerada una de las soluciones más prometedoras para el reconocimiento de voz fiable, particularmente cuando el audio está dañado por el ruido. Sin embargo, la selección cautelosa de las características sensoriales es crucial para lograr un alto rendimiento de reconocimiento. En [310], gracias a técnicas de aprendizaje profundo, se consigue, por un lado, adquirir características de audio robustas al ruido, usando pares de características de audio deterioradas y las características correspondientes, limpias. Por otro lado, una red neuronal convolucional (CNN) es empleada para extraer características visuales de las imágenes de la zona de la boca, a fin de obtener como salida las etiquetas de los fonemas correspondientes. Finalmente, se aplica un modelo de Markov de múltiples flujos para integrar los modelos auditivos y visuales independientemente adquiridos, con sus respectivas características. La aplicación de este tipo de técnica ha demostrado una ganancia de velocidad de reconocimiento de palabra de aproximadamente 65% frente a los enfoques tradicionales de reconocimiento.

Sí bien es cierto que en los últimos años han emergido un gran número de tecnologías avanzadas, basadas en la Inteligencia Artificial, para el reconocimiento de características recibidas por entradas sensoriales, se ha observado que la mayor parte de estudios en la materia, así como de innovaciones, servicios y soluciones tecnológicas, están orientadas a la aplicación en interfaces de interacción entre hombre-máquina, o en el aprendizaje específico en un único tipo de dimensión. Por lo general, visual, seguido de auditiva. Este tipo de aprendizaje avanzado es utilizado posteriormente en sistemas de reconocimiento de características en imágenes, vídeo o sonido, habilitando posibilidades como la recuperación de información en base a contenido multimedia y en base a etiquetas textuales, o la clasificación y generación automática de etiquetas textuales para catalogación de grandes colecciones de archivos, entre otras. También en el caso de la información textual (aquella cuyo contenido está basado en el texto, como por ejemplo las páginas de la World Wide Web), que podría ser considerada como otro tipo de entrada sensorial (diálogo escrito que no hay que procesar a nivel visual, auditivo o táctil para ser interpretado), se han dado numerosos avances, por ejemplo, en el área de reconocimiento de lenguaje natural [311]–[313] y aplicaciones derivadas, como el análisis de sentimiento (*sentiment analysis*) [314], [315]. Por desgracia, no se han encontrado estudios

orientados al aprendizaje multimodal holístico, considerando como entradas perceptuales todas las posibles (existentes y venideras), bajo un marco común de representación. Como se ha estudiado en el presente y anterior capítulo, muchos autores consideran que un marco así de complejo, capaz de cubrir el área de percepción (sentidos), conocimiento (representación) y memoria (almacenamiento), es habilitador de otras tareas avanzadas de la cognición, como el razonamiento y la inteligencia.

3.2.3. Tendencias de mercado, e innovación de las empresas más importantes del sector

Según Gartner [316], nos encontramos al principio de “*the perceptual smart machine age*”, o la era de las máquinas perceptuales inteligentes (ver Ilustración 14). Tal y como comentan en su informe sobre tendencias clave tecnológicas, de Julio de 2016, ésta será la clase de tecnologías más disruptivas en los próximos 10 años debido, entre otros factores, al incremento radical de poder computacional, a las cantidades casi infinitas de datos generados e identificados, y a los avances sin precedentes en las redes neuronales profundas, los cuales permitirán a las organizaciones dotadas de tecnologías relacionadas con las máquinas inteligentes, aprovechar los datos para adaptarse a nuevas situaciones y resolver problemas que nadie ha conseguido resolver previamente. Gartner recomienda considerar una serie de tecnologías específicas a las empresas que pretendan aprovechar esta área. Algunas de ellas son: Smart Dust, Aprendizaje Automático, Asistentes Personales Virtuales, Asesores Expertos Cognitivos, Descubrimiento Inteligente de Datos, Espacio de Trabajo Inteligente, Interfaces de Usuario Conversacional, Robots Inteligentes, Drones comerciales, Vehículos Autónomos, Respuesta de preguntas mediante lenguaje natural, generación de analíticas personales, administración de taxonomía empresarial y gestión ontológica, entre otros.



Source: Gartner (July 2016)

Ilustración 14. Ciclo de "hype" para tecnologías emergentes (2016) según Gartner [316]

Hecho que refuerza la relevancia de la presente Tesis Doctoral, y su importancia en un contexto tecnológico de vanguardia.

Las principales empresas del sector, como IBM, Intel, Microsoft, Google, Amazon o NVIDIA, entre otras, han invertido durante los últimos años, grandes sumas de dinero en el área de la inteligencia artificial y de la percepción computacional [317]–[319], formando incluso alianzas para promover el desarrollo del sector a través de la fundación de organizaciones no gubernamentales (*“Partnership on Artificial Intelligence to Benefit People and Society”*) [320]. Por separado, dichas empresas han hecho importantes avances en el área de la percepción computacional y del aprendizaje automático. Por ejemplo, IBM desarrolló Watson [321], un sistema computacional orientado a la resolución de preguntas formuladas través del lenguaje natural, bajo el proyecto DeepQA [322]. Dicho proyecto pretende ilustrar cómo gracias a la vasta y creciente accesibilidad de los contenidos presentados en lenguaje natural, el avance de las técnicas para procesar el lenguaje natural, recuperación de información, aprendizaje automático, representación y razonamiento del conocimiento y el paralelismo masivo, pueden conducir al desarrollo de tecnologías orientadas a resolución de preguntas en dominios abiertos hasta un punto en el que pueda rivalizarse clara y consistentemente con el mejor rendimiento humano. La versión inicial de Watson, con acceso a más de 200 millones de páginas de contenido estructurado y desestructurado (unos 4 Terabytes de espacio de almacenamiento) almacenado localmente, y sin acceso a Internet, fue presentado al concurso americano *“Jeopardy”*, donde venció casi impecablemente a los dos mejores jugadores humanos hasta la fecha [323]. Dicha versión inicial de Watson ha evolucionado notablemente, y donde al principio consistía únicamente en una cabina

de servidores, en el presente, se extiende a través de una nube de servidores de estándar abierto, que ejecutan varios cientos de instancias del sistema a la vez. Watson proporciona servicio a clientes simultáneos en cualquier parte del mundo, que pueden acceder a él usando sus teléfonos inteligentes, computadores de escritorio o sus propios servidores de datos. Este tipo de sistema se puede escalar horizontal y verticalmente bajo demanda. Debido a que la inteligencia artificial del sistema mejora a medida que la utiliza la gente, Watson está en constante evolución, transmitiendo los conceptos aprendidos en todo el sistema. Además, al tratarse de un conjunto de servicios distribuidos, su módulo de deducción lógica puede estar escrito en un lenguaje de programación y ejecutarse en una serie de servidores, mientras que, por ejemplo, su módulo de análisis de lenguaje puede estar escrito en otro lenguaje totalmente diferente, y estar instanciado en otra serie de servidores diferentes. A fecha actual, Watson puede analizar e interpretar datos de tipo: texto no estructurado, imágenes, audio y vídeo. Puede desarrollar tareas específicas de aprendizaje automático, para ser utilizadas en aplicaciones y sistemas. Puede desarrollar tareas de recomendación inteligente, mediante la comprensión de la personalidad del usuario, el tono y la emoción, y puede ser utilizado en *bots* (agentes automáticos) para establecer interacción con los usuarios a través del lenguaje natural. Finalmente, Watson se encuentra disponible como un conjunto de APIs (*application programming interfaces*) abiertas y productos SaaS (*Software as a Service*).

Por otra parte, tanto Google como Microsoft han desarrollado una serie de APIs muy potentes, orientadas a la computación perceptual y al aprendizaje automático, mediante el uso de sus plataformas de computación en la nube (*Google Cloud Platform*, productos [324]; *Microsoft Cognitive Services*, APIs [325]). De *Google Cloud Platform* caben destacar los siguientes servicios, totalmente integrados e interconectados en sus respectivos ecosistemas:

- **Cloud Machine Learning (beta)**: Es un servicio gestionado que permite construir fácilmente modelos de aprendizaje automático sobre cualquier tipo de datos, de cualquier tamaño. Permite el uso de *TensorFlow* [326] (potente biblioteca de software de código abierto para el cálculo numérico utilizando gráficos de flujo de datos). Dicho servicio es utilizado también por otros productos de Google, como *Google Fotos*, o *Google Cloud Speech*. Gracias a la pantagruélica y escalable infraestructura de Google, permite la generación de modelos de prácticamente cualquier tamaño. La plataforma de predicción a su vez, puede soportar miles de usuarios y Terabytes de datos. El servicio, además, está integrado con *Google Cloud Dataflow* [327] para el pre-procesamiento de grandes volúmenes de datos, lo que le permite acceder a datos de *Google Cloud Storage* [328] y *Google BigQuery* [329], entre otros.
- **Natural language API**: revela la estructura y el significado del texto, ofreciendo poderosos modelos de aprendizaje automático a través de una API REST [330]–[332] fácil de utilizar. Puede ser usado para extraer información sobre personas, lugares, eventos y otro extenso conjunto de categorías, presentes en documentos de texto, artículos de noticias o entradas de blog. También permite realizar análisis de sentimiento en las redes sociales o analizar la intención de las conversaciones entre usuarios (por ejemplo, en un centro de llamadas o una aplicación de mensajería). Soporta la integración con otras soluciones de

Google, como el módulo de reconocimiento óptico de caracteres, la API de reconocimiento vocal y la API de traducción.

- **Speech API:** permite a los desarrolladores convertir audio a texto gracias a la aplicación de potentes modelos de redes neuronales. La API reconoce más de 80 idiomas y variantes, para dar soporte a un conjunto global de usuarios. Entre otras funcionalidades, permite el dictado a través del micrófono, transcribiendo a texto el discurso del usuario, ejecutar comandos mediante control de voz, o transcribir archivos de audio, entre muchos otros casos de uso.
- **Vision API:** permite a los desarrolladores analizar el contenido de una imagen mediante la encapsulación de potentes modelos de aprendizaje automático, mediante una API REST fácil de utilizar. El servicio permite clasificar las imágenes en miles de categorías (por ejemplo, "velero", "León", "Torre Eiffel"), detecta objetos individuales y caras dentro de las imágenes, y encuentra y lee las palabras impresas contenidas dentro de las imágenes. Puede crear metadatos en catálogos de imágenes, ser utilizado para la moderación de contenido ofensivo, o empleado para habilitar nuevos escenarios de marketing a través del análisis del sentimiento de la imagen.

Microsoft a su vez, y de manera análoga, proporciona los siguientes servicios, agrupados por categorías:

- **Visión:**
 - **Computer Vision API (Preview):** funcionalidad que devuelve información sobre el contenido visual encontrado en una imagen. Permite el uso de etiquetas, descripciones y modelos específicos de dominio para identificar el contenido. De manera análoga a la API de visión de Google, permite la identificación de contenido ofensivo / para adultos, para crear filtros automatizados de restricción de contenido. También identificar los tipos de imagen y esquemas de color presentes en las imágenes.
 - **Emotion API (Preview):** la API de emociones reconoce expresiones faciales en una imagen de entrada, y devuelve el valor estadístico de confianza asociado a dichas expresiones, respecto a un conjunto determinado de emociones. Las emociones detectadas son la ira, el desprecio, el disgusto, el miedo, la felicidad, la neutralidad, la tristeza y la sorpresa. Se entiende que estas emociones se comunican transcultural y universalmente mediante expresiones faciales particulares. Esta API puede interactuar con la *API Face*, orientada al reconocimiento facial en imágenes.
 - **Face API (Preview):** permite la detección de una o más caras humanas en una imagen, obteniendo los rectángulos de posición correspondientes a cada cara detectada, junto con los atributos de la cara, que contienen predicciones basadas en el aprendizaje automático de rasgos faciales. Las características de atributos de cara disponibles son: edad, sexo, postura, sonrisa y pelo facial junto con 27 puntos de referencia para cada cara de la imagen (posición de ojos, boca, etc.).

- **Video API (Preview):** tecnología que detecta inestabilidad en el vídeo, por ejemplo, debido a traqueteo o vibraciones durante la grabación, y automáticamente lo estabiliza.
- **Habla:**
 - **Bing Speech API (Preview):** encargada de convertir diálogos hablados (audio) a texto. Permite reconocer el audio procedente del micrófono en tiempo real, o el audio procedente de una fuente sonora diferente, así como reconocer diálogo en audio almacenado en un archivo.
 - **Custom Recognition Intelligent Service (CRIS - Private Preview):** funcionalidad que permite la personalización de los modelos de lenguaje, la creación de modelos acústicos específicos y el despliegue de dichos modelos en aplicaciones finales dotadas de capacidad de reconocimiento de audio.
 - **Speaker Recognition API (Preview):** funcionalidad que habilita el uso de la voz como mecanismo autenticador (biometría). El servicio solicita registrar tres muestras de voz correspondientes a la lectura de una frase de paso elegida en una lista de frases pre-determinada (entrenamiento). Una vez el usuario ha cargado las muestras, es posible iniciar el proceso de verificación utilizando una frase diferente a la elegida en el proceso de entrenamiento.
- **Idioma:**
 - **Language Understanding Intelligent Service (LUIS - Preview):** servicio diseñado para proporcionar una manera sencilla de creación de modelos de reconocimiento de lenguaje natural, a fin de mejorar la interacción entre los usuarios y las aplicaciones que utilizan servicios cognitivos.
 - **Linguistic Analysis API (Preview):** utiliza herramientas avanzadas de análisis lingüístico para el procesamiento del lenguaje natural, permitiendo identificar los conceptos y acciones en el texto, junto con el etiquetado de parte del habla, así como encontrar frases y conceptos utilizando analizadores de lenguaje natural.
 - **Text Analytics API (Preview):** funcionalidad que permite extraer información sobre el lenguaje, sentimiento y frases clave del texto pasado como parámetro.
 - **Translator API (Preview):** soporta la traducción de texto hablado entre nueve idiomas, y entre texto escrito en sesenta idiomas. Previo a la traducción, es capaz de determinar automáticamente el lenguaje origen. Finalmente, permite la interacción con los usuarios (refuerzo) a fin de aprender, y mejorar las traducciones ofrecidas.
- **Conocimiento:**
 - **Academic Knowledge API (Preview):** permite interpretar búsquedas en lenguaje natural, orientadas al ámbito académico, generar entidades (artículos, autores), estadísticas avanzadas sobre dichas entidades (como la distribución de citas por año), y evaluar la similitud entre dos entradas textuales (no solo palabras, sino artículos completos, identificando los conceptos del contenido).
 - **Entity Linking Intelligence Service (Preview):** servicio que permite discriminar el tipo de entidad de un texto en base al contexto en el que

aparece (p. ej. el término “*times*” en “*New York Times*” y “*Times Square*”).

- **Knowledge Exploration Service (Preview)**: servicio que permite la integración en aplicaciones de búsqueda interactiva sobre datos estructurados mediante el uso de lenguaje natural.

Microsoft, al igual que Google, también proporciona APIs de búsqueda asociados a texto, imagen y vídeo. Cabe destacar que la mayor parte de servicios ofrecidos, tanto por Google como por Microsoft son tan novedosos que aún se encuentran en fase de desarrollo (*Beta / Preview*).

Finalmente, y orientado a la parte de la percepción computacional más relacionada con el diseño de interfaces, se pueden destacar diversos avances desarrollados a lo largo de los años, que han ido incrementando su complejidad a medida que se han perfeccionado técnicas algorítmicas, e incrementado la potencia computacional. Además de las interfaces de entrada convencionales, a las que estamos acostumbrados, como el ratón o el teclado (que “sienten” de manera táctil la posición o letra que el usuario quiere accionar, o el movimiento que quiere que el cursor realice por la pantalla), existen otros muchos tipos de interfaz, aplicados sobre todo al mundo del entretenimiento digital. Por ejemplo Nintendo lanzó al mercado el 18 de febrero de 1984 una interfaz de control con forma de pistola (*NES Zapper*) [333]. El modo de funcionamiento es el siguiente: al pulsar el gatillo en la interfaz, el juego genera un fotograma completamente negro que cubre toda la pantalla. En el siguiente fotograma, todos los blancos válidos se dibujan en blanco, y el resto de la pantalla permanece en negro. La interfaz detecta el cambio de iluminación (brillante / no brillante) y determina si alguno de los objetivos está en la línea directa de visión. La discriminación del objetivo se hace en base a la duración de la iluminación, ya que cada objetivo se ilumina durante una cantidad diferente de tiempo. Seguidamente, se siguen presentando los fotogramas de manera normal. Todo el proceso es casi imperceptible para el ojo humano, aunque se puede notar un ligero parpadeo de la imagen. [334] Aunque dicha interfaz sólo detecta luz, su aplicabilidad se limita a las pantallas CRT. No funciona en pantallas actuales, como las de cristal líquido o plasma debido al retardo de visualización. Sin embargo, *Zapper* no fue la primera ni última interfaz que aprovecha la percepción (en su caso, visual) para determinar la intención del usuario. Años más tarde, Nintendo lanzó al mercado la consola Wii, con una interfaz inalámbrica denominada *Wii Mote (Wii Remote)*. Dicha interfaz, dotada de diferentes sensores, es capaz de determinar la aceleración e inclinación del dispositivo en un espacio tridimensional, junto con la determinación de su posición en base a un mecanismo de triangulación, basado en luz infrarroja [335]. Sin embargo, Nintendo no es la única empresa que ha hecho avances y desarrollos en el área de la percepción computacional aplicada a las interfaces de usuario. Empresas como Sony han desarrollado familias de interfaces, como *EyeToy (PlayStation 2)*, *PlayStation Eye (PlayStation 3)* y *PlayStation Camera (PlayStation 4)*, consistentes en una cámara más o menos avanzada, dependiendo de la generación, comenzando como una webcam de baja resolución con un micrófono, continuando como una cámara de mayor resolución y sensibilidad dotada de un *array* de micrófonos, y culminando como una cámara dual de alta resolución, capaz de capturar imágenes estéreo y calcular la profundidad de los elementos de la escena, y dotada de un *array* de micrófonos,

similares al de la generación anterior [336]. Dichas cámaras con micrófono son empleadas como interfaz de control en videojuegos, y utilizadas para realizar seguimiento de objetos y reconocimiento de audio. Microsoft también ha contribuido al desarrollo de interfaces alternativas, basadas en la percepción mediante la visión. Tal es el caso de la familia de cámaras Microsoft Kinect [337], con capacidades de seguimiento en 3D, identificación biométrica, detección automática de extremidades y articulaciones, y reconocimiento de voz, entre otras, con aplicación no sólo en su familia de videoconsolas, sino en computadores, a través del SDK (*Software Development Kit*) desarrollado a tal efecto [338]. Recientemente, Microsoft ha desarrollado una interfaz avanzada que combina la percepción con la realidad aumentada, permitiendo interactuar de manera inmersiva con el entorno. Dicha interfaz, HoloLens [339] cuenta un numeroso conjunto de sensores, que le permite escanear el entorno que le rodea en 3D, determinar la posición y movimiento del usuario por dicho entorno, determinar interacciones gestuales sencillas (movimiento de las manos), y generar la interacciones apropiadas mediante una pantalla especial que genera imágenes a modo de hologramas, y un complejo y novedoso *array* de altavoces que permiten generar sonido posicional. Además, todo el procesamiento perceptual sobre la detección e interacción con el entorno se realiza en la propia unidad HoloLens.

Por otro lado, empresas como HTC, también han desarrollado complejas interfaces perceptuales, como VIVE [340]. Este último, es un conjunto de dispositivos compuesto por un casco dotado de una gran variedad de sensores, entre los que se incluyen cámaras (para proporcionar una experiencia de realidad aumentada y realizar detección de gestos) y sensores de detección de movimiento de alta precisión. También se dispone de dos mandos inalámbricos con 24 sensores de movimiento, y dos estaciones base que se deben situar en posiciones específicas de la habitación, para determinar y hacer seguimiento de la posición del usuario en un entorno de 360°, completamente en 3D. La combinación de todos los sensores físicos del sistema junto con el procesamiento llevado a cabo en el computador, permiten determinar con mucha precisión los datos sobre posición y movimiento del usuario, generando las experiencias visuales y hápticas correspondientes, y creando de esta manera un entorno altamente inmersivo.

Por último, Intel también ha hecho una gran inversión (y promoción) hacia la computación perceptual. Uno de los más notables avances introducidos por la firma es *Intel RealSense* [341]. *RealSense* consiste en un conjunto de cámaras avanzadas, desarrolladas para ser integradas en computadores de sobremesa, portátiles y *tablets*. Cada dispositivo *RealSense* está compuesto por tres cámaras que actúan como una: una cámara HD 1080p, una cámara de infrarrojos y un proyector de láser infrarrojo, que emulan el proceso de visión del ojo humano, para percibir la profundidad y rastrear el movimiento. Intel también proporciona un completo SDK, con el que se pueden realizar tareas de escaneado en 3D de objetos (incluyendo sus texturas), reconocimiento de gestos, y reconocimiento auditivo del discurso.

3.3. Cloud Computing

Como se ha observado en los apartados anteriores del capítulo, a fin de proporcionar un sistema que pueda escalar con la demanda dinámica de recursos, la mayor parte de soluciones (sobre todo las de tipo *software*) hacen uso de la computación distribuida. Específicamente, del “*cloud computing*”.

El “*cloud computing*” o computación en la nube, es una tendencia tecnológica que afecta a un diverso número de entornos de negocio, puesto que permite optimizar recursos, y gestionarlos de manera eficiente y bajo demanda. De acuerdo con la definición propuesta por el NIST (*National Institute of Standard and Technology*), “*La computación en la nube es un modelo habilitador de acceso ubicuo, conveniente y bajo demanda a través de una red de comunicaciones a un conjunto de recursos computacionales configurables (p. ej. redes, servidores, almacenamiento, aplicaciones y servicios) que pueden ser desplegados y liberados rápidamente, con un esfuerzo de mantenimiento y una interacción por parte del proveedor mínimos*” [342].

La computación en nube representa una convergencia de dos tendencias principales en las Tecnologías de la Información (TI): la agilidad de negocio, por medio de la cual las TI pueden ser utilizadas como una herramienta competitiva a través del despliegue rápido, el uso de analíticas computacionalmente intensivas aplicadas al ámbito empresarial, el procesamiento por lotes paralelo y el soporte a aplicaciones interactivas móviles capaces de responder en tiempo real a los requisitos del usuario, y la eficiencia de las TI, donde la potencia de las computadoras se utiliza de manera más eficiente a través de recursos de hardware y software altamente escalables. [346].



Ilustración 15. Definiciones de NIST para la computación en la nube

NIST ha definido de manera compleja y abstracta [342] algunos de los ámbitos para caracterizar distintos entornos de computación en la nube (ver Ilustración 15). Los modelos de despliegue y de servicio son las categorías más frecuentemente utilizadas. Acorde con su modelo de despliegue, es posible encontrar:

- **Nube pública:** donde la infraestructura es proporcionada para el uso abierto por el público general. Este tipo de nube se encuentra en las instalaciones del proveedor del servicio, y puede estar en propiedad, y ser administrada y operada por una o varias partes.
- **Nube privada:** donde la infraestructura es proporcionada específicamente para el uso de una sola organización, compuesta por múltiples consumidores. Puede estar en propiedad, y ser administrada y operada por la organización, por un tercero o por una combinación de ambos. La infraestructura puede residir dentro o fuera de las instalaciones de la organización.
- **Nube comunitaria:** donde la infraestructura es proporcionada para el uso de una comunidad específica de consumidores, pertenecientes a organizaciones que tienen relaciones comunes, como políticas, requisitos de seguridad o consideraciones de cumplimiento. Puede estar en propiedad, y ser operada y administrada por uno o varios miembros de la comunidad, un tercero o una combinación de ambos. Además, la infraestructura puede residir dentro o fuera de las instalaciones de alguna o varias de las organizaciones.
- **Nube híbrida:** se aplica a la combinación de dos o más infraestructuras de nube distintas (públicas, privadas o comunitarias) que siguen siendo entidades únicas, pero pueden interactuar entre sí mediante una tecnología que permite la portabilidad de datos y aplicaciones.

El último modelo de despliegue, la nube híbrida, ofrece varias ventajas a las entidades que necesitan mantener partes de la lógica y los datos empresariales en un entorno privado y seguro, pero que también necesitan proporcionar servicios públicos a los clientes a través de Internet. Servicios que pueden requerir el acceso a alguna funcionalidad o datos de la infraestructura de la nube privada.

En cuanto a los modelos de servicio, se definen las siguientes tres categorías:

- **IaaS (*Infrastructure as a Service*):** en el modelo de infraestructura como servicio, el cliente dispone de recursos informáticos fundamentales (potencia de procesamiento, almacenamiento, redes, etc.) sobre los que desplegar y ejecutar sistemas operativos y aplicaciones. El cliente no gestiona ni controla la infraestructura subyacente, sino que tiene control sobre los sistemas desplegados, de acuerdo con sus necesidades (por lo general, máquinas virtuales).
- **PaaS (*Platform as a Service*):** en el modelo de plataforma como servicio se le facilita al cliente un entorno de computación pre-configurado, en donde desplegar aplicaciones creadas por el cliente, o adquiridas. Aplicaciones desarrolladas utilizando diversos lenguajes de programación, servicios, bibliotecas y herramientas soportadas por el entorno informático citado. El cliente no controla la infraestructura subyacente, sino que toma el control sobre las aplicaciones desplegadas y determinados parámetros de configuración para

el entorno de alojamiento de aplicaciones (p. ej. un servidor web, o un sistema gestor de base de datos).

- **SaaS (Software as a Service):** en el modelo de *software* como servicio, el cliente puede utilizar las aplicaciones del proveedor, que se ejecutan en una infraestructura en la nube. Las aplicaciones son accesibles por diferentes dispositivos a través de Internet, a través de interfaces específicas (navegadores web, APIs y/o clientes de escritorio/móviles). El cliente no controla la infraestructura subyacente ni el software de soporte, pero en algunos casos, es posible realizar ajustes para aplicaciones específicas, limitadas a los usuarios (p. ej. sistemas de almacenamiento de datos en la nube basados en servicios web).

NIST también define cinco características esenciales, o atributos de servicio, que los sistemas de computación en la nube deben ofrecer:

- **Agrupación de recursos:** los recursos se agrupan en un sistema que admite el uso por múltiples clientes. Los sistemas físicos y virtuales se asignan dinámicamente, y se reasignan según sea necesario. Bajo este concepto de agrupación se encuentra la idea de abstracción, que oculta la ubicación real de los recursos, tales como la potencia de procesamiento, la cantidad de memoria, el tamaño de almacenamiento, el ancho de banda de red, la topología de red y, por ejemplo, las máquinas virtuales alojadas en la infraestructura.
- **Elasticidad rápida:** los recursos se pueden aprovisionar rápida y elásticamente. El sistema puede agregar recursos a través de dos procesos principales, mediante la adición de unidades de cómputo más potentes, o mediante la adición de más instancias computacionales de iguales características a las existentes en la infraestructura. El escalado puede realizarse de manera automática o manual. Desde la perspectiva del cliente, los recursos de computación en la nube deben parecer ilimitados y deben poderse comprar en cualquier momento y en cualquier cantidad.
- **Auto servicio bajo demanda:** los clientes deben tener la capacidad de desplegar recursos informáticos sin necesidad de interactuar con el personal del proveedor de servicios en la nube.
- **Servicio medido:** el uso de los recursos del sistema en la nube es auditado, medido y notificado al cliente en base a un sistema establecido. Se puede cobrar a un cliente en base a una métrica conocida, como la cantidad de potencia de procesamiento utilizada, el número de transacciones, la cantidad de almacenamiento utilizado, el ancho de banda de la red, etc. A dichos clientes se les cobra en base al nivel de servicios prestados.
- **Acceso amplio de red:** el acceso a los recursos en la nube está disponible a través de la red de datos, utilizando métodos estándar, de manera que se proporcione acceso independiente a la plataforma para clientes de todo tipo (computadores de escritorio, computadores portátiles, *smartphones*, *tablets*, etc.).

Algunos beneficios derivados de la adopción de la computación en nube, teniendo en cuenta la características presentadas, son: una reducción significativa de coste, porque las redes en la nube operan con mayor eficiencia y con un mayor grado de

utilización, optimizando el uso de recursos y proporcionando una mayor fiabilidad, ya que la escala de red de dichos sistemas es suficientemente grande como para proporcionar mecanismos de distribución (balanceo) de carga y mecanismos de conmutación por error (costoso o incluso inalcanzable para organizaciones individuales), una mejora en la facilidad de uso en función del servicio contratado. Por ejemplo, un servicio SaaS suele implicar que el cliente final no requiere licencias de software adicionales para implementar su servicio. La capacidad para definir e implementar políticas de QoS (*Quality of Service*, calidad de servicio) para el acceso y utilización de recursos. Capacidad de externalizar la gestión de TI, por lo que una entidad externa puede encargarse de gestionar toda la infraestructura informática, reducir los costes de personal de TI y trasladar la gestión a grupos cualificados de expertos en TI y especialistas en diversas áreas como la administración y la seguridad, los cuales pueden ocuparse de mantener la infraestructura y aplicar parches y actualizaciones a medida que van apareciendo.

3.3.1 Virtualización

La virtualización es el proceso por el cual a través de un *software* específico es posible crear una representación (virtual) de uno o varios recursos físicos [343]. La aplicación generalizada del software de virtualización es la de ejecutar múltiples sistemas operativos y múltiples aplicaciones en el mismo servidor al mismo tiempo. La virtualización puede emplearse con aplicaciones, servidores, almacenamiento y redes, y es una forma extremadamente eficaz de reducir los gastos de TI en entornos empresariales, permitiendo una amortización efectiva de la inversión realizada en el *hardware*.

Sin embargo, la virtualización no es lo mismo que la computación en la nube. Esencialmente, la virtualización se refiere a un tipo de *software* que manipula *hardware*, mientras que la computación en nube se refiere al servicio resultante de dicha manipulación [344]. Por tanto, se puede considerar que la virtualización es un elemento fundamental habilitador para la computación en la nube.

Hasta hace no mucho, las organizaciones estaban sujetas a las limitaciones impuestas por sus servidores, diseñados para ejecutar un único sistema operativo, con un conjunto de aplicaciones y servicios, a la vez. Como resultado, incluso pequeños centros de datos tenían que desplegar muchos servidores, operando cada uno entre el 5 y el 15 por ciento de su capacidad, que como puede observarse, es muy ineficiente. Con la inclusión de la virtualización se dotó de la capacidad de ejecutar múltiples máquinas virtuales dentro de una única física, repartiendo los recursos físicos reales entre las distintas máquinas virtuales, y con el consiguiente incremento del aprovechamiento de recursos.

Un sistema computacional virtual, también conocido como “máquina virtual” (VM, *virtual machine*) es un contenedor de software herméticamente aislado, con un sistema operativo y una serie de aplicaciones dentro. Cada VM autónoma es completamente independiente. Poner varias máquinas virtuales en un solo ordenador permite que varios sistemas operativos y aplicaciones se ejecuten en un solo servidor

físico, o "host". La aplicación gestora de máquinas virtuales (o monitor de máquinas virtuales), utilizada para virtualización "de escritorio" (p. ej. *Oracle VirtualBox* [345], *Parallels Workstation* [346], o *VMware Workstation/Fusion* [347], entre otros), o el hypervisor, utilizado en virtualización "de servidores" (p. ej. *Citrix XenServer* [348], *VMware vSphere* [349], o *Proxmox Virtual Environment* [350], entre otros), son las encargadas de desacoplar las máquinas virtuales del host y de asignar dinámicamente los recursos computacionales a cada máquina virtual, según sea necesario.

Las principales características y ventajas derivadas de las máquinas virtuales son las siguientes:

- **Aislamiento:** gracias a la ejecución en contenedores de software que abstraen el hardware físico, se consigue aislamiento ante fallos y mayor seguridad a nivel de hardware. Además, la gestión de dichos contenedores puede configurarse con una elevada granularidad, lo cual permite mejorar el rendimiento al establecer un control avanzado de los recursos. Igualmente, la ejecución en contenedores independientes deriva en un aislamiento entre máquinas virtuales, incrementando la seguridad al evitar el conocimiento y acceso de unas máquinas virtuales con otras.
- **Encapsulación:** la abstracción de recursos físicos mediante software permite características extendidas respecto al hardware, como, por ejemplo, guardar todo el estado de una máquina virtual en uno o varios archivos. Esto puede facilitar despliegues masivos de infraestructuras (basados en plantillas), o mover y copiar máquinas virtuales entre servidores de virtualización de una forma tan sencilla como mover y copiar archivos. De esta manera, pueden conseguirse sistemas redundantes, clústeres de alta disponibilidad o migraciones de servidores de una forma muy económica y sencilla.
- **Particionamiento:** como se ha comentado anteriormente, al dividir los recursos físicos de la máquina física entre distintas máquinas virtuales es posible aprovechar mucho mejor los recursos hardware disponibles, reduciendo costes de mantenimiento, incrementando la velocidad de despliegue de sistemas, y reduciendo la energía consumida, entre otras ventajas.
- **Independencia del hardware:** gracias a la configuración establecida en la máquina virtual sobre los recursos hardware virtuales que tiene asignados, va a ser posible desplegar dicha máquina virtual en cualquier tipo de servidor de virtualización, siempre y cuando éste tenga más recursos físicos disponibles que los requeridos por la máquina virtual. Esta característica facilita la migración y despliegue de las máquinas virtuales, en cualquier tipo de infraestructura (desde grandes nubes de computación que ofrecen servicios de virtualización, como Microsoft Azure [351], o Amazon WS EC2 [352], hasta servidores hospedados en empresas pequeñas).

3.3.2 Ejecución mediante contenedores

La “*containerización*”, o ejecución mediante contenedores, es una alternativa ligera a la virtualización de máquinas completas, que implica el encapsulamiento de una aplicación en un contenedor, con su propio entorno operativo (p. ej. bibliotecas del sistema). Esto proporciona muchos de los beneficios análogos a los de la ejecución de aplicaciones en máquinas virtuales, ya que la aplicación ejecutada en un contenedor puede desplegarse en cualquier máquina física adecuada, sin tener que preocuparse por las dependencias de la misma.

El principio básico de un contenedor es que permite aislar los procesos y sus recursos, sin necesitar una emulación de *hardware* o de recursos *hardware* (ver Ilustración 16). Los contenedores proporcionan una especie de plataforma de virtualización en la que cada contenedor puede ejecutar su propio sistema operativo, pero todos comparten el *kernel*. Así que cada contenedor tiene su propio sistema de archivos y pila de red, y cada contenedor puede ejecutar su propia distribución de Linux. Por ejemplo, un host Debian puede ejecutar Arch, Ubuntu, e incluso otros contenedores Debian, simultáneamente. Estas abstracciones hacen que un contenedor se comporte como una máquina virtual con un sistema de archivos separado, capacidades de red y otros recursos del sistema operativo, aunque en realidad no se trate de máquinas virtuales al no existir emulación de hardware.

La tecnología de contenedores existe desde hace casi una década. Su origen proviene de los *Linux Containers* (LXC), sucesores de *VServer* y *OpenVZ* [353], que proporcionan una virtualización ligera del sistema operativo. Sin embargo, gracias a Docker [354] y el ecosistema de aplicaciones y servicios creados a su alrededor en los últimos años, es cuando dicha tecnología ha ganado popularidad.

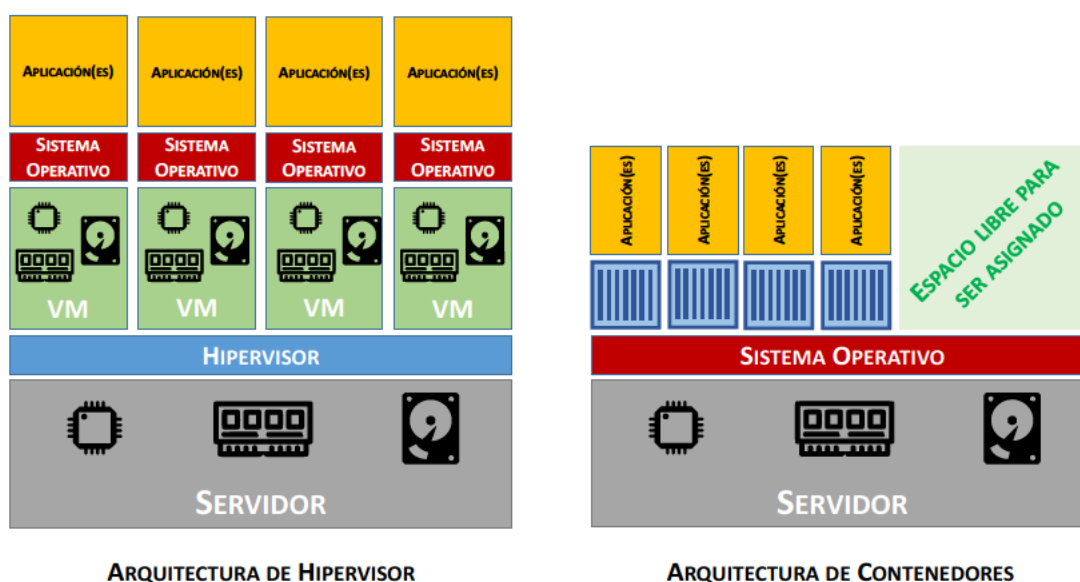


Ilustración 16. Arquitectura de Hipervisor (Máquinas Virtuales) frente a Arquitectura de Contenedores

Docker es una herramienta software basada en estándares abiertos, que facilita el empaquetado, despliegue y ejecución de una aplicación y todas sus dependencias en un contenedor. Para ello, proporciona un conjunto de herramientas y una API unificada para administrar tecnologías a nivel de *kernel*, como, por ejemplo, contenedores LXC, un sistema de ficheros *copy-on-write*, y *cgroups*. Docker emplea AuFS (*Advanced Multi-Layered Unification Filesystem*) [355] como sistema de archivos para contenedores. AuFS es un sistema de archivos en capas, que puede superponerse de forma transparente a uno o más sistemas de archivos existentes. AuFS permite a Docker utilizar ciertas imágenes como base (a modo de plantillas) para los contenedores. De esta forma, es posible disponer de una única imagen de un *software* específico que pueda utilizarse como base para muchos contenedores diferentes. Gracias a AuFS, sólo se requiere una copia de la imagen de dicho software, lo que se traduce en un ahorro de espacio de almacenamiento y consumo de memoria, así como un despliegue más rápido de los contenedores. Otra ventaja de usar AuFS es la capacidad de facilitar un sistema de versiones para las imágenes. Cada nueva versión contiene únicamente los cambios realizados sobre la versión anterior, manteniendo los archivos de la imagen al mínimo. Esta característica es muy útil, pues de ella extienden funcionalidades inherentes como la posibilidad de realizar una auditoría de cambios entre versiones, o la posibilidad de volver a un estado anterior en caso necesario.

Los contenedores Docker están diseñados para funcionar en prácticamente cualquier sistema, desde computadoras físicas a máquinas virtuales, servidores *bare-metal*, clústeres tipo *OpenStack* para computación en la nube, etcétera. Tal es la versatilidad de dicho sistema, que la comunidad Docker consiguió tras la DockerCon 2015, no sólo ejecutar el *stack* completo de Docker en una Raspberry Pi 2, sino ejecutar de forma concurrente en la misma, 2334 instancias del servidor web “*nginx*”, cada una en su contenedor independiente [356]. Seguidamente, y con la extensión del sistema operativo para dispositivos embebidos, HyprIoTOS, basado en Linux, se ha dotado de soporte a un gran número de otros dispositivos con diferentes arquitecturas (p. ej. ARM de 32 y 64 bits), utilizados en el ámbito del IoT, como por ejemplo Raspberry Pi 1, 2 y 3, Banana Pi, Beagle Bone, NVIDIA Shield y Pine A64 [357].

Con la enorme popularidad y adopción de la tecnología de contenedores, ha surgido una extensión de los modelos de servicio en la computación en la nube (ver Ilustración 15). Este nuevo modelo de servicio se denomina *CaaS*, o *Containers-as-a-Service* (contenedores como servicio) [358].

Dicho modelo está diseñado para ayudar a los desarrolladores y departamentos de TI a crear, ejecutar y administrar aplicaciones mediante contenedores. En el modelo *CaaS*, los contenedores y clústeres se proporcionan como un servicio que puede ser implementado tanto en centros de datos locales o sobre la nube. Para los desarrolladores, la adopción de un modelo *CaaS* simplifica el proceso de construcción de contenedores escalables y el despliegue de las aplicaciones, además de facilitar un muy robusto entorno de pruebas, puesto que los contenedores mantienen su estructura interna independientemente de la plataforma en la que se desplieguen, con lo que se puede pasar de entornos de producción a desarrollo y viceversa casi instantáneamente. Finalmente, para los departamentos de TI, *CaaS* ofrece un servicio de implementación de contenedores con un completo control de seguridad y de

gobierno, para la administración de las TI. En el corazón de un modelo *CaaS* se encuentra la plataforma de orquestación de contenedores, diseñada para manejar operaciones como el despliegue de dichos contenedores y la administración de clústeres. Ejemplos populares de estas plataformas incluyen *Docker Swarm* [359], *Google Kubernetes* [360], *Apache Mesos* [361], *Nova-Docker* para *OpenStack* [362], y *Carina* de *RackSpace* [363], entre otros. Las plataformas y proveedores de modelos *CaaS* suelen facilitar a sus clientes: motores de contenedores, orquestación y recursos computacionales. Estos recursos por lo general son proporcionados como servicio a través de una nube pública. Los principales proveedores de *CaaS* en infraestructuras de nube pública / híbrida actuales incluyen *Google* [364], *Amazon EC2 Container Service (ECS)* [365], *IBM Bluemix* [366], *Rackspace* [367], y recientemente *Microsoft Azure* [351].

3.3.3 Gobierno de la nube, o *cloud governance*

Como se ha observado en las secciones anteriores, la computación en la nube es una tendencia tecnológica muy significativa, altamente adoptada gracias a sus numerosos beneficios. Sin embargo, no todas las características de la computación en la nube son beneficiosas. También hay algunas barreras significativas en cuanto a su adopción. La interoperabilidad y la portabilidad de la información entre nubes públicas y privadas son factores críticos para la migración a entornos de computación en la nube por parte de algunas empresas [368]. Varias empresas han avanzado en este sentido mediante la normalización de sus procesos, datos y sistemas a través del uso de ERPs. Gracias a las infraestructuras escalables y a los altos niveles de conectividad entre instancias, la coherencia entre los datos maestros y de transacción es más manejable, y la información resultante es consolidada y confiable. Sin embargo, con estas plataformas mejoradas, la velocidad a la que cambian las empresas puede superar a la capacidad de las organizaciones de TI para reflejar dichos cambios. En cuanto a la fiabilidad, las aplicaciones organizativas son tan críticas que deben proporcionar operaciones 24/7 fiables y disponibles en la mayoría de los casos. Para combatir fallas o cortes, los planes de contingencia deben estar muy cuidados y no suponer problemas, y ante fallos desastrosos o catastróficos, los planes de recuperación deben aplicarse con un número mínimo de interrupciones. Muchas organizaciones están más cómodas ante el riesgo cuando tienen un mayor control sobre los procesos y equipos involucrados. Un alto grado de control ofrece la opción de sopesar alternativas, establecer prioridades y actuar con decisión, en pro del mejor interés de la organización frente a un incidente. A este respecto, la gestión del riesgo es el proceso de identificación y evaluación del riesgo para las operaciones, activos o individuos de la organización resultante del funcionamiento de un sistema de información, tomando las medidas necesarias para reducirlo a un nivel aceptable [369]. Cada aspecto relacionado con la fiabilidad debe ser cuidadosamente considerado a la hora de seleccionar un CSP (*Cloud Service Provider*, proveedor de servicios en la nube), deber ser negociado como parte del SLA (*Service Level Agreement*, acuerdo de nivel de servicios), y probados en simulacros error. Se pueden asociar costes adicionales acorde a los niveles requeridos de fiabilidad, sin embargo, las organizaciones pueden llevar a cabo una evaluación del riesgo, implementar una estrategia de mitigación del riesgo y emplear técnicas y

procedimientos para la monitorización continua del estado de seguridad del sistema de información [370]. Además, en el entorno de computación en nube, existe una variabilidad en términos de localización geográfica en la que residen los datos físicos, en la que se realiza el procesamiento y desde donde se accede a los datos. Debido a esta variabilidad, pueden aplicarse diferentes legislaciones, reglas y regulaciones de privacidad, con la posibilidad de problemas políticos debido a los límites globales [371].

La ubicación de datos ofrece uno de los problemas de cumplimiento más comunes [372]. El uso de un centro de datos interno permite a las organizaciones estructurar y conocer en detalle en qué lugar se almacenan los datos, y qué salvaguardias se están utilizando para la protección de los mismos. Por el contrario, en muchos servicios de computación en la nube, los datos se almacenan de forma redundante en múltiples ubicaciones físicas, distribuidas a lo largo del planeta, y la información detallada sobre la ubicación de los datos de una organización específica, generalmente no está disponible para el consumidor del servicio. Esta situación hace difícil determinar si existen salvaguardias suficientes y si se cumplen los requisitos legales de cumplimiento. Un ejemplo considerando la compleja regulación alemana sobre el área sanitaria está expuesto en [373].

Actualmente, algunos grandes poderes tecnológicos y políticos mundiales están elaborando leyes que pueden tener un impacto negativo en el desarrollo de la nube global. Por ejemplo, como resultado de la Ley Patriota de los Estados Unidos, el gobierno de Canadá no está usando computadoras en la red global que operan dentro de las fronteras de los Estados Unidos, temiendo por la confidencialidad y privacidad de los datos canadienses almacenados en dichos computadores [374]. Para hacer frente a esta situación, varios CSPs, como Google, Microsoft y Amazon, están ofreciendo servicios en la nube con la posibilidad de especificar límites geográficos. Esto es, garantizando que los datos serán almacenados y procesados dentro de áreas geográficas concretas. El precio de los servicios, sin embargo, puede variar dependiendo del lugar en el que residan los datos. Otra serie de problemas de seguridad y privacidad relacionados con la adopción de la nube, así como directrices para afrontarlos se pueden encontrar en [375].

Los citados problemas y la búsqueda y aplicación de soluciones están ampliando el papel del gobierno en las TI. Según el ITGI (*IT Governance Institute*), el gobierno corporativo es *"un conjunto de responsabilidades y prácticas ejercidas por el consejo y la dirección ejecutiva, con el objetivo de proporcionar dirección estratégica, asegurar que los objetivos se logren, verificar que los riesgos se gestionan adecuadamente y verificar que Los recursos de la empresa se utilizan de manera responsable"* [376]. El gobierno de las TI es una disciplina perteneciente al subconjunto del gobierno corporativo, centrada en el funcionamiento y la gerencia de riesgo de sistemas de TI. El gobierno de las tecnologías de la información busca asegurar que las inversiones en TI generen valor comercial y mitiguen los riesgos asociados con TI. El gobierno de la nube, por tanto, requiere definir políticas e implementar una estructura organizativa con roles bien definidos para la responsabilidad de la gestión de las TI, los procesos empresariales y las aplicaciones, puesto que estos elementos se trasladan del entorno TI tradicional a la nube.

Existen diferentes estándares y marcos de referencia que se aplican al gobierno de las TI. Algunos de los principales se pueden observar en la Ilustración 17 [377] junto con las relaciones entre sus controles. Como puede observarse, ninguno de ellos cubre por separado todo el espectro del gobierno de las TI, pero en su conjunto sí lo hacen (el qué, el cómo y el alcance).

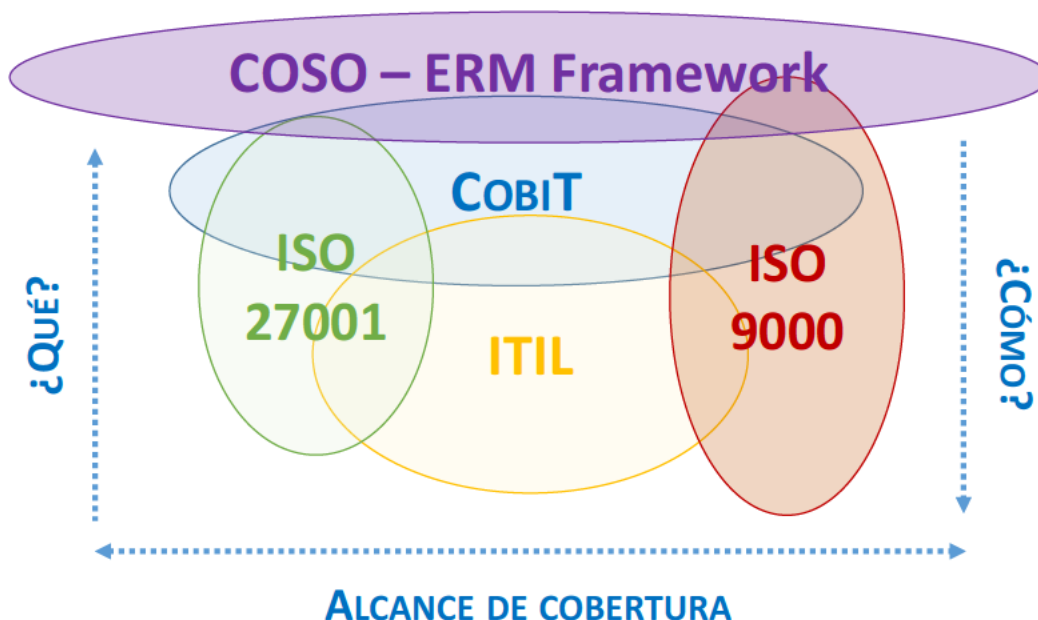


Ilustración 17. Relaciones entre estándares y marcos de referencia para el gobierno de las TI

Respecto a los estándares y marcos de referencia mostrados en la ilustración:

- **CobiT (*Control Objectives for Information and related Technology*)** [378]: publicado en 2011 por ISACA (*Information Systems Audit and Control Association*), es un marco de control de gobierno de las TI que ayuda a las organizaciones para abordar las áreas de cumplimiento normativo, gestión de riesgos y alineación de la estrategia de TI con las metas de la organización. CobiT identifica 34 procesos relacionados con las TI, sus objetivos de control de alto nivel y presenta unas guías de gestión al respecto (p. ej. para el proceso correspondiente a "asistir y dar soporte a los clientes de TI", los objetivos podrían definirse como: establecer un servicio de *help-desk*, almacenar un registro de consultas de clientes, gestionar el escalado de soporte, controlar las incidencias y presentar informes). Los objetivos de alto nivel definidos por CobiT pueden ser implementados a través de ITIL.
- **ITIL (*IT Infrastructure Library*)** [379]: publicado por la *Central Computer and Telecommunications Agency* de Reino Unido, proporciona un conjunto de buenas prácticas que se han convertido en el enfoque aceptado más ampliamente para la gestión de servicios de TI en el mundo. ITIL es esencialmente una colección de documentos, recogidos en cinco volúmenes (Estrategia de Servicio, Diseño de Servicios, Transición de Servicio, Operación de Servicio y Mejora Continua del Servicio), que se utilizan para ayudar a la implementación de un marco de ciclo de vida para la Gestión de Servicios de

TI. Este marco personalizable define cómo aplicar la gestión de servicios en una organización. Además, ITIL proporciona un vocabulario común (semántica) que permite la discusión de las etapas y actividades necesarias para construir y mantener sistemas TI a través de un lenguaje común.

- **COSO-ERM Framework (Committee of Sponsoring Organizations of the Treadway Commission - Enterprise Risk Management Framework)** [380], [381]: consiste en un marco de referencia para la gestión del riesgo en la empresa, aplicable a los entornos de computación en la nube. La aplicación de *COSO-ERM Framework* a los procesos de negocio, o a los dominios de aplicaciones soportados por los proveedores de computación en la nube, facilita a la administración una visión más completa de los riesgos asociados, los beneficios y las opciones de respuesta al riesgo. Este marco se centra en siete aspectos principales:
 - **Entorno Interno**, cómo se perciben los riesgos y los controles.
 - **Establecimiento de objetivos**, alineación de los objetivos de la organización.
 - **Identificación del evento**, identificación de oportunidades o riesgos
 - **Evaluación del riesgo**, determinación del impacto de los riesgos
 - **Respuesta al riesgo**, mitigación del riesgo
 - **Actividades de control**, asignación de responsabilidad de control (organización o proveedor de servicios en la nube)
 - **Información y comunicación**, establecer flujos de comunicación oportunos y precisos.

Dependiendo de la combinación de opciones de nube elegidas para cubrir las necesidades específicas (procesos de negocio, implementación y modelos de entrega de servicios), las preocupaciones sobre el riesgo, seguridad y cumplimiento variarán. Según los autores del *ERM Framework*, “*es una buena práctica incorporar el gobierno de la nube en las etapas iniciales (cuando se está definiendo una estrategia de computación en la nube), antes de adoptar una solución basada en la nube. Para las organizaciones que ya han adoptado la computación en la nube sin seguir las buenas prácticas descritas en el ERM, sigue siendo prudente realizar una evaluación de riesgos y establecer el gobierno de la nube*” [380].

- **ISO 27001/9000** [382], [383]: son dos familias de normas de estandarización publicadas por la *International Standards Organization* (ISO). ISO 27001 es el estándar internacional de buenas prácticas para un Sistema de Gestión de la Seguridad de la Información. ISO 9000 se compone de una serie de estándares internacionales de calidad, cuyo objetivo es la prevención de defectos mediante la planificación y aplicación de buenas prácticas en todas las fases del negocio, desde el diseño hasta la instalación y el mantenimiento.

Como puede observarse, existen numerosos modelos y marcos de referencia para el gobierno de las TI, muchos de los cuales se derivan de las normas y prácticas recomendadas por las principales entidades de estandarización de tecnologías de la información. La elección de un modelo particular o una mezcla de diferentes aspectos de varios modelos, depende de muchos factores, pero, en cualquier caso, la adopción

del gobierno de la nube es necesaria a fin de cumplir los objetivos propuestos, obtener dirección estratégica, gestionar eficientemente los riesgos y optimizar el uso de los recursos de manera adecuada.

3.4. IoT. *Internet of Things*

Aprovechando el crecimiento de la computación en nube, también hay otra tendencia tecnológica que se beneficia directamente de la gigantesca expansión de Internet, de los protocolos de comunicación de bajo consumo, de la alta reducción de costos computacionales y de la miniaturización del hardware. El Internet de las Cosas (IoT, también conocida como Internet de Objetos) se refiere a la presencia omnipresente en el entorno, de elementos inteligentes tales como sensores, actuadores, etiquetas de identificación por radiofrecuencia (RFID), etc., capaces de proporcionar funcionalidades inteligentes, reaccionando de forma autónoma, informando y aprendiendo del contexto, de los seres humanos y del entorno. Para alcanzar sus objetivos, los elementos inteligentes pueden interactuar entre sí, cooperando en ocasiones con sus vecinos para alcanzar objetivos comunes [7].

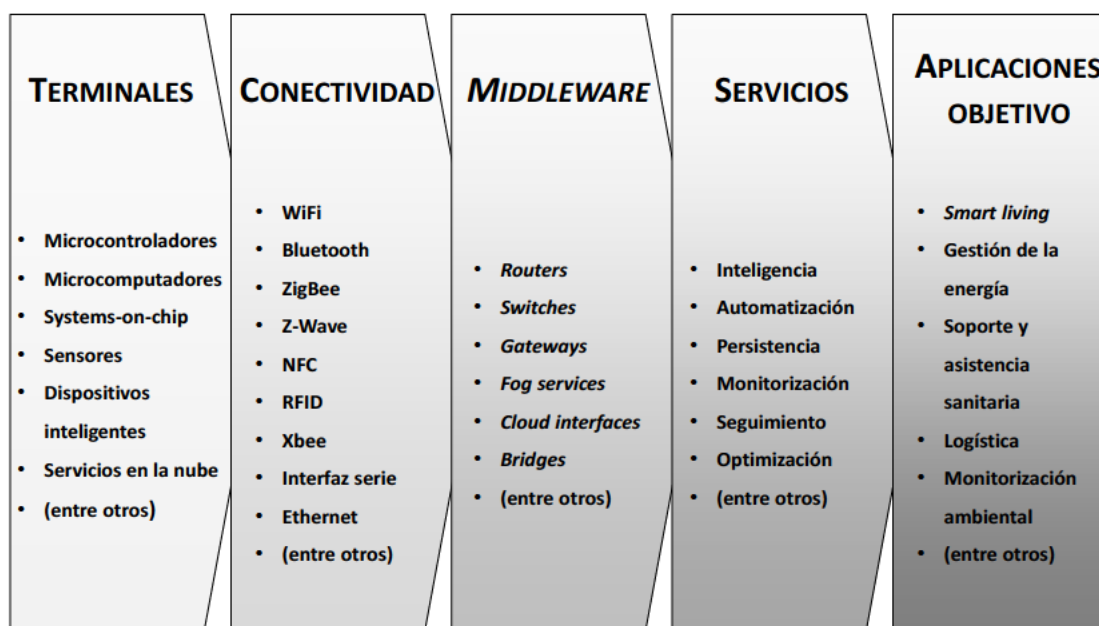


Ilustración 18. Ecosistema del Internet de las Cosas

El IoT puede ser considerado como un facilitador para mejorar el estilo de vida, al ofrecer un vasto número de servicios novedosos, tanto en dominios relacionados con la salud y el ámbito sanitario [384], [385], como en entornos inteligentes [386] (p. ej. en el hogar, en museos, en gimnasios, etc.), y mecanismos de seguridad mejorados [387], [388].

A pesar de ser una cifra difícil de determinar, se estima que hay alrededor de 15 billones de dispositivos en el ecosistema del IoT [389], sin embargo, un informe generado por CISCO predice que habrá cerca de 50 billones de dispositivos hacia el 2020 [390]. Un despliegue exitoso del IoT depende de garantizar la seguridad y la privacidad de acuerdo con sus capacidades de procesamiento, uso de recursos y requisitos establecidos. Debido al hecho de que la mayoría de los sensores y dispositivos inteligentes involucrados en IoT son inalámbricos, las soluciones y servicios de IoT son vulnerables al espionaje y a un gran número de ataques inherentes a este tipo de mecanismos y protocolos de comunicación.

Además, por lo general, los “objetos desatendidos” son vulnerables también a los ataques físicos, y la mayoría de los componentes del IoT están limitados por el consumo de energía y por un poder de procesamiento reducido. Estos factores deben ser considerados cuidadosamente cuando se implementan sistemas complejos de soporte en el área de la seguridad [391]. Los principales problemas relacionados con la seguridad, tienen que ver con la integridad de la información y la autenticación [7]. La autenticación tiende a ser difícil, ya que requiere infraestructuras específicas adecuadas y servidores que actúan como intermediarios para proporcionar seguridad adicional a toda la infraestructura. En algunos escenarios del IoT, este enfoque no es factible, ya que existen algunos componentes pasivos, como las etiquetas RFID, que no pueden llevar a cabo estos complejos protocolos de autenticación. Los algoritmos criptográficos típicos consumen gran cantidad de recursos en términos de energía y ancho de banda, tanto en la fuente como en el destino, por lo que tales soluciones no son viables para escenarios que involucran ese tipo de componentes pasivos, seriamente restringidos en términos de consumo de energía, recursos de comunicación y capacidades de cómputo. Si bien, existen algunas soluciones capaces de proporcionar un nivel satisfactorio de seguridad independientemente de la escasez de recursos. Han surgido algunas soluciones para proporcionar criptografía de clave simétrica ligera, aplicables en escenarios basados en tecnología RFID [392]–[394] y escenarios basados en redes de sensores [395]–[397].

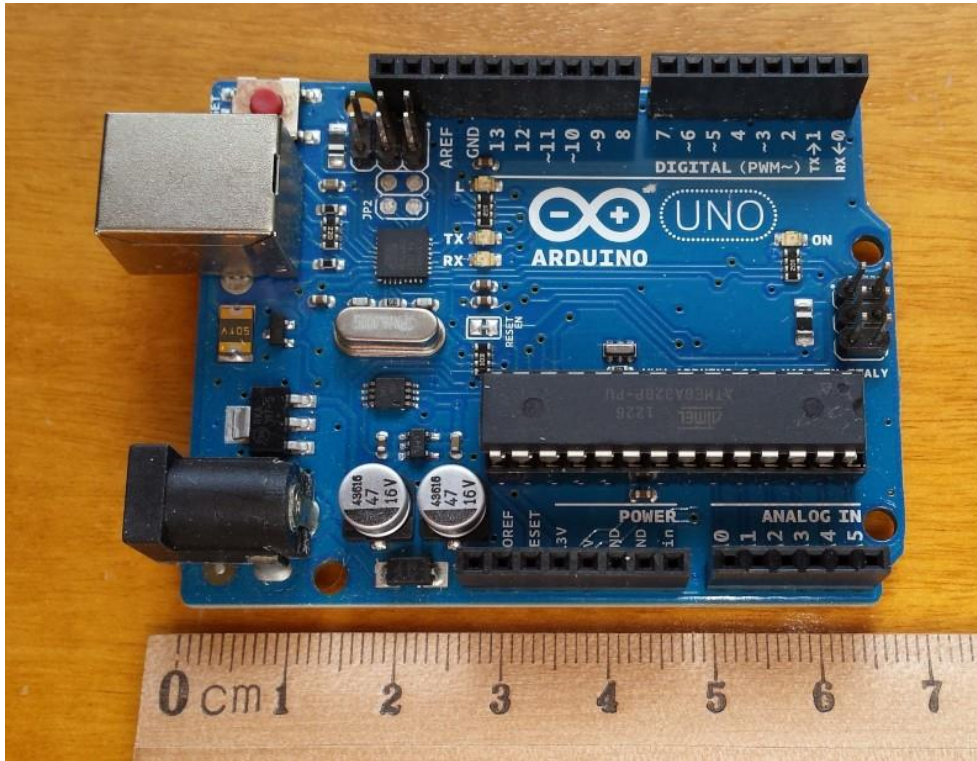


Ilustración 19. Microcontrolador Arduino UNO con escala de medida

Afortunadamente, en los últimos años han surgido notables avances en el área de los microcontroladores y *systems-on-chip* (SoC, sistemas integrados en un único circuito o chip, que presentan los distintos módulos que componen un computador), bajo un paradigma de hardware abierto. Dispositivos tales como placas Arduino, la familia de Raspberry Pi, Intel® Edison SoC y así sucesivamente [398]–[401], los cuales cuentan con interfaces *hardware* estándar y bien documentadas (tales como puertos GPIO, *General Purpose Input-Output*, o entradas y salidas de propósito general), que les permiten conectarse a un muy amplio conjunto de sensores y dispositivos inteligentes. Por otra parte, los sistemas citados son realmente pequeños (van desde el tamaño de una tarjeta de crédito, hasta el tamaño de un botón. Ver ejemplos en Ilustración 19 e Ilustración 20) y requieren una cantidad bastante reducida de energía para funcionar. En el caso de la familia Raspberry Pi, el consumo medio apenas roza los 2.2 vatios, en contraste con el promedio de 82 vatios requeridos por un computador de escritorio funcionando con el monitor apagado [402]. En cuanto a la potencia de cálculo y los recursos, en el caso de la plataforma de desarrollo Intel Edison, las especificaciones técnicas muestran que viene equipado con una CPU de doble núcleo de doble hilo Intel Atom a 500 MHz y con un microcontrolador de 32-bit Intel Quark funcionando a 100 MHz. También tiene 1 GB LPDDR3 de memoria RAM, 4 GB de memoria flash eMMC, una interfaz inalámbrica WiFi integrada de banda dual, Wi-Fi 802.11a/b/g/n y una interfaz de comunicación Bluetooth 4.0. Todo ello integrado en una placa de tamaño 35,5x25,0x3,9 mm (aproximadamente el tamaño de un pequeño sello postal) [400]. Esta completa plataforma es capaz de ejecutar una distribución de código abierto basada en GNU Linux, denominada Yocto, que cuenta con características de tiempo real, es actualizada frecuentemente (con las consiguientes mejoras en seguridad y estabilidad) y soporta la mayor parte de paquetes de Linux, los cuales

podrían necesitar ser compilados para la plataforma a fin de proporcionar funcionalidades adicionales, como es el caso de algunos protocolos VPN. Gracias a la compatibilidad de paquetes, la plataforma soporta el desarrollo en multitud de lenguajes de programación bien conocidos (como Python o Java). De igual manera, aprovechando paradigma de código abierto seguido por la plataforma, es posible reducir el tiempo de desarrollo ampliando el código existente, incrementando las funcionalidades del sistema (por ejemplo, a través del desarrollo de servicios complejos y de mecanismos avanzados de seguridad) y permitiendo el soporte y e integración sencilla ante protocolos y tecnologías futuros.

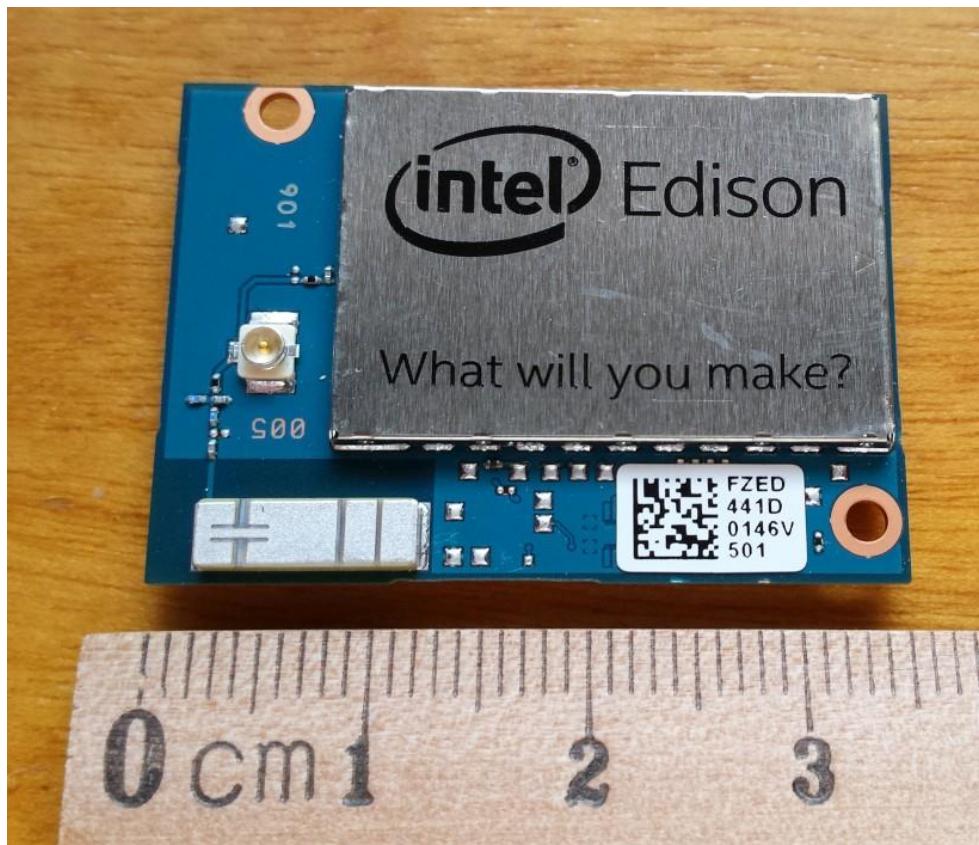


Ilustración 20. System-on-chip Intel Edison con escala de medida

El uso de estos dispositivos baratos, eficientes, de bajo consumo, dotados (o con la capacidad de ser dotados) de múltiples sensores, versátiles capacidades de comunicación y con un reducido coste de mantenimiento, pueden contribuir notablemente en el modelo propuesto al ocuparse de realizar las tareas de captura perceptual a través de dichos sensores, un pre-procesamiento inicial de la información, y su envío posterior, considerando la seguridad, a un entorno de computación distribuido mucho más potente (por ejemplo, basado en *cloud*), en el que realizar las tareas convenientes.

3.5. Seguridad de la información y de las comunicaciones

Hoy en día, las interfaces cliente basadas en web y en dispositivos móviles son dos de los métodos más comunes para acceder a los servicios albergados en la nube. Además, algunos de estos servicios basan su funcionamiento en los datos recopilados por sensores, ya sea en entornos específicos o globales (por ejemplo, seguimiento personal de la salud, ejercicio, o el pronóstico del tiempo). Con el avance de las tecnologías web y las capacidades de los dispositivos móviles, incluso es posible beneficiarse de los sensores presentes en los dispositivos personales al acceder a un servicio específico, convirtiendo un cliente pasivo en uno activo (por ejemplo, mediante el envío de registros de salud o actividad, capturados por los sensores del *smartphone*, o mediante el envío de la posición geográfica para acceder al pronóstico específico del área de presencia). Con este aumento de la conectividad y la presencia en Internet, las amenazas a la seguridad han aumentado de manera alarmante, tanto en número de casos como en diversidad de amenazas.

Según Symantec [403], las circunstancias y las implicaciones de los ataques a la web varían ampliamente. Este tipo de ataques pueden dirigirse a organizaciones, negocios específicos o simplemente pueden ser ataques generalizados de oportunidad tratando de explotar vulnerabilidades de día cero (*zero-day vulnerabilities*), o vulnerabilidades recientemente descubiertas, divulgadas y arregladas, que requieren actualización inmediata por parte de los usuarios, entre otras. La creciente omnipresencia de las aplicaciones para navegadores web, junto con la de las aplicaciones web, cada vez más comunes, fácilmente explotadas y extremadamente versátiles, ha dado lugar al crecimiento generalizado de amenazas basadas en la web. Los atacantes que quieren aprovechar las vulnerabilidades del lado del cliente ya no necesitan comprometer activamente redes específicas para tener acceso a esas computadoras. Mientras que los ataques mayores pueden tener importancia individual y, a menudo, recibir una atención significativa después de ocurrir, el análisis de la situación general muestra una visión sobre cómo los patrones de ataque están cambiando. El análisis de dicha tendencia puede proporcionar algunas percepciones acerca de cambios potenciales en la ocurrencia de ataques, y actuar como una medida preventiva acerca del tipo de ataques que pueden suceder en el futuro. Conforme a los informes de Symantec, el número aproximado de actividades maliciosas relacionadas con sitios web durante el año 2015 fue en torno a 1.1 millones eventos diarios, aumentando el número de eventos en un 117% respecto al año anterior [404].

A medida que las organizaciones buscan implementar niveles de control adecuados para minimizar el riesgo de acceso total a la web, es importante comprender el nivel de amenaza asociado a ciertos tipos y categorías de sitios web, siendo susceptibles de ser comprometidas y por lo tanto exponer a sus usuarios a mayores niveles de riesgo. Dependiendo del área de aplicación, algunas amenazas pueden ocasionar consecuencias catastróficas, como es el caso del área sanitaria. Las actividades de hacking han seguido una tendencia creciente en lo relacionado con las brechas de seguridad y datos. En 2014, Norton Cybercrime Index [405] registró 312 incidentes significativos de violación de datos, y un total de 348 millones de identidades expuestas. Según Symantec [404], en 2015 las cifras se han incrementado a: 318 grandes

incidentes de violación de información, con un total de 429 millones de identidades expuestas. La siguiente tabla muestra los diez sectores más afectados considerando el número de incidentes [403]:

PUESTO	SECTOR	# INCIDENTES	PORCENTAJE
1	Sanitario / salud	120	39.3%
2	Negocios	20	6.6%
3	Educación	20	6.6%
4	Seguros	17	5.6%
5	Hoteles	14	4.6%
6	Comercio	10	3.3%
7	Restauración y alimentación	9	3.0%
8	Ejecutivo, legislativo y general	9	3.0%
9	Instituciones depositarias	8	2.6%
10	Servicios sociales	6	2.0%

Tabla 2. Sectores afectados en base al número de incidentes

Como se puede observar, los sectores de salud, negocios y educación, fueron los más afectados por brechas de seguridad de datos durante 2015, lo que representa alrededor del 52.5 por ciento de todas las brechas de datos. Es importante destacar que más de un tercio del global de todos los incidentes son correspondientes únicamente al área de la asistencia sanitaria. Teniendo en cuenta las causas de las brechas de datos, se han identificado cuatro grupos principales: atacantes (hackers), datos accidentalmente hechos públicos, robo o pérdida de equipo o dispositivo y presencia de robos de información privilegiada desde dentro de las compañías.

En cuanto al IoT, la diversidad de amenazas refleja la variedad de dispositivos. Durante tanto en 2014 como en 2015 se ha incrementado el número de ataques experimentales en una amplia gama de dispositivos, incluyendo algunos ataques con serias repercusiones [404]. El mercado de dispositivos IoT está muy fragmentado, presentando un gran número de plataformas *hardware* de bajo costo y sistemas operativos embebidos. Existen ataques capaces de explotar vulnerabilidades en sistemas operativos Linux, integrados en dispositivos IoT y *routers*; Sin embargo, a medida que aumente y madure el mercado, el número y variedad de ataques contra algunos dispositivos específicos se verán incrementados. Aparentemente este fenómeno seguirá una tendencia similar a la observada en los ataques contra la plataforma Android, tras su enorme crecimiento en popularidad y adopción durante los últimos años [406]–[410].

Internet se compone de un gran número de dispositivos interconectados a través de concentradores, conmutadores y *routers*, que intercambian información entre ellos.

Estos dispositivos, desde *routers* domésticos hasta sistemas NAS (*Network Attached Storage*), cámaras IP, pueden considerarse parte del ecosistema IoT. Presentan conectividad a Internet, potencia de procesamiento, almacenamiento y, a veces, incorporan sistemas operativos complejos. Todas estas familias de dispositivos han sido atacadas en los últimos años, y pueden considerarse como precursores de lo que está por venir en el ecosistema IoT en cuanto a seguridad se refiere. Por ejemplo, los dispositivos NAS de Synology se han visto afectados por *ransomware* [411], [412], han surgido nuevos tipos de *malware* tipo gusano, que opera en sistemas Linux, como el caso de Darloz [413] diseñado para atacar tanto *routers* domésticos, como *set-top boxes* y cámaras de seguridad (capaz de operar en diversas arquitecturas hardware, como Intel x86, ARM, MIPS y PowerPC). Darloz se encontró se encontró en más de treinta mil dispositivos infectados en 2014 [414]. Según el análisis de Symantec, el autor del gusano actualiza continuamente su código y añade nuevas características, en particular centrándose en ganar dinero mediante el gusano, con aplicaciones como minería de moneda electrónica (*cryptocurrency*). Los atacantes de red pueden utilizar herramientas dispuestas libremente, como el motor de búsqueda Shodan [415], o como las distribuciones Linux específicamente diseñadas para realizar pruebas de penetración y auditoría avanzada, como Kali [416], con los que buscar dispositivos con conectividad de red / a Internet, como cámaras de seguridad o *smartphones*, y realizar ataques con los que acceder a los mismos. Comprometer estos dispositivos puede dar lugar a importantes problemas de privacidad y, además, derivar amenazas como la extorsión y el chantaje.

Por último, y en relación con el ámbito sanitario (área más afectada según los informes), se ha demostrado que los dispositivos médicos son relativamente inseguros y fáciles de *hackear*, como se ha comprobado en el caso de marcapasos [417], dispositivos quirúrgicos, dispositivos de anestesia, ventiladores, bombas de infusión, desfibriladores, sistemas de monitorización de pacientes [418], [419]. y en bombas de insulina [420] entre otros. Estos incidentes de ciberseguridad en dispositivos médicos pueden dañar gravemente a los pacientes afectados. Debido a que los dispositivos médicos están tan estrechamente vinculados en el proceso de cuidado sanitario, las amenazas también pueden afectar negativamente en los procesos de atención sanitaria y en los casos de uso hospitalarios. Como se analizó en el Informe sobre amenazas a la seguridad en Internet de 2016 [373] expuesto por Symantec, en 2015 las amenazas en éste área han superado el tercio del volumen total de amenazas globales. Afortunadamente, las organizaciones gubernamentales [421], las regulaciones legales existentes, las renovadas y las recientes [391], [392] en diferentes países [422, Sec. BOE nº 298], [423], y normas emergentes para la aplicación de códigos de buenas prácticas, marcos, técnicas de seguridad y controles de seguridad en las tecnologías relacionadas [382], [424], [425], pueden ayudar a mitigar esos graves problemas cuando son aplicadas adecuadamente.

Proteger un ecosistema con tanta diversidad de dispositivos interconectados, como es el caso de IoT, requiere el mismo enfoque holístico que otras áreas de seguridad de TI. Por desgracia, tanto los ecosistemas industriales de IoT (p. ej. *Industrial Internet Consortium*, IIC), como los ecosistemas de IoT de consumo (p. ej. *AllSeen Alliance*), todavía están en una fase muy temprana de definición de estándares en esta área tan cambiante. Como medida de soporte, Symantec ha publicado su “Arquitectura de

Seguridad de Referencia” [426], contribuyendo con los esfuerzos de la ICC y *AllSeen*, así como con el “Marco de Confianza para IoT” [427] de la *Online Trust Alliance* (OTA) y con los “Principios de Seguridad para Sistemas Críticos Integrados” [428] definidos por el Departamento de Seguridad Nacional de Estados Unidos. Cabe destacar que una seguridad eficaz requiere de distintas capas de seguridad incorporadas tanto en los dispositivos y como la infraestructura que las administra. Estas capas incluyen la autenticación, la firma de código y la propia seguridad física y lógica del dispositivo. La monitorización, las tareas de auditoría y la evaluación de las analíticas de funcionamiento también son fundamentales para comprender la naturaleza de las amenazas que surgen en esta área. Por último, y respecto a las comunicaciones, es fundamental el uso de tecnologías como SSL/TLS con cifrados fuertes a la hora de realizar tareas como la autenticación y la protección de datos.

Al igual que con otros aspectos de la seguridad en Internet, algunas amenazas son más peligrosas que otras, y aunque el *hackeo* de un monitor de actividad física puede quedarse en una “simple” molestia, una vulnerabilidad en millones de automóviles puede representar un peligro mucho más grave [429]. Del mismo modo, una puerta trasera en un dispositivo médico puede facilitar el acceso a los registros médicos a los atacantes, o en casos peores, conducir a lesiones graves o incluso la muerte. Las soluciones al problema están bien definidas, con lo que los fabricantes deben dar prioridad a la seguridad y encontrar el equilibrio adecuado entre la innovación, la facilidad de uso y las restricciones de tiempo en el mercado. Es fundamental disponer de las medidas adecuadas de seguridad en los dispositivos IoT para evitar presentes y futuros problemas derivados de la ausencia o mala implementación de las mismas.

Considerando que el ser humano es una “*caja negra*” en el sentido de que resulta muy difícil alterar directamente y de manera arbitraria partes específicas del conocimiento adquirido y utilizado, así como de la información sensorial que interpreta, una vez recogida por sus sensores, es necesario disponer de garantías para que un proceso análogo pueda ser realizado en el entorno computacional. Por ende, la seguridad de la información y el uso de unas buenas prácticas en el diseño de los componentes, y la manera en la que se interconectan y comunican es fundamental para garantizar la estabilidad del sistema y la coherencia del conocimiento que se persigue.

3.6. Modelos de representación del conocimiento

Tras analizar las propiedades y requisitos de la cognición, analizar los enfoques y soluciones propuestas por otros grupos de investigación, evaluar las diferentes tendencias en el área de la computación perceptual, ofrecidas por las principales empresas del sector de las TI, estudiar las diferentes soluciones tecnológicas capaces de soportar un conjunto distribuido tan grande y complejo de tareas que cooperen entre sí, junto con sus ventajas, inconvenientes y buenas prácticas, queda investigar sobre los principales modelos existentes de representación de conocimiento y sus características esenciales.

Dada la inherente naturaleza matemática de la informática, los principales modelos surgidos inicialmente para el tratamiento y representación de conocimiento se basaron en la lógica matemática clásica [430]. Dichos modelos parten del uso de la lógica formal como lenguaje de representación y como principales características presentan mecanismos de inferencia potentes, bien conocidos y formalizados, como por ejemplo el “*modus ponens*”, considerado como regla de inferencia primitiva [431].

Dentro de la lógica matemática han surgido diferentes variantes partiendo de ramas clásicas como la lógica proposicional y la lógica de predicados de primer orden. Variantes como son la lógica modal, encargada del estudio del comportamiento deductivo de las expresiones “es necesario que” y “es posible que”, relacionadas con la verificación de los juicios de verdad [432], o la lógica difusa, basada en la teoría de conjuntos difusos [433] y en el grado probabilístico de pertenencia (grado de veracidad) de un elemento a uno de éstos conjuntos.

Los modelos formales basados en la lógica se centran principalmente en la representación de hechos y características verificables de los elementos desde un punto de vista idealizado (generalizado) más que en la definición de las categorías fundamentales y relaciones básicas que determinan la existencia de los conceptos [434].

Por otra parte, aunque guardando cierto grado de relación con los modelos formales, existen una serie de modelos estructurados, cuyos mecanismos de inferencia, aunque limitados, son por lo general muy eficientes a la vez que específicos. Modelos como, las redes semánticas, basadas en el concepto de memoria asociativa y vinculada a la idea de que la memoria se estructura a través de la asociación entre conceptos. Dichas redes cuentan con su base formal en los trabajos de Cayley y Sylvester [435], aunque su expansión no llega hasta que Quillian las introduce en el área de la Inteligencia Artificial al desarrollar un sistema sobre significados de palabras [436]. Entre estos modelos se pueden destacar:

- Redes Is-As
- Marcos
- Redes semánticas
- Redes bayesianas
- Ontologías

Y dos lenguajes principales, muy versátiles y potentes, para la representación de algunos de los modelos anteriores:

- SGML (como modelo tipo clave-valor y marcado)
- JSON y JSON-LD

Por lo general, con redes semánticas se suele hacer referencia a los esquemas redes Is-As y a las redes de marcos, ya que comparten algunas características fundamentales, como es la herencia por defecto. En una red semántica, los conceptos están organizados en una red en la que existe un nodo superior (nodo *top*, representado con el carácter T) al que se le asigna uno o varios hijos que a su vez tienen otros conceptos hijos y así sucesivamente hasta que se alcanza el final de la red (nodos *bottom*, representados con el carácter ⊥), cuyos nodos son instancias en lugar de conceptos. A continuación, se analizarán los modelos y lenguajes citados anteriormente.

3.6.1. Redes Is-As

Este tipo de red está formada por una jerarquía taxonómica, cuya base principal es un sistema de enlaces de herencia entre los conceptos de representación o nodos. Las redes Is-As intentan particularizar las instancias en base a clases más generales, que a su vez son subconjunto de otras clases todavía más generales. Como ejemplo, las taxonomías clásicas son buenos ejemplos: una araña es un insecto, un insecto es un invertebrado, un invertebrado es un animal.

Aunque la representación empleando lógica de predicados es suficiente para expresar relaciones entre elementos, la estructuración jerárquica proporcionada por las redes Is-As facilita que la asignación de propiedades a una determinada categoría en la red se reduzca únicamente a aquellas que son específicas de la misma, heredando las propiedades de sus categorías superiores en la jerarquía.

Estas redes, aunque muy flexibles, presentan una serie de problemas que limitan su uso en diversos ámbitos. Por ejemplo, la elección de los conceptos de representación y de los arcos que los unen es crucial en la fase de análisis, pues una vez se ha decidido una estructura determinada resulta muy complicado cambiarla. Por otra parte, es complicado expresar cuantificación (“*algunos gatos son pardos*”) o intencionalidad (“*Ricardo cree que Vladimir tiene coche*”).

3.6.2. Marcos

Son un modelo de representación del conocimiento propuesto por Marvin Minsky en 1974, cuya idea inicial se expresaba de la siguiente manera [437]: “*Un Marco es una estructura de datos para representar una situación estereotipada [...] como una red de nodos y relaciones. Los niveles superiores de un Marco son fijos y representan cosas*”

que son siempre ciertas en la situación supuesta. Los niveles inferiores tienen muchos terminales que tienen que llenarse con instancias o datos específicos [...] Las suposiciones por defecto son asociadas débilmente con sus terminales, de forma tal que puedan ser reemplazadas por nuevos elementos que se ajusten mejor a la situación actual”.

La representación de un marco suele realizarse por medio de un conjunto de *slots* o casillas de información relativas a dicho marco. Cada *slot* contiene la información sobre un atributo del objeto que se modela, y se le pueden asociar varios tipos de información [438] entre las cuales están: *value* (valor), *procedure* (un procedimiento para calcular el valor), *demon* (procedimientos que se ejecutan al cambiar el valor almacenado en *value*), *default* (valor inicial o por defecto) y *restrictions* (conjunto de condiciones lógicas que deben ser verdaderas para el valor almacenado en *value*), *explanation* (descripción del marco) y *univalued* (valor booleano en función de si el atributo tiene un único valor o puede presentar varios).

Las principales características de los marcos atendiendo a [437] son:

- **Precisión:** los objetos, sus propiedades y relaciones con otros objetos son descritas de una forma precisa. Igualmente es posible restringir los valores de los mismos. En caso de que un valor no pertenezca al rango se puede utilizar el valor definido por defecto.
- **Modularidad:** la base de conocimiento está organizada en componentes claramente diferenciados.
- **Herencia no-monotónica:** por defecto los marcos están conceptualmente relacionados, permitiendo que los atributos de los objetos sean heredados de otros objetos predecesores en la jerarquía.
- **Triggering:** o activación dinámica de procesos. Es posible ejecutar los procesos contenidos en un marco de forma automática ante el cambio de alguna propiedad o valor.
- **Consistencia:** en la sintaxis para referenciar objetos del área, que a su vez es fácil de leer.

Los marcos se suelen emplear para representar conocimiento estereotipado o construido a partir de experiencias y características bien conocidas. En el ser humano, se dispone de una gran cantidad de conocimiento, sentido común y experiencias almacenadas en el cerebro que permiten analizar un objeto o experiencia nueva al resolver un problema. Análogamente al ser humano, los marcos presentan mecanismos de razonamiento que tratan de colocar un nuevo marco en el lugar más apropiado dentro de la jerarquía existente. Para ello se requiere que el mecanismo de reconocimiento pueda recibir información de la situación en formato marco y realice una búsqueda del marco más adecuado entre los existentes en la base de conocimiento.

El problema de este mecanismo reside en la conceptualización inicial del ámbito que se está representando y la estructuración de dicha información en una red de marcos. En un nivel fundamental, los elementos del ámbito se codificarán como prototipos con una serie de valores por defecto en una jerarquía taxonómica, pero puede resultar

complicado determinar los prototipos (sus características) así como la taxonomía (o reglas de clasificación).

3.6.3. Redes semánticas como modelo de visualización y presentación

Las redes semánticas son esquemas de representación de conocimiento basados en gráficos dirigidos. Los nodos representan objetos o conceptos y los enlaces representan relaciones entre nodos. Los nodos y los enlaces están generalmente etiquetados. Se puede apreciar un ejemplo en la siguiente figura:

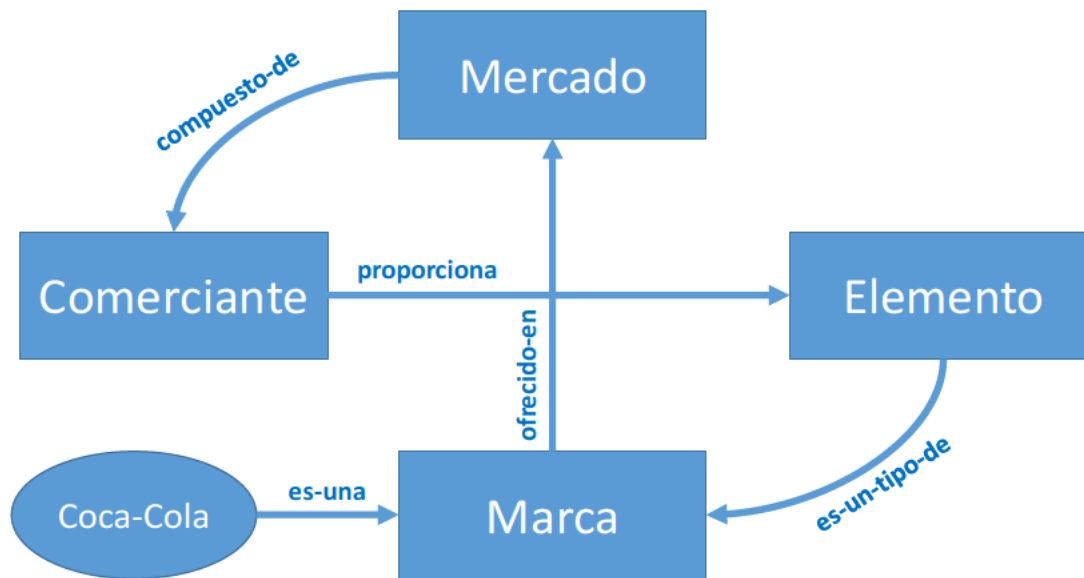


Ilustración 21. Ejemplo de red semántica

Algunos de los primeros usos de la formulación basada en nodos-y-enlaces se empleó para modelar cómo el lenguaje natural es entendido y cómo los significados de las palabras se pueden capturar en una máquina [439], pero a lo largo del tiempo, dicho modelo ha sido empleado para distintos usos, tales como modelar relaciones complejas entre sistemas [440] mediante redes semánticas.

Las redes semánticas presentan algunos problemas complicados en el área de la representación de conocimiento, relacionadas con la expresividad. Por ejemplo, respecto a la negación y a la disyunción, que no se pueden representar de forma sencilla utilizando dicho modelo. Lo mismo ocurre con la cuantificación [441].

3.6.4. Redes bayesianas como modelo estocástico

Una red bayesiana es una herramienta para modelar y razonar con hechos inciertos. Una red bayesiana consiste en dos partes: un componente cualitativo en forma de grafo acíclico dirigido (DAG, *directed acyclic graph*), y un componente cuantitativo en forma de probabilidades condicionales. Los nodos del grafo representan las variables de interés y los enlaces representan influencias directas entre dichas variables. Las probabilidades condicionadas de una red bayesiana cuantifican las dependencias entre las variables y sus padres en el grafo acíclico dirigido. Sin embargo, formalmente una red bayesiana se interpreta como la especificación de una distribución de probabilidad única sobre sus variables.

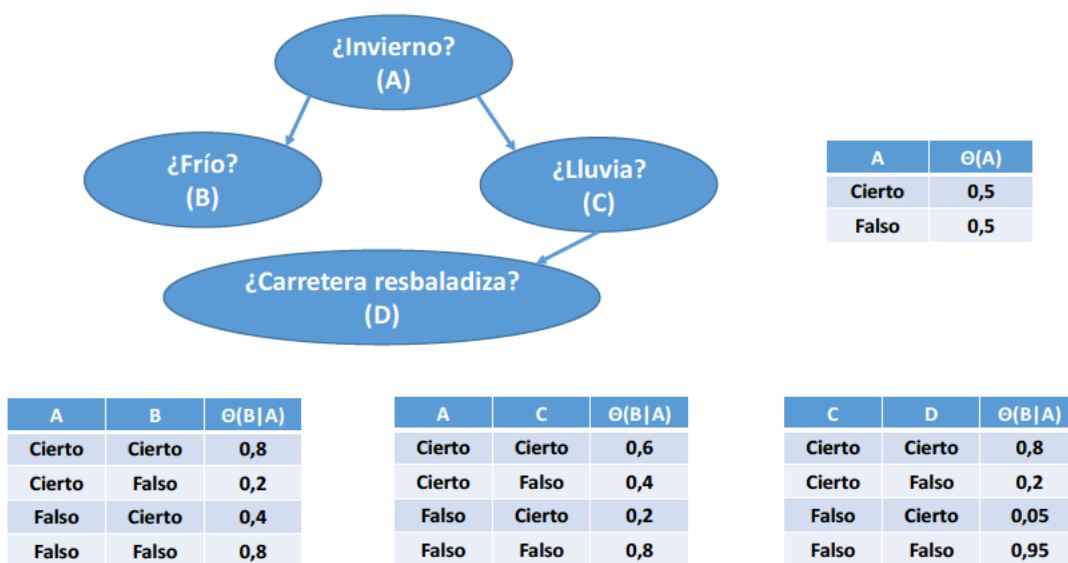


Ilustración 22. Ejemplo de red bayesiana

En las Ciencias de la Computación, las redes bayesianas son utilizadas para un gran número de aplicaciones, tales como la recuperación de información [442], servicios de razonamiento para la web semántica [443] y el soporte a la toma de decisiones [444].

3.6.5. Ontologías como modelo de representación

El término “ontología” (adoptado de la filosofía, definido por primera vez por Aristóteles como “estudio de lo que existe”) comenzó a ser empleado de forma generalizada en la comunidad de la Ingeniería del Conocimiento a partir del trabajo de Gruber [445] en el que se presentaba una definición del término como “una especificación formal y explícita de una conceptualización compartida”. Aunque la definición sea algo vaga, una ontología refleja una conceptualización, o lo que es lo mismo, una visión del mundo respecto a un cierto dominio, estando formada por un conjunto de conceptos (entidades, procesos, atributos, etc.) junto con sus definiciones e interrelaciones [446].

Dada la complejidad y completitud de estos modelos, la definición propuesta por Gruber ha ido evolucionando a través de otros trabajos, incluyendo ideas tales como *conjunto de términos jerárquicamente estructurados* [447], *representación explícita de conceptos en un dominio* [448] y *teoría que utiliza un vocabulario específico* [449].

De esas definiciones e ideas propuestas por los autores citados, es posible identificar algunos aspectos generales de las ontologías:

- Las ontologías son utilizadas para describir un dominio específico.
- El significado de los términos utilizados en una ontología específica debe ser consistente entre sus usuarios.
- Las relaciones y los términos utilizados por una ontología deben ser definidos claramente.
- Los términos son organizados siguiendo un mecanismo específico (tales como una estructura jerárquica. Ver Ilustración 23).
- Los términos y axiomas representados en la ontología restringen la interpretación de dichos términos [450].

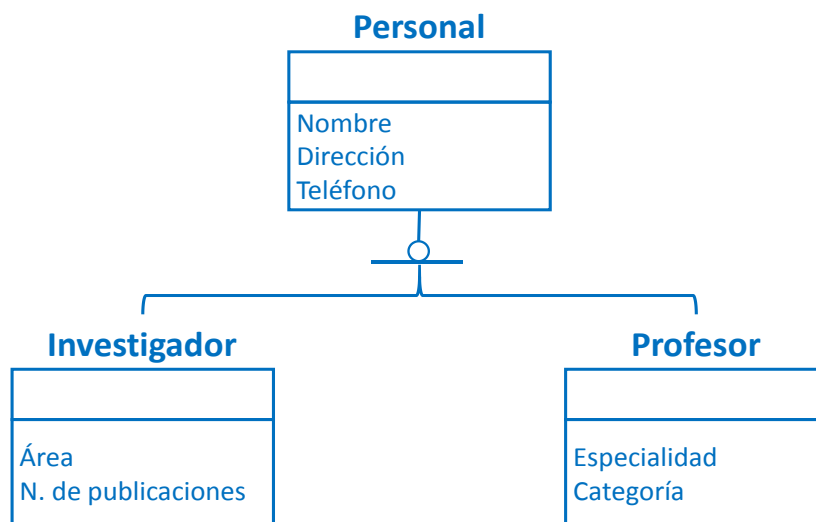


Ilustración 23. Ejemplo básico de una ontología

La representación del conocimiento recogida en una ontología es explícita. Esto es, que tiene entidad como tal, de forma aislada de cualquier uso que se haga de ella. Desde el punto de vista de la informática, se podría ver como el conocimiento y el código del programa: el modelo existe, aunque no haya ninguna aplicación que lo utilice.

Una amplia variedad de tareas, en diferentes áreas de investigación, están soportadas por las ontologías. Aparte de la representación de conocimiento, algunas tareas relevantes facilitadas por las ontologías son: compartir conocimiento en sistemas multi-agente, gestión del conocimiento, adquisición del conocimiento y recuperación de la información [451], [452].

En lo que respecta a la gestión del conocimiento, las ontologías permiten la intercomunicación con recursos de información entre humanos y/o agentes software. Las relaciones en las ontologías (las cuales describen la información semántica de los datos) son legibles por humanos y sistemas computacionales, permitiendo habilidades tales como crear enunciados y contestar peticiones bajo un dominio específico, debido a la conceptualización que describe a las entidades y a sus relaciones. La utilidad de las ontologías en sistemas basados en agentes puede ser vista como habilitadora de interoperabilidad a nivel del conocimiento [453]. De manera similar, las ontologías soportan compartir el entendimiento, la interoperabilidad entre herramientas, la ingeniería de sistemas, la reusabilidad y la especificación declarativa [454].

Las ontologías conforman las bases de conocimiento, las cuales están compuestas por la ontología en sí y un conjunto de instancias individuales de sus clases. [448]. Este tipo de almacenamiento del conocimiento puede ser utilizado por agentes inteligentes a fin de enriquecerse, reutilizarse y mantenerse. De acuerdo con [455], las bases de conocimiento están formadas por información dependiente del estado, mientras las ontologías concentran información independiente del estado.

Con respecto a la adquisición de conocimiento, las ontologías pueden ser utilizadas como una herramienta útil. Por ejemplo, los equipos de trabajo pueden emplear ontologías como un soporte común para clasificar el conocimiento de una organización, ya que las ontologías permiten a los usuarios reutilizar el conocimiento en nuevos sistemas.

En lo referente a las tareas de recuperación de información, las ontologías pueden ser utilizadas para elaborar taxonomías completas de términos a fin de mejorar la precisión de los resultados y para desambiguar las peticiones realizadas por los usuarios [455]. Igualmente, las ontologías basadas en las interacciones con los usuarios pueden ser extendidas por medio de técnicas de aprendizaje automático.

Actualmente, las ontologías y otros modelos son utilizados a fin de conseguir la representación del conocimiento. Al tratarse de modelos muy versátiles, permiten la conceptualización de entidades dentro de un dominio, y la caracterización de relaciones entre ellos. Aun así, dichos modelos presentan algunas desventajas, tales como la heterogeneidad entre los distintos lenguajes basados en ontologías [456] y las dificultades combinando el lenguaje existente ya conceptualizado [457], aunque en la última década han surgido distintos enfoques para emparejar, mezclar e integrar ontologías. Enfoques basados tanto en la interacción humana (supervisados y semi-supervisados) [458] como en técnicas autónomas (sin supervisión) [459]. En cualquier caso, el uso común de las ontologías es el de representar áreas específicas de conocimiento, tales como la biomédica o comercio electrónico, pero no un dominio global de conocimiento combinado. Esto puede ser un problema si se trata de representar una combinación de dominios heterogéneos relacionados con la entrada sensorial de distintos dispositivos. A fin de percibir el entorno y categorizarlo de manera análoga a como lo hacen los seres humanos, la solución debería ser capaz de unificar dichos dominios para permitir la interrelación de dichos conceptos y crear una representación global de conocimiento, unificando la información en cada dominio, a fin de tener una conceptualización completa de entidades. Y para ello, sería deseable

tener una conceptualización de las entidades basada en múltiples y diferentes aspectos cualitativos, tales como el tamaño, el color, el sonido, aroma, tamaño, etc.

Los diccionarios, tesauros, etc. son considerados ontologías en algunos ámbitos. Igualmente, se han definido términos como ontologías ligeras (*lightweight ontologies*) para definir una jerarquía simple de conceptos. De este modo, dependiendo del tipo de construcciones conceptuales que contienen, se pueden encontrar diferentes tipos de “*estructuras parecidas a ontologías*” [460]. La noción o versión más simple de ontología sería una lista de términos, como por ejemplo un catálogo. El siguiente paso en complejidad sería el glosario (una lista de términos con su significado *expresado en lenguaje natural*). Otro paso más serían los tesauros, que contienen información adicional de tipo semántico (por ejemplo, una taxonomía de la especificidad de los términos). Todos estos sistemas están pensados para ser usados por un agente humano, no por un agente computacional. Para que un ordenador pueda utilizar sistemas como los anteriormente presentados, las definiciones deben ser in-ambiguas y estar expresadas por medio de algún formalismo. El tipo de ontología más simple procesable (computacionalmente hablando) consistiría en una jerarquía que recoja la relación de generalización/especialización entre los términos.

Podría resumirse la diferenciación entre vocabularios controlados, taxonomías y tesauros frente a las ontologías con que los primeros están orientados a recoger y organizar términos, mientras que las segundas están dirigidas a representar el conocimiento de forma independiente del agente que lo utilice. Los modelos de datos y de objetos están dirigidos a representar la información en el ámbito computacional.

Existen numerosos sistemas de representación ontológicos, como por ejemplo OWL, los marcos (OKBC), XML+DTDs, esquemas de bases de datos, modelos de datos (UML, STEP), etc. cada uno de ellos orientado a representar conceptos específicos pertenecientes a distintos ámbitos.

3.6.6. SGML como modelo tipo clave-valor y marcado

El Language Estandar Generalizado de Marcado (SGML, *Standard Generalized Markup Language*) es un lenguaje estándar para la descripción de documentos (ISO8879:1986) propuesto por la Organización Internacional para la Estandarización (ISO, *International Organization for Standardization*). Está diseñado para habilitar intercambio de texto y orientado para ser usado tanto en el ámbito de la publicación como en áreas de oficina e ingeniería. Los documentos basados en SGML tienen una estructura descrita rigurosamente, lo cual permite la realización de dos tareas fundamentales en una forma muy sencilla: el tratamiento computacional de la información y el análisis y entendimiento del contenido por seres humanos [461].

SGML es considerado un meta-lenguaje, ya que ha permitido generar muchos lenguajes bien conocidos y extensivamente utilizados (subconjuntos de SGML), tales como el Lenguaje de Marcado de Hipertexto (HTML, *Hyper Text Mark-up Language*) y el Lenguaje de Marcado Extensible (XML, *eXtensible Mark-up Language*). Este último diseñado inicialmente para la implementación de un *parser* comparado con la

especificación completa de SGML. Dicha característica entre otras restricciones adicionales (tales como el cierre forzado de cada etiqueta abierta) ha permitido a XML ser más utilizado que la especificación completa SGML. La aplicación de este lenguaje y otras variantes derivadas del mismo están presentes incluso en la definición de ciertos lenguajes de ontologías, tales como el Lenguaje Web de Ontologías (OWL, *Ontology Web Language*), [462].

3.6.7. JSON y JSON-LD

JavaScript Object Notation (JSON) [463] es un formato de intercambio de datos ligero, basado en texto e independiente del idioma. Se deriva del estándar de lenguaje de programación ECMAScript (basado a su vez en JavaScript). JSON define un pequeño conjunto de reglas de formato para la representación portátil de datos estructurados.

Los objetivos de diseño de JSON incluyen el minimizar su complejidad, hacerlo portátil, textual y que constituya un subconjunto de JavaScript. JSON puede representar cuatro tipos primitivos (cadenas, números, booleanos y nulo) y dos tipos estructurados (objetos y matrices). Los términos "objeto" y "matriz" provienen de las convenciones de JavaScript. Una cadena es una secuencia de cero o más caracteres *Unicode* [464]. Un objeto es una colección no ordenada de cero o más pares clave/valor, donde la clave es una cadena, y un valor puede estar representado por una cadena, número, booleano, nulo, objeto o matriz. Finalmente, una matriz es una secuencia ordenada de cero o más valores.

JavaScript Object Notation for Linking Data (JSON-LD) es una extensión de JSON para vincular datos y que proporciona información sobre el contexto. El contexto vincula propiedades de objeto en un documento JSON a conceptos en una ontología. Un contexto puede ser embebido directamente en el documento JSON-LD, o puesto en un archivo separado, y referenciado desde diferentes documentos (desde documentos JSON tradicionales a través de un enlace HTTP). Al extender a JSON presenta características similares, como su simplicidad en ser generado e interpretado tanto por humanos como por máquinas. Su principal objetivo es el de ayudar a los datos JSON a interactuar a escala web, dotando de información semántica adicional. Estas propiedades adicionales convierten a JSON-LD en un lenguaje de intercambio de datos vinculados ideal para entornos JavaScript, servicios Web y bases de datos no estructuradas como MongoDB (*Humongous Database*) [465].

JSON-LD fue desarrollado por el W3C *JSON-LD Community Group* [466], pasando a formar parte de las recomendaciones de la W3C en enero de 2014 [467]. Actualmente su desarrollo y acceso son abiertos y públicos a través de sus repositorios en GitHub [468], pero todavía coordinados por el *JSON-LD Community Group*.


```
{
  "@context": "http://json-ld.org/contexts/person.jsonld",
  "@id": "http://dbpedia.org/page/Nikola_Tesla",
  "name": "Nikola Tesla",
  "born": "1856-07-10",
  "patents": "around 278"
}
```

Ilustración 24. Ejemplo de entidad codificada mediante JSON-LD

La codificación JSON-LD se emplea en servicios importantes, como es el caso de *WordNet* (base léxica para el idioma Inglés) [469], o el caso de la API del gráfico de conocimiento de Google [470], para devolver entradas de información que cumplan los requisitos solicitados.

3.7. Conclusiones del capítulo

El presente capítulo ha comenzado con un estudio sobre las principales líneas de investigación paralelas, y los tipos de soluciones o conclusiones presentadas, evaluando las tendencias y resultados para poder determinar qué enfoques resultan más adecuados a la hora de proseguir con el desarrollo del modelo de representación del conocimiento. Se han identificado dos tipos principales de soluciones que aprovechan las características de la cognición o de la percepción para conseguir sus objetivos: “las soluciones aplicables a sistemas físicos”, por lo general, de naturaleza cognitiva y no sólo perceptual, utilizadas fundamentalmente en sistemas cibernéticos, y “las soluciones aplicables a servicios”, que suelen ser más limitadas en cuanto a prestaciones respecto a los sistemas cibernéticos, pero resultan mucho más potentes y eficientes en cuanto a tareas específicas se refiere (p. ej. tareas de aprendizaje automático basadas en identificación visual o auditiva, o comprensión del lenguaje natural). El primer tipo de soluciones cubre un amplio abanico de opciones, muy complejas en cuanto a funcionalidad, al tratar de emular sistemas generalistas completos que responden a los estímulos (por lo general, aplicados a la robótica). Sin embargo, dicha complejidad limita la capacidad de los mismos, y hace que los estímulos percibidos (sentidos) sean escasos en número (por lo general, vista, en ocasiones oído, y en pocas, propiocepción y tacto). En el caso del segundo tipo de soluciones, al tener una aplicabilidad en otros servicios (por lo general servicios *software*, facilitados a través de la web), se han desarrollado sistemas menos completos, pero cuya complejidad reside en la elevada especialización y servicios que los componen. Estos servicios por lo general requieren mucha potencia de cómputo, y trabajan sobre estímulos específicos (sentidos) de manera individual (p. ej. reconocimiento de elementos en imágenes, reconocimiento de sentimiento e intención

en texto, reconocimiento de elementos en pistas de audio, interpretación de la voz y del discurso, etc.). Por desgracia, no se han encontrado estudios orientados al aprendizaje multimodal holístico, considerando como entradas perceptuales todas las posibles (existentes y venideras), bajo un marco común de representación. Como se ha estudiado en el presente y anterior capítulo, muchos autores consideran que un marco así de complejo, capaz de cubrir el área de percepción (sentidos), conocimiento (representación) y memoria (almacenamiento), es habilitador de otras tareas avanzadas de la cognición, como el razonamiento y la inteligencia.

Tras el análisis de las investigaciones paralelas, y teniendo en cuenta las conclusiones derivadas del estudio de la cognición, se determinó que un modelo con la ambición en término de objetivos como el propuesto, requerirá de una serie de características (cómputo, almacenamiento, resiliencia, entre otras), que a fecha actual sólo pueden conseguirse mediante grandes redes de servidores interconectados, trabajando en paralelo. A este respecto, se ha considerado el campo de la computación en la nube, tendencia tecnológica actual, como muy prometedora para satisfacer dichos requisitos impuestos por el modelo. Así pues, el siguiente paso ha sido el de analizar las características de la computación en la nube, los distintos tipos y enfoques, y los problemas asociados, junto con los estándares y marcos de referencia a tener en cuenta para minimizar todos esos problemas (gobierno de la nube).

El siguiente aspecto a considerar es la entrada sensorial del modelo, o, dicho de otra forma, la percepción de los elementos del entorno en base a sus características fundamentales. Gracias al movimiento del Internet de las cosas, existen infinidad de dispositivos interconectados, dotados de multitud de sensores. Dichos dispositivos pueden ser aprovechados como entradas sensoriales, pero teniendo en cuenta una serie de consideraciones que se expondrán en el siguiente capítulo. Considerando la versatilidad del modelo propuesto, bien podría enmarcarse entre la ciberéntica / sistemas cognitivos (hace uso de hardware para capturar información, y puede aplicarse a sistemas robóticos o automáticos para mejorar sus capacidades) y entre las soluciones software de reconocimiento y aprendizaje (ver Ilustración 25), pues a través de un conjunto de algoritmos y servicios software interconectados proporcionará ambas características.

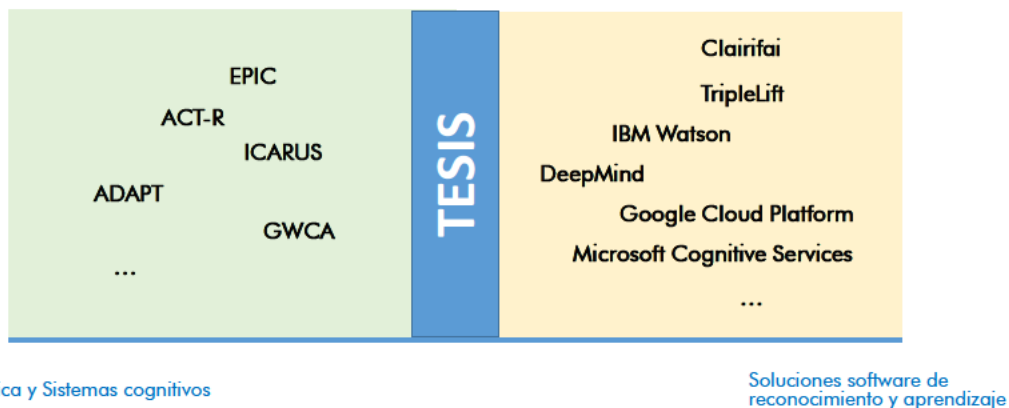


Ilustración 25. Posición del marco de referencia propuesto en la Tesis

Finalmente, y para concluir con el capítulo, se analizaron una serie de modelos de representación del conocimiento, interesantes y relevantes para el problema a resolver, como es el caso de los marcos, las redes bayesianas o las ontologías. Respecto a estas últimas y dada su versatilidad, parece prometedor orientar el modelo de representación del conocimiento en base a ellas. Sin embargo, dado que las ontologías tienden a almacenar información semántica en formato textual, y están centradas en dominios relativamente cerrados aun presentando diversas dimensiones [471], será necesario adaptar los modelos existentes para permitir la representación del conocimiento a nivel global, o lo que es lo mismo, en un dominio abierto. En ésta última parte también se analizaron estructuras de información que pueden ayudar a realizar el intercambio eficiente del conocimiento entre distintos sub-sistemas y entidades de la solución final (p. ej. XML, JSON, o JSON-LD).

Capítulo 4. Propuesta (marco teórico)

4.1. Introducción

Como se ha descrito a lo largo de la tesis doctoral, gracias a la inmensa capacidad de interconexión entre dispositivos y sistemas gracias a redes como Internet, al abaratamiento y mejora en prestaciones de la tecnología, y a la popularización masiva del uso de dichas tecnologías debido al conjunto de servicios desarrollados para dotarlas de utilidad en casi todas las áreas posibles, el volumen de información generado es muy difícil de tratar. Por otra parte, como los datos generados proceden de diversas fuentes heterogéneas (computadores de escritorio, *smartphones*, servicios en la nube, estaciones de seguimiento meteorológico, puestos de monitorización sísmica y un creciente etcétera), y se presentan en diferentes naturalezas (texto, imágenes, flujos de vídeo, fragmentos de sonido, etc.), resultan muy difíciles de analizar y estructurar, dotándoles de una semántica fundamental para comprender las relaciones y estado conceptual del entorno. Esta información semántica extendida sobre los conceptos, podría ayudar a dotar a sistemas de inferencia y planificadores a obtener relaciones desconocidas hasta la fecha entre elementos, o a generar planes de ejecución más eficientes y avanzados.

En este capítulo se plantea una solución al problema de la heterogeneidad y disgregación entre la información, en forma de un marco de referencia para la creación de un modelo global de representación de conocimiento basado en la percepción humana, y modelada mediante un conjunto de modelos de representación interconectados. Este enfoque tiene diversas ventajas mostradas en el estudio, al igual que algunos inconvenientes relacionados con las tecnologías encontradas en el estado de la cuestión. Por ello, el capítulo se estructura de la siguiente manera: en primer lugar, se presenta un conjunto de Principios Fundamentales (PFs), inferidos de los requisitos y características asociados a la cognición, y a los tres principales pilares sobre los que se fundamentará el modelo, esto es, la percepción (y los sentidos perceptuales), el conocimiento y la memoria. Seguidamente, se presentarán un conjunto de Características Técnicas de diseño (CTs) sobre las que construir un modelo arquitectónico de la solución. Nótese que tanto los PFs como las CTs mostradas en el capítulo son las obtenidas tras el proceso de validación con los expertos, que se detallará en el capítulo 5. Finalmente, se propone un patrón de análisis sensorial, módulo básico de la arquitectura lógica, y base de la arquitectura física, definida en base a un módulo inicial de visión, y seguidamente extendido en base a cualquier tipo de entrada sensorial, aprovechando el ecosistema IoT.

A lo largo del capítulo se exponen características esenciales que debe cumplir el sistema, y las soluciones tecnológicas que a fecha actual pueden cubrirlas de manera adecuada. Sin embargo, cabe destacar que tanto los PFs como las CTs han sido definidos independientemente de la tecnología, con el objetivo de que en caso de que en un futuro surjan nuevas soluciones tecnológicas más eficientes, que se aproximen mejor a los requisitos específicos definidos, puedan ser directamente aplicadas. Así

pues, se pretende mediante un diseño abstracto y modular, una gran capacidad evolutiva y un incremento en el mantenimiento y escalabilidad de las diferentes partes del sistema, al no estar sujetas a soluciones concretas, pudiendo cambiar y mejorar junto con el progreso tecnológico.

4.2. Planteamiento del problema

Como se comentó en el apartado 1.2, tras realizar un estudio del estado del arte en el ámbito tecnológico, y tras consultar a distintos expertos en la materia, no se han encontrado evidencias sobre la existencia de un sistema singular capaz de aprender los conocimientos de cualquier disciplina y realizar un proceso de razonamiento que combine los conocimientos de diferentes áreas. Además, al ser capaz de aprender a partir de diferentes fuentes, tal y como lo pueden hacer los seres humanos, puede garantizar una reducción dramática de la redundancia en la base de datos de conocimiento, pues según la teoría de prototipos [14] y de espacios conceptuales [472], consideradas en la psicología como modelos de representación conceptual humano, la clasificación que se realiza de un concepto es única, pudiendo presentar dicho prototipo diferentes características diferenciadoras que guardan una relación de proximidad. No sólo se garantiza una reducción en el tamaño de almacenamiento, sino que se aumenta la precisión en la recuperación de información al caracterizar mejor a los conceptos del dominio. Esto podría ayudar a crear motores de búsqueda semánticos precisos, útiles en Internet; aplicables a la investigación en nuevas áreas; el desarrollo de sistemas expertos en dominios abiertos, como por ejemplo el diagnóstico médico o la asistencia a la toma de decisiones en la empresa. Desde un punto de vista a largo plazo, un sistema capaz de aprender y almacenar la información de forma análoga a como lo hacen los seres humanos ayudaría a mejorar el nivel de vida de todo ser humano, al tener aplicación directa en todas las áreas en las que operan los mismos.

En la actualidad se utilizan principalmente las ontologías como modelo de representación de conocimiento. Este tipo de modelos presenta algunos inconvenientes, como la heterogeneidad entre los lenguajes de ontologías [456] o la dificultad a la hora de combinar conocimientos de diferentes áreas [473]. Además, las ontologías suelen emplearse para representar áreas específicas e individuales de conocimiento, como el área biomédica o área de negocio de comercio electrónico, pero no un área global combinada de conocimiento [474]. El modelo propuesto en la tesis doctoral, basado en la percepción para la representación del conocimiento, está destinado a romper esta barrera, acabando, o mermando los inconvenientes de los modelos actuales.

El desarrollo de un modelo de representación del conocimiento implica desarrollar un sistema de almacenamiento que soporte dicho modelo, y un sistema mediante el que introducir datos en dicho modelo. El sistema que de soporte al modelo propuesto tendrá que ser eficiente y escalable, pues como ya se ha mencionado, el objetivo principal es poder almacenar el conocimiento de manera global, que, como se ha visto

previamente, sigue una tendencia de crecimiento exponencial [2]. Finalmente, dicho sistema debe cubrir las necesidades de los servicios que lo utilicen, con lo que también se vaticina un enorme número de accesos concurrentes por parte de diferentes sistemas externos, como motores de búsqueda, con lo que deberá soportar adecuadamente la concurrencia. A este respecto, también es importante mencionar la seguridad, fundamental para el buen funcionamiento del mismo.

Resulta conveniente asentar una serie de bases y principios fundamentales, tanto a nivel psicológico y neurocientífico, como a nivel tecnológico, sobre el que poder desarrollar un sistema estable que cumpla las características deseadas.

4.3. Características y requisitos generales de la cognición. Principios Fundamentales

Sobre la cognición y sus diferentes subsistemas pueden inferirse las siguientes características y requisitos generales, que determinarán las decisiones de diseño del modelo:

PF-1	DESCRIPCIÓN	LA COGNICIÓN SE BASA EN LA INTERACTUACIÓN ENTRE DISTINTOS SISTEMAS COMPLEJOS: EL RAZONAMIENTO, LA INTELIGENCIA (FUERA DEL OBJETIVO DE LA PRESENTE TESIS, PERO A TENER EN CUENTA EN LÍNEAS FUTURAS Y EXTENSIÓN DEL TRABAJO), LA PERCEPCIÓN, EL CONOCIMIENTO Y LA MEMORIA.
	CATEGORÍA	Cognición
	COMENTARIO	Lo cual recuerda la cooperación entre múltiples sistemas, presentes en arquitecturas distribuidas, para resolver problemas y proporcionar servicios.

Tabla 3. PF-1

PF-2	DESCRIPCIÓN	LA PERCEPCIÓN PERMITE RECONOCER ELEMENTOS DEL ENTORNO EN BASE A DISTINTOS SENTIDOS Y A LA EXPERIENCIA PREVIA.
	CATEGORÍA	Percepción
	COMENTARIO	Característica útil que relaciona la memoria, el aprendizaje y el tipo de información de entrada en el sistema. Información de tipo perceptual, percibida por los distintos sentidos.

Tabla 4. PF-2

PF-3	DESCRIPCIÓN	LOS PRINCIPALES SENTIDOS EN EL SER HUMANO SON: LA VISTA, EL OÍDO, EL OLFATO, EL GUSTO, EL TACTO, LA PROPIOCEPCIÓN, LA EQUILIBRIOCEPCIÓN, LA CRONOCEPCIÓN, LA TERMOCEPCIÓN Y LA NOCIOCEPCIÓN.
	CATEGORÍA	Percepción
	COMENTARIO	Desde un punto de vista computacional, dichos sentidos pueden recrearse mediante el uso y combinación de diferentes sensores. Véase: cámaras, micrófonos de diversos tipos, sensores de partículas gaseosas en suspensión, sensores de propiedades químicas, acelerómetros, giróscopos, sensores de temperatura, y en el caso de la nocioccepción, puede tenerse en cuenta el estado (buen funcionamiento) de los sensores involucrados.

Tabla 5. PF-3

PF-4	DESCRIPCIÓN	LOS ANIMALES AMPLÍAN EL ABANICO DE LOS SENTIDOS, INCLUYENDO ENTRE OTROS: LA ECOLOCALIZACIÓN, LA MAGNETOCEPCIÓN, Y LA ELECTROCEPCIÓN.
	CATEGORÍA	Percepción
	COMENTARIO	Dichos sentidos pueden ser cubiertos por otra gama de sensores específicos, como los sónares y los sensores de campo eléctrico y magnético. La importancia de esta característica es que refleja que el abanico de sentidos no tiene por qué ser limitado y cerrado al conjunto de sentidos provisto del ser humano. Si bien, puede ser extendido e incluso mejorado (mayor nivel de sensibilidad, detección extendida de rangos perceptuales, etc.).

Tabla 6. PF-4

PF-5	DESCRIPCIÓN	ALGUNAS DE LAS FUNCIONALIDADES HABILITADAS POR LA SINERGIA PERCEPCIÓN-CONOCIMIENTO SON: EL RECONOCIMIENTO, LA INTERPRETACIÓN, LA DISCRIMINACIÓN PERCEPTUAL ENTRE MIEMBROS DE UNA MISMA CATEGORÍA Y EL ACCESO A SOLUCIONES PREVIAMENTE APRENDIDAS.
	CATEGORÍA	Cognición
	COMENTARIO	De donde se deduce la importancia de realizar la evaluación de las entradas contra una base de conocimiento aprendido previamente.

Tabla 7. PF-5

PF-6	DESCRIPCIÓN	CARACTERÍSTICAS PRINCIPALES DE LA PERCEPCIÓN SON: LA CONSTANCIA (RECONOCIMIENTO DE UN ELEMENTO DESDE DISTINTAS ENTRADAS SENSORIALES), EL AGRUPAMIENTO (FORMA EN LA QUE SON PERCIBIDOS LOS PATRONES ASOCIADOS A LOS ELEMENTOS) Y LOS EFECTOS DE CONTRASTE (CUALIDADES PERCIBIDAS, AFECTADAS POR EL CONTEXTO).
	CATEGORÍA	Continuidad
	COMENTARIO	De la constancia puede inferirse que reconocer múltiples características de un elemento desde diferentes sentidos perceptuales ayuda notablemente a tener una representación más completa y útil del mismo. El agrupamiento proporciona una serie de pistas heurísticas sobre cómo abordar los problemas de reconocimiento de patrones desde una perspectiva biológica. Finalmente, los efectos de contraste determinan cómo puede variar la percepción en base a las condiciones del entorno y de otros estímulos externos que pueden afectar al reconocimiento. Conocer dichos factores y sus implicaciones, puede ayudar a corregir o mitigar problemas derivados de tales efectos sobre la percepción.

Tabla 8. PF-6

PF-7	DESCRIPCIÓN	EN LA PERCEPCIÓN Y EL RECONOCIMIENTO ENTRAN EN JUEGO DOS TIPOS DE PROCESOS QUE ACTÚAN EN PARALELO, Y ES EN LA CONVERGENCIA ENTRE AMBOS DONDE RADICA EL RECONOCIMIENTO. EL PRIMER TIPO DE PROCESOS ES EL DENOMINADO “ <i>BOTTOM-UP</i> ” (JERARQUÍA DE ANÁLISIS SOBRE ELEMENTOS PERCIBIDOS SENSORIALMENTE). EL SEGUNDO TIPO ES EL DENOMINADO “ <i>TOP-DOWN</i> ” (COMPARAN EL ANÁLISIS DE LOS PROCESOS <i>BOTTOM-UP</i> CON EL CONOCIMIENTO MEMORIZADO).
	CATEGORÍA	Percepción
	COMENTARIO	El complejo proceso de reconocimiento es a su vez producto de diferentes y complejos sub-procesos. Recuerda, desde una perspectiva relativamente simplista, al enfoque algorítmico “divide y vencerás”, donde para resolver problemas complejos, dichos problemas se sub-dividen en tareas más sencillas que son resueltas independientemente (de manera secuencial o paralela) y luego se combina el resultado de las mismas, obteniendo la solución general o final del problema.

Tabla 9. PF-7

PF-8	DESCRIPCIÓN	LOS PROCESOS “ <i>TOP-DOWN</i> ” PUEDEN COMPLETAR INFORMACIÓN AUSENTE UTILIZANDO LA EXPERIENCIA BASADA EN EL CONTEXTO.
	CATEGORÍA	Conocimiento
	COMENTARIO	Gracias a un histórico de posibles interacciones con y entre los elementos del mismo contexto es posible inferir parte de la información ausente a fin de dotar de coherencia al conjunto de información percibida.

Tabla 10. PF-8

PF-9	DESCRIPCIÓN	EXISTEN PROCESOS HEURÍSTICOS RELACIONADOS CON LOS PROCESOS "TOP-DOWN", QUE ACELERAN NOTABLEMENTE EL RECONOCIMIENTO, PERO PUEDEN LLEVAR A CABO A RECONSTRUCCIÓN ERRÓNEA DE HECHOS ADQUIRIDOS PARCIALMENTE. ESTAS RECONSTRUCCIONES SUELEN CORRESPONDERSE CON LAS DENOMINADAS ILUSIONES PERCEPTUALES (P. EJ. ILUSIONES ÓPTICAS, O AUDITIVAS).
	CATEGORÍA	Percepción
	COMENTARIO	Como se ha observado, la mayor parte de las ilusiones ópticas están ligadas a las limitaciones físicas de los sistemas perceptuales. Ya sea por limitaciones en cuanto a precisión o en cuanto a volumen de procesamiento. También se ha observado que dichas ilusiones pueden ser resultado de interacciones incorrectas entre distintos sentidos que cooperan para dotar de coherencia al contexto. La incapacidad biológica del ser humano para "escalar" sus capacidades cognitivas bajo demanda, y de controlar consciente e independientemente sus sistemas perceptuales, hace que tengan que utilizarse estrategias algorítmicas a nivel biológico. Sin embargo, en un sistema computacional, podrían desarrollarse arquitecturas y sistemas de tal manera que no fuesen afectados de la misma manera ante ilusiones ópticas como el ser humano (p. ej. incluyendo sensores adicionales, más precisión en las mediciones, mayor capacidad de cálculo y procesamiento, una red de comunicación más veloz, etcétera).

Tabla 11. PF-9

PF-10	DESCRIPCIÓN	EL CONOCIMIENTO PREVIO DEL ENTORNO (CONTEXTO) ES CONSIDERADO UNA DE LAS BASES DEL RECONOCIMIENTO. EL CONTEXTO Y SU IDENTIFICACIÓN PREVIA TIENE UN IMPACTO MUY POSITIVO Y ACELERADOR EN LAS TAREAS DE RECONOCIMIENTO. LOS ELEMENTOS SITUADOS EN CONTEXTOS FRECUENTES SON MÁS FÁCILMENTE RECONOCIBLES.
	CATEGORÍA	Memoria
	COMENTARIO	El concepto de contexto recuerda al concepto de cachés y jerarquías de memoria, utilizadas en los entornos computacionales para acelerar los procesos de acceso a datos.

Tabla 12. PF-10

PF-11	DESCRIPCIÓN	LA CAPACIDAD DE RECONOCIMIENTO NO ESTÁ LIGADA A UN ÓRGANO SENSORIAL ESPECÍFICO. SI BIEN, DEPENDE FUERTEMENTE DE LA COMBINACIÓN DE LOS SENTIDOS DISPONIBLES.
	CATEGORÍA	Percepción
	COMENTARIO	Característica importante que refleja la importancia de combinar información de diferentes naturalezas a la hora de identificar elementos del entorno. Parece que un mayor nivel de representatividad de la información adquirida, refuerza y extiende el conocimiento sobre los elementos involucrados, y afecta a la precisión en la recuperación, tanto para el reconocimiento como para determinar posibles inferencias con y entre dichos elementos.

Tabla 13. PF-11

PF-12	DESCRIPCIÓN	LA LIMITACIÓN EN LA CAPACIDAD NEURAL TIENE UNA SERIE DE EFECTOS ADVERSOS SOBRE EL SISTEMA COGNITIVO. POR EJEMPLO, SE PUEDEN CITAR LOS EFECTOS “ES DEMASIADO” Y “NO ES SUFICIENTE”, RELACIONADOS CON LA SATURACIÓN O ESCASEZ DE INFORMACIÓN PERCEPTUAL.
	CATEGORÍA	Percepción
	COMENTARIO	Como se comentó previamente en PF-9, este es otro problema derivado de la incapacidad biológica del ser humano para “escalar” sus capacidades cognitivas bajo demanda. Al igual que en el caso anterior, en un sistema computacional, es posible mitigar dichos efectos gracias a la inclusión de recursos a medida que sean necesarios, previniendo la saturación, o ampliando la capacidad de reconocimiento de información.

Tabla 14. PF-12

PF-13	DESCRIPCIÓN	ANTE LA LIMITACIÓN EN LA CAPACIDAD NEURAL, SE HA OBSERVADO UNA DISTRIBUCIÓN BALANCEADA DE LOS RECURSOS ENTRE LOS DISTINTOS SISTEMAS PERCEPTUALES Y SUS ÁREAS DE RECONOCIMIENTO ASOCIADAS. SI EL BALANCEO NO ES ADECUADO, SE PUEDEN PRODUCIR PROBLEMAS PERCEPTUALES (COMO POR EJEMPLO LA COMPENSACIÓN DE INTENSIDAD PERCIBIDA POR LOS OJOS).
	CATEGORÍAS	Percepción, Conocimiento
	COMENTARIO	Lo cual recuerda a los sistemas de balanceo dinámico de carga entre servidores y sistemas a fin de distribuir de manera homogénea el trabajo, a fin de prevenir la saturación y cubrir la demanda.

Tabla 15. PF-13

PF-14	DESCRIPCIÓN	SE HA OBSERVADO LA COMPETICIÓN POR LOS RECURSOS ENTRE DISTINTOS SISTEMAS PERCEPTUALES (P. EJ. ENTRE LOS DEDOS DE LA MANO, O ENTRE LA VISTA Y EL OÍDO, POR CITAR ALGUNOS).
	CATEGORÍA	Percepción
	COMENTARIO	Todas las tareas relacionadas con la percepción, en el ser humano se llevan a cabo en un sistema cerrado. Esto es, no extensible o aumentable por medios naturales. Dada la naturaleza biológica, de carácter optimizador basado en la evolución, las estrategias para utilizar los recursos “computacionales” existentes han derivado en sistemas interconectados y balanceados que aprovechan tanto como les es posible los recursos disponibles en función de la demanda o carga de trabajo a la que están sometidos. Por eso, en determinadas ocasiones, a fin de adquirir mayor información sensorial, se puede recurrir a la “desactivación” temporal y consciente de otros sentidos que requieren mayor capacidad de procesamiento (p. ej. cerrar los ojos al escuchar una canción en un dictado musical).

Tabla 16. PF-14

PF-15	DESCRIPCIÓN	EXISTE UN FENÓMENO CONOCIDO COMO “ADAPTACIÓN” O “FATIGA”, ENCARGADO DE EVITAR LA SATURACIÓN DE ESTÍMULOS, TANTO AL SISTEMA SENSORIAL COMO AL CEREBRO. LA ADAPTACIÓN TIENE UNA VELOCIDAD VARIABLE, ACORDE AL SISTEMA RECEPTOR IMPLICADO.
	CATEGORÍA	Percepción
	COMENTARIO	La adaptación es una estrategia para minimizar uso de recursos perceptuales dependiente del tiempo. Ante la ausencia en la variabilidad (o cambios) de la entrada perceptual, se tiende a filtrar dicha información invariante de entrada para dotar de más recursos a otros sentidos que los puedan necesitar.

Tabla 17. PF-15

PF-16	DESCRIPCIÓN	EL CORRECTO PROCESAMIENTO SENSORIAL IMPLICA LA REALIMENTACIÓN ENTRE DISTINTAS ÁREAS DEL CEREBRO, ASÍ COMO CONEXIONES BIDIRECCIONALES ENTRE DISTINTAS ÁREAS.
	CATEGORÍA	Percepción
	COMENTARIO	De aquí se deduce que una buena estrategia para replicar el comportamiento sensorial debería estar basada en un conjunto de diferentes sistemas independientes, pero interconectados.

Tabla 18. PF-16

PF-17	DESCRIPCIÓN	LA VISIÓN EN LOS BEBÉS, DURANTE SU PRIMER AÑO DE VIDA, CONSISTE EN UN APRENDIZAJE INTENSIVO BASADO EN UNAS 10 MILLONES DE MIRADAS, QUE CONFORMAN UN APRENDIZAJE BÁSICO DE ELEMENTOS, COMO CARAS Y OBJETOS, DESDE DIFERENTES PERSPECTIVAS, EN DIFERENTES CONTEXTOS Y CON DISTINTAS CONDICIONES DE ILUMINACIÓN. DICHA INFORMACIÓN APRENDIDA, COMBINADA Y REFORZADA, FACILITA EL RECONOCIMIENTO POSTERIOR.
	CATEGORÍAS	Conocimiento, Memoria
	COMENTARIO	Esta característica demuestra la importancia del proceso de entrenamiento y aprendizaje a la hora de llevar a cabo tareas de reconocimiento.

Tabla 19. PF-17

PF-18	DESCRIPCIÓN	SE CONSIDERA QUE LA VISIÓN UTILIZA DIVERSOS MODELOS DE REPRESENTACIÓN PARA EVITAR PROBLEMAS COMO LA DEPENDENCIA DEL PUNTO DE VISTA O LA VARIACIÓN DEL ELEMENTO. ALGUNOS DE LOS MODELOS IDENTIFICADOS MÁS IMPORTANTES SON: COINCIDENCIA DE CARACTERÍSTICAS, COINCIDENCIA CON PLANTILLAS, MODELOS BASADOS EN COMPONENTES Y MODELOS DE CONFIGURACIÓN.
	CATEGORÍA	Conocimiento
	COMENTARIO	La representación de conceptos adquiridos a través de la visión requiere de la combinación de varios modelos diferentes, con diferentes características.

Tabla 20. PF-18

PF-19	DESCRIPCIÓN	LOS MODELOS DE VISIÓN BASADOS EN COINCIDENCIA DE CARACTERÍSTICAS PRESENTAN LAS SIGUIENTES CUALIDADES PRINCIPALES: REQUIEREN POCO ESPACIO DE ALMACENAMIENTO AL REDUCIR LA REDUNDANCIA DE CARACTERÍSTICAS ENTRE ELEMENTOS Y SUB-ELEMENTOS, FUNCIONAN MUY BIEN SI LAS CARACTERÍSTICAS BÁSICAS SON IDENTIFICABLES, SON PARALELIZABLES Y SON DEPENDIENTES DEL NIVEL DE DETALLE.
	CATEGORÍAS	Conocimiento, Memoria
	COMENTARIO	El reconocimiento visual puede ser realizado mediante procesos de descomposición en características fundamentales. Dicha descomposición reduce la redundancia de almacenamiento entre miembros de la categoría y sub-categorías. Se intuye que es posible paralelizar el proceso de descomposición y reconocimiento individual de características siempre y cuando se combinen los resultados de los distintos procesos.

Tabla 21. PF-19

PF-20	DESCRIPCIÓN	LOS MODELOS DE VISIÓN BASADOS EN COINCIDENCIA CON PLANTILLAS PRESENTAN LAS SIGUIENTES CUALIDADES PRINCIPALES: SE BASAN EN LA COMPARACIÓN CON PROTOTIPOS PREVIAMENTE DEFINIDOS O GENERADOS, SUELEN SER RÁPIDOS Y PRECISOS Y RESULTAN POCO FLEXIBLES. SIN EMBARGO, A FIN DE PALIAR SU ESCASA FLEXIBILIDAD, SE HAN ESTUDIADO VARIANTES CONSISTENTES EN EL ESTABLECIMIENTO DE JERARQUÍAS DE COMPONENTES BÁSICOS. DICHS MODELOS SON APLICADOS, POR EJEMPLO, EN LAS ÚLTIMAS TÉCNICAS BIOMÉTRICAS.
	CATEGORÍAS	Conocimiento, Memoria
	COMENTARIO	La coincidencia con plantillas parece seguir un comportamiento similar al de comparar la distancia con un prototipo. Si bien, la ausencia de flexibilidad hace más atractivo el uso de jerarquías de componentes, obteniendo un modelo que recuerda a la combinación de coincidencia de características, donde las características son las correspondientes a los componentes básicos de los que se han generado a su vez prototipos.

Tabla 22. PF-20

PF-21	DESCRIPCIÓN	LOS MODELOS DE VISIÓN BASADOS EN COMPONENTES PRESENTAN LAS SIGUIENTES CUALIDADES PRINCIPALES: SE BASAN EN LA DESCOMPOSICIÓN DE LOS ELEMENTOS EN PARTES FUNCIONALES Y RELACIONES ESPACIALES (GEONES Y RELACIONES GEOMÉTRICAS DE PERTINENCIA O ADYACENCIA), RESULTAN ÚTILES PARA EL RECONOCIMIENTO DE ELEMENTOS EN 3D, SOBRE TODO, CUANDO EL RECONOCIMIENTO ES DE OBJETOS MANUFACTURADOS (NO NATURALES), O CUANDO OCURRE <i>PRIMING</i> VISUAL (ELEMENTOS VISTOS POR SEGUNDA VEZ, O SUCESIVAS).
	CATEGORÍAS	Conocimiento, Memoria
	COMENTARIO	Una conclusión muy importante derivada de los modelos basados en componentes es que resulta fundamental conocer no sólo los componentes que conforman un objeto identificado, sino las relaciones (geométricas) entre dichos componentes. También se puede deducir que una “caché” visual puede ayudar a facilitar el reconocimiento de elementos identificados anteriormente.

Tabla 23. PF-21

PF-22	DESCRIPCIÓN	LOS MODELOS DE VISIÓN BASADOS EN MODELOS DE CONFIGURACIÓN PRESENTAN LAS SIGUIENTES CUALIDADES PRINCIPALES: MEJORAN EL MODELO DE COMPONENTES AL ESTAR BASADOS EN LA COMPOSICIÓN DE SUB-OBJETOS, JUNTO CON UNA ESTRUCTURA EN LA QUE SE ALMACENAN LAS RELACIONES ENTRE DICHOS SUB-OBJETOS Y LA DESVIACIÓN DE UN PROTOTIPO ESTABLECIDO O GENERADO. SE APLICAN DE MANERA ARTIFICIAL EN LOS PROCESOS DE RECONOCIMIENTO DE CARAS. SE HA OBSERVADO QUE ES UN MODELO COMÚN ENTRE LOS SERES HUMANOS Y LOS PRIMATES.
	CATEGORÍAS	Conocimiento, Memoria
	COMENTARIO	Extendiendo a los modelos basados en componentes, y la conclusión derivada de los mismos, no sólo es importante almacenar las relaciones entre componentes, sino que es conveniente emplear prototipos (combinación con modelos basados en coincidencia de plantillas), y las características de las mismas. Así pues, una combinación con las características de los cuatro modelos parece ser suficiente para llevar a cabo un reconocimiento visual más o menos completo, similar al llevado a cabo por seres humanos y primates.

Tabla 24. PF-22

PF-23	DESCRIPCIÓN	EL OÍDO ES MUY SENSIBLE Y PRESENTA UN ELEVADO RANGO DINÁMICO. TRANSFORMA LAS ONDAS SONORAS (DE PRESIÓN) EN SEÑALES ELÉCTRICAS.
	CATEGORÍA	Percepción
	COMENTARIO	Al igual que la visión, el oído es un receptor remoto muy sensible. El procesamiento de la información, al igual que en el resto de sentidos, se basa en el procesamiento de señales, como las que podrían recogerse en sistemas computacionales a través de sensores.

Tabla 25. PF-23

PF-24	DESCRIPCIÓN	EL OÍDO NO SÓLO SE ENCARGA DE LA AUDICIÓN, SINO QUE SOPORTA DE MANERA FUNDAMENTAL AL SENTIDO DEL EQUILIBRIO.
	CATEGORÍA	Percepción
	COMENTARIO	El oído actúa como una combinación de sensores que dan soporte a distintos sistemas perceptuales.

Tabla 26. PF-24

PF-25	DESCRIPCIÓN	EL OÍDO SE COMPORTA COMO UN ANALIZADOR DE FRECUENCIAS. LA SENSACIÓN AUDITIVA SE BASA EN EL RECONOCIMIENTO DEL TONO, LA LOCALIZACIÓN SUBJETIVA, LA INTENSIDAD (VOLUMEN) Y EL TIMBRE.
	CATEGORÍA	Percepción
	COMENTARIO	La sensación auditiva puede clasificarse utilizando una serie de características fundamentales. Aunque muchas de ellas son unidimensionales, y relativamente fáciles de medir (p. ej. el tono en Hz, o la intensidad en dB), otras, como el timbre, pueden resultar muy complejas, y requerir diversos sistemas computacionales especializados, para su análisis.

Tabla 27. PF-25

PF-26	DESCRIPCIÓN	EL TIMBRE ES LA ÚNICA CARACTERÍSTICA MULTIDIMENSIONAL RECONOCIDA, PUES NO EXISTE UNA ESCALA ÚNICA PARA DETERMINARLO, Y SE CONSIDERA CRUCIAL PARA LA IDENTIFICACIÓN DE DISTINTOS ELEMENTOS. EL TIMBRE ES LA CARACTERÍSTICA QUE PERMITE, POR EJEMPLO, DETERMINAR SI UNA MISMA NOTA HA SIDO EMITIDA POR UN VIOLÍN O POR UN PIANO.
	CATEGORÍA	Memoria (*reconocimiento)
	COMENTARIO	Para la determinación del timbre es necesario realizar un complejo análisis multidimensional a fin de determinar los distintos sonidos parciales (o fundamentales) superpuestos que componen el sonido percibido.

Tabla 28. PF-26

PF-27	DESCRIPCIÓN	LAS FLUCTUACIONES DE DISTRIBUCIÓN DE ENERGÍA SOBRE LA FRECUENCIA A LO LARGO DEL TIEMPO (COMPORTAMIENTO TEMPORAL DEL TIMBRE) SON FUNDAMENTALES PARA LA PERCEPCIÓN AUDITIVA.
	CATEGORÍA	Percepción
	COMENTARIO	La percepción auditiva está fuertemente ligada a una continuidad temporal. Se deduce que, para un reconocimiento completo auditivo, es necesario disponer de las características correspondientes al sonido analizado, junto con su comportamiento temporal.

Tabla 29. PF-27

PF-28	DESCRIPCIÓN	EL RECONOCIMIENTO DEL TIMBRE SE BASA EN SU PERIODICIDAD, EN LA FORMA DE LA ONDA (CONSTANTE O FLUCTUANTE, Y EN DICHO SUPUESTO, CÓMO FLUCTÚA), EN LA FORMA DEL ESPECTRO Y SU VARIACIÓN Y EN LOS SONIDOS ANTECEDENTES Y PRECEDENTES.
	CATEGORÍAS	Percepción, Memoria
	COMENTARIO	El timbre presenta patrones, así como se pueden encontrar patrones en las imágenes, o en texturas reconocidas mediante el tacto.

Tabla 30. PF-28

PF-29	DESCRIPCIÓN	LA IDENTIFICACIÓN ENTRE DIFERENTES ELEMENTOS SE CONSIGUE A TRAVÉS DE UNA SERIE DE FUNCIONES INDEPENDIENTES PERO INTERRELACIONADAS: LA AGRUPACIÓN PERCEPTUAL, EL <i>PARSING</i> , EL ANÁLISIS AUDITIVO DE LA ESCENA (ASIGNACIÓN DE SONIDOS A FUENTES) Y LA SEGREGACIÓN AUDITIVA (SEPARACIÓN DE LOS SONIDOS ENTRE VARIAS FUENTES).
	CATEGORÍA	Conocimiento
	COMENTARIO	El oído, al igual que otros sentidos, se basa en la interacción de distintos módulos específicos interrelacionados, y especializados en el reconocimiento de características concretas.

Tabla 31. PF-29

PF-30	DESCRIPCIÓN	LA DETERMINACIÓN DE LA FRECUENCIA FUNDAMENTAL (ARMONICIDAD) PERMITE LA DISCRIMINACIÓN DE TONOS PERCIBIDOS A LA VEZ (SI SUS FRECUENCIAS FUNDAMENTALES SON DIFERENTES). ESTA CARACTERÍSTICA MEJORA EL RECONOCIMIENTO, POR EJEMPLO, DE LAS VOCALES.
	CATEGORÍAS	Percepción, Memoria
	COMENTARIO	Esta característica tiene que ver con PF-26. El análisis y separación de frecuencias e intensidades que componen un sonido es fundamental para el reconocimiento auditivo y la discriminación de elementos.

Tabla 32. PF-30

PF-31	DESCRIPCIÓN	AL IGUAL QUE EN OTROS SENTIDOS, COMO ES EL CASO DE LA VISTA O EL TACTO, LOS CAMBIOS DE INFORMACIÓN CON EL TIEMPO, DESTACAN SOBRE EL RESTO.
	CATEGORÍA	Percepción
	COMENTARIO	La continuidad temporal, y el almacenamiento de los sucesos teniendo en cuenta el momento y duración de los mismos influyen notablemente en la percepción (siendo fundamental para algunos sentidos, como el oído).

Tabla 33. PF-31

PF-32	DESCRIPCIÓN	TRAS UN PROCESO DE ADAPTACIÓN, SI SE PRODUCE UNA VARIACIÓN EN LA INFORMACIÓN PERCIBIDA, REPERCUTE EN UN INCREMENTO DE LA DETECCIÓN (P. EJ. ADAPTACIÓN A ZUMBIDO DE CONDENSADORES Y SENSACIÓN CUANDO SE APAGA EL DISPOSITIVO QUE LOS CONTIENE).
	CATEGORÍA	Percepción
	COMENTARIO	Como ya se apreció en PF-15, la adaptación, o fatiga, es un proceso subconsciente responsable del ahorro de recursos perceptuales. Ante el cambio o variabilidad de los estímulos de entrada, la atención reactiva el proceso de reconocimiento a fin de percibir e identificar cómo y de qué manera se ha producido el cambio.

Tabla 34. PF-32

PF-33	DESCRIPCIÓN	GRACIAS AL RETARDO ENTRE LA SEÑAL PERCIBIDA ENTRE LOS OÍDOS (EN TORNO AL MILLISEGUNDO) SE CONSIGUE EL POSICIONAMIENTO PARCIAL DE LA FUENTE.
	CATEGORÍAS	Conocimiento, Percepción
	COMENTARIO	La medición de la misma fuente de sonido desde dos puntos bien conocidos y calibrados previamente, permite la localización subjetiva (respecto a los receptores) de la fuente de sonido, basándose en el retardo percibido, el conocimiento de la velocidad de propagación del sonido en el medio y la posición de cada receptor. Así pues, cuantos más sensores localizados en posiciones conocidas capten la misma información, más precisión se obtendrá a la hora de determinar la posición de la fuente.

Tabla 35. PF-33

PF-34	DESCRIPCIÓN	EL OLFATO ES UN SENTIDO TELERRECEPTOR QUÍMICO. LAS MOLÉCULAS DE OLOR SON DISUELTAS EN LAS MUCOSAS Y DISCRETIZADAS EN EL EPITELIO OLFATORIO.
	CATEGORÍA	Percepción
	COMENTARIO	Para el reconocimiento de los elementos del entorno no sólo se utilizan telerreceptores, basados en la recepción de energía en distintas formas (luz, sonido). También se utilizan otros receptores, que pueden requerir el contacto directo con el elemento a percibir, y de diversas naturalezas, como, por ejemplo, en el caso del olfato, de tipo químico.

Tabla 36. PF-34

PF-35	DESCRIPCIÓN	EL OLFATO ES DE LOS POCOS SENTIDOS CON UN RECONOCIMIENTO INMEDIATO, EN PARTE GRACIAS A LA COMUNICACIÓN DIRECTA DE LOS SISTEMAS NEURALES CON EL CEREBRO.
	CATEGORÍA	Memoria
	COMENTARIO	La variabilidad en la complejidad de sistemas perceptuales, tanto a nivel estructural como funcional, hace que la velocidad de reconocimiento sea distinta.

Tabla 37. PF-35

PF-36	DESCRIPCIÓN	AL IGUAL QUE EN OTROS SENTIDOS, EN EL OLFATO EXISTE UNA ADAPTACIÓN SENSORIAL EN CASO DE EXPOSICIÓN A UN OLOR CONSTANTE.
	CATEGORÍA	Percepción
	COMENTARIO	En relación a PF-32 y PF-15, a fin de saturar el sistema perceptual con información no relevante una vez reconocida, la adaptación cubre un importante papel economizando recursos de manera dinámica.

Tabla 38. PF-36

PF-37	DESCRIPCIÓN	EL SISTEMA OLFATIVO HUMANO ES CAPAZ DE IDENTIFICAR ENTRE MIL Y CUATRO MIL OLORES E INTENSIDADES DISTINTOS. SI BIEN, ES CIERTO QUE, EN ALGUNOS ANIMALES, DICHA CAPACIDAD ESTÁ MUY AUMENTADA. DENTRO DE LOS TIPOS DE OLOR, SE HAN ESTABLECIDO SEIS CATEGORÍAS PRINCIPALES: FLORAL, AFRUTADO, PICANTE, RESÍNICO, QUEMADO Y PÚTRIDO.
	CATEGORÍA	Conocimiento
	COMENTARIO	Al igual que con el resto de sentidos, la capacidad de percepción está determinada con la precisión y rango de detección del sistema perceptual. En un sistema computacional dotado de sensores químicos capaces de discretizar las características odorantes de las moléculas, la precisión sería correspondiente a la de sus sensores. En caso de ser superiores a las de seres humanos y/o animales, la información olfativa podría dar luz a un sinnúmero de nuevas aplicaciones (p. ej. detección de cáncer en etapas tempranas [475], [476], o de hipoglucemia en pacientes de diabetes tipo 1 [477] de manera análoga a la forma en que pueden hacerlo algunos perros).

Tabla 39. PF-37

PF-38	DESCRIPCIÓN	EL GUSTO ES UN SENTIDO BASADO EN LA DETECCIÓN QUÍMICA, Y DEPENDIENTE DE OTROS SENTIDOS, COMO: EL TACTO, LA TERMOCEPCIÓN Y LA NOCIOCEPCIÓN DE LA MUCOSA ORAL. TAMBIÉN ESTÁ FUERTEMENTE INFLUENCIADO POR LA CANTIDAD DE LA SALIVA CONTENIDA EN LA BOCA, LA CUAL AYUDA A DILUIR LAS MOLÉCULAS RESPONSABLES DEL SABOR, PARA SU PROCESAMIENTO POR LOS RECEPTORES GUSTATIVOS (A MAYOR SEQUEDAD BUCAL, MENOR PERCEPCIÓN).
	CATEGORÍA	Percepción
	COMENTARIO	Como en el caso del olfato, el gusto se trata de un receptor químico de contacto directo. Además, también entra en juego la disolución de los compuestos a analizar en saliva (en lugar de mucosa). Finalmente, cabe destacar la importancia de combinar otros receptores perceptuales a la hora de determinar el gusto (tacto, termocepción y nociocepción). Parece que se refuerza la hipótesis correspondiente a la importancia de combinar estímulos de diferentes fuentes y naturalezas a la hora de percibir y almacenar la información relativa a los elementos del entorno.

Tabla 40. PF-38

PF-39	DESCRIPCIÓN	EN EL GUSTO SE HAN IDENTIFICADO CINCO TIPOS DE RECEPTORES DIFERENTES: DULCE, ÁCIDO, SALADO, AMARGO Y UMAMI.
	CATEGORÍA	Percepción
	COMENTARIO	De manera análoga al olfato, se puede determinar el sabor en base a la intensidad relativa de los estímulos percibidos, respecto a una clasificación inicial de diferentes matices (dulce, ácido...).

Tabla 41. PF-39

PF-40	DESCRIPCIÓN	CADA RECEPTOR DEL GUSTO RESPONDE ANTE LOS ESTÍMULOS CORRESPONDIENTES A LOS DEMÁS TIPOS, PERO DE MANERA DIFERENTE (MEJOR O PEOR A ALGUNOS).
	CATEGORÍA	Percepción
	COMENTARIO	Parece que el sistema gustativo está compuesto de sub-sistemas expertos en la identificación de tipos concretos de matices, pero capaces de determinar parcialmente las características del resto.

Tabla 42. PF-40

PF-41	DESCRIPCIÓN	LA PERCEPCIÓN DEL GUSTO ESTÁ ASOCIADA A UN PATRÓN DE ACTIVACIÓN COMPLEJO EN LOS RECEPTORES GUSTATIVOS AFECTADOS.
	CATEGORÍA	Percepción
	COMENTARIO	Para determinar el gusto, se requiere conocer el conjunto o patrón correspondiente a los estímulos percibidos por los distintos elementos de la matriz sensorial. El gusto, por tanto, no se corresponde a una activación puntual por un detector, sino al conjunto de activaciones en una superficie.

Tabla 43. PF-41

PF-42	DESCRIPCIÓN	LOS SISTEMAS SENSITIVOS SOMÁTICOS, PESE A DETECTAR CARACTERÍSTICAS DIFERENTES, ESTÁN FUERTEMENTE RELACIONADOS. DICHOS SISTEMAS COMPONEN LOS SIGUIENTES SENTIDOS: TACTO, TERMOCEPCIÓN, PROPIOCEPCIÓN Y NOCIOCEPCIÓN.
	CATEGORÍA	Percepción
	COMENTARIO	Como en el caso de otros sistemas sensitivos, la combinación de las respuestas producidas por varios de ellos ayuda a la detección de características del entorno, o del propio sistema de detección, lo cual puede ayudar a comprender mejor el entorno (p. ej. las correcciones visuales aplicadas de manera automática y subconsciente en base a la inclinación de la cabeza).

Tabla 44. PF-42

PF-43	DESCRIPCIÓN	LOS SISTEMAS SENSITIVOS SOMÁTICOS SE ENCARGAN DE COORDINAR EL EQUILIBRIO Y EL MOVIMIENTO.
	CATEGORÍA	Conocimiento
	COMENTARIO	El conocimiento sobre el estado y características del propio sistema es fundamental para la correcta aplicación y funcionamiento de otros sistemas sensoriales.

Tabla 45. PF-43

PF-44	DESCRIPCIÓN	EL DESARROLLO DE LOS SENTIDOS VIENE DETERMINADO POR EL GRADO DE COMPETITIVIDAD POR LOS RECURSOS. EN CASO DE AUSENCIA DE ALGÚN SENTIDO CON MAYOR PESO, COMO, POR EJEMPLO, LA VISTA, LOS RECURSOS DISPONIBLES SE REASIGNAN, PROPORCIONANDO MÁS SENSIBILIDAD EN SENTIDOS QUE INICIALMENTE TENÍAN MENOS PESO A NIVEL PERCEPTUAL (P. EJ. LA CAPACIDAD DE LOS CIEGOS PARA LEER BRAILLE A TRAVÉS DEL TACTO).
	CATEGORÍA	Percepción
	COMENTARIO	Como ya se comentó en PF-14, las tareas relacionadas con la percepción, en el ser humano se llevan a cabo en un sistema cerrado y finito. La falta de extensibilidad natural deriva en un reparto de recursos “computacionales” en base a la importancia del sentido perceptual. En caso de fallo, los recursos que inicialmente estaban asociados a dicho sentido, pasan a distribuirse entre el resto. La ventaja de replicar el modelo de representación de conocimiento y técnicas de aprendizaje, a un entorno computacional es que muchas de esas restricciones son mitigadas o eliminadas, al tener la capacidad de incrementar bajo demanda tanto el número de sensores reconocedores, como la precisión de los mismos, y las entidades computacionales encargadas del análisis, clasificación y almacenamiento de los conceptos y entidades.

Tabla 46. PF-44

PF-45	DESCRIPCIÓN	EL TACTO PERMITE PERCIBIR: ESTIRAMIENTO MUSCULAR O DE LA PIEL, GRADOS DE PRESIÓN CONSTANTES, MOVIMIENTO DE GOLPETEO O ALETEO, VIBRACIONES Y TEXTURAS.
	CATEGORÍA	Percepción
	COMENTARIO	De la característica PF-45 se infiere que el tacto puede reconocer seis tipos diferentes de estímulo, junto con sus intensidades, cadencias y otras propiedades específicas. Dichos estímulos pueden ser reconocidos en paralelo, lo cual incita a pensar que, como en otros sentidos analizados, la acción de reconocimiento táctil puede ser derivada a un conjunto de sub-sistemas expertos en el reconocimiento de cada uno de los estímulos específicos.

Tabla 47. PF-45

PF-46	DESCRIPCIÓN	LOS RECEPTORES DEL TACTO SE ENCUENTRAN EN EL CUERPO A NIVEL SUPERFICIAL (PIEL) Y A NIVEL INTERNO (MÚSCULOS, ÓRGANOS Y ARTICULACIONES).
	CATEGORÍA	Percepción
	COMENTARIO	El tacto es uno de los sistemas perceptuales más extendidos, cubriendo el cuerpo completo, no sólo externamente, sino también internamente.

Tabla 48. PF-46

PF-47	DESCRIPCIÓN	LOS TERMOCEPTORES AFECTAN A LA PERCEPCIÓN DEL GUSTO Y DEL OLFATO.
	CATEGORÍA	Percepción
	COMENTARIO	Al igual que en muchos otros sentidos, se requiere de la interacción entre varios de ellos para alcanzar un nivel de percepción más completo sobre un mismo tipo de características.

Tabla 49. PF-47

PF-48	DESCRIPCIÓN	LOS TERMOCEPTORES EN EL SER HUMANO ACTÚAN EN UN RANGO COMPRENDIDO DENTRO DE LOS MÁRGENES DE “SEGURIDAD” DE TEMPERATURA. DICHOS RANGOS ESTÁN APROXIMADAMENTE COMPRENDIDOS, EN EL FRÍO DE 8°C A 20°C, Y EN EL CALOR DE 25°C A 40°C. FUERA DE DICHOS RANGOS, (INFERIOR Y SUPERIOR) LOS TERMOCEPTORES SE BLOQUEAN Y ENTRA EN JUEGO LA NOCIOCEPCIÓN PARA PREVENIR POSIBLE DAÑO A LOS TEJIDOS.
	CATEGORÍA	Conocimiento
	COMENTARIO	A modo de protección, existen sistemas perceptuales que alertan sobre el estado propio, y sobre posibles condiciones de peligro, a fin de proteger el sistema. La integración de dichas características en un entorno computacional puede resultar muy útil para mantener al mismo en buen estado en base a las condiciones que puede soportar.

Tabla 50. PF-48

PF-49	DESCRIPCIÓN	EL SENTIDO DEL EQUILIBRIO ESTÁ DIRECTAMENTE RELACIONADO CON EL OÍDO, LA VISTA Y LA PROPIOCEPCIÓN, SIENDO EL OÍDO EL SENTIDO QUE TIENE MÁS PESO EN EL CONJUNTO.
	CATEGORÍA	Percepción
	COMENTARIO	Como se comentó en PF-47, existen sentidos, o categorías de percepción que dependen de la integración de múltiples respuestas sensoriales parciales distintas. La combinación de dichas respuestas es lo que conforma la percepción completa sobre la categoría sensorial (en este caso, sobre el equilibrio).

Tabla 51. PF-49

PF-50	DESCRIPCIÓN	EL EQUILIBRIO SE CATEGORIZA EN DOS GRANDES GRUPOS: ESTÁTICO, CORRESPONDIENTE A LA POSICIÓN, INCLINACIÓN Y DIRECCIÓN DE LA CABEZA, Y DINÁMICO, CORRESPONDIENTE A LAS SENSACIONES ORIGINADAS POR EL MOVIMIENTO A TRAVÉS DEL ESPACIO A LO LARGO DEL TIEMPO.
	CATEGORÍAS	Percepción, Conocimiento
	COMENTARIO	La percepción del equilibrio ayuda a la interacción y al reconocimiento por parte de otros sistemas perceptuales sobre el medio. La aplicación de sensorización, análoga a la empleada por el equilibrio, en sistemas artificiales que puedan moverse, puede aportar los mismos beneficios que en el ser humano a la hora de interactuar con el medio, así como de aplicar acciones orientadas a corregir la percepción de otros sentidos (p. ej., aplicar rotaciones en imágenes antes de procesarlas, o ayudar a localizar fuentes de sonido teniendo en cuenta la posición propia del sistema reconecedor).

Tabla 52. PF-50

PF-51	DESCRIPCIÓN	EL EQUILIBRIO ESTÁTICO PERMITE LA DETECCIÓN DEL MOVIMIENTO, Y ESTÁ RELACIONADO CON LA ACTIVACIÓN DE ESTÍMULOS PARA COMPENSAR LA PÉRDIDA DE ESTABILIDAD.
	CATEGORÍA	Percepción
	COMENTARIO	En sistemas móviles, la aplicación de los sensores adecuados puede contribuir a determinar la posición relativa entre componentes, y respecto al entorno. De esta forma, ante un movimiento que cambie la posición del sistema, se pueden tomar las medidas adecuadas para evitar daños, y para continuar percibiendo el medio de manera correcta.

Tabla 53. PF-51

PF-52	DESCRIPCIÓN	EL CONOCIMIENTO ESTÁ COMPUESTO POR HECHOS, TÉCNICAS Y PROCEDIMIENTOS QUE PROPORCIONAN INFORMACIÓN SOBRE EL MUNDO QUE PUEDE SER CIERTA, ES COHERENTE Y ESTÁ JUSTIFICADO CREERLA.
	CATEGORÍA	Conocimiento
	COMENTARIO	El conocimiento se corresponde al conjunto de información previamente aprendida que da sentido a la realidad. Dicha información tiene diferentes características, y se encuentra interrelacionada a través de diferentes propiedades.

Tabla 54. PF-52

PF-53	DESCRIPCIÓN	EL CONOCIMIENTO POSIBILITA EL FUNCIONAMIENTO DE OTROS SISTEMAS, COMO LA MEMORIA, EL LENGUAJE, EL PENSAMIENTO, LA PERCEPCIÓN Y LA ATENCIÓN.
	CATEGORÍA	Cognición
	COMENTARIO	Como se ha observado anteriormente, la cognición depende del buen funcionamiento y cooperación de una serie de sistemas. El conocimiento es uno de los más importantes, al ser un habilitador de funciones tan fundamentales como la memoria (almacenamiento), el lenguaje (comunicación), el pensamiento (inferencias), la percepción (reconocimiento y aprendizaje) y la atención (discriminación de información relevante y/o significativa).

Tabla 55. PF-53

PF-54	DESCRIPCIÓN	EL CONOCIMIENTO SE MODELA MEDIANTE DIFERENTES REPRESENTACIONES QUE ACTÚAN DE MANERA CONJUNTA. POR UN LADO, ESTÁN LAS REPRESENTACIONES DE MODALIDAD ESPECÍFICA (COMO SON LAS IMÁGENES O LOS REGISTROS DE CARACTERÍSTICAS) Y POR OTRO, LAS REPRESENTACIONES AMODALES (PROPIEDADES Y RELACIONES ENTRE ENTIDADES DEL ENTORNO).
	CATEGORÍA	Conocimiento
	COMENTARIO	Al igual que ocurre con las modalidades sensitivas, el conocimiento requiere de la actuación de múltiples modelos que actúan en cooperación a fin de representar la información. Por un lado, representaciones basadas en las características y por otro, representaciones basadas en las propiedades y relaciones entre entidades.

Tabla 56. PF-54

PF-55	DESCRIPCIÓN	PARA QUE UNA REPRESENTACIÓN PASE A FORMAR PARTE DE LA MEMORIA A MODO DE RECUERDO, DEBEN DARSE DOS FACTORES, EL CRITERIO DE INTENCIONALIDAD (UNA REPRESENTACIÓN DEBE CONSTRUIRSE DELIBERADAMENTE PARA REPRESENTAR ALGO) Y EL CRITERIO DE TRANSMISIÓN DE INFORMACIÓN (COMPORTAMIENTOS Y CARACTERÍSTICAS DE ELEMENTOS PUEDEN UTILIZARSE PARA CONJETURAR COMPORTAMIENTOS Y CARACTERÍSTICAS DE OTROS EN BASE A SUS PROPIEDADES COMUNES).
	CATEGORÍA	Memoria
	COMENTARIO	Para almacenar una representación derivada del conocimiento, debe ser útil (representar algo) y la información asociada a la representación debe tener la propiedad de poder ser utilizada posteriormente para inferir comportamientos y características de elementos con propiedades comunes.

Tabla 57. PF-55

PF-56	DESCRIPCIÓN	EL CONOCIMIENTO SE BASA EN LA CATEGORIZACIÓN DE ELEMENTOS EN BASE A SU RELACIÓN CON EXPERIENCIAS ANTERIORES. DICHA CATEGORIZACIÓN ES ÚTIL A LA HORA DE REALIZAR PROCESOS DEDUCTIVOS.
	CATEGORÍAS	Conocimiento, Memoria
	COMENTARIO	La experiencia, o histórico de hechos y elementos, ayuda a la clasificación y categorización de elementos nuevos en base a su relación con los ya aprendidos. Dicha experiencia adquirida también ayuda a la hora de realizar procesos deductivos, haciendo el papel de un módulo predictor basado en las acciones ocurridas anteriormente, o en características no observadas, pero aprendidas previamente.

Tabla 58. PF-56

PF-57	DESCRIPCIÓN	LA CATEGORIZACIÓN EN EL CONOCIMIENTO SE REALIZA A TRAVÉS DE MODELOS ESTADÍSTICOS. LOS ELEMENTOS QUE COMPARTEN MUCHAS UNIDADES BÁSICAS DE INFORMACIÓN FORMAN AUTOMÁTICAMENTE UNA CATEGORÍA.
	CATEGORÍA	Memoria
	COMENTARIO	Los elementos de información pertenecientes al conocimiento, son categorizados en base al parecido estadístico con otros elementos previamente analizados. La clasificación se realiza en base a las unidades básicas de información detectadas o inferidas.

Tabla 59. PF-57

PF-58	DESCRIPCIÓN	LAS CATEGORÍAS AYUDAN AL ACCESO Y RECUPERACIÓN DE MIEMBROS SIMILARES. TAMBIÉN FACILITA LA GENERACIÓN DE PROTOTIPOS.
	CATEGORÍAS	Memoria, Conocimiento
	COMENTARIO	Las categorías recuerdan a una jerarquía en la que cada nodo presenta los atributos comunes que tendrán todos sus hijos, y cada hijo podrá contener atributos adicionales que lo diferencian del resto de hermanos, y lo hacen "extender" del padre.

Tabla 60. PF-58

PF-59	DESCRIPCIÓN	EL INCREMENTO EN LA EXPERIENCIA RECONOCIENDO ELEMENTOS DE UNA MISMA CATEGORÍA REFUERZA Y AMPLÍA EL CONOCIMIENTO DE LA MISMA.
	CATEGORÍA	Memoria
	COMENTARIO	Cuanto mayor es el entrenamiento para reconocer elementos de un mismo tipo, más preciso es el reconocimiento posterior de nuevos elementos correspondientes a dicho tipo. Este enfoque se lleva aplicando mucho tiempo en sistemas de aprendizaje automático, donde un buen entrenamiento es determinante para el correcto funcionamiento de los reconocedores y clasificadores.

Tabla 61. PF-59

PF-60	DESCRIPCIÓN	EL CONOCIMIENTO DE UNA CATEGORÍA COMIENZA DESDE MUCHOS ORÍGENES DIFERENTES (SISTEMA PERCEPTUAL VISUAL, AUDITIVO, TÁCTIL...). ESTO PERMITE REALIZAR PROCESOS DEDUCTIVOS ÚTILES Y REALIZAR FUNCIONES INTELIGENTES.
	CATEGORÍA	Percepción
	COMENTARIO	Las categorías guardan propiedades de distinta naturaleza. El conocimiento asociado, tan completo y variado, puede facilitar el desarrollo de sistemas de inferencia que aprovechen un modelo de conocimiento basado en estas características, al igual que ocurre en el ser humano de forma subconsciente.

Tabla 62. PF-60

PF-61	DESCRIPCIÓN	NO TODAS LAS CATEGORÍAS REQUIEREN O ESTÁN BASADAS EN LOS MISMOS TIPOS DE ESTÍMULOS (P. EJ. PLATO DE COMIDA VS. PIANO).
	CATEGORÍAS	Percepción, Conocimiento
	COMENTARIO	Aunque se tenga la capacidad de identificar un elemento en una categoría a través de múltiples estímulos sensitivos (olfato, gusto, tacto, vista), no implica que sea necesario el uso de todos los estímulos para reconocer dicho elemento. Si bien, en sistemas artificiales, dotados de gran capacidad de almacenamiento y potencia de cómputo, la categorización empleando tantos estímulos sensitivos (sensores) como sea posible, podría dar lugar a la identificación de nuevas inferencias o relaciones que al ser humano se le han pasado por alto, debido a las limitaciones perceptuales inherentes a su naturaleza biológica.

Tabla 63. PF-61

PF-62	DESCRIPCIÓN	UNA ETIQUETA O NOMBRE PERMITE EL ACCESO A CONTENIDO DE DICHA CATEGORÍA. LA INTEGRACIÓN EN EL CEREBRO DEL NOMBRE DE CATEGORÍA Y LA REFERENCIA AL CONTENIDO DE LA MISMA ES UN FACTOR CLAVE EN EL CONOCIMIENTO.
	CATEGORÍA	Memoria
	COMENTARIO	El funcionamiento de las etiquetas y acceso al contenido recuerda al funcionamiento de las estructuras de datos tipo <i>hash table</i> o <i>hash map</i> (matriz asociativa o tabla de dispersión), donde mediante el uso de una clave es posible acceder al valor correspondiente, con un orden de complejidad temporal promedio, constante.

Tabla 64. PF-62

PF-63	DESCRIPCIÓN	LAS ETIQUETAS TIENEN UNA IMPORTANCIA MUY GRANDE EN LAS REPRESENTACIONES AMODALES DEL CONOCIMIENTO.
	CATEGORÍA	Conocimiento
	COMENTARIO	La capacidad de aceleración de acceso a la información sobre las características correspondientes a las categorías mediante el uso de etiquetas potencia las representaciones Amodales, al desacoplar la información de la etiqueta, pero permitiendo el acceso rápido a dicha información. De esta manera, para las representaciones Amodales pueden utilizarse en conjunción: estructuras que representen las relaciones entre categorías, y estructuras que contengan las características de dichas categorías.

Tabla 65. PF-63

PF-64	DESCRIPCIÓN	SE PUEDEN CONSIDERAR CUATRO GRANDES MODELOS PARA DEFINIR LAS CATEGORÍAS: EJEMPLARES Y REGLAS, PROTOTIPOS, CONOCIMIENTO DE BASE Y ESQUEMAS Y REPRESENTACIONES DINÁMICAS.
	CATEGORÍA	Conocimiento
	COMENTARIO	Como se ha observado en otros sistemas cognitivos, la interacción entre diferentes modelos de representación permite una definición más compleja y completa de la información, que posteriormente puede ser utilizada para mejorar las características de los mismos sistemas (p. ej. en procesamiento " <i>top-down</i> "), o de otros diferentes (p. ej. deducciones).

Tabla 66. PF-64

PF-65	DESCRIPCIÓN	EN LOS MODELOS BASADOS EN EJEMPLARES Y REGLAS, LOS EJEMPLARES SE CORRESPONDEN A LOS MIEMBROS PERTENECIENTES A LAS CATEGORÍAS Y LAS REGLAS SE CORRESPONDEN AL CONJUNTO DE CRITERIOS QUE DEFINEN A LAS CATEGORÍAS.
	CATEGORÍA	Conocimiento
	COMENTARIO	Este modelo pretende asociar cada instancia, de manera individual, con las características correspondientes a las categorías que definen.

Tabla 67. PF-65

PF-66	DESCRIPCIÓN	EN LOS MODELOS BASADOS EN PROTOTIPOS SE REPRESENTAN LAS PROPIEDADES MÁS PROBABLES PERTENECIENTES A UNA CATEGORÍA DE MANERA AGRUPADA. DICHAS PROPIEDADES NO SON ABSOLUTAS, ES LO QUE SE CONOCE COMO GRADIENTES DE TIPISMO.
	CATEGORÍA	Conocimiento
	COMENTARIO	Mediante rangos de valores asociados a características individuales, y el agrupamiento de dichas características en categorías pueden definirse "prototipos difusos", más flexibles a la hora de comparar con elementos durante las labores de reconocimiento y clasificación.

Tabla 68. PF-66

PF-67	DESCRIPCIÓN	EN LOS MODELOS BASADOS EN EL CONOCIMIENTO DE BASE Y ESQUEMAS, PARA RECORDAR AL EJEMPLAR, SE ALMACENA CON LOS OBSERVADOS PREVIAMENTE (REFUERZO). LOS ESQUEMAS CONFORMAN UNA REPRESENTACIÓN ESTRUCTURADA CON INFORMACIÓN SOBRE UNA SITUACIÓN O SUCESO, O VISTO DESDE OTRA PERSPECTIVA, UN CONJUNTO COHERENTE DE RELACIONES QUE ASOCIAN PROPIEDADES.
	CATEGORÍAS	Conocimiento, Memoria
	COMENTARIO	Este tipo de modelos está más orientado a la categorización y clasificación de experiencias o hechos mediante un conjunto de relaciones entre los distintos elementos involucrados en el suceso.

Tabla 69. PF-67

PF-68	DESCRIPCIÓN	EN LOS MODELOS BASADOS EN LA REPRESENTACIÓN DINÁMICA SÓLO SE ACTIVA LA INFORMACIÓN ASOCIADA A UNA CATEGORÍA OPORTUNA ACORDE AL CONTEXTO O AL ENTORNO.
	CATEGORÍA	Memoria
	COMENTARIO	Este tipo de modelos es conservador en cuanto al uso de recursos de procesamiento, activando únicamente la información asociada a la categoría en caso de ser requerida por la situación contextual.

Tabla 70. PF-68

PF-69	DESCRIPCIÓN	LA MEMORIA ES CONSIDERADA LA FORMA FUNDAMENTAL DE COGNICIÓN QUE GUÍA LA CONDUCTA. LA INFORMACIÓN SIGUE UN FLUJO DE CODIFICACIÓN, CONSOLIDACIÓN Y RECUPERACIÓN.
	CATEGORÍA	Memoria
	COMENTARIO	El ser humano, como se ha comentado anteriormente, tiene una serie de capacidades limitadas. Entre ellas, la memoria. La capacidad no aumenta a lo largo de la vida del ser humano, mientras que la cantidad de información sí que lo hace, cada momento que pasa. A fin de tratar con semejante volumen de información, se requieren una serie de procesos mentales para codificar de forma conveniente la información a modo de recuerdos, que pueden ser almacenados de una u otra manera en base a su importancia, relevancia y significación. Dependiendo de estas características, la recuperación (total o parcial) de dicha información, podrá ser más o menos sencilla. De manera análoga se pueden considerar los grandes centros de computación, en los que se opera con un volumen determinado de información relevante (con un acceso muy rápido a la misma), y se dispone de silos de almacenamiento de información menos relevante (copias de seguridad e históricos), necesaria únicamente en determinados casos (y con tiempo de acceso notablemente más lento).

Tabla 71. PF-69

PF-70	DESCRIPCIÓN	LA CODIFICACIÓN EN LA MEMORIA ES FUERTEMENTE DEPENDIENTE DE LA ATENCIÓN Y DEL ESPACIAMIENTO TEMPORAL.
	CATEGORÍAS	Percepción, Memoria
	COMENTARIO	Como se ha comentado, a través de procesos subconscientes de optimización, a fin de no saturar la memoria disponible, gracias a la atención se discrimina el contenido relevante, y gracias al espaciamiento se refuerza su importancia.

Tabla 72. PF-70

PF-71	DESCRIPCIÓN	LA ATENCIÓN JUEGA UN PAPEL DETERMINANTE EN MUCHAS DE LAS ÁREAS COGNITIVAS. DESDE EL RECONOCIMIENTO HASTA LA CONSOLIDACIÓN DE RECUERDOS EN LA MEMORIA.
	CATEGORÍA	Conocimiento
	COMENTARIO	En sistemas limitados y no escalables, el fenómeno de la atención, o de la discriminación de contenido relevante basado en su variabilidad, parece muy importante en cuanto a la optimización de recursos físicos (almacenamiento y procesamiento). Es similar a los sistemas de vigilancia de vídeo que sólo graban (almacenamiento en disco) en caso de detectar movimiento, en lugar de grabar permanentemente. En dichos sistemas, la adaptación hace que la ausencia de movimiento establezca el sistema, dejando de grabar, y en el momento de detección de cambio, la atención se incrementa, comenzando la grabación. Este tipo de estrategia permite ahorrar una gran cantidad de espacio de almacenamiento.

Tabla 73. PF-71

PF-72	DESCRIPCIÓN	LA RECUPERACIÓN EN LA MEMORIA TIENE UNA FUERTE DEPENDENCIA DEL CONTEXTO. PARA EL ACCESO A LA MISMA SE HACE USO DE PISTAS Y CLAVES.
	CATEGORÍA	Conocimiento
	COMENTARIO	De manera análoga a las categorías en el conocimiento, el acceso a la información se realiza mediante asociaciones tipo clave-valor. Igualmente, las claves parecen estar agrupadas y relacionadas en base al contexto en el que se adquirieron / generaron, habilitando un acceso más eficiente a información sobre elementos que podrían estar relacionados o presentes en el contexto detectado.

Tabla 74. PF-72

PF-73	DESCRIPCIÓN	LA MEMORIA A LARGO PLAZO SE PUEDE DIVIDIR EN DOS GRANDES GRUPOS: MEMORIA DECLARATIVA O EXPLÍCITA Y MEMORIA EPISÓDICA.
	CATEGORÍA	Memoria
	COMENTARIO	La memoria no es un sistema que almacena un único tipo de información. Debido a la diversidad de sistemas asociados a la percepción y al conocimiento, la naturaleza de la información percibida y recordada varía en forma y propósito. Así pues, la memoria puede clasificarse en dos grandes grupos, con objetivos diferentes.

Tabla 75. PF-73

PF-74	DESCRIPCIÓN	LA MEMORIA DECLARATIVA REFIERE A LA INFORMACIÓN QUE PUEDE RECORDARSE DE FORMA CONSCIENTE Y DESCRIBIRSE A OTRAS PERSONAS (HECHOS, ACONTECIMIENTOS E IDEAS).
	CATEGORÍAS	Memoria, Conocimiento
	COMENTARIO	El primer bloque principal de la memoria está asociado al conocimiento de hechos, acontecimientos e ideas. Contiene información que puede ser transmitida y descrita.

Tabla 76. PF-74

PF-75	DESCRIPCIÓN	LA MEMORIA DECLARATIVA INCLUYE LA MEMORIA EPISÓDICA (CONOCIMIENTO CONSCIENTE DE ACONTECIMIENTOS O EPISODIOS DATADOS TEMPORALMENTE) Y LA MEMORIA SEMÁNTICA (CONOCIMIENTO GENERAL RELATIVO A LOS ELEMENTOS DE UN ENTORNO, PALABRAS Y CONCEPTOS, PROPIEDADES E INTERRELACIONES). LA DIFERENCIA FUNDAMENTAL ENTRE AMBOS TIPOS ES EL CONTEXTO O SU AUSENCIA.
	CATEGORÍAS	Memoria, Conocimiento
	COMENTARIO	La memoria declarativa, comentada en PF-74, parece ser la responsable de cubrir las necesidades principales de los sistemas de conocimiento y percepción, pues almacena información semántica relativa al aprendizaje llevado a cabo durante la existencia del ser humano. Además, se puede dividir en dos modalidades, cuya diferencia es la dependencia del contexto. Esto puede ser determinante a la hora de definir sistemas que cubran las características descritas en PF-68 y PF-72.

Tabla 77. PF-75

PF-76	DESCRIPCIÓN	LA MEMORIA DECLARATIVA ES MUY FLEXIBLE E IMPLICA ASOCIAR MÚLTIPLES ELEMENTOS DE INFORMACIÓN FORMANDO UNA REPRESENTACIÓN DE MEMORIA UNIFICADA.
	CATEGORÍA	Memoria
	COMENTARIO	La información almacenada por la memoria declarativa debe presentar ciertos aspectos homogéneos que permitan la integración de elementos básicos de información de diferentes naturalezas. Así pues, es importante buscar una representación adecuada para las partes fundamentales de conocimiento, de forma que puedan ser combinadas y utilizadas de forma unificada.

Tabla 78. PF-76

PF-77	DESCRIPCIÓN	LA MEMORIA NO DECLARATIVA, O IMPLÍCITA ESTÁ RELACIONADA CON FORMAS NO CONSCIENTES DE MEMORIA, COMO LA ADQUISICIÓN GRADUAL DE HABILIDADES MOTORAS (P. EJ. APRENDER A MONTAR EN BICI O DESARROLLO DE HÁBITOS).
	CATEGORÍAS	Memoria, Conocimiento
	COMENTARIO	La información almacenada en la memoria no declarativa parece tener relación con las acciones a llevar a cabo en base a los estímulos percibidos por el sistema sensorial, permitiendo la interacción con los elementos del entorno y el descubrimiento y afianzamiento sobre las posibles maneras de interactuar con el mismo.

Tabla 79. PF-77

PF-78	DESCRIPCIÓN	LA CONTINUIDAD EN LA COGNICIÓN PERMITE EVALUAR COHERENTEMENTE EL ENTORNO ACORDE A LOS ESTÍMULOS SENSORIALES PERCIBIDOS Y LA INFORMACIÓN PREVIA ASOCIADA A DICHOS ESTÍMULOS, ALMACENADA COMO CONOCIMIENTO EN LA MEMORIA.
	CATEGORÍA	Continuidad
	COMENTARIO	La correcta interpretación del entorno, basándose en los estímulos percibidos, tiene fuerte dependencia de la continuidad perceptiva.

Tabla 80. PF-78

4.4. Características técnicas de diseño para un modelo computacional de representación del conocimiento bio-inspirado

En base a las conclusiones observadas en el estado de la cuestión correspondiente al área de la psicología y neurociencia, los distintos sistemas de la percepción en el ser humano son sobre-estimulados continuamente. Esto ha concluido en que evolutivamente, cada uno de los sentidos filtre únicamente aquellos estímulos que resultan significativos. Dicho procesamiento se realiza tanto en el cerebro como por parte de los sensores nerviosos asociados a los receptores, descartando todo aquello que no resulta de utilidad y relacionando los distintos estímulos relevantes con experiencias pasadas relacionadas o que tengan cierto grado de parecido. Esto permite, desde evocar recuerdos hasta generar otros nuevos, dependiendo del parecido de los estímulos con experiencias previas almacenadas tanto individual como colectivamente.

Desde un punto de vista más concreto, y considerando los Principios Fundamentales identificados en el apartado anterior, se pueden identificar las siguientes características técnicas de diseño que deberá reflejar el modelo:

CT-1	DESCRIPCIÓN	LA COGNICIÓN SE FUNDAMENTA EN LA INTERRELACIÓN DE TRES MÓDULOS PRINCIPALES, INDEPENDIENTES PERO INTERCONECTADOS Y COOPERANTES. LA PERCEPCIÓN (CAPTURA DE INFORMACIÓN), EL CONOCIMIENTO (LA REPRESENTACIÓN DE DICHA INFORMACIÓN, DE MANERA CONJUNTA E INTERRELACIONADA) Y LA MEMORIA (EL ALMACENAMIENTO Y RECUPERACIÓN DE DICHA INFORMACIÓN).
	CATEGORÍAS	Procesamiento, Modularidad, Conectividad, Memoria
	PFs ASOCIADOS	PF-1, PF-5, PF-6, PF-10, PF-53, PF-78

Tabla 81. CT-1

CT-2	DESCRIPCIÓN	LA PERCEPCIÓN SE BENEFICIA DE LA EXPERIENCIA PREVIA (CONOCIMIENTO + MEMORIA). TAMBIÉN SE BENEFICIA DEL RECONOCIMIENTO DE LOS ELEMENTOS MEDIANTE EL USO DE MÚLTIPLES ENTRADAS SENSORIALES SIMULTÁNEAS.
	CATEGORÍAS	Modularidad
	PFs ASOCIADOS	PF-2, PF-3, PF-4, PF-10, PF-11, PF-29, PF-35, PF-53, PF-56, PF-60, PF-68, PF-71, PF-72, PF-78

Tabla 82. CT-2

CT-3	DESCRIPCIÓN	EL RECONOCIMIENTO SE BASA EN DOS TIPOS DE PROCESOS QUE SE EJECUTAN DE MANERA SIMULTÁNEA Y CONVERGENTE, LOS PROCESOS "BOTTOM-UP" (JERARQUÍA DE ANÁLISIS SOBRE ELEMENTOS PERCIBIDOS SENSORIALMENTE) Y LOS PROCESOS "TOP-DOWN" (UTILIZAN EL CONOCIMIENTO PREVIO PARA BUSCAR CORRELACIONES CON ESTOS ELEMENTOS PERCIBIDOS). LOS PROCESOS "TOP-DOWN", AL BASARSE EN LA EXPERIENCIA PREVIA, PUEDEN AYUDAR A COMPLETAR INFORMACIÓN AUSENTE EN BASE AL CONTEXTO, Y A ACELERAR LAS TAREAS DE RECONOCIMIENTO, PERO UN USO INCORRECTO PUEDE LLEVAR A RECONSTRUCCIONES ERRÓNEAS DE LA INFORMACIÓN (ILUSIONES ÓPTICAS).
	CATEGORÍAS	Modularidad, Procesamiento
	PFs ASOCIADOS	PF-5, PF-7, PF-8, PF-9, PF-53, PF-60, PF-71, PF-78

Tabla 83. CT-3

CT-4	DESCRIPCIÓN	EL CONOCIMIENTO DEL ENTORNO, O CONTEXTO, TIENE UN PAPEL FUNDAMENTAL EN EL PROCESO DE RECONOCIMIENTO, EL CUAL SE VE AFECTADO PRINCIPALMENTE POR LA MEMORIA Y LA PERCEPCIÓN.
	CATEGORÍAS	Memoria, Procesamiento
	PFs ASOCIADOS	PF-10, PF-16, PF-17, PF-29, PF-35, PF-42, PF-53, PF-56, PF-59, PF-60, PF-68, PF-71, PF-72, PF-78

Tabla 84. CT-4

CT-5	DESCRIPCIÓN	LA CAPACIDAD PERCEPTUAL DEL SISTEMA DEBE ESTAR DISTRIBUIDA Y SOPORTADA CON LOS RECURSOS COMPUTACIONALES NECESARIOS BAJO DEMANDA, A FIN DE PREVENIR EFECTOS ADVERSOS, COMO LOS DETECTADOS EN ORGANISMOS BIOLÓGICOS (P. EJ. RECONSTRUCCIONES INCORRECTAS DE REPRESENTACIONES, O EFECTOS OPTO/AUDITIVOS PRODUCIDOS POR CULPA DE LIMITACIONES DE CAPACIDAD).
	CATEGORÍAS	Modularidad, Procesamiento
	PFs ASOCIADOS	PF-2, PF-9, PF-11, PF-12, PF-13, PF-14, PF-16, PF-33, PF-36, PF-44, PF-49, PF-71, PF-78

Tabla 85. CT-5

CT-6	DESCRIPCIÓN	LAS DIFERENTES ÁREAS SENSORIALES DEBEN PERMITIR LA COOPERACIÓN ENTRE SÍ, A MODO DE REFORZAR Y MEJORAR LA CAPACIDAD DE RECONOCIMIENTO, Y LA CALIDAD DE LA INFORMACIÓN RELATIVA A LOS ELEMENTOS IDENTIFICADOS.
	CATEGORÍAS	Conectividad
	PFs ASOCIADOS	PF-2, PF-3, PF-4, PF-11, PF-12, PF-13, PF-14, PF-16, PF-23, PF-33, PF-34, PF-37, PF-38, PF-39, PF-41, PF-43, PF-45, PF-46, PF-47, PF-48, PF-49, PF-51, PF-60, PF-78

Tabla 86. CT-6

CT-7	DESCRIPCIÓN	EL ENTRENAMIENTO PREVIO DEL SISTEMA (ANÁLOGO A LOS BEBÉS) ES CRUCIAL PARA CONSEGUIR UNA BASE CONCEPTUAL CRÍTICA, DESDE LA QUE PODER ACTUAR PROGRESIVAMENTE DE MANERA INDEPENDIENTE, MEDIANTE UN PROCESO RECURSIVO QUE MEJORA LA COGNICIÓN.
	CATEGORÍAS	Memoria
	PFs ASOCIADOS	PF-10, PF-17, PF-29, PF-59, PF-71

Tabla 87. CT-7

CT-8	DESCRIPCIÓN	LA REEVALUACIÓN Y MONITORIZACIÓN DEL PROCESO DE ADQUISICIÓN Y FORMALIZACIÓN DE CONOCIMIENTO (REGISTROS / LOGS) PUEDEN AYUDAR AL AJUSTE Y EVALUACIÓN DE LA CALIDAD DE DICHO CONOCIMIENTO.
	CATEGORÍAS	Procesamiento, Memoria
	PFs ASOCIADOS	PF-6, PF-7, PF-9, PF-16, PF-17, PF-18, PF-25, PF-33, PF-43, PF-51, PF-57, PF-58

Tabla 88. CT-8

CT-9	DESCRIPCIÓN	LOS DIFERENTES ÓRGANOS SENSORIALES ACTÚAN COMO MÓDULOS INDEPENDIENTES, PERO INTERCONECTADOS (AL IGUAL QUE LOS PROCESOS SUPERIORES DE LA COGNICIÓN). ADEMÁS, UTILIZAN MÚLTIPLES MODELOS ABSTRACTOS SIMULTÁNEAMENTE PARA REPRESENTAR LOS DIFERENTES ELEMENTOS PERCIBIDOS. ASÍ PUES, POR UN LADO, SE RECONOCEN LAS CARACTERÍSTICAS BÁSICAS O ESENCIALES, POR OTRO SE COMPARAN LAS CARACTERÍSTICAS CON ELEMENTOS PROTOTÍPICOS O JERARQUÍAS DE LOS MISMOS, Y POR OTRO SE DETERMINAN LAS RELACIONES ENTRE LOS DISTINTOS ELEMENTOS RECONOCIDOS (P. EJ. EN LA VISIÓN, RELACIONES GEOMÉTRICAS O DE POSICIÓN).
	CATEGORÍAS	Procesamiento, Conectividad
	PFs ASOCIADOS	PF-11, PF-16, PF-18, PF-19, PF-20, PF-21, PF-22, PF-23, PF-25, PF-26, PF-28, PF-29, PF-33, PF-34, PF-37, PF-38, PF-39, PF-40, PF-41, PF-45, PF-46, PF-49, PF-50, PF-57, PF-58

Tabla 89. CT-9

CT-10	DESCRIPCIÓN	LAS CARACTERÍSTICAS PERCIBIDAS POR SISTEMAS SENSORIALES DE DIVERSAS NATURALEZAS (P. EJ. VISTA - OÍDO) PUEDEN SER COMPLETAMENTE DIFERENTES ENTRE SÍ, PRESENTANDO UN MÍNIMO (O NULO) NÚMERO DE ATRIBUTOS COMUNES.
	CATEGORÍAS	Procesamiento
	PFs ASOCIADOS	PF-18, PF-23, PF-24, PF-25, PF-26, PF-28, PF-33, PF-34, PF-37, PF-38, PF-39, PF-40, PF-41, PF-42, PF-43, PF-45, PF-51, PF-57, PF-60, PF-61

Tabla 90. CT-10

CT-11	DESCRIPCIÓN	ES IMPORTANTE CONSIDERAR EL PROPIO ESTADO DEL SISTEMA SENSORIAL A LA HORA DE ANALIZAR LA INFORMACIÓN RECOGIDA POR EL MISMO (P. EJ. SI LA CABEZA ESTÁ INCLINADA, LA VISIÓN SE CORRIGE EN BASE A LOS ÁNGULOS DE DICHA INCLINACIÓN). ÉSTA INFORMACIÓN TAMBIÉN PUEDE SER DETERMINANTE PARA EVALUAR LA CALIDAD DE LOS DATOS ADQUIRIDOS (ESTADO DE SENSORES, CARACTERÍSTICAS DE LOS MISMOS), Y PARA LA PROTECCIÓN DEL SISTEMA DE MEDIDA (P. EJ. SI EL RANGO DE TEMPERATURA DETECTADO PUEDE DAÑAR EL SISTEMA O A DETERMINADOS SENSORES DEL MISMO).
	CATEGORÍAS	Procesamiento, Modularidad, Conectividad
	PFs ASOCIADOS	PF-16, PF-24, PF-36, PF-38, PF-40, PF-41, PF-42, PF-43, PF-46, PF-48, PF-50, PF-51, PF-78

Tabla 91. CT-11

CT-12	DESCRIPCIÓN	EXISTEN CARACTERÍSTICAS DEL ENTORNO QUE PUEDEN AFECTAR A LA PERCEPCIÓN DE OTRAS CARACTERÍSTICAS, COMO POR EJEMPLO LA TEMPERATURA, QUE AFECTA AL OLFATO Y AL GUSTO. ASÍ PUES, EL REFUERZO ENTRE SISTEMAS SENSORIALES SERÁ BENEFICIOSO PARA MODELAR NO SÓLO EL CONOCIMIENTO EXPLÍCITO DE LAS CARACTERÍSTICAS, SINO LAS INTERACCIONES ENTRE ELLAS.
	CATEGORÍAS	Procesamiento, Conectividad
	PFs ASOCIADOS	PF-24, PF-29, PF-30, PF-33, PF-36, PF-38, PF-40, PF-41, PF-43, PF-47, PF-48, PF-50, PF-51, PF-57, PF-58, PF-68, PF-78

Tabla 92. CT-12

CT-13	DESCRIPCIÓN	EL CONOCIMIENTO ES MODELABLE A UN CONJUNTO DE HECHOS, TÉCNICAS Y PROCEDIMIENTOS, LOS CUALES GUARDAN RELACIÓN ENTRE ELLOS, Y PROPORCIONAN INFORMACIÓN PARCIAL O COMPLETA SOBRE EL ENTORNO.
	CATEGORÍAS	Conectividad, Memoria
	PFs ASOCIADOS	PF-52, PF-53, PF-54, PF-60, PF-78

Tabla 93. CT-13

CT-14	DESCRIPCIÓN	EL CONOCIMIENTO SE MODELA MEDIANTE REPRESENTACIONES QUE ACTÚAN DE MANERA CONJUNTA: LAS REPRESENTACIONES DE MODALIDAD ESPECÍFICA (CONJUNTOS DE CARACTERÍSTICAS) Y LAS REPRESENTACIONES AMODALES (PROPIEDADES Y RELACIONES ENTRE ENTIDADES DEL ENTORNO). LA CATEGORIZACIÓN DE ELEMENTOS EN BASE A ESTAS REPRESENTACIONES FACILITA LA REALIZACIÓN DE PROCESOS DEDUCTIVOS.
	CATEGORÍAS	Memoria
	PFs ASOCIADOS	PF-18, PF-19, PF-20, PF-21, PF-22, PF-23, PF-24, PF-25, PF-26, PF-28, PF-29, PF-30, PF-33, PF-37, PF-38, PF-39, PF-40, PF-41, PF-43, PF-50, PF-53, PF-54, PF-57, PF-60, PF-78

Tabla 94. CT-14

CT-15	DESCRIPCIÓN	EL INCREMENTO DE LA EXPERIENCIA RECONOCIENDO ELEMENTOS POTENCIA Y AMPLÍA LAS CARACTERÍSTICAS DE RECONOCIMIENTO DE LA MISMA.
	CATEGORÍAS	Procesamiento, Memoria
	PFs ASOCIADOS	PF-17, PF-29, PF-30, PF-33, PF-35, PF-50, PF-51, PF-53, PF-54, PF-56, PF-58, PF-59, PF-62, PF-71, PF-78

Tabla 95. CT-15

CT-16	DESCRIPCIÓN	NO TODOS LOS ELEMENTOS RECONOCIDOS PRESENTAN EL MISMO TIPO DE CARACTERÍSTICAS (P. EJ. LA LUZ NO TIENE SABOR, AL MENOS CONOCIDO).
	CATEGORÍAS	Procesamiento, Memoria
	PFs ASOCIADOS	PF-18, PF-23, PF-24, PF-25, PF-26, PF-28, PF-33, PF-34, PF-37, PF-39, PF-43, PF-45, PF-50, PF-57, PF-58, PF-60, PF-61

Tabla 96. CT-16

CT-17	DESCRIPCIÓN	AL IGUAL QUE LA INFORMACIÓN ESPECÍFICA DE CADA SISTEMA SENSORIAL, EL GRUPO GENERAL DEL CONOCIMIENTO CORRESPONDIENTE AL GRUPO SENSORIAL COMPLETO SE PUEDE ESTRUCTURAR MEDIANTE CATEGORÍAS REPRESENTADAS MEDIANTE EL USO DE DIFERENTES MODELOS QUE ACTÚAN EN COOPERACIÓN. MODELOS COMO LOS BASADOS EN EJEMPLARES, REGLAS, PROTOTIPOS O REPRESENTACIONES DINÁMICAS.
	CATEGORÍAS	Conectividad, Memoria
	PFs ASOCIADOS	PF-18, PF-19, PF-20, PF-21, PF-22, PF-25, PF-28, PF-33, PF-37, PF-39, PF-40, PF-41, PF-43, PF-45, PF-54, PF-56, PF-57, PF-58, PF-60, PF-62, PF-64, PF-65, PF-66, PF-67, PF-68

Tabla 97. CT-17

CT-18	DESCRIPCIÓN	ES IMPORTANTE EL USO DE ETIQUETAS O NOMBRES PARA DEFINIR CATEGORÍAS DE ELEMENTOS. EN EL CEREBRO, LA ASOCIACIÓN DE ETIQUETAS AL RESPECTIVO CONTENIDO ES UN FACTOR CLAVE EN EL ÁMBITO DEL CONOCIMIENTO, TENIENDO UN IMPACTO MUY GRANDE SOBRE LAS REPRESENTACIONES AMODALES, AL PERMITIR RELACIONAR DE MANERA DIRECTA DIFERENTES ELEMENTOS SIN TENER QUE UTILIZAR TODO EL CONTENIDO COGNITIVO ASOCIADO A CADA ELEMENTO.
	CATEGORÍAS	Procesamiento, Memoria
	PFs ASOCIADOS	PF-35, PF-56, PF-58, PF-62, PF-63, PF-72

Tabla 98. CT-18

CT-19	DESCRIPCIÓN	ES IMPORTANTE CONSIDERAR EL TIEMPO COMO FACTOR A LA HORA DE ALMACENAR REPRESENTACIONES EN LA MEMORIA (RECUERDOS) Y PARA RECONOCER ELEMENTOS.
	CATEGORÍAS	Procesamiento, Memoria
	PFs ASOCIADOS	PF-15, PF-27, PF-28, PF-31, PF-32, PF-36, PF-50, PF-55, PF-69, PF-70, PF-78

Tabla 99. CT-19

CT-20	DESCRIPCIÓN	LA RECUPERACIÓN DE ELEMENTOS EN LA MEMORIA SE REALIZA MEDIANTE PISTAS Y CLAVES (COMO, POR EJEMPLO, LAS ETIQUETAS MENCIONADAS ANTERIORMENTE, O UN ESTÍMULO ESPECÍFICO ASOCIADO A LA INFORMACIÓN QUE QUIERE RECUPERARSE.
	CATEGORÍAS	Memoria
	PFs ASOCIADOS	PF-35, PF-56, PF-58, PF-62, PF-63, PF-71, PF-72

Tabla 100. CT-20

CT-21	DESCRIPCIÓN	LA MEMORIA PUEDE DIVIDIRSE EN DOS GRANDES GRUPOS, LA MEMORIA DECLARATIVA, QUE A SU VEZ INCLUYE LA MEMORIA EPISÓDICA (CONOCIMIENTO CONSCIENTE DE ACONTECIMIENTOS O EPISODIOS DATADOS TEMPORALMENTE) Y LA MEMORIA SEMÁNTICA (CONOCIMIENTO GENERAL RELATIVO A LOS ELEMENTOS DE UN ENTORNO, PALABRAS Y CONCEPTOS, PROPIEDADES E INTERRELACIONES), Y LA MEMORIA IMPLÍCITA O NO DECLARATIVA, QUE ESTÁ RELACIONADA CON FORMAS NO CONSCIENTES DE MEMORIA, COMO POR EJEMPLO LA ADQUISICIÓN DE HABILIDADES MOTORAS (P. EJ. MONTAR EN BICI O TOCAR EL PIANO).
	CATEGORÍAS	Memoria
	PFs ASOCIADOS	PF-18, PF-19, PF-20, PF-21, PF-22, PF-43, PF-49, PF-50, PF-55, PF-61, PF-64, PF-69, PF-73, PF-74, PF-75, PF-76, PF-77

Tabla 101. CT-21

CT-22	DESCRIPCIÓN	LOS ELEMENTOS DE CONOCIMIENTO EN LA MEMORIA PUEDEN MIGRAR O EXTENDER A OTRA LOCALIZACIÓN. ESTO ES, PASAR O REPLICARSE DE UN TIPO A OTRO, COMO EN EL CASO DE LA MEMORIA IMPLÍCITA, QUE ORIGINALMENTE ESTÁ COMPUESTA POR CONOCIMIENTO DE TIPO DECLARATIVO.
	CATEGORÍAS	Memoria
	PFs ASOCIADOS	PF-15, PF-17, PF-48, PF-53, PF-54, PF-55, PF-64, PF-69, PF-70, PF-74, PF-75, PF-76, PF-77

Tabla 102. CT-22

CT-23	DESCRIPCIÓN	DE MANERA SUCINTA PUEDEN DEFINIRSE CUATRO TIPOS GENERALES DE MEMORIA EN BASE A SU CONTENIDO: HISTÓRICOS DEL SISTEMA, REPRESENTACIÓN DE CARACTERÍSTICAS DE OBJETOS Y SUB-OBJETOS, REPRESENTACIÓN DE RELACIONES ENTRE OBJETOS Y SUB-OBJETOS, Y FINALMENTE INTERACCIÓN ENTRE OBJETOS DE DIFERENTE CLASE (P. EJ. CÓMO SE USA).
	CATEGORÍAS	Memoria
	PFs ASOCIADOS	PF-18, PF-19, PF-20, PF-21, PF-22, PF-25, PF-52, PF-54, PF-56, PF-57, PF-58, PF-64, PF-65, PF-66, PF-67, PF-68, PF-69, PF-73, PF-74, PF-75, PF-76, PF-77, PF-78

Tabla 103. CT-23

Con el objetivo de mejorar la mantenibilidad y facilitar la aplicación, tanto de PFs como de CTs, se muestra a continuación la matriz de trazabilidad que relaciona ambos grupos. Dicha matriz se presenta dividida en dos tablas, a fin de mejorar su legibilidad.

4.5. Aproximación del modelo biológico de representación del conocimiento a la computación

En otros campos de la computación se han conseguido notables avances al aplicar enfoques basados en la biología para resolver problemas de forma algorítmica. Así bien, un problema que inicialmente presenta una complejidad computacional de tipo NP, han podido ser aproximados mediante algoritmos evolutivos que consiguen soluciones optimas (o sub-optimas) en tiempos relativamente pequeños.

Adaptar el sistema de representación y obtención de la información de los humanos, basado en la percepción, al ámbito computacional puede suponer una revolución similar al del desarrollo de la computación evolutiva, pero respecto al tratamiento semántico de la información y a su clasificación. Si bien puede que la representación inicial de los conceptos no sea perfecta o absoluta, pero el propio modelo se encargará de refinar los conceptos almacenados a medida que se perciban nuevas características de los mismos.

Desde un punto de vista modular, y considerando una perspectiva simplificada del funcionamiento de la percepción humana, una posible representación computacional del conocimiento basada en entradas sensitivas podría ser construida teniendo en cuenta (pero no limitado a) los siguientes requisitos:

1. Diversas ontologías pueden ser utilizadas, cada una para representar distintos dominios asociados a estímulos particulares (visual, auditivo, termal, etc.). Existen diferentes ontologías y técnicas de razonamiento para procesar distintos sentidos. Por ejemplo, relacionado con la visión, [478] propone utilizar una ontología de conceptos visuales para guiar a los expertos en la descripción visual de los objetos de su dominio integrando interpretación semántica de las imágenes. Igualmente, [479] propone un marco de detección capaz de identificar instancias de objetos compuestos, y es validada mediante la identificación de laterales de vehículos. Ni que decir tiene que cada sentido debería ser representado de forma específica e independiente.
2. Las ontologías definidas para cada dominio deben tener la habilidad de ser expandidas y modificadas acorde a nuevas características percibidas y no clasificadas previamente. Es decir, su contenido debe ser abierto y extensible, pero manteniendo la consistencia y estructura fundamental.
3. Es necesario el procesamiento de cada estímulo percibido por cada sistema sensorial, teniendo en cuenta el conocimiento almacenado previamente a fin de determinar el grado de significatividad del estímulo (relevancia). Esto permite un filtrado inicial de los estímulos recibidos acorde al conocimiento existente, y la clasificación de nuevos (y desconocidos) estímulos, incrementando la base de conocimiento de las áreas sensoriales específicas relacionadas.

4. Además de realizar un procesamiento individual de cada sentido, la combinación de conocimiento percibido a través de otros sentidos debería ser considerada a fin de definir los conceptos o entidades de una forma completa, compleja y estructurada en una base de conocimiento global. Como ejemplo, una manzana de plástico, visualmente puede ser similar a una natural, y ambas pertenecen a la categoría “manzana”. Para diferenciarlas y darse cuenta de que no son exactamente del mismo tipo, sería necesario recurrir al tacto / olfato / gusto. O aplicar una acción, como golpear ambas manzanas y escuchar el sonido percibido por cada una. Esta combinación global de estímulos puede ayudar a diferenciar y percibir características en un amplio rango de elementos del entorno.

5. El enriquecimiento de la base de datos de conocimiento global debería ser realizado mediante un proceso análogo, pero más complejo, que el procesado y filtrado de estímulos citado en el punto 3, ya que, en este caso, no se trata de estímulos individuales, sino de la combinación de todos los percibidos y su histórico.

4.5.1. Patrón de análisis sensorial

Respetando fielmente los distintos requisitos principales mostrados anteriormente, junto con sus interrelaciones, se propone el siguiente patrón de análisis sensorial:

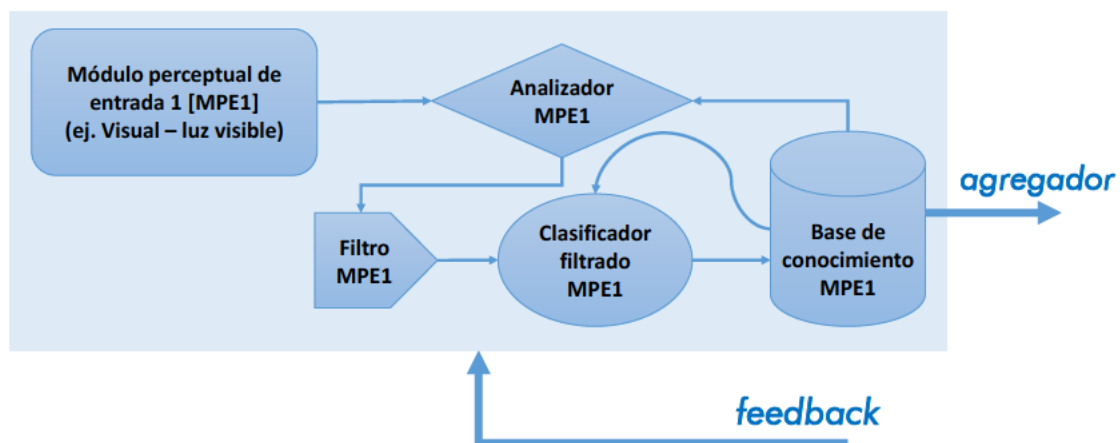


Ilustración 26. Patrón de análisis sensorial

4.5.2. Arquitectura lógica

La Ilustración 27 muestra el diagrama correspondiente a la propuesta de una arquitectura inicial, combinando diferentes instancias del patrón de análisis sensorial junto con los diferentes elementos que conforman el modelo propuesto.

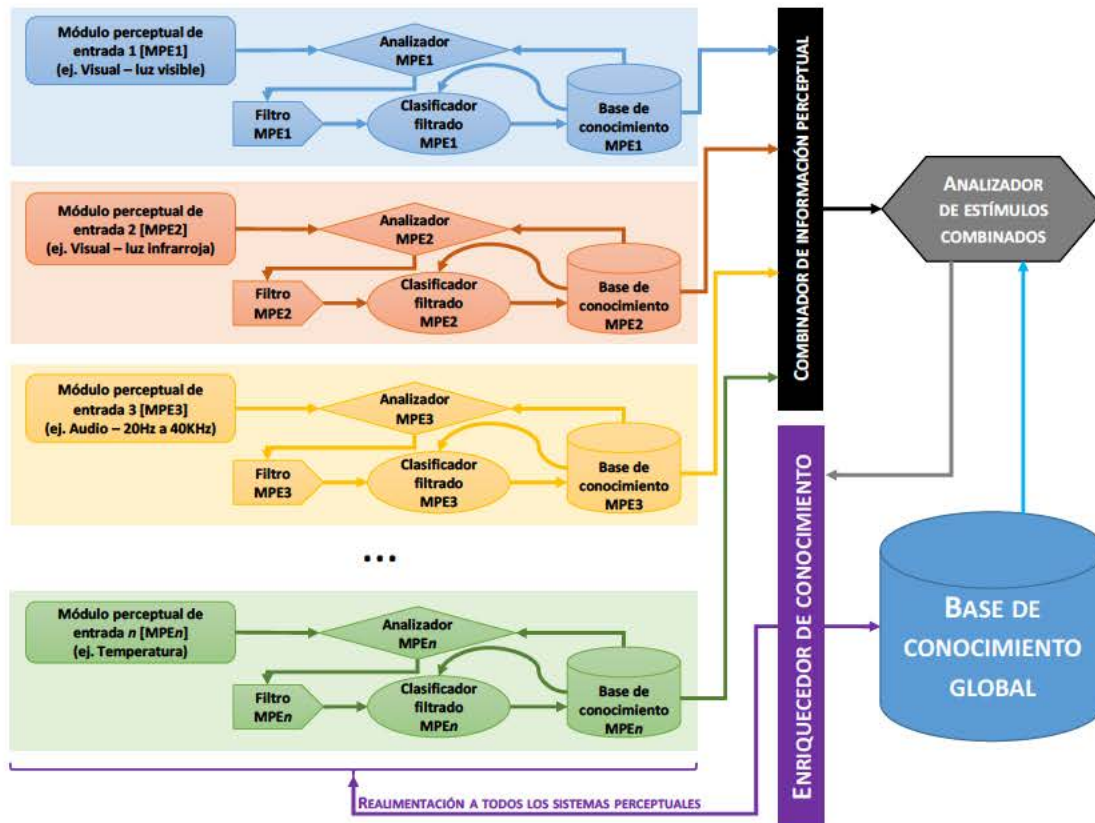


Ilustración 27. Propuesta de arquitectura lógica

Como primera aproximación, se puede pensar en un sistema de representación que tenga en cuenta y de forma relacionada, la información recogida por diferentes receptores artificiales. Dichos receptores no tienen por qué ser únicamente los correspondientes a cualquier ser humano, ni estar calibrados acorde a como el ser humano lo está. Por ejemplo, se podría dotar de cámaras infrarrojas al sistema encargado de alimentar al modelo, o de sensores de temperatura mucho más precisos que la piel, de micrófonos con rangos de sensibilidad mucho mayores que el oído humano, etc. Además, intuitivamente cuanto más precisos y más tipos de estímulos reciba el sistema de aprendizaje, mayor definición y completitud tendrá la conceptualización del conocimiento. Los distintos sistemas de percepción deberán ser filtrados mediante un *analizador*, a fin de evitar sobrecarga del sistema y redundancia con estímulos previos que indiquen los mismos conceptos. Igualmente, dichos estímulos se analizarán tanto de forma combinada como individual, y se contrastarán

con información similar y relevante almacenada previamente en la base de conocimiento. El modelo de representación deberá tener en cuenta los estímulos por separado y en conjunto, al igual que las relaciones que se establezcan entre ellos. Finalmente, los estímulos destinados a integrarse en la base de conocimiento serán previamente analizados por un *combinador*, encargado de mezclar adecuadamente esta información para alimentar a un sistema *enriquecedor de conocimiento* capaz de determinar si existe algún concepto previo en la base de conocimiento que pueda ser refinado en base a las nuevas entradas, o bien, debe crearse un concepto nuevo para dichas entradas. Este módulo también proporcionará refuerzo a los sistemas perceptuales a fin de reforzar el proceso de detección, facilitando, entre otros aspectos, la respuesta de categoría reconocida, de manera que los sistemas perceptuales puedan ajustar y mejorar acorde a dicho *feedback*, su proceso de detección, mejorando el rendimiento global del sistema.

La parte correspondiente a la base de conocimiento podría partir de algún modelo actual, como por ejemplo las ontologías multidimensionales, aunque dichos modelos deberán ser refinados y adaptados a fin de integrarse con el sistema analizador de estímulos y permitir la realimentación del mismo para completar la fase de filtrado de información no relevante. Es importante distinguir, al menos, tres categorías diferentes de elementos a almacenar:

- Clases y prototipos representativos de las mismas, junto con sus características esenciales obtenidas y determinadas mediante el enriquecimiento perceptual.
- Relaciones entre etiquetas de clase y sus características (punto anterior). De esta manera, será posible asociar de forma abstracta los elementos (mediante sus etiquetas) con sus relaciones (p. ej. posición relativa, o interacciones). Esta característica también facilita el uso de múltiples idiomas para el acceso y categorización a los elementos, puesto que se puede asociar un conjunto de etiquetas (perro, dog, chien, Hund, собака, 犬...), con las características y prototipos correspondientes al concepto que representan.
- Relaciones e interacciones entre clases (como la pertenencia)
- Contexto, o grupo de clases más frecuentes en determinados entornos. Actúa como caché para acelerar el acceso a las clases potenciales a las que podría pertenecer un elemento durante el proceso de reconocimiento.
- Histórico de instancias representativas empleadas para el reconocimiento de las categorías o subcategorías.

Como ya se comentó en el planteamiento del problema, el almacén de la información debe soportar grandes volúmenes de datos, además debe ser eficiente a fin de poder recuperar rápidamente información almacenada y hacer frente a muchas peticiones simultáneas. Por razones de seguridad, escalabilidad y eficacia, y teniendo en cuenta el estado actual de la tecnología y sus tendencias, sería apropiado para este modelo ser distribuido entre varias máquinas interconectadas, lo cual dirige la solución a la adopción de un enfoque de computación en la nube. Las ventajas adicionales de la

utilización de este enfoque son, por ejemplo, el acceso simultáneo de diferentes terminales o aplicaciones, la asignación dinámica de los recursos, una mejor capacidad de recuperación de la base de datos de conocimiento mediante la distribución de los datos y una mayor resiliencia y tolerancia a fallos. Además, se pueden aplicar técnicas de proximidad de información – proximidad geográfica. Sin embargo, como se observó en el apartado 3.3.3, la adopción de tecnología basada en la computación en la nube requiere, para su aplicación a nivel empresarial, el desarrollo de un plan de gobierno [368] que facilite su implementación e implantación, su desarrollo, prevenga contingencias y problemas o, defina los acuerdos de servicio, entre otros.

No es obvio definir las diferentes partes de esta arquitectura, tales como el combinador de información perceptual, el cual debería analizar cada entrada sensorial de manera conjunta, y considerando el conocimiento previamente almacenado, determinar si la pieza resultante de información puede enriquecer la base de conocimiento global a través del módulo enriquecedor de conocimiento.

Respecto a los modelos a utilizar para la representación del conocimiento, y conforme a lo expuesto en secciones anteriores, las ontologías, en su forma más abstracta, suponen un potente método para representar y compartir conocimiento, pero como se citó anteriormente, a menudo cada ontología definida está orientada hacia un dominio específico. Sin embargo, la percepción humana está compuesta por complejas series de sistemas sinérgicos interconectados, incluyendo órganos receptores de estímulos y la memoria. Desde un punto de vista simplificado, cada sentido proporciona (al menos) una característica del entorno, la procesa y la filtra teniendo en cuenta su nivel de significatividad basada en la experiencia (memoria). Los estímulos irrelevantes (aquellos que son filtrados) se pierden debido a la memoria a corto plazo (como un *buffer* en términos de memoria digital). Por otra parte, los estímulos relevantes son procesados de forma global. Esto es, teniendo en cuenta la información adquirida a través de otros sentidos en ese preciso momento. Durante dicho procesamiento se realizan también tareas de razonamiento, acceso a memoria y refinado de la información conocida sobre los conceptos almacenados.

Crear un sistema como el propuesto, capaz de proporcionar acceso eficiente a una representación global de conocimiento abstracta no-convencional, y dotarlo de técnicas de aprendizaje automático e interfaces de adquisición (percepción) y presentación de datos adecuadas, podría convertirse eventualmente en un sistema que acabara aprendiendo autónomamente sobre el entorno y sus características, de forma análoga a dicho proceso en los seres humanos: en un estado inicial, de forma semi-supervisada, pero transicionalmente cada vez más autónoma. Esto permitirá tener una base de conocimiento extremadamente extensa, con una clasificación de elementos basada en la información obtenida por las diferentes entradas sensoriales, y una reducción de la redundancia en el almacenamiento de los mismos gracias a un proceso de filtrado y generalización. Gracias a la clasificación y reorganización de los conceptos en base a la diversidad de naturalezas de la información percibida, la precisión de definición conceptual será mucho más amplia que la de otros sistemas convencionales existentes, al reflejar un mayor número de características. Fenómeno que puede ayudar a descubrir nuevas (y desconocidas) inferencias entre elementos existentes. En un estado posterior, dotar al sistema de razonadores que puedan

aprovechar la base de conocimiento global, podría mejorar y enriquecer muy notablemente aplicaciones existentes tales como motores de búsqueda web o sistemas de soporte de decisiones, y generar un amplio abanico de nuevas aplicaciones inteligentes.

4.5.3. Entradas perceptuales - La visión

Para dotar de información al modelo, que será transformada en conocimiento de manera análoga a como lo hacen los seres humanos, es necesario recurrir al uso de sistemas perceptuales. Uno de los más importantes, como se observó en el estado del arte, es la visión. En el presente apartado, se propone un modelo de análisis y almacenamiento semántico multi-nivel considerando las diferentes características reconocibles mediante la visión, considerando múltiples dimensiones (imágenes bidimensionales, tridimensionales y sus análogos en vídeo) a fin de completar la parte correspondiente al módulo de visión del modelo representado en la Ilustración 27. Con este módulo se pretende proporcionar un nuevo enfoque sobre cómo procesar y evaluar la información visual en el ámbito de las ciencias de la computación para entornos multidimensionales (como se comentaba, imagen y vídeo en 2D/3D). Para ello es necesario la extracción de distintos tipos de características requeridos por los distintos modelos generales de reconocimiento expuestos en el estado del arte, así como la inferencia de información semántica entre ellas, tales como las relaciones de pertenencia o posición. El modelo está basado en la percepción humana, considerando los distintos sistemas biológicos que hacen la visión posible, y su adaptación al medio computacional mediante la combinación de estructuras, sistemas de comunicación, algoritmos y plataformas existentes. El modelo proporciona las bases para una solución completa y abierta orientada a la detección genérica de elementos y el aprendizaje a través de la visión, siendo capaz de adaptar y extender los conceptos previamente aprendidos de acuerdo a la nueva información recibida y las características detectadas en la misma.

Para comprender la relevancia de la visión, desde un punto de vista global, se estima que el volumen de datos generados por los usuarios cada minuto (inicios de 2015) es de aproximadamente 1.25 millones de imágenes nuevas [480]. Sin embargo, la infraestructura de la red está construida en torno al texto. Incluso los principales motores de búsqueda dependen del texto. En el caso de la búsqueda de imágenes, utilizan meta-etiquetas de información asociadas, en caso de estar disponibles. Dichas meta-etiquetas proporcionan una representación textual de las imágenes. Sin embargo, este enfoque está cambiando a un esquema mejor y más sofisticado. Muchos modelos de negocio se aprovechan del uso de imágenes (por ejemplo, *marketing* y sitios de venta on-line [481]), al igual que muchas aplicaciones de diferentes áreas (seguridad [482], salud [452] y entretenimiento, entre otros). Esta necesidad está incrementando la investigación y el desarrollo de nuevos métodos para incrementar los beneficios y las capacidades ofrecidas. Debido a los avances tecnológicos en computación, avanzados y potentes algoritmos basados en la inteligencia artificial pueden ser finalmente implementados, con el objetivo de detectar y clasificar automáticamente elementos contenidos en imágenes. Sin embargo, los algoritmos

existentes por sí mismos se encuentran todavía lejos de las capacidades perceptuales humanas.

La arquitectura propuesta para el módulo de visión ha sido diseñada a fin de evaluar el entorno percibido a través de la visión, y permitir su integración en el modelo global propuesto en la presente tesis. Antes de describir la arquitectura, es importante definir el concepto de visión desde una perspectiva dimensional y sus equivalencias entre el ser humano y la visión artificial (ver Ilustración 28).

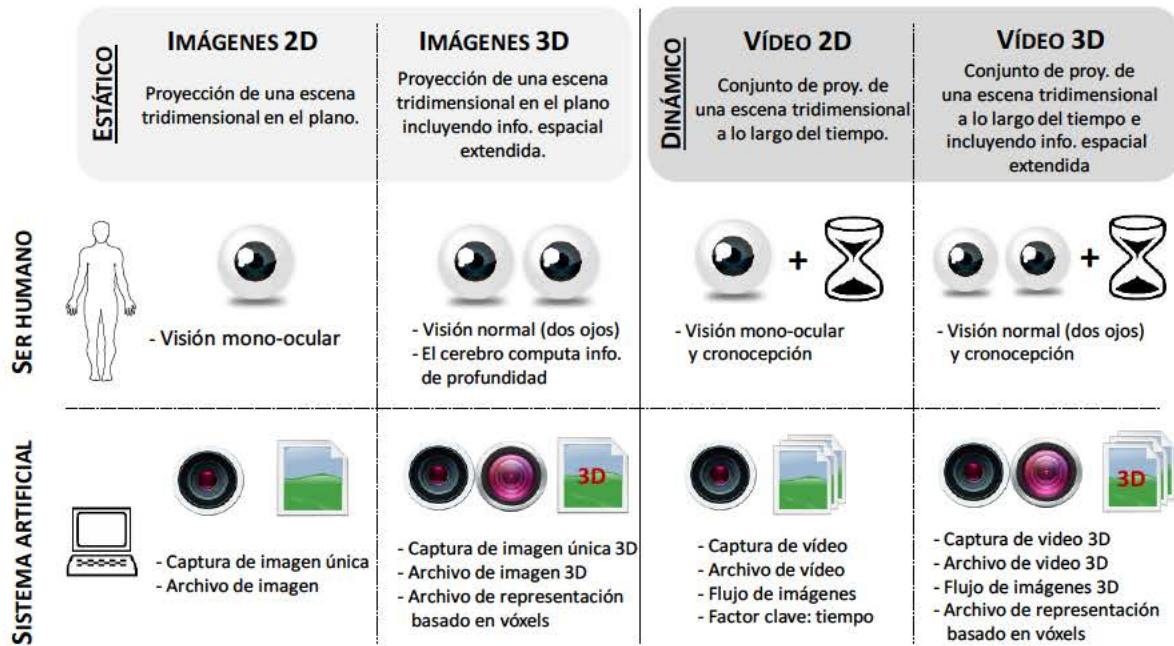


Ilustración 28. Varios escenarios asociados a la visión. Analogía hombre-máquina

Actualmente las imágenes y vídeos en 2D componen el conjunto de formatos más comunes y utilizados en sistemas artificiales debido a su simplicidad en la captura, almacenamiento y presentación [483]. El contenido 2D representa básicamente una proyección en el plano de un entorno tridimensional, por ello, una gran parte de la información (tal como la distancia entre objetos en distintos planos) no puede ser inferida. Más aún, debido a la proyección, algunos escenarios tridimensionales podrían derivar en ilusiones ópticas (por ejemplo, con objetos de diferentes tamaños a diferentes distancias podría ocurrir que los objetos pequeños aparentasen tener mayor tamaño que los grandes). En el caso de los vídeos, algunas de dichas ilusiones pueden ser mitigadas debido al efecto del movimiento. En caso de animaciones o conjuntos de imágenes capturados a lo largo del tiempo, la información extendida asociada a las acciones que pasan entre los elementos representados puede ser inferida (por ejemplo, dos objetos que se están moviendo en direcciones opuestas, o un objeto que cruza frente a otro debido al solapamiento). Las imágenes 3D aumentan la complejidad computacional al proporcionar información de profundidad. Los beneficios de incluir dicha información adicional se presentan al facilitar el cálculo

preciso de distancias entre objetos de la escena, y la determinación correcta de su tamaño, siempre y cuando las cámaras o el dispositivo de captura en cuestión esté calibrado correctamente. Para obtener información de profundidad a partir de visión 2D, es necesario combinar y analizar dos o más imágenes correspondientes a la misma escena, pero observadas desde diferentes puntos de vista conocidos y pre-calibrados. Mediante el uso de geometría Epipolar, entre los distintos puntos equivalentes en las distintas imágenes, y utilizando la información de calibración de los dispositivos de captura, es posible calcular las relaciones entre los puntos que componen las imágenes. Por otra parte, existen otros mecanismos de captura 3D basados en diferentes tecnologías, tales como las cámaras *Microsoft Kinect* o *Creative Senz3D*, las cuales proporcionan diferentes formatos de representación de la escena que observan, incluyendo información 3D (por ejemplo, representando la imagen mediante *voxels* en lugar de *pixels*).

El sistema perceptual humano es extremadamente dependiente de la cronocepción para percibir el paso del tiempo. Aplicado a las imágenes, esto lleva a la percepción continua del entorno en base a la visión tridimensional o bidimensional. Cabe destacar la cooperación en el proceso por parte de algunos sistemas perceptuales, tales como la propiocepción y la equilibriocepción, los cuales proporcionan información adicional e inherente que ayuda a interpretar o corregir posibles fallos en la visión [484]. En los sistemas artificiales, este tipo de información (e incluso más cantidad de información y mucho más precisa) puede ser proporcionada por algunos dispositivos de captura, como, por ejemplo, al considerar la meta-información EXIF almacenada en las imágenes por algunas cámaras. Dicha información EXIF puede contener datos sobre geo-localización y su precisión (lugar en el que fue capturada la imagen con un margen de error en metros), la rotación de la cámara, las características de la lente, si la imagen se disparó con flash y en qué modo, etc. Esta meta-información puede ser muy relevante para aplicar ciertas correcciones (como rotación o reducción de ruido) a la hora de representar e interpretar la imagen, de una forma similar a como lo hace el cerebro de manera automática.

Considerando las características perceptuales presentadas, la aproximación artificial al modelo de visión humana se ha diseñado como puede observarse en la Ilustración 29 [485].

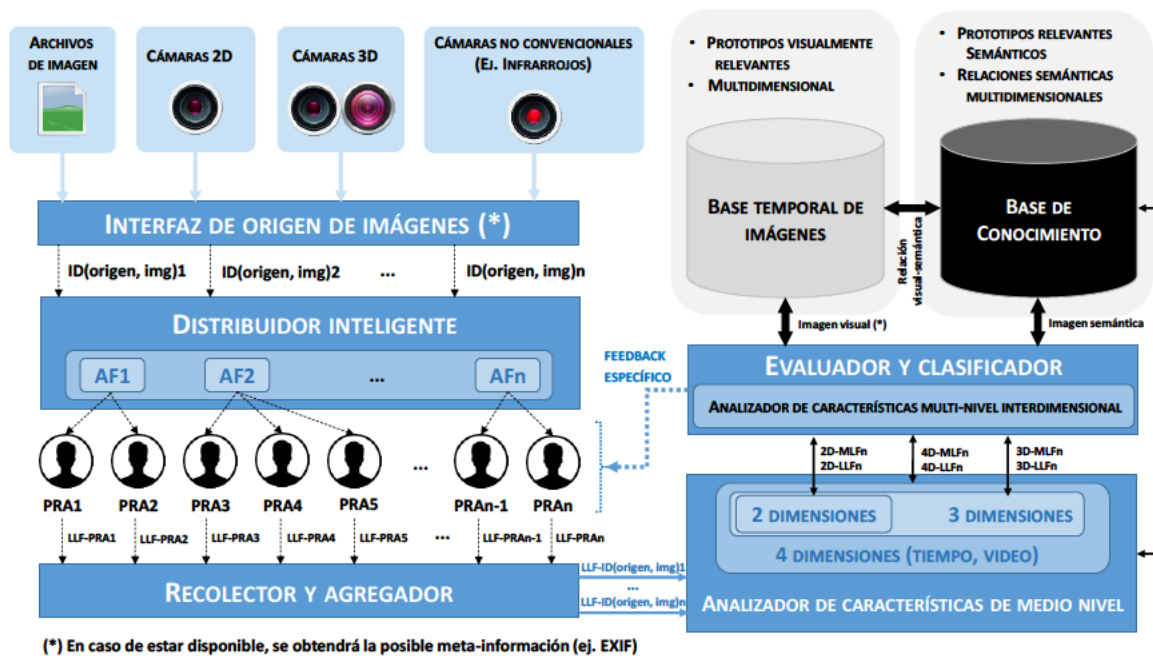


Ilustración 29. Arquitectura física del modelo para el módulo de visión

El modelo está compuesto por distintos módulos interconectados, de manera análoga a los procesos identificados en la cognición humana, a fin de facilitar su mantenimiento, incrementar su extensibilidad y su escalabilidad, permitiendo cambios futuros, como la integración de posibles mejores algoritmos de reconocimiento, sistemas adicionales de entrada (captura), futuras tecnologías de clasificación, de almacenamiento, o incluso la distribución entre múltiples servidores para incrementar bajo demanda las necesidades computacionales del sistema.

El primero de los módulos del modelo, la “interfaz de origen de imágenes”, actúa como un conector común entre los sistemas de captura u orígenes de datos con el resto del sistema. Dicha interfaz genérica proporciona la información de imagen desde diferentes formatos y fuentes. Como puede verse en la ilustración, los formatos pueden variar: desde archivos de imagen estáticas contenidas en una unidad de almacenamiento, imágenes recuperadas desde sitios web, fuentes de vídeo originados en cámaras web y fuentes de vídeo originadas en otro tipo de cámaras (3D, infrarrojos, etc.). Este módulo de entrada proporciona igualmente información relativa a las distintas entradas, y en caso de que esté disponible, meta-información, como información EXIF relativa a la cámara y la captura, o información específica de la calibración de la cámara / parámetros de la lente al siguiente módulo, marcando cada una de las entradas de información con un identificador (ID) específico. El módulo que recibe este conjunto de información es un “distribuidor inteligente”, cuyo objetivo es el de instanciar, distribuir y asignar las tareas de reconocimiento, en función del formato recibido, a diferentes agentes paralelos de reconocimiento (PRAs, *parallel recognition agents*), especializados en reconocimiento de patrones. Dichos agentes pueden trabajar de forma colaborativa, esto es, analizando cada uno diferentes fragmentos la misma imagen a la vez, buscando el mismo tipo de patrón (o características de bajo

nivel), o independientemente, analizando la misma imagen completa, pero empleando cada uno una técnica diferente, permitiendo reducir el tiempo de extracción de características. Por otra parte, los algoritmos empleados por los agentes pueden ser determinados y ejecutados en base al tipo de tarea de reconocimiento que debe ser llevado a cabo. En cualquier caso, los algoritmos de reconocimiento a emplear por el modelo final están fuera del estudio y quedan como futuras líneas de investigación, ya que pueden variar con el tiempo a medida que emerjan nuevas técnicas más refinadas y potentes. Una ventaja de la modularidad mostrada en el modelo propuesto, como se comentaba anteriormente, es la sencillez de modificar las partes que lo componen para añadir dichas mejoras a posteriori.

El resultado de los PRAs es un conjunto de características de bajo nivel (LLF, *low-level features*) en diferentes categorías para la imagen identificada. Estos LLFs proporcionan información semántica básica sobre los componentes reconocidos en la imagen. Algunos de los LLFs son: características cromáticas (histograma de color, momentos de color...), características de la textura (co-ocurrencias, momento y transformada wavelet...) y características estructurales (distribución espacial y puntos salientes). Este análisis semántico visual pretende abordar un enfoque “*bottom-up*” [22]. Entonces, dichas LLFs son enviadas a un módulo “recolector y agregador” que agrupa la información semántica teniendo en cuenta la fuente de origen de la que fue inferida, y envía los resultantes grupos agregados al módulo siguiente, el “analizador de características de medio nivel”.

El analizador de características de medio nivel está compuesto por una compleja distribución de algoritmos de clasificación y regresión, que, operando de manera distribuida, comprueban las relaciones entre los LLFs entrantes y los LLFs y MLFs (*mid-level features*, características de medio nivel) reconocidas y almacenadas previamente, a fin de encontrar una posible relación (como pertenencia o interacción) considerando las diferentes dimensiones para cada caso (imágenes / vídeos en 2D / 3D). Una vez que la información extendida ha sido inferida, los MLFs y LLFs correspondientes a las fuentes de entrada son propagados a un módulo “analizador de características multi-nivel interdimensional”, el cual está compuesto por otro conjunto de métodos de clasificación y regresión, y cuyo objetivo es encontrar relaciones entre los elementos analizados previamente y sus posibles distintas representaciones (por ejemplo, determinar si una imagen es una proyección de una escena 3D, o la relación entre los diferentes elementos que componen un objeto complejo). Este analizador de características multi-nivel es también el responsable de la inferencia del contexto. Como se estudió en la literatura, el contexto tiene un impacto enorme a la hora de percibir e interpretar el entorno. Desde el punto de vista computacional, el entorno puede considerarse como una caché evolutiva que contiene los elementos reconocidos de forma más frecuente (o probable) en el mismo. Cuanto mayor sea la experiencia reconociendo elementos para un entorno determinado, mejor será el tiempo de respuesta promedio a la hora de reconocer visualmente elementos en el mismo.

La sinergia resultante entre el analizador de características multi-nivel interdimensional y el analizador de características de medio nivel deriva en un reconocimiento semántico de alto nivel para el objeto (etiqueta), definido por medio de una compleja jerarquía de características y sus relaciones. El objetivo de esta ontología es

convertirse en parte de una estructura mucho más compleja, es decir, en parte de la estructura definida para el modelo global de representación de conocimiento planteado en la presente tesis doctoral, la cual es capaz de integrar las características percibidas mediante diferentes mecanismos sensoriales, expandiendo el conocimiento disponible para elementos previamente reconocidos y ayudando al reconocimiento de nuevos elementos.

Esto representa una alternativa completamente novedosa para el enfoque CBIR [486] en lo que concierne a la transformación de entradas visuales desde características de bajo nivel hasta información semántica de alto nivel. Un evaluador puede determinar si el objeto identificado es suficientemente relevante como para generar una nueva categoría, o es sólo una variante de un objeto reconocido previamente, que conforma un subconjunto de características para el mismo objeto (p. ej. Una manzana verde y una manzana roja). En cualquier caso, los resultados pueden ser almacenados o descartados (en caso de no ser suficientemente relevantes). Por otra parte, el proceso de aprendizaje puede ser adaptado, es decir, puede configurarse un aprendizaje supervisado, semi-supervisado o completamente automatizado (como en sistemas avanzados basados en *deep-learning*). La información semántica y la representación visual del objeto detectado son almacenadas en dos bases de datos diferentes. Por un lado, una base de imágenes temporal se ocupa de almacenar las correspondencias visuales más representativas de un grupo semántico de elementos (p. ej. el modelo visual de un gato) y, por otro lado, una base de datos global de conocimiento visual almacena las diferentes características utilizadas en la representación semántica de la imagen. Mientras se analicen nuevos prototipos visuales, la base de imágenes temporal podrá ser actualizada en consecuencia, a fin de almacenar el prototipo más relevante posible para cada categoría (y sub-categorías), o lo que es lo mismo, el centroide del clúster para cada caso. La base de conocimiento, por otra parte, almacenará las diferentes características aprendidas a lo largo del tiempo, creciendo en representatividad y capacidad de reconocimiento de elementos complejos, actuando conjuntamente con los módulos “analyzer de características de medio nivel” y “evaluador y clasificador” en un proceso de refuerzo “*top-down*” [22].

4.5.4. Otras entradas perceptuales – Beneficios del IoT

La visión, aunque importante, no es la única entrada perceptual disponible. Gracias al Internet de las cosas, y al abaratamiento de los sensores y su integración en diversas plataformas hardware (p. ej. *tablets* y *smartphones*), la capacidad potencial de “sentir” el entorno crece cada día. Sin embargo, no sólo los dispositivos de consumo presentan sensorización. También es posible desarrollar plataformas específicas dotadas de sensores, para cubrir necesidades de detección específicas. Un ejemplo al respecto puede verse en la Ilustración 30. Dicha ilustración muestra un Intel Edison (SoC) montado en la placa de desarrollo, a través de la que se proporciona conectividad con (de izquierda a derecha) un sensor de reconocimiento de vibraciones, un sensor táctil de contacto, un sensor de temperatura y un acelerómetro con detección en los tres ejes (X, Y, Z).



Ilustración 30. Intel Edison montado en placa de desarrollo, y conectado a sensores específicos

Sin embargo, existe una gran diversidad de sensores que pueden utilizarse para ampliar las capacidades perceptuales de un sistema. En la Ilustración 31 se muestra un pequeño subconjunto de sensores. Concretamente, de arriba hacia abajo y de izquierda hacia derecha (sobre las bolsitas verdes), un sensor de luminosidad, un sensor de sonido, un botón (táctil), un sensor de vibraciones, un sensor de contacto, un sensor de temperatura y un acelerómetro. Algunos otros sensores existentes comprenden giróscopos, módulos GPS, detectores de partículas en el aire, detectores de emisiones químicas, detectores de campo magnético y detectores de radiación y radiofrecuencia, entre otros.



Ilustración 31. Algunos sensores utilizables en plataformas de desarrollo

La integración de este tipo de sensores, representa de manera individual un nuevo tipo de sentido perceptual, que puede actuar en colaboración con otros, o de manera independiente. La integración de cada módulo supone el desarrollo de las interfaces correspondientes para transformar los datos de los sensores en información interpretable por el componente “distribuidor inteligente”, y la integración de modelos de reconocimiento específicos para tales sentidos en los PRAs. En la Ilustración 32 se puede apreciar una extensión lógica de la arquitectura para visión, presentada en la Ilustración 29, a fin de cubrir las necesidades de otras categorías sensoriales.

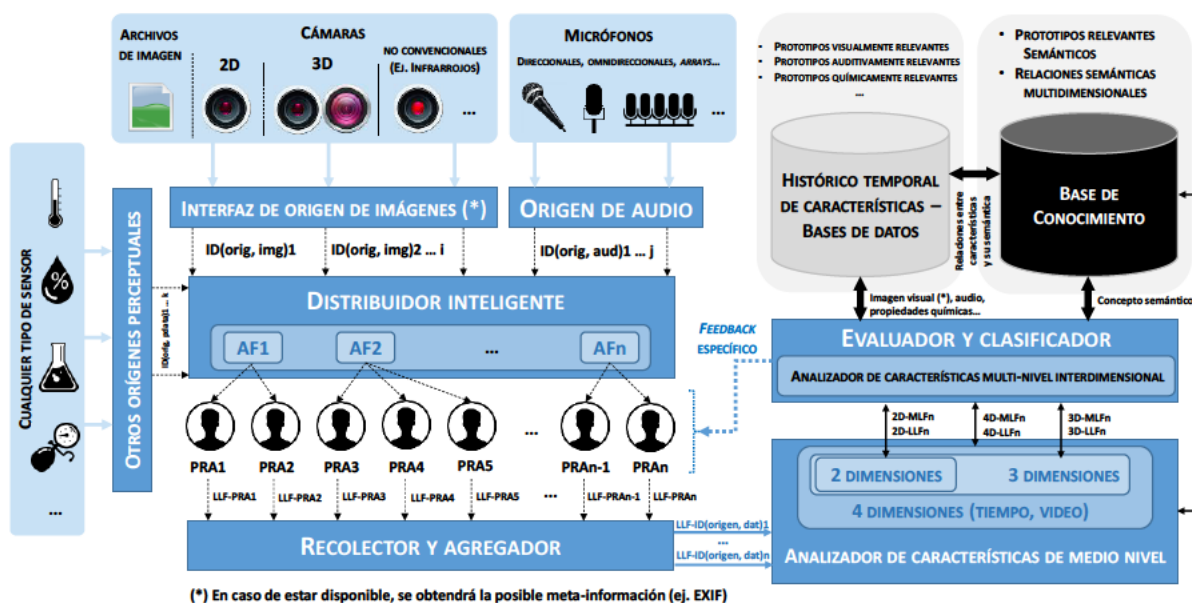


Ilustración 32. Arquitectura física del modelo para percepción multi-sensorial

Como puede apreciarse, la adición de sentidos adicionales implica una modificación mínima en la arquitectura, de carácter modular. La adaptación de los distintos módulos se realiza acorde a los nuevos sensores integrados, teniendo un impacto relativamente bajo en los existentes. Si bien es cierto que a priori se podría intuir que, en la base de conocimiento, y en el registro histórico temporal podría ser necesario crear estructuras adicionales para almacenar la información correspondiente a los nuevos sentidos, en realidad no es necesario, puesto que el uso de tecnologías de almacenamiento, basadas en paradigmas como *NoSQL* pueden resolver este problema de manera transparente. Para facilitar el intercambio y procesamiento de la información entre los distintos módulos, y en la base de conocimiento, se intuye una buena opción hacer uso de JSON, tras el estudio realizado al respecto. Además, el uso de ciertos sistemas de almacenamiento de datos, como *MongoDB* [465], codifican y trabajan con la información empleando JSON como modelo de representación, lo cual impulsa todavía más a utilizar dicho modelo de codificación de información.

La integración en el modelo global (citado en [487]) se realiza mediante la integración de la base de conocimiento que contiene la información semántica de los prototipos, variantes y sub-categorías, mediante el módulo “combinador de información perceptual” (ver Ilustración 27) del modelo global. Este módulo analiza cada una de las entradas sensoriales de forma conjunta, para proporcionar una representación semántica más compleja (p. ej. visual, olfativa, táctil, de temperatura, etc.) para cada elemento detectado, teniendo en cuenta el conocimiento global previamente almacenado. Igualmente comprueba si la pieza de información generada es válida o no para enriquecer la base global de conocimiento, que combina la información de cada sub-base de conocimiento correspondiente a los diferentes módulos perceptuales. Por ejemplo, para integrar los sub-sistemas de visión y sonido en el modelo general, podrían considerarse diferentes herramientas, como es el caso de SABUMO [488], pero extendiendo el vocabulario “controlado” del mismo, de forma que

nuevas características semánticas puedan ser integradas con facilidad en la base de representación de conceptos. Además, considerando la escalabilidad del sistema y la posibilidad de distribuir los módulos operativos entre diferentes dispositivos orientados a computación *cloud*, el uso de un lenguaje de interoperabilidad basado en tecnologías semánticas para la integración *PaaS* [489] ayudaría a distribuir la carga de trabajo entre los componentes computacionales de una forma automática e inteligente, de acuerdo a los requisitos algorítmicos y siguiendo los enfoques presentados en [490]. Otra opción factible, y posiblemente más sencilla de implementar sería la del uso de una arquitectura *cloud* orientada a microservicios, y el uso de contenedores (p. ej. Docker) con la implementación y despliegue específicos para cada módulo. Esto permitiría una escalabilidad horizontal enorme, pudiendo llegar a cubrir prácticamente cualquier tipo de demanda (suponiendo la existencia y accesibilidad a los recursos *cloud* necesarios) de forma automática, facilitando muchísimo las tareas de mantenimiento y actualización, y proporcionando una alta resiliencia y tolerancia ante fallos gracias a la gestión inteligente de los recursos. Finalmente, y respecto a la seguridad y fiabilidad, tanto de la información como del sistema en sí mismo, sería deseable que cada módulo perteneciente a la arquitectura validase mediante una firma la información recopilada o procesada, pudiendo verificar la integridad, y detectar entradas inválidas para evitar corrupción o manipulación intencionada. Igualmente sería deseable que todas las comunicaciones utilizarasen los estándares de seguridad de más alto nivel asumibles por cada componente, manteniendo una coherencia y equilibrio entre seguridad, rendimiento y comunicación. Un sistema de clave pública con certificados y conexiones entre los módulos con TLS podría ser una solución viable teniendo en cuenta el marco tecnológico actual.

4.6. Conclusiones del capítulo

En este capítulo se han propuesto una serie de Principios Fundamentales y Características Técnicas de diseño esenciales, que debe presentar un modelo computacional de representación del conocimiento basado en la percepción. Para ello se han realizado dos recapitulaciones. Por un lado, sobre las distintas características y requisitos observados en el área de la psicología y neurociencia relacionados con la cognición. Por el otro, sobre las distintas características y requisitos observados en el marco tecnológico. Tras la identificación y definición de características, así como de los requisitos asociados, se ha definido una aproximación abstracta y generalista del modelo biológico de representación del conocimiento a la computación, mediante el planteamiento, tanto de un patrón de análisis sensorial, como de una arquitectura abierta, modular y extensible, presentada por niveles (de general, marco global, a particular, considerando los sentidos), capaz de cubrir los requisitos específicos del modelo. La evolución del capítulo queda reflejada en la Ilustración 33.

El desarrollo de este nuevo modelo está orientado a proporcionar un enfoque novedoso, modular, extensible y completo del procesamiento de información, considerando diferentes representaciones y dimensionalidad (imágenes 2D, imágenes 3D, vídeos, secuencias de audio, mediciones de temperatura, inclinación,

coordenadas GPS y demás), por medio de la aplicación paralela de diferentes algoritmos existentes (o venideros) así como sean requeridos, en un amplio conjunto de entornos computacionales, soportados por una arquitectura basada en la nube.

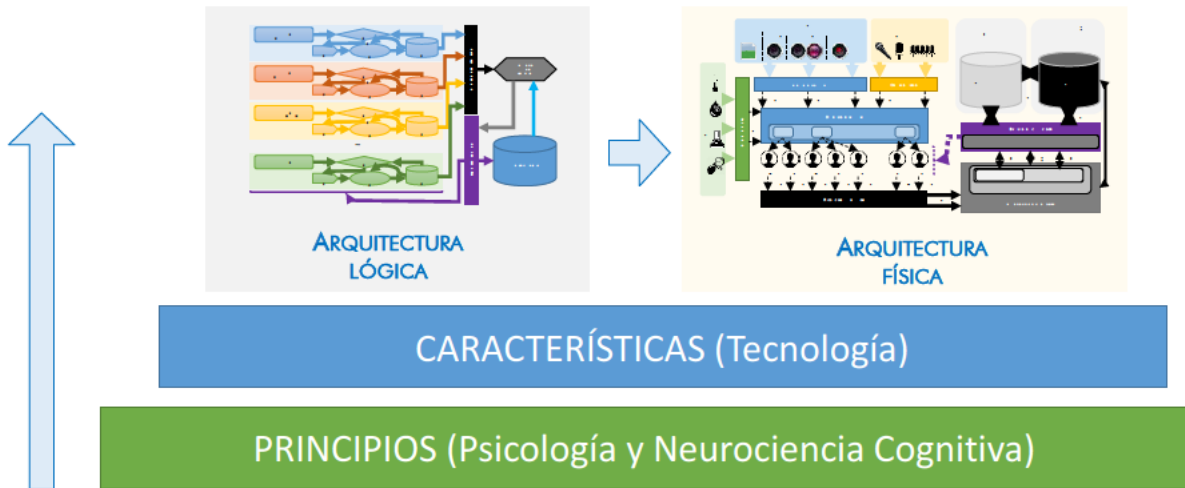


Ilustración 33. Esquema de flujo del desarrollo del marco de referencia

La motivación tras el modelo es la de emular la forma en la que los seres humanos almacenan la información, lo cual facilita el aprendizaje, a fin de desarrollar un nuevo espectro de aplicaciones basadas en un sistema inteligente capaz de llegar a un aprendizaje autónomo desde el entorno que le rodea, asimilando el conocimiento existente en todas sus formas y combinándolo en una base de conocimiento global y abstracta, para desarrollar nuevos conceptos, ideas y teoremas en un entorno eficiente y escalable, expandiendo las capacidades de las ontologías al representar múltiples dominios pero relacionados entre sí (información recogida de los sistemas sensoriales).

Capítulo 5. Validación

5.1. Introducción

Una vez estudiadas las diferentes características y requisitos necesarios para la creación de un modelo de representación basado en el conocimiento, el siguiente paso lógico consiste en probar la validez del mismo. A lo largo de este capítulo se desarrolla el proceso de validación, comenzando con la validación teórica de los Principios Fundamentales y Características Técnicas de diseño observadas en el capítulo anterior. Se prosigue realizando una implementación incremental de la arquitectura propuesta, a fin de probar su validez y evaluar su versatilidad y eficiencia reconociendo elementos. Finalmente, se analizan los resultados a fin de validar las hipótesis y el objetivo propuestos en la tesis.

5.2. Validación de la solución

5.2.1. Validación teórica. Principios Fundamentales y Características Técnicas de diseño

Tanto el desarrollo de los Principios Fundamentales inferidos de las características de la percepción, el conocimiento y la memoria, dentro del marco cognitivo, como la formulación de las Características Tecnológicas de diseño propuestas siguieron un proceso iterativo de refinamiento. En dicho proceso se contó con la ayuda de dos expertos en sus respectivas áreas. En el caso de los PFs, se contó con la ayuda de D. Roque Viñuelas Paredes, Licenciado en Psicología. Para las CTs, la asistencia fue proporcionada por el Prof. Dr. D. Alejandro Calderón Mateos, Doctor en Ciencias de la Computación.

El proceso de creación de los PFs y CTs siguió el siguiente flujo de desarrollo:

1. Análisis de los estudios del estado del arte correspondientes al capítulo 2 y al capítulo 3.
2. Síntesis de las características más significativas del capítulo 2 en forma de PFs.
3. Entrevistas individuales periódicas con el experto en el área de Psicología a fin de refinar y validar los PFs definidos.
4. Síntesis de las características más significativas del capítulo 3 y adaptación de los PFs definidos al marco tecnológico en base a dichas características. Resultado, lista de CTs.
5. Entrevistas individuales periódicas con el experto en el área de Computación a fin de refinar y validar los CTs definidos.

6. Mesa redonda con los dos expertos a fin de encontrar algún PF o CT no definido, o de separar alguno complejo en varios más simples.

El flujo de desarrollo puede apreciarse en la Ilustración 34.

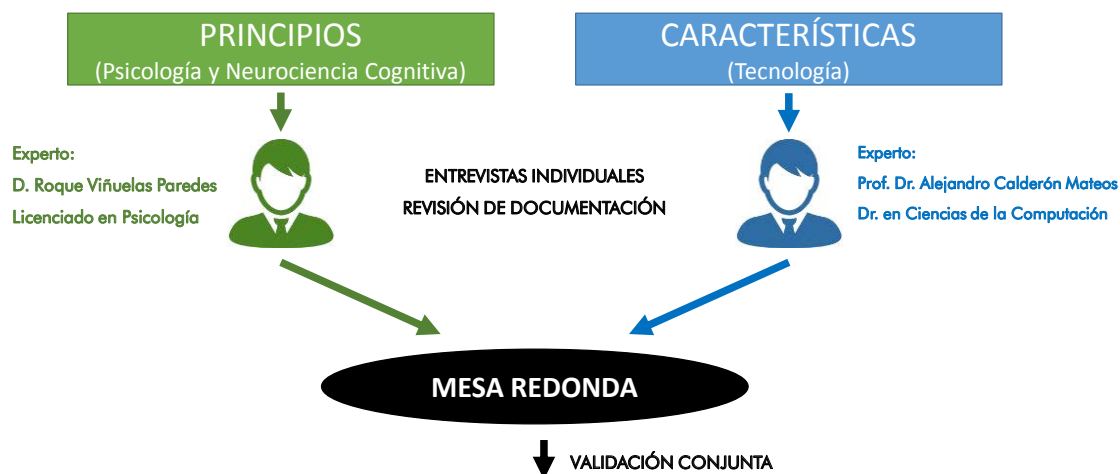


Ilustración 34. Validación de los conceptos teóricos

Como resultado del proceso de validación teórica se obtuvieron una serie de beneficios, entre ellos:

- Catalogación de PFs y CTs con ayuda de los expertos.
- Verificación de las relaciones entre unos y otros.
- Definición de un nuevo PF (PF-78).
- Aceptación de PFs y CTs.

5.2.2. Validación empírica, o experimental. Arquitectura del modelo

Con el objetivo de comprobar la viabilidad y capacidades del modelo propuesto, se ha llevado a cabo una sucesión de implementaciones incrementales del mismo, creciendo en complejidad y características disponibles. La versión inicial y simplificada del modelo, incluye un distribuidor de carga inteligente, agentes inteligentes de reconocimiento de patrones y un módulo recolector y agregador (ver Ilustración 35). El objetivo de esta primera implementación es comprobar la factibilidad de la solución al aplicarse sobre un área específica de la percepción. Concretamente, la visión. Otro objetivo principal, es evaluar el comportamiento del modelo en un entorno de computación distribuida, evaluando el comportamiento en base a la variación del número de instancias computacionales. Los agentes inteligentes de reconocimiento de patrones han sido programados para identificar instancias de objetos muy simples y sus características formantes (características de bajo nivel), siendo más específicos, identificando formas cuadradas en archivos de imágenes bidimensionales,

proporcionando la posición central de dichos cuadrados (coordenadas los píxeles), la longitud de sus lados (en píxeles) y su color individual en formato 24-bits.

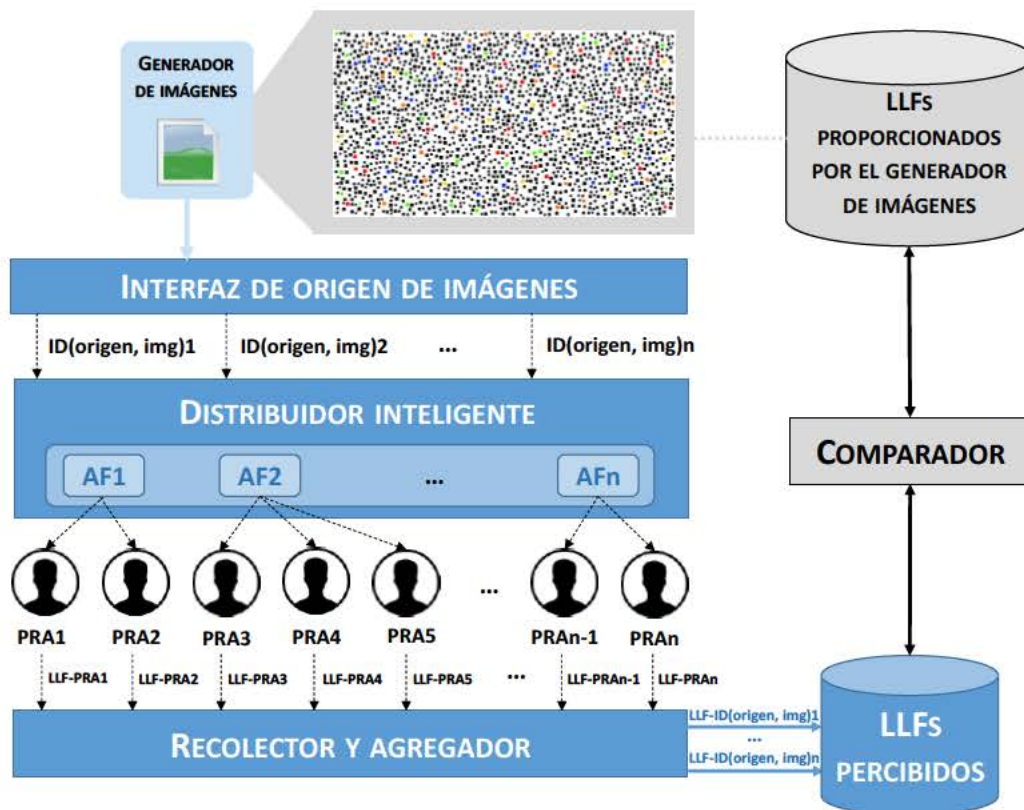


Ilustración 35. Parte implementada del modelo con fines de prueba

A fin de maximizar la eficiencia de los algoritmos distribuidos y simplificar el desarrollo para esta versión inicial del modelo, la implementación del sistema se ha realizado en el lenguaje de programación C, haciendo uso de la biblioteca MPI [491]. Para simplificar el desarrollo, el algoritmo ejecutado por los agentes inteligentes en paralelo ha sido diseñado bajo el paradigma del enfoque por fuerza bruta, pero podría ser fácilmente reemplazado por redes neuronales entrenadas para reconocer patrones específicos (cuadrados, en el test propuesto. Ver Ilustración 36).

El distribuidor inteligente es responsable de dividir y distribuir la imagen recibida a los agentes paralelos, los cuales se ejecutan en diferentes servidores interconectados en la misma subred, por medio de mensajes MPI.

El algoritmo general para la distribución y el reconocimiento de patrones ha sido adaptado desde una versión inicial secuencial a una versión paralela optimizada, siguiendo un enfoque clásico en cuatro fases (descomposición, asignación, orquestación y mapeo) [492], y se puede describir como sigue:

1. El proceso inicial (proceso 0) lee la imagen especificada como argumento, aunque es posible su invocación como una función de un sistema más grande y general, que podría abastecer de imágenes recuperadas de sitios web (*spider, crawler*) o incluso desde cámaras digitales.

2. El proceso 0 distribuye la imagen entre los otros procesos (agentes) que se encuentran en ejecución en los diferentes servidores del *grid*.
3. El proceso 0 comprueba el tamaño de partición a ser evaluado (tamaño del grano) para cada uno de los agentes, y les envía dicha información.
4. Los procesos distribuidos (agentes) ejecutan el algoritmo de reconocimiento de patrones en su fragmento específico de la imagen. Tras finalizar el reconocimiento de los elementos, cada proceso distribuido envía la información correspondiente a los objetos reconocidos al proceso inicial.
5. El proceso inicial (0) combina toda la información semántica de bajo nivel (LLFs) recibida de los demás agentes y lo almacena en un fichero.

Para este evaluador, en lugar de procesar los distintos archivos de imagen a la vez, se ha decidido procesar un archivo de imagen por diferentes agentes paralelos al mismo tiempo. Esto puede ser ajustado a fin de evaluar múltiples archivos a la vez por diferentes agentes, o realizar un enfoque híbrido: múltiples agentes para múltiples imágenes, donde subconjuntos de agentes pueden trabajar en la misma imagen. Al final, el recolector y agregador generarán tantos archivos como imágenes hayan sido analizadas, conteniendo la información semántica de bajo nivel en cada uno de ellos.

Otro programa ha sido desarrollado a fin de generar automáticamente las imágenes de prueba (ver Ilustración 36). Dichas imágenes contienen cuadrados de distintos tamaños y colores siguiendo una distribución aleatoria homogénea o geométrica a lo largo de la superficie del lienzo. Este programa genera también un archivo con información similar a la que se espera recibir del proceso inicial tras el reconocimiento (LLFs de la imagen), a fin de comprobar la precisión del algoritmo.

Para la prueba se generaron diferentes imágenes con dimensiones fijas de 1920x1080 píxeles (resolución *FullHD*) y 1976 centroides (Ver Ilustración 36). A fin de simplificar la ejecución del algoritmo, no se solaparon cuadrados ni parcial ni totalmente. En caso de aplicar algoritmos más avanzados de reconocimiento de patrones, como redes Hopfield entrenadas [493], algunas de esas restricciones podrían evitarse.

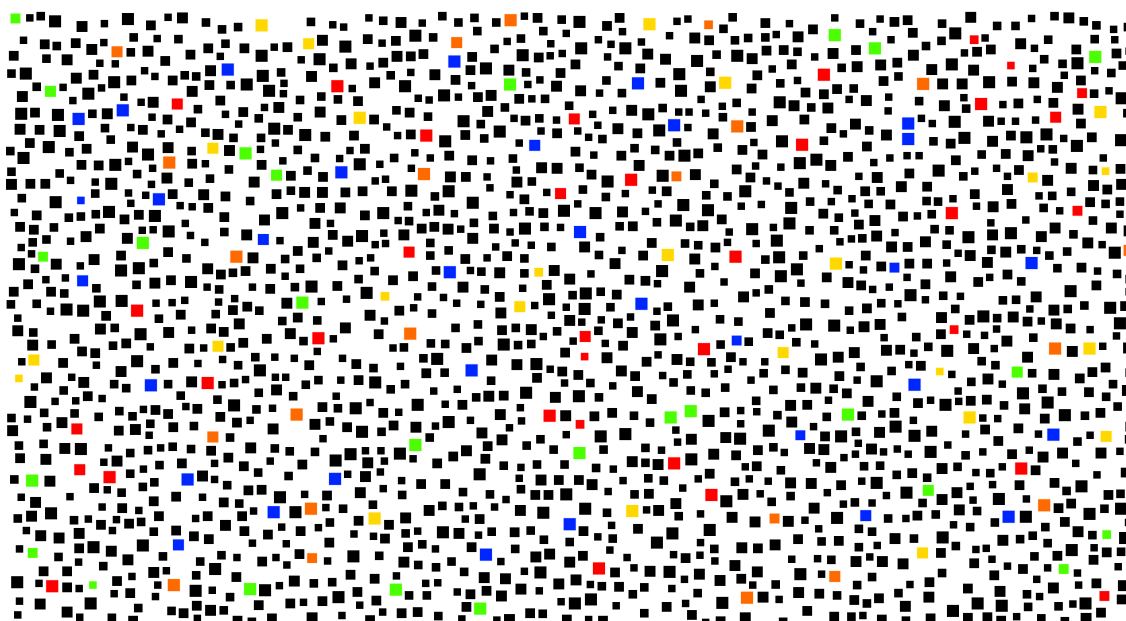


Ilustración 36. Ejemplo de imagen con una distribución uniforme de cuadrados

Las imágenes generadas pueden ser posteriormente sustituidas por conjuntos de entrenamiento reales, pertenecientes a múltiples escenarios de aplicación (p. ej. sector sanitario, o sector de seguridad de la información).

Los procesos distribuidos encargados de analizar las fuentes de imagen pueden ser considerados agentes inteligentes independientes especializados en detección de patrones concretos, y la información inferida por ellos es posteriormente evaluada y combinada por el proceso receptor (inicial). Para este caso específico, este proceso principal únicamente recolecta la información, la combina, la almacena en un archivo y comprueba su validez respecto a la solución proporcionada por la aplicación generadora de imágenes de prueba. En la versión final del modelo, mediante la implementación de las interfaces adecuadas, el recolector se comportará como se ha descrito anteriormente, proporcionando valiosa información para llevar a cabo posteriores tareas avanzadas de clasificación.

El proceso de validación fue llevado a cabo en un clúster compuesto por doce equipos dotados con un Intel Pentium IV, conectados entre sí a través de una red Gigabit Ethernet. Las pruebas se realizaron incrementando el número de agentes, comenzando con un único proceso y acabando con doce (uno por cada equipo de la red). Respecto a los tiempos de ejecución, procesamiento y comunicación de los agentes distribuidos del sistema, es posible apreciar un incremento del *speedup* en el algoritmo cuantos más agentes son utilizados. La tabla inferior muestra los resultados promedio obtenidos tras el proceso de validación.

NÚMERO DE AGENTES DISTRIBUIDOS	TIEMPO MEDIO (EN SEGUNDOS)			SPEEDUP
	PROCESAMIENTO	COMUNICACIÓN	TOTAL	
1	11.583594	0.000000	11.584	---
4	4.186458	0.244535	4.431	2.614
8	2.034395	0.347599	2.382	4.863
12	1.448336	0.410354	1.859	6.232

Tabla 106. Resultados promedio de la ejecución en el procesamiento de imágenes individuales

Por otra parte, comparando los resultados del análisis con la solución proporcionada por el generador de imágenes de prueba, se revela que la precisión es muy cercana al 100%. Posiblemente los pequeños desajustes sean resultado del algoritmo de particionado, que podría ser extendido para considerar el análisis de la vecindad correspondiente a las fronteras de particionado de la imagen, a fin de detectar posibles cuadrados divididos en el proceso.

En cualquier caso, como se comentó anteriormente, el modelo soporta el ajuste y cambio de los algoritmos ejecutados por los agentes, siendo posible aprovechar las ventajas ofrecidas por hardware específico, como las GPUs y MPUs, expandir las capacidades de reconocimiento de patrones, reducir los tiempos de detección e incrementar la variedad (complejidad) de los elementos analizados. Dada la escalabilidad del sistema propuesto, podría ser posible utilizarlo para detectar patrones en vídeo de alta resolución tras ajustar los parámetros necesarios, tales como los algoritmos de reconocimiento empleados y la estrategia de distribución carga de trabajo paralela.

La segunda iteración en el proceso de implementación del modelo se centró en agregar características y funcionalidades, cubriendo la mayor parte de aspectos generales, comentados en la definición del modelo teórico, y aplicando tecnologías relacionadas con el estado del arte en el ámbito tecnológico. En este caso, en lugar de analizar la factibilidad y desempeño en entornos paralelos, el objetivo fue el de determinar las capacidades representativas del modelo, el factor de mejora de reconocimiento de elementos frente a soluciones existentes, y comprobar la aplicabilidad en entornos basados en computación en la nube, y escalables. Debido al incremento de la complejidad del sistema, se decidió optar por el uso del sistema de control de versiones Git [494], mediante una serie de repositorios privados gestionados en un servidor local a través de la herramienta de código abierto GitLab [495].

El primer cambio significativo en la nueva implementación, fue la migración de un sistema paralelo basado en el uso de la biblioteca MPI y codificado en lenguaje C, a un sistema distribuido en red, eliminando las restricciones de lenguaje de programación y arquitectura computacional. Para eliminar dichas restricciones, se decidió orientar el sistema hacia una arquitectura de microservicios, empleando un sistema de contenedores basado en Docker. El escenario de prueba en este caso,

consistió en el reconocimiento de cuatro tipos de cítrico: pomelos, naranjas, limones y limas (ver Ilustración 37).



Ilustración 37. Elementos prototípicos para cada categoría del escenario II

Se desarrollaron una serie de imágenes para Docker (*docker-images* de ahora en adelante, para evitar la confusión con la representación de una fotografía en un formato digital, o imagen) desde las que instanciar sus respectivos contenedores, cada una de ellas encargada de una o varias tareas específicas, correspondientes a un subconjunto de los elementos mostrados en la Ilustración 29. Más específicamente:

- *Docker-image-1*: Entrada de imagen en 2D, transformación (conversión y redimensión) de la misma en base a los parámetros especificados, codificación de la misma junto con otros parámetros significativos (como el id asociado a la entrada) a código en JSON, y envío del código JSON generado al contenedor encargado de determinar el tipo de elemento reconocido.
- *Docker-image-2*: Recogida de entradas perceptuales, aplicación de una red de neuronas con retropropagación previamente entrenada para reconocer un tipo específico de elemento basado en su representación visual en 2D, generación de la salida, codificada en JSON, junto con otros parámetros significativos (id, propagación del elemento de entrada), y envío del código JSON generado al contenedor encargado de combinar la información relativa a otros reconocedores, y emitir un veredicto sobre el elemento reconocido.
- *Docker-image-3*: Recogida de resultado de reconocedores, recuperación de candidatos de la base de conocimiento en base a sus características esenciales, emisión de veredicto sobre identidad de elemento reconocido.

Para *Docker-image-1* se codificó una sencilla aplicación en el lenguaje de programación Java (versión 8), que a través de los parámetros definidos (ruta a un archivo correspondiente a una fotografía de imagen, dimensiones de imagen a enviar, identificador de origen y dirección IP del siguiente contenedor), abre la imagen a reconocer, la convierte en escala de grises, la redimensiona en base a las dimensiones especificadas, genera una estructura codificada en JSON con el identificador de origen, la fecha y hora, y la fotografía codificada en base64. Finalmente, se envía la estructura en JSON al siguiente contenedor.

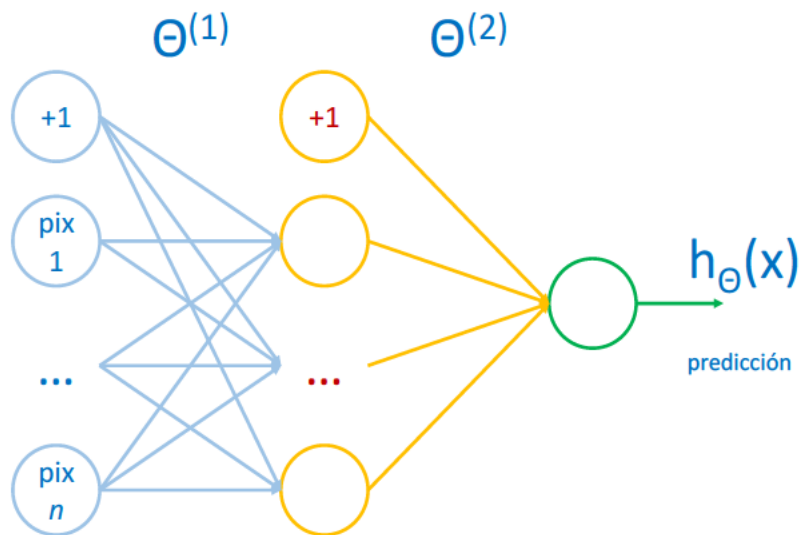


Ilustración 38. Ejemplo de red de neuronas

Para *Docker-image-2* se definieron cuatro variantes, correspondientes a la misma implementación de una red de neuronas con retropropagación (BPNN, *backpropagation neural network*), en lenguaje Java (versión 8), dotada de tres niveles (uno de entrada, uno oculto y uno de salida, ver Ilustración 38), donde las entradas se corresponden a los píxeles de la imagen. La red de neuronas admite imágenes de 30x30 píxeles en escala de grises (por motivos de complejidad computacional), lo que genera una capa de entrada con 900 unidades (excluyendo la unidad *bias*, que siempre devuelve 1). La red de neuronas cuenta con una única unidad de salida que devuelve el porcentaje correspondiente al reconocimiento de la entrada, respecto al entrenamiento de la red. Cada una de las cuatro BPNNs se entrenó específicamente mediante un conjunto de cien imágenes por categoría (pomelos, naranjas, limones y limas), recopiladas y ajustadas manualmente, a los requisitos de la red (dimensiones de 30x30 píxeles y escala de grises). Parte de cada conjunto de entrenamiento utilizado en las diferentes BPNNs puede observarse en la Ilustración 39.

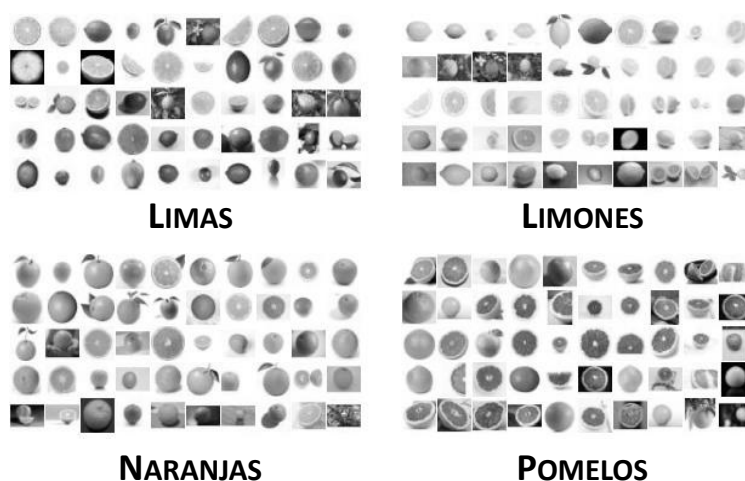


Ilustración 39. Parte del conjunto de entrenamiento para cada BPNN

Una vez evaluada la imagen de entrada, se codifica el resultado estadístico devuelto por la BPNN correspondiente a la *Docker-image-2*, junto con el id del elemento en el que está especializada, el id de origen de la imagen, la imagen en base64 (tal cual entró en el contenedor) y el id de instancia del reconocedor específico en una estructura JSON. Dicha estructura es enviada al siguiente contenedor, ya que su dirección IP es conocida previamente y pasada por parámetro.

Docker-image-3 contiene la implementación en Java (versión 8) del módulo “evaluador y clasificador” mostrado en la Ilustración 29 e Ilustración 32. En esta versión implementada del modelo, es esta imagen la que recibe los diferentes resultados correspondientes al análisis y/o normalización de las entradas perceptuales, y determina un veredicto sobre el tipo de elemento correspondiente a dichas entradas. En este caso, la imagen almacena los identificadores asociados a los reconocedores, y las etiquetas correspondientes al elemento en el que han sido entrenados. El resultado es la etiqueta correspondiente al elemento representado en la entrada que tenga mayor valor resultado de la BPNN.

Tras realizar el ensayo con 15 imágenes de cada categoría, donde cada elemento de entrada es conocido a priori, se observaron los siguientes resultados:

ENTIDADES	CATEGORÍA	SALIDA	CANTIDAD
15	Naranja	Naranja	11
		Pomelo	4
		Limón	0
		Lima	0
15	Pomelo	Naranja	2
		Pomelo	13
		Limón	0
		Lima	0
15	Limón	Naranja	0
		Pomelo	0
		Limón	12
		Lima	3
15	Lima	Naranja	0
		Pomelo	0
		Limón	2
		Lima	13

Tabla 107. Resultados de reconocimiento al aplicar la segunda implementación del modelo

Como se aprecia en la Tabla 107, se obtuvieron errores de reconocimiento en determinadas instancias de categorías visualmente similares. Aparentemente parece un problema relacionado con la percepción visual, identificado en el apartado 2.2.2, el problema “no es suficiente” (información sensorial). Veremos cómo se solucionó el problema, con la siguiente iteración en la implementación del modelo.

La última implementación del modelo, realizada en el proceso de validación, incorporó un mayor número de funcionalidades, entre las que cabe destacar la inclusión de datos desde interfaces de captura de imagen (webcams) conectadas a dispositivos Raspberry Pi 3, la adaptación de las *docker-images* existentes, e implementación de otras nuevas para tratar la información relativa a otras entradas perceptuales, la adición de un sistema gestor de base de datos (MongoDB) como contenedor de Docker, a fin de realizar las labores de base de conocimiento global y la extensión de la definición de las estructuras JSON correspondientes a los diferentes mensajes y piezas de conocimiento correspondientes al modelo, incluyendo aspectos como la seguridad y veracidad de la información.

En primer lugar, se desarrollaron una serie de extensiones del *Docker-image-1*. El objetivo de dichas extensiones fueron las siguientes:

- Versión A: obtener imagen bidimensional del objeto a reconocer mediante una cámara web y aplicar proceso análogo a la versión anterior (conversión, redimensión, codificación, envío).
- Versión B: obtener imagen desde dos cámaras web pre-calibradas, y mediante un conjunto de bibliotecas externas (*OpenCV* [496]) obtener las medidas del elemento a reconocer. Codificar la información en una estructura JSON y enviarla al siguiente contenedor.
- Versión C: obtener el valor químico del PH del elemento a reconocer. Codificar la información en una estructura JSON y enviarla al siguiente contenedor.

Las versiones A y B se desarrollaron en el lenguaje de programación Python, y se orientaron a la ejecución en sistemas de la familia *Raspberry Pi*. Tras las pruebas iniciales con la versión B (ver Ilustración 40) se obtuvieron buenos resultados de medida, pero debido a la dificultad para calibrar las cámaras, llevada a cabo mediante un sistema manual (por falta de recursos), se decidió adaptar el módulo B para recibir las medidas de forma manual. Nótese que para el correcto funcionamiento del sistema de reconocimiento fue necesario definir una entrada en la base de conocimiento relativa a las características dimensionales (medidas: mínima, máxima, promedio, desviación típica y varianza) para cada categoría de elemento, en base a las mediciones tomadas manualmente sobre un conjunto representativo de candidatos de dichas categorías.



Ilustración 40. *Raspberry Pi* conectada a dos cámaras iguales (visión estéreo)

Respecto a la versión C, tras buscar sensores adecuados (como [497]) para determinar de manera digital el pH de las frutas a reconocer, se observó que debido a su elevado precio, sería conveniente utilizar otra solución. Concretamente se optó por el uso de tiras de papel indicador de pH tornasol (baratas y fáciles de conseguir). Sin embargo, tras tomar las primeras medidas, se observó que, debido al grado de acidez tan similar de los cítricos, resultaba muy complicado discernir sobre un valor concreto en base al color obtenido. La solución final consistió en la generación de valores aleatorios típicos para cada categoría, que fueron introducidos de manera manual. Para la obtención de los valores de referencia correspondientes a cada categoría, se accedió al listado proporcionado por la *U.S. Food and Drugs Administration* [498] (ver Tabla 108).

CATEGORÍA	ACIDEZ (PH)
Naranja	3.1 - 4.1
Pomelo	3.0 - 3.3
Limón	2.2 - 2.4
Lima	1.8 - 2.0

Tabla 108. Rangos de acidez promedio para categorías cítricas (fuente: FDA [498])

Seguidamente se desarrolló una *docker-image* correspondiente al módulo “analyzer de características de medio nivel” (ver Ilustración 32), también codificado en Java (versión 8). Dicho módulo recibe entradas en formato JSON de diferentes contenedores reconocedores, extrae las características recogidas, y busca en el modelo de conocimiento contenido en una instancia de MongoDB [465], los diferentes elementos que contengan una característica cuyos rangos contengan el valor del elemento reconocido (p. ej. pH de entrada 3.2, entidades asociadas, id correspondiente a pomelo e id correspondiente a naranja). Una vez obtenidos los diferentes elementos potenciales, se genera una estructura en JSON codificando la información sobre los mismos, y se envía al siguiente contenedor, que hace las tareas de evaluador y clasificador (*Docker-image-3*).

La extensión sobre *Docker-image-3* fue realizada para dar soporte a la evaluación en base a múltiples entradas perceptuales. El contenedor recibe información en JSON sobre los identificadores más prometedores que definen a la categoría del elemento analizado. Se realiza un proceso de clasificación a fin de estimar cuál de los identificadores tiene un mayor número de ocurrencias, y cuál es el valor asociado a las distintas probabilidades de predicción por cada entrada sensorial para dicho identificador. Una vez determinado, se accede a la base de conocimiento y se busca la etiqueta correspondiente al identificador. Se realiza una agregación de las distintas características recogidas en el proceso, y transmitidas por los diferentes contenedores, se codifican en una estructura JSON, preservando el formato recogido, y se devuelven como resultado del proceso, junto con el identificador y su etiqueta asociada.

El *pipeline*, o proceso de identificación correspondiente a la implementación final del modelo y de los sistemas perceptuales de entrada se puede observar en la Ilustración 41. Nótese que en el caso del contenedor correspondiente al módulo de la visión 2D, el cual utiliza una BPNN no comparte su resultado con el “analizador de características de medio nivel”. Esto es consecuencia de que las BPNN almacenan las características detectadas de manera automática y abstracta en sus niveles ocultos, por lo que se considera el resultado final de la misma como entrada a ser procesada por el módulo “clasificador y evaluador”. Además, aunque se muestre un único módulo, correspondiente a una BPNN entrenada para reconocer naranjas, en realidad la imagen bidimensional se está procesando en cuatro módulos paralelos, con cuatro redes de neuronas independientes y de características similares, pero cada una con un entrenamiento específico. A fin de simplificar la representación en la ilustración sólo se ha reflejado un módulo. Para el resto de características, (dimensiones y niveles de pH) los valores prototípicos y sus respectivos rangos fueron definidos manualmente e introducidos previamente en la base de conocimiento global.

Cabe destacar que, para proporcionar una mayor seguridad al sistema, en los procesos de comunicación entre contenedores, se aplicó el protocolo TLS, utilizando certificados auto-firmados, y forjados específicamente para el caso. Igualmente, se extendió el número de campos correspondientes a las estructuras JSON para introducir, entre otros, una propiedad denominada “signature” o firma, correspondiente a la firma digital del contenido recogido por los contenedores, junto con la marca temporal. La idea que motivó la aplicación de esta mejora es la prevención de falsificación de entradas en caso de comprometer la red mediante la que se comunica el sistema, al ser posible determinar si el contenido de la estructura JSON relativo a la información perceptual es válido o no.

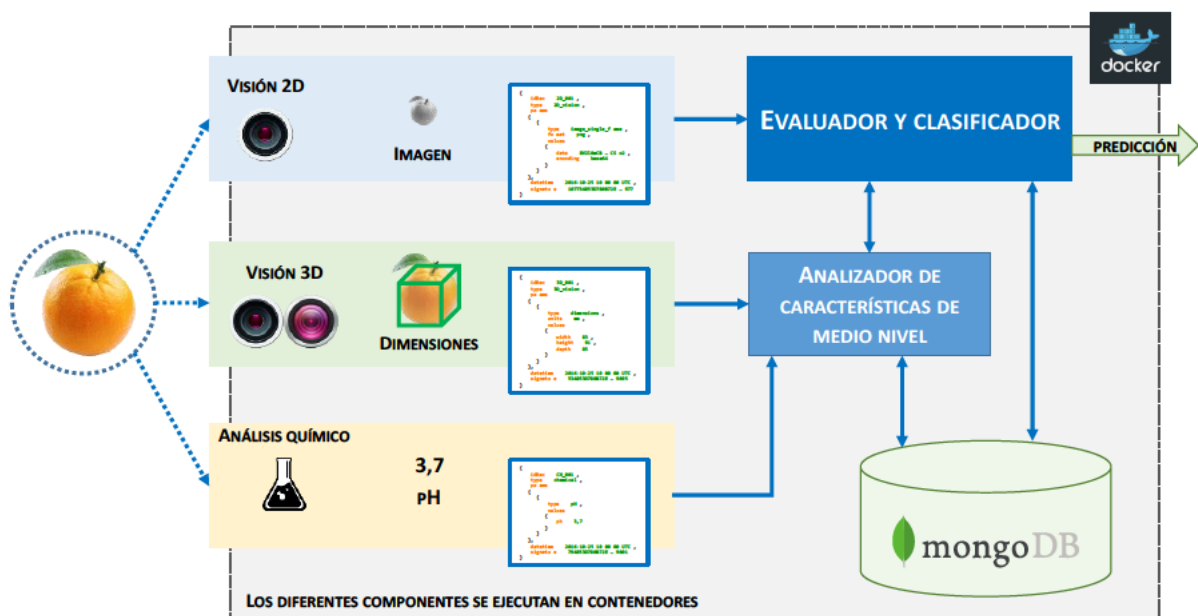


Ilustración 41. Pipeline de reconocimiento utilizando el modelo propuesto

Tras repetir el análisis con otras 15 entidades de cada categoría, se obtuvieron los resultados mostrados en la Tabla 109

ENTIDADES	CATEGORÍA	SALIDA	CANTIDAD
15	Naranja	Naranja	15
		Pomelo	0
		Limón	0
		Lima	0
15	Pomelo	Naranja	0
		Pomelo	15
		Limón	0
		Lima	0
15	Limón	Naranja	0
		Pomelo	0
		Limón	15
		Lima	0
15	Lima	Naranja	0
		Pomelo	0
		Limón	0
		Lima	15

Tabla 109. Resultados de reconocimiento al aplicar la implementación final del modelo

Como puede apreciarse, el resultado del reconocimiento, haciendo uso del modelo de representación de conocimiento basado en múltiples entradas perceptuales, parece haber mitigado el efecto del problema “no es suficiente”, observado en la implementación anterior, consiguiendo un índice impecable de aciertos en la predicción de categoría.

5.3. Análisis de resultados. Validación de hipótesis y objetivo

En base a los resultados del proceso de validación, tanto teóricos como experimentales, se puede proceder a la evaluación de las hipótesis, y objetivo asociado, definidos en el apartado 1.3.

Sub-Hipótesis H₁	DESCRIPCIÓN	LA INTEGRACIÓN DE ESPACIOS DE CONOCIMIENTO CORRESPONDIENTES A DIFERENTES ÁREAS SE VERÁ FACILITADA MEDIANTE LA ADOPCIÓN DE UN MODELO DE REPRESENTACIÓN DEL CONOCIMIENTO BASADO EN LA PERCEPCIÓN.
	EVALUACIÓN	En base a los Principios Fundamentales identificados y las características asociadas a los mismos, se puede afirmar que la adopción de un modelo de representación del conocimiento inspirado en el funcionamiento del subconjunto cognitivo percepción-conocimiento-memoria del ser humano facilitará la integración de espacios de conocimiento correspondientes a diferentes áreas. Una de las claves para conseguir dicho objetivo es la combinación simultánea de varios modelos de representación diferentes, con características específicas que en su combinación proporcionan mayor complejidad semántica a la entidad representada.

Tabla 110. Validación de la sub-hipótesis H₁

Sub-Hipótesis H₂	DESCRIPCIÓN	ES POSIBLE MEJORAR LA CALIDAD DEL APRENDIZAJE AUTOMÁTICO AL REDUCIR LA REDUNDANCIA DE TÉRMINOS Y CONCEPTOS ALMACENADOS EN UNA BASE DE CONOCIMIENTO GLOBAL, MEDIANTE LA APLICACIÓN DE UN MODELO DE REPRESENTACIÓN DEL CONOCIMIENTO BASADO EN LA PERCEPCIÓN.
	EVALUACIÓN	Como se ha observado en capítulos anteriores, varios de los modelos de representación que actúan en el proceso cognitivo del ser humano se ocupan de reducir la redundancia en el almacenamiento del conocimiento mediante la representación de los elementos a través de prototipos, categorías y sub-categorías. Sí bien es cierto que la redundancia en las entradas sensoriales refuerza el proceso de aprendizaje y definición de los conceptos, a nivel biológico, dicha redundancia no se almacena, principalmente por motivos de capacidad. En el modelo propuesto se sugiere mantener una base temporal (histórica) para almacenar elementos representativos reconocidos, facilitando la re-evaluación del conocimiento almacenado en caso de detectar nuevas sub-categorías.

Tabla 111. Validación de la sub-hipótesis H₂

Sub-Hipótesis H₃	DESCRIPCIÓN	EL TRATAMIENTO SEMÁNTICO DE LA INFORMACIÓN RESULTARÁ MÁS PRECISO AL APLICAR DE UN MODELO DE REPRESENTACIÓN DEL CONOCIMIENTO BASADO EN LA PERCEPCIÓN.
	EVALUACIÓN	Como se ha podido comprobar en la última fase de experimentación, la aplicación del modelo de representación de conocimiento basado en múltiples características perceptuales, o para ser más específico, en reconocimiento visual de imágenes 2D, reconocimiento de dimensiones en 3D y reconocimiento de índice de pH ha conseguido una tasa de éxito absoluta frente al uso de modelos que consideran únicamente un tipo perceptual (BPNN, correspondiente al experimento anterior).

Tabla 112. Validación de la sub-hipótesis H₃

Sub-Hipótesis H₄	DESCRIPCIÓN	SERÁ POSIBLE ALCANZAR MAYOR PRECISIÓN EN LA RECUPERACIÓN DE INFORMACIÓN AL CONTENER ÉSTA MAYOR DIVERSIDAD EN LA NATURALEZA DE LOS DATOS QUE FORMAN SU DEFINICIÓN (MÁS DESCRIPTIVA). EL TRATAMIENTO DE LA MISMA SERÁ MÁS COMPLETO Y LA GENERACIÓN DE NUEVOS CONCEPTOS O RELACIONES ENTRE LOS MISMOS PODRÁ ADECUARSE A MODELOS DE RAZONAMIENTO SIMILARES A LOS HUMANOS. TODO GRACIAS A LA APLICACIÓN DE UN MODELO DE REPRESENTACIÓN DEL CONOCIMIENTO BASADO EN LA PERCEPCIÓN.
	EVALUACIÓN	Según se aprecia en el estudio sobre el área de la psicología y neurociencia cognitiva, dos áreas no tratadas en la presente tesis, pero directamente relacionadas con la simbiosis percepción-conocimiento-memoria son la inteligencia y el razonamiento. Según dichos estudios, se ven fuertemente influenciadas y son dependientes de las tres áreas anteriores, y completan el marco cognitivo tal cual está definido para el ser humano. Debido a la complejidad del trabajo y de la gran extensión del mismo al tratarse de una aproximación holística a un modelo de representación de conocimiento, no se ha llegado a implementar un modelo de razonamiento o inferencia avanzado que se beneficie del modelo propuesto. Si bien, los éxitos experimentales preliminares del modelo, junto con la validación teórica de los PFs y CTs, indican que la definición de razonadores o motores de inferencia que utilicen el modelo propuesto pueden conseguir resultados bastante prometedores en la emulación del comportamiento cognitivo humano. Respecto a la naturaleza de la información, efectivamente se ha definido y demostrado una mayor capacidad representativa al combinar diferentes modelos de características, correspondientes a diferentes entradas perceptuales, para un mismo elemento o categoría de elementos.

Tabla 113. Validación de la sub-hipótesis H₄

Sub-Hipótesis H₅	DESCRIPCIÓN	ES MUY POSIBLE QUE LA COMPLEJIDAD COMPUTACIONAL AUMENTE CONSIDERABLEMENTE AL SER UN MODELO COMPLEJO BASADO EN DIFERENTES MÓDULOS INTERRELACIONADOS.
	EVALUACIÓN	Efectivamente, al hacer uso de diferentes tecnologías de reconocimiento, acceso a sensores, algoritmos de aprendizaje automático y algoritmos de filtrado y pre-procesamiento de información, entre otros, se ha observado un aumento en la complejidad computacional. Sin embargo, gracias a la aplicación de una arquitectura paralela, basada en computación en la nube, y optimizada mediante el uso de contenedores (Docker), se ha conseguido distribuir dicha complejidad entre un conjunto de servidores interconectados, permitiendo incluso la escalabilidad de la solución, al facilitar más recursos computacionales bajo demanda. Así pues, aunque la complejidad sea elevada, la solución es tratable computacionalmente.

Tabla 114. Validación de la sub-hipótesis H₅

Considerando la evaluación positiva para cada una de las sub-hipótesis definidas, se puede determinar que la hipótesis formulada en la tesis es válida.

5.4. Conclusiones del capítulo

El capítulo se ha estructurado presentando, en primer lugar, el procedimiento de validación teórico de los Principios Fundamentales inspirados en los procesos que componen la cognición en el ser humano, y de las Características Técnicas de diseño inferidas de dichos PFs. Seguidamente, y teniendo como propósito el validar la arquitectura definida acorde dichos PFs y CTs, se ha propuesto una estrategia de validación experimental y su correspondiente implementación, de la cual, se han obtenido unas conclusiones satisfactorias.

La implementación del modelo de representación del conocimiento basado en la percepción ha seguido tres fases, en las que se ha incrementado su complejidad y ámbito (características abarcadas). La versión final, más cercana en cuanto a funcionalidades y cercanía con el modelo, se ha basado en un sistema de contenedores tipo Docker, que pueden ser utilizados tanto en entornos de computación local, como en entornos de computación en la nube. En este último caso, se pueden aprovechar las ventajas relacionadas con la escalabilidad, disponibilidad, seguridad y mantenibilidad derivadas de dicho entorno. Si bien, es fundamental tener en cuenta que el paso a producción de un modelo como el propuesto requiere del desarrollo de un plan de gobierno detallado, a fin de minimizar riesgos, cubrir los requisitos establecidos por la ley, y definir los diferentes acuerdos de servicio necesarios.

Para el modelado del conocimiento se han empleado estructuras basadas en JSON, analizado previamente en el capítulo anterior. Se ha optado por JSON en lugar de JSON-LD puesto que, aunque el último permita relaciones con datos enlazados, el enlace se realiza mediante direcciones web. En el modelo propuesto, las piezas de información que componen el conocimiento no forman parte de sitios web, sino que son estructuras abstractas compuestas por diferentes campos (ver los ejemplos mostrados en la Ilustración 42). Por tanto, resulta más conveniente establecer los enlaces entre dichas piezas a través de identificadores únicos, de forma análoga a como se realiza en las bases de datos estructuradas.

Para almacenar los esquemas de información propuestos, definidos a continuación:

- *Chunks*, o componentes esenciales de información sobre una entidad, normalmente asociados a características recogidas mediante sensores (percepción)
- Relaciones nombre de clase / información de clase, útiles para mantener la abstracción conceptual, facilitando el uso de conceptos en otras estructuras
- Jerarquía ontológica de relaciones entre clases y sub-clases
- Conjunto contextual de elementos (categorías que pertenecen a un mismo entorno. Como se analizó previamente, se comporta como una caché, que permitirá acelerar ciertas tareas de reconocimiento, todavía no implementadas)

Se ha utilizado también JSON debido a su enorme versatilidad. Dichos esquemas se recogen mediante instancias de MongoDB. La elección de un sistema de almacenamiento como MongoDB, tipo NoSQL se realiza por varios motivos. Las bases de datos NoSQL son más escalables y proporcionan un rendimiento superior, y su modelo de datos aborda varios problemas que el modelo relacional no está diseñado para tratar: permiten manejar grandes volúmenes de datos estructurados, semi-estructurados y no estructurados que cambian rápidamente, permiten la iteración rápida de esquemas, tienen una naturaleza relacionada con la programación orientada a objetos, fácil de usar y flexible y su arquitectura es de alta escalabilidad, con distribución geográfica, en lugar de monolítica [499]. MongoDB aborda todos estos problemas, siendo extremadamente eficiente y escalable. Teniendo en cuenta el marco tecnológico actual, otra opción prometedora para el almacenamiento de información y su rápida recuperación, se basa en el uso de almacén de datos *Redis (Remote Dictionary Server)* [500], el cual es distribuido, concurrente y presenta un enorme rendimiento en las búsquedas, pues se basa en el almacenamiento tipo clave-valor, siguiendo el paradigma de estructuras de datos como los diccionarios o tablas *hash*. Comportamiento observado también en la recuperación de recuerdos a través de claves en la memoria biológica.

<pre>{ idRec : 2D_001 , type : 2D_vision , params : { { type : image_single_frame , format : png , values : { data : 8V210aCB ... CSFo2 , encoding : base64 } } } , datetime : 2016-10-25 10:00:00 UTC , signature : 107734853D780B72E ... 977 }</pre>	<pre>{ "idRec": "3D_001", "type": "3D_vision", "params": { { "type": "dimensions", "units": "mm", "values": { "width": "83", "height": "91", "depth": "85" } } } , "datetime": "2016-10-25 10:00:00 UTC", "signature": "534853D780B72E ... 9095" }</pre>	<pre>{ "idRec": "CH_001", "type": "chemical", "params": { { "type": "pH", "values": { "ph": "3,7" } } } , "datetime": "2016-10-25 10:00:00 UTC", "signature": "794853D780B72E ... 9001" }</pre>
EJ. DE INFORMACIÓN PERCEPTUAL ASOCIADA A VISIÓN EN 2D	EJ. DE INFORMACIÓN PERCEPTUAL ASOCIADA A VISIÓN EN 3D	EJ. DE INFORMACIÓN PERCEPTUAL ASOCIADA A DETECCIÓN QUÍMICA

Ilustración 42. Ejemplos de representación de información en el sistema mediante JSON

Los resultados de las pruebas iniciales muestran la viabilidad de la solución. La evaluación de la última implementación realizada del modelo ha devuelto resultados muy prometedores, superando la capacidad de reconocimiento del modelo anterior, basado en un único tipo de entrada perceptual y su reconocimiento mediante BPNN. Sin embargo, la implementación realizada no cubre el extenso número de requisitos y características que debería contener el modelo final. Una vez se implemente la solución completa, se realizarán más pruebas para obtener resultados con respecto a la precisión de reconocimiento y aprendizaje en la detección de objetos más complejos.

Finalmente, en el apartado 5.3 se describe la validación de la hipótesis y objetivo propuestos en la Tesis, mediante la evaluación de cada una de las sub-hipótesis asociadas.

Capítulo 6. Epílogo

6.1. Recapitulación

Imaginemos por un momento un sistema inteligente y autónomo, capaz de adquirir información a través de un sinnúmero de sensores y entradas diferentes, distribuidas por todo el entorno, y conectadas con una serie de complejos elementos computacionales encargados de filtrar, procesar y analizar dicha información, de manera individual y colectiva, transformándola en conocimiento. Un conocimiento muy complejo y rico, semánticamente hablando. Un conocimiento que trasciende más allá de la capacidad cognitiva humana, presentando características que sólo podrían adquirirse a través de sentidos ampliados, o percepción extendida. Si el conocimiento del ser humano está basado en un conjunto de hechos adquiridos en base a la información que percibe del entorno gracias a sus (limitadas) capacidades sensoriales, ¿qué clase de conocimiento e inferencias serían resultantes de tal amplitud perceptual? Y no sólo eso. El ser humano, pese a ser un organismo muy evolucionado comparado con el resto de animales, tiene limitaciones notables de capacidad. Su sistema nervioso central no es extensible bajo demanda, y esto juega un papel determinante en la percepción, y por ende, en el proceso cognitivo (ej, ilusiones ópticas, competición sensorial, etc.). Sin embargo, este tipo de limitaciones no aplica al sistema mencionado anteriormente, donde la escalabilidad computacional sólo depende de los recursos disponibles, pudiendo añadir más bajo demanda. El resultado: no sólo un conocimiento mucho más rico y preciso, sino una capacidad cognitiva extendida.

La motivación tras esta tesis doctoral es la de emular el modo en el que los seres humanos asimilan la información, esto es, a través de la percepción de sus sentidos, pero minimizando las limitaciones inherentes a dicho sistema biológico, con el objetivo de desarrollar un amplio rango de nuevas aplicaciones basadas en un sistema inteligente capaz del aprendizaje autónomo desde el entorno, de refinar y extender el conocimiento existente y de combinar la información aprendida en una base de conocimiento global, incluyendo otras propiedades no convencionales en el ser humano para los elementos reconocidos (como visión en rangos fuera del espectro visible, detección precisa de medidas y distancias, detección de campos electromagnéticos, características térmicas, químicas y un largo etcétera), o un incremento en la sensibilidad de las propiedades correspondientes a los sentidos biológicos (mayor precisión en la detección de matices de color, rangos ampliados de frecuencia en la detección de sonido, etc.). Dicha base de conocimiento global será utilizada para desarrollar conceptos más complejos, ideas y teoremas de manera eficiente en base a la información contenida en ella.

El desarrollo de este nuevo modelo tiene como objetivo proporcionar una representación global del conocimiento de una manera no-convencional abstracta. Estos modelos deben tener en cuenta la percepción de los conceptos como un factor importante, que ayudará a clasificarlos y reconocerlos. Este nuevo mecanismo, además, refinará dichos conceptos con el incremento en la experiencia del

reconocimiento, reducirá las redundancias en el modelo de características asociadas a los elementos representados, y facilitará el descubrimiento de relaciones y nuevas inferencias entre dichos elementos.

Mediante el uso de este modelo, también se pretende ampliar las capacidades de las ontologías mediante la representación de múltiples dominios diferentes, pero relacionados, de una manera eficiente, combinando múltiples modelos diferentes de representación bajo una base de conocimiento global abstracta común.

La investigación y desarrollo de la idea inicial comienza orientada hacia la percepción. Gracias a la guía de un experto en el área de la Psicología, se descubre que es necesario incorporar más procesos para poder emular el comportamiento biológico. De hecho, se descubre que la percepción no es sino un componente de la cognición, la cual diferencia al ser humano como entidad inteligente y consciente. Con ayuda del mismo experto, se comienza un estudio sistemático del estado del arte, y se acaban identificando las tres áreas principales a modelar (percepción, conocimiento y memoria), formulando una serie de Principios Fundamentales que debería contener un sistema computacional que trate de emular el comportamiento cognitivo biológico. Seguidamente, se procede a evaluar el estado del arte en el área tecnológica, estudiando los principales sistemas tecnológicos relacionados con la percepción y la cognición. Dicho proceso es soportado por varios expertos en diferentes áreas del marco tecnológico. De dicho estudio se infieren dos grandes grupos principales que guardan relación tecnológica con la percepción y la cognición: los sistemas y arquitecturas cognitivas orientadas a la cibernética (y sistemas físicos), y servicios *software* avanzados y específicos, especializados en el reconocimiento de patrones complejos (desde composición conceptual en imágenes, hasta reconocimiento de fragmentos de audio, pasando por reconocimiento de lenguaje natural y detección de sentimiento). De este segundo estudio se definen, en base a los Principios Fundamentales definidos anteriormente, una serie de Características Técnicas de diseño que deberán estar presentes en la arquitectura del modelo de representación de conocimiento basado en la percepción. Tanto los Principios Fundamentales como las Características Técnicas son refinados iterativamente, y validados por expertos en las áreas correspondientes. Además, ambos son independientes del estado tecnológico, a fin de poder aplicar nuevos avances y soluciones futuras para mejorar el modelo de representación de conocimiento. Es entonces cuando se define un patrón de análisis sensorial, y una arquitectura lógica que extiende dicho patrón, asentándose en las bases de los Principios y Características. Finalmente, para comprobar la validez de la arquitectura propuesta, se realiza una implementación incremental en tres fases, donde se comprueba la factibilidad, escalabilidad y precisión en el reconocimiento, al aplicar un modelo complejo basado en la representación de múltiples características obtenidas mediante la percepción.

Comparando el modelo propuesto con otros existentes, como por ejemplo las arquitecturas de tipo *deep learning*, mientras que en el caso de éstas últimas, el conocimiento está modelado en las capas ocultas de la red, o en autocodificadores, de manera abstracta, y normalmente considerando un conjunto perceptual de entrada limitado, cuyo cambio implica reconfigurar la arquitectura y re-entrenar la red, el modelo propuesto proporciona mucha más flexibilidad y escalabilidad, pudiendo integrar nuevos sentidos perceptuales con un mínimo impacto en las interfaces de

entrada y en el modelo de tratamiento de conocimiento, al estar basado en una combinación de diferentes modelos de representación, que actúan en conjunto, y son extensibles (en la implementación física actual, basados en una extensión a medida de JSON, y en bases de datos *MongoDB*). Además, debido a la naturaleza de tratamiento de información a través de un patrón de análisis sensorial, ajustado al sentido a analizar, es posible aplicar las propias arquitecturas tipo *deep learning* dentro de dichos patrones, reforzando la potencia del modelo propuesto.

6.2. Aportaciones a la Ciencia

De manera sucinta, las principales aportaciones a la Ciencia reflejadas en la Tesis se exponen a continuación:

- Definición y validación de Principios Fundamentales basados en las áreas correspondientes a la psicología y la neurociencia cognitiva sobre los que desarrollar un modelo tecnológico que pretenda emular el funcionamiento del modelado de conocimiento basado en la cognición humana.
- Definición y validación de Características Técnicas de diseño, independientes del estado tecnológico, basadas en los Principios Fundamentales anteriores, sobre los que desarrollar una arquitectura que soporte el modelo tecnológico que pretenda emular el funcionamiento del modelado de conocimiento basado en la cognición humana.
- Definición de un patrón abstracto de análisis sensorial, aplicable a la adquisición y refinado de información a través de sensores u otro tipo de entrada perceptual.
- Definición de una arquitectura lógica que extiende el uso del patrón abstracto de análisis sensorial, a fin de integrar características de diferentes naturalezas en el modelo de representación de conocimiento.
- Definición, implementación y validación experimental exitosa de una arquitectura física basada en la arquitectura lógica descrita en el punto anterior.
- Definición de un modelo de representación de conocimiento que permite homogeneización del mismo, junto con mayor representatividad semántica de las categorías representadas. Mejora de reconocimiento frente a otros modelos actuales, como BPNN (redes neurales con retropropagación).

Nótese que la definición de las bases fundamentales (Principios Fundamentales y Características Técnicas de diseño) conforma junto con el patrón y la arquitectura, un marco de referencia que facilitara la expansión y trabajos futuros sobre la presente investigación.

6.3. Líneas futuras de investigación

Aunque los resultados preliminares son muy prometedores, al tratarse de una aproximación holística muy compleja es conveniente extender el estudio sobre las distintas partes que conforman el modelo, así como realizar más experimentación a medida que se implementen más partes de la arquitectura.

La cantidad de posibilidades a considerar como trabajo futuro son enormes. Por citar algunas:

- Respecto a la implementación de la arquitectura:
 - Implementar los módulos y sistemas de representación del contexto. Su importancia según los estudios en cognición es muy elevada, pues favorece la velocidad del reconocimiento. Para dicha implementación se sugiere, respecto al marco tecnológico actual, el uso de almacén de datos *Redis*, al ser distribuido, concurrente y presentar un enorme rendimiento en las búsquedas (comportamiento clave-valor, como las tablas *hash*, y las claves en la memoria biológica).
 - Almacenar estado de entrenamiento de redes neurales aplicadas en el patrón de análisis sensorial. Recientemente *Redis* ha incorporado una interfaz a tal efecto en su repositorio. El almacenamiento del estado, además, permite la propagación de los cambios a todas las redes neurales basadas en la misma instancia, y el re-entrenamiento en paralelo de las mismas, incrementando la eficiencia global del sistema.
 - Implementar el sistema de refuerzo en el aprendizaje en los módulos perceptuales.
 - Automatizar el despliegue de contenedores con módulos específicos bajo demanda, e implementación de un registro, para mejorar las capacidades inspiradas en los micro-servicios, presentes en la arquitectura definida.
 - Diseñar una plataforma e interfaz para interactuar con seres humanos (p. ej. expertos), que puedan proporcionar *feedback* en el proceso de aprendizaje (de forma análoga a la implementación de Microsoft en LUIS).
 - Integrar más tipos de sensores, y evaluación de características relevantes a almacenar.
 - Considerar la información del estado de los sensores a fin de aplicar correcciones o evaluar la calidad de la información recogida.
- Definir una guía precisa y conjunto de buenas prácticas en la adaptación de la arquitectura (independiente de la tecnología) a una solución técnica, aplicando soluciones técnicas existentes, o basadas en la tecnología actual.

- Realizar estudios específicos de aplicabilidad, por ejemplo, en áreas como la medicina (modelado, detección y tratamiento de patologías), la seguridad (biometría, identificación y prevención de amenazas para la seguridad pública, etc.), la domótica, o cualquier otra área que requiera una interacción y conocimiento de nivel humano o superior.
- Establecer un plan de gobierno adecuado previo a la implantación del sistema.
- Realizar un estudio sobre la relación entre el coste que supone una mejora en la calidad de la información representada y en la fiabilidad del tratamiento de la misma en base a los modelos y soluciones tecnológicas aplicadas en la implementación de la arquitectura propuesta.
- Cubrir otros aspectos de la cognición como la inteligencia y razonamiento, a fin de generar motores de inferencia avanzados que aprovechen el *framework* presentado en la tesis, completando la emulación del proceso cognitivo en el ser humano, pero sin las limitaciones derivadas de la componente biológica.

Chapter 6. Epilogue (English version)

6.1. Recapitulation

Let's imagine for a moment an intelligent and autonomous system, capable of acquiring information through an endless number of different sensors and inputs, distributed throughout the environment, and connected to a series of complex computational elements responsible for filtering, processing and analyzing such information, individually and collectively, transforming it into knowledge. Semantically speaking, a very complex and rich knowledge. A knowledge that transcends beyond human cognitive abilities, presenting characteristics that could only be acquired through extended senses, or through extended perception. If the knowledge of the human being is based on a set of facts inferred from the information he perceives from the environment, via his (limited) sensorial capabilities, what kind of knowledge and inferences would result from such a perceptual amplitude? And that is not all. The human being, despite being a highly evolved organism compared to the rest of animals, has remarkable capacity limitations. Their central nervous system is not extensible on demand, and this plays a determining role in perception, and hence in the cognitive process (e.g., optical illusions, sensory competition, etc.). However, the aforementioned system will not be hindered by this limitations as computational scalability depends only on the available resources, and can add more on demand. The result: not only in a much richer and more precise knowledge, but in an extended cognitive ability.

The motivation behind this doctoral thesis is to emulate the way in which human beings assimilate information, that is, through the perception of their senses, but minimizing the inherent limitations of such biological system, with the aim of developing a wide range of new applications based on an intelligent system, capable of performing autonomous learning from the environment, of refining and extending existing knowledge and of combining information learned in a global knowledge base, including other properties for recognized elements not available in humans (such as vision in ranges outside the visible spectrum, precise detection of measurements and distances, detection of electromagnetic fields, thermal and chemical characteristics, etc.), or an increase in the sensitivity of properties corresponding to the biological senses (improved precision on the detection of color hues, extended frequency ranges in sound detection, etc.). This global knowledge base will be used to develop more complex concepts, ideas and theorems in an efficient way, based on the information contained therein.

The development of this new model aims to provide a global representation of knowledge in an abstract non-conventional way. These models must take into account the perception of concepts as an important factor, which will help to classify and recognize them. This new mechanism will also refine these concepts with the increase in the recognition experience, it will reduce the redundancies in the characteristics

model associated with the represented elements, and it will facilitate the discovery of relationships and new inferences between these elements.

Through the use of this model, it is also intended to extend the capabilities of ontologies by representing multiple but related domains in an efficient way, combining multiple different models of representation under a common global abstract knowledge base.

The research and development of the initial idea started looking at perception. Thanks to the guidance of an expert in the area of Psychology, the necessity to incorporate more processes in order to emulate the biological behavior is discovered. In fact, what is discovered is that perception is but a component of cognition, which differentiates the human being as an intelligent and conscious entity. With the help of the same expert, a systematic study of the state of the art began, identifying the three main areas to be modelled (perception, knowledge and memory), ending in the formulation of a series of Fundamental Principles that a computer system that tries to emulate Biological cognitive behavior should contain. Next, an evaluation in the state of the art in the technological area is conducted, studying the main technological systems related to perception and cognition. This process is supported by several experts in different areas of the technological framework. From this study, two major groups with a technological relationship with perception and cognition are inferred: cybernetic oriented systems and cognitive architectures (plus physical systems), and advanced and specific software services, specialized in the recognition of complex patterns (from conceptual composition in images, to recognition of audio fragments, through natural language recognition and sentiment analysis). This second study defines, based on the Fundamental Principles defined above, a series of Technical Design Characteristics that must be present in the architecture of the knowledge representation model based on perception. Both the Fundamental Principles and the Technical Characteristics are iteratively refined, and validated by experts in the corresponding areas. In addition, both are independent of the current technological state, in order to be able to apply new advances and future solutions to improve the knowledge representation model. It is then when a pattern of sensorial analysis is defined, as well as a logical architecture that extends this pattern, based on the bases of the Principles and Characteristics. Finally, to verify the validity of the proposed architecture, an incremental implementation in three phases is carried out, where feasibility, scalability and precision in the recognition are verified, applying a complex model based on the representation of multiple characteristics obtained through the perception.

Comparing the proposed model with other existing ones, such as deep learning architectures, while in the case of the latter, knowledge is modeled in the hidden layers of the network, or by means of self-encoders, in an abstract way, and usually considering a limited input perceptual set whose change involves reconfiguring the architecture and re-training the network, the proposed model provides much more flexibility and scalability, being able to integrate new perceptual senses with minimal impact on the input interfaces and the treatment model of knowledge, since it is based on a combination of different models of representation, that act together, and are extensible (in the current physical implementation, based on a custom extension of JSON, and in MongoDB databases). In addition, due to the nature of information processing through a sensorial analysis pattern, adjusted to the sensory input to be

analyzed, it is possible to apply deep learning architectures within these patterns, reinforcing the power of the proposed model.

6.2. Contributions to Science

Briefly, the main contributions to Science reflected in the Thesis are as follows:

- Definition and validation of Fundamental Principles based on the areas corresponding to psychology and cognitive neuroscience, pillars to develop a technological model that seeks to emulate the operation of knowledge modelling based on human cognition.
- Definition and validation of Technical Design Characteristics, independent of the technological state, based on previous Fundamental Principles, on which to develop an architecture that supports the technological model that seeks to emulate the operation of knowledge modelling based on human cognition.
- Definition of an abstract pattern for sensory analysis, applicable to the acquisition and refining of information through sensors or other type of perceptual input.
- Definition of a logical architecture that extends the use of the abstract pattern of sensory analysis, in order to integrate the different characteristics of the knowledge representation model.
- Definition, implementation and successful experimental validation of a physical architecture based on the logical architecture described in the previous point.
- Definition of a knowledge representation model that allows the homogenization of knowledge, providing also greater semantic representation of the categories represented. Improvement of recognition against other current models, such as BPNN (backpropagation neural networks).

Note that the definition of the fundamental bases (Fundamental Principles and Technical Design Characteristics) together with the pattern and architecture, define a framework that facilitates the expansion and future work on the present research.

6.3. Future research lines

Although the preliminary results are very promising, as it is a very complex holistic approach, it is advisable to extend the study of the different parts that make up the model, as well as to perform more experimentation as more parts of the architecture are implemented.

The amount of possibilities to consider as future work are enormous. Just to mention some of them:

- Regarding the implementation of the architecture:
 - To implement modules and systems for representing the context. Its importance according to studies in cognition is very high, because it favors the speed of recognition. For this implementation, and considering the state of the technological framework, the usage of Redis cache is proposed, as it is distributed, concurrent and it presents a great performance in the information retrieval process (key-value behavior, such as hash tables and keys in biological memory).
 - To store training state of neural networks applied in the sensory analysis pattern. Recently Redis has incorporated an interface to that effect in its repository. State storage, in addition, allows the propagation of changes to all neural networks based on the same instance, and the parallel re-training of the same, increasing the overall efficiency of the system.
 - To implement the reinforcement learning mechanism for the perceptual modules.
 - To automate the deployment of containers with specific modules on demand, and to implement a registry, in order to improve the capabilities inspired by micro-services, present in the defined architecture.
 - To design of a platform and interface to interact with human beings (e.g. experts), who can provide feedback in the learning process (similarly to the implementation of Microsoft in LUIS).
 - To integrate more types of sensors, and evaluation of relevant characteristics to be stored.
 - To consider sensor status information in order to apply corrections or to evaluate the quality of the information collected.
- To define an accurate guide and set of best practices to adapt the architecture (independent of technology) to a technical solution, by applying existing technical solutions, or based on current technology.
- To conduct specific applicability studies, for example in areas such as medicine (modeling, detection and treatment of pathologies), security (biometrics, identification and prevention of threats to public safety, etc.), home automation,

or any other area which requires interaction and knowledge at the human level or above.

- To establish a proper governance plan prior to the implementation of the system.
- To perform a study on the relationship between the cost of improving the quality of the information represented and the reliability of its processing, based on the models and technological solutions applied in the implementation of the proposed architecture.
- To cover other aspects of cognition such as intelligence and reasoning, in order to generate advanced inference engines that take advantage of the framework presented in the thesis, completing the emulation of the cognitive process in the human being, but without the limitations derived from the biological component.

Referencias bibliográficas

- [1] M. G. Miniwatts, “Internet World Stats,” Jun. 2012.
- [2] M. J. Eppler and J. Mengis, “The Concept of Information Overload: A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines,” *Inf. Soc.*, vol. 20, no. 5, pp. 325–344, Nov. 2004.
- [3] R. A. E. R.A.E., *Diccionario de la lengua española*, 23rd ed. Real Academia Española, 2013.
- [4] J. Bengson and M. A. Moffett, *Knowing How: Essays on Knowledge, Mind, and Action*. Oxford University Press, 2011.
- [5] T. Gruber, “Collective knowledge systems: Where the Social Web meets the Semantic Web,” *Web Semant. Sci. Serv. Agents World Wide Web*, vol. 6, no. 1, pp. 4–13, Feb. 2008.
- [6] F. Xia, L. T. Yang, L. Wang, and A. Vinel, “Internet of things,” *Int. J. Commun. Syst.*, vol. 25, no. 9, p. 1101, 2012.
- [7] L. Atzori, A. Iera, and G. Morabito, “The Internet of Things: A survey,” *Comput. Netw.*, vol. 54, no. 15, pp. 2787–2805, Oct. 2010.
- [8] J. Gubbi, R. Buyya, S. Marusic, and M. Palaniswami, “Internet of Things (IoT): A vision, architectural elements, and future directions,” *Future Gener. Comput. Syst.*, vol. 29, no. 7, pp. 1645–1660, Sep. 2013.
- [9] C. Romero, F. Vázquez, and C. De Castro, “Domótica e Inmótica. Viviendas y edificios inteligentes,” *RA-MA 2ª Ed Del*, 2006.
- [10] H. Lasi, P. Fettke, H.-G. Kemper, T. Feld, and M. Hoffmann, “Industry 4.0,” *Bus. Inf. Syst. Eng.*, vol. 6, no. 4, p. 239, 2014.
- [11] J. Lee, B. Bagheri, and H.-A. Kao, “A cyber-physical systems architecture for industry 4.0-based manufacturing systems,” *Manuf. Lett.*, vol. 3, pp. 18–23, 2015.
- [12] “How does Google Maps predict traffic?,” *HowStuffWorks*, 31-Oct-2014. [Online]. Available: <http://electronics.howstuffworks.com/how-does-google-maps-predict-traffic.htm>. [Accessed: 24-Nov-2016].
- [13] S. Hyuga, M. Ito, M. Iwai, and K. Sezaki, “Estimate a User’s Location Using Smartphone’s Barometer on a Subway,” in *Proceedings of the 5th International Workshop on Mobile Entity Localization and Tracking in GPS-less Environments*, New York, NY, USA, 2015, p. 2:1–2:4.
- [14] E. H. Rosch, “Natural categories,” *Cognit. Psychol.*, vol. 4, no. 3, pp. 328–350, 1973.
- [15] A. Amarasingam, *The Stewart/Colbert effect: Essays on the real impacts of fake news*. McFarland, 2011.
- [16] R. Marchi, “With Facebook, blogs, and fake news, teens reject journalistic ‘objectivity,’” *J. Commun. Inq.*, vol. 36, no. 3, pp. 246–262, 2012.

- [17] P. Jarvis, *Towards a comprehensive theory of human learning*. Routledge, 2012.
- [18] U. Neisser, “Without perception, there is no knowledge: Implications for artificial intelligence,” *Nat. Artif. Minds*, pp. 174–164, 1993.
- [19] D. L. Schacter, D. T. Gilbert, and D. M. Wegner, *Introducing psychology*. Worth Publishers, 2011.
- [20] E. B. Goldstein, *Sensation and perception*. Wadsworth Publishing Company, 2010.
- [21] R. L. Gregory, “Oliver Louis Zangwill. 29 October 1913-12 October 1987,” *Biogr. Mem. Fellows R. Soc.*, vol. 47, pp. 517–524, 2001.
- [22] D. Bernstein, *Essentials of psychology*, 6th ed. Wadsworth Publishing, 2013.
- [23] M. Bach and C. Poloschek, “Optical illusions,” *Adv. Clin. Neurosci. Rehabil.*, vol. 6, no. 2, pp. 20–21, 2006.
- [24] T. J. Slakey, “Aristotle on sense perception,” *Philos. Rev.*, vol. 70, no. 4, pp. 470–484, 1961.
- [25] I. Rock, “Perception and knowledge,” *Acta Psychol. (Amst.)*, vol. 59, no. 1, pp. 3–22, 1985.
- [26] H. McGurk and J. MacDonald, “Hearing lips and seeing voices,” *Nature*, pp. 746–748, 1976.
- [27] D. M. Small and J. Prescott, “Odor/taste integration and the perception of flavor,” *Exp. Brain Res.*, vol. 166, no. 3–4, pp. 345–357, 2005.
- [28] C. Osborne, “Aristotle, De anima 3. 2: How do we perceive that we see and hear?,” *Class. Q. New Ser.*, vol. 33, no. 2, pp. 401–411, 1983.
- [29] F. Macpherson, *The Senses: Classic and Contemporary Philosophical Perspectives*, 1st ed. Oxford: Oxford University Press, 2011.
- [30] R. Le Poidevin, “The Experience and Perception of Time,” in *The Stanford Encyclopedia of Philosophy*, E. N. Zalta, Ed. Stanford University: The Metaphysics Research Lab, 2015.
- [31] I. Phillips, “Perceiving temporal properties,” *Eur. J. Philos.*, vol. 18, no. 2, pp. 176–202, 2010.
- [32] L. A. Pogorzala, S. K. Mishra, and M. A. Hoon, “The Cellular Code for Mammalian Thermosensation,” *J. Neurosci.*, vol. 33, no. 13, pp. 5533–5541, 2013.
- [33] R. Suzuki and A. H. Dickenson, “1 Nociception: basic principles,” *Cancer Pain Assess. Manag.*, p. 3, 2003.
- [34] C. B. Currie, G. W. McConkie, L. A. Carlson-Radvansky, and D. E. Irwin, “The role of the saccade target object in the perception of a visually stable world,” *Percept. Psychophys.*, vol. 62, no. 4, pp. 673–683, Jan. 2000.
- [35] R. DeVere and M. Calvert, *Navigating smell and taste disorders*, 1st ed. Demos Health, 2010.
- [36] R. West, J. Wilding, D. French, R. Kemp, and A. Irving, “Effect of low and moderate doses of alcohol on driving hazard perception latency and driving speed,” *Addiction*, vol. 88, no. 4, pp. 527–532, 1993.
- [37] Jphotonics, Inc., “Distributed Temperature Sensing System (DTS),” 2013.

-
- [38] E. R. Hilgard, L. Atkinson, R. C. Atkinson, E. Smith, D. J. Bem, and S. Nolen-Hoeksema, *Introduction to Psychology (13th Edition)*. Wadsworth Publishing Co Inc, 1999.
- [39] E. B. Goldstein, *Encyclopedia of perception*, vol. 1. SAGE Publications, Incorporated, 2009.
- [40] T. Sonderegger, *Cliffs Quick Review Psychology*. 2004.
- [41] S. Yantis, *Visual perception: Essential readings*. Psychology Pr, 2001.
- [42] J. C. Banerjee, “Gestalt Theory of Perception,” *Encyclopaedic Dictionary of Psychological Terms*. pp. 107–108, 1994.
- [43] P. O. Gray, *Psychology (5th ed.)*. New York: Worth, 2006.
- [44] J. M. Wolfe et al., “Gestalt Grouping Principles,” *Sensat. Percept.*, 2008.
- [45] W. Weiten, *Psychology: themes and variations (4th ed.)*. Brooks/Cole Pub. Co., 1998.
- [46] S. Coren and J. S. Girgus, “Principles of perceptual organization and spatial distortion: The gestalt illusions,” *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 6, no. 3, pp. 404–412, 1980.
- [47] R. J. Watt, *Understanding vision*. Academic Press, 1991.
- [48] R. J. Corsini, *The dictionary of psychology*. Psychology Press, 2002.
- [49] L. H. Kushner, *Contrast in judgments of mental health*. ProQuest, 2008.
- [50] S. Plous, *The psychology of judgment and decision making*. McGraw-Hill, 1993.
- [51] A. N. Popper, *Music Perception*. Springer, 2010.
- [52] M. C. Potter and E. I. Levy, “Recognition memory for a rapid sequence of pictures,” *J. Exp. Psychol.*, vol. 81, no. 1, p. 10, 1969.
- [53] H. Intraub, “Presentation rate and the representation of briefly glimpsed pictures in memory,” *J. Exp. Psychol. [Hum. Learn.]*, vol. 6, no. 1, p. 1, 1980.
- [54] J. J. DiCarlo, D. Zoccolan, and N. C. Rust, “How does the brain solve visual object recognition?,” *Neuron*, vol. 73, no. 3, pp. 415–434, 2012.
- [55] P. R. Huttenlocher, “Morphometric study of human cerebral cortex development,” *Neuropsychologia*, vol. 28, no. 6, pp. 517–527, 1990.
- [56] P. R. Huttenlocher, *Neural plasticity*. Harvard University Press, 2002.
- [57] T. N. Wiesel and D. H. Hubel, “Single-cell responses in striate cortex of kittens deprived of vision in one eye,” *J. Neurophysiol.*, vol. 26, no. 6, pp. 1003–1017, 1963.
- [58] T. N. Wiesel and D. H. Hubel, “Comparison of the effects of unilateral and bilateral eye closure on cortical unit responses in kittens,” *J. Neurophysiol.*, vol. 28, no. 6, pp. 1029–1040, 1965.
- [59] S. Le Vay, T. N. Wiesel, and D. H. Hubel, “The development of ocular dominance columns in normal and visually deprived monkeys,” *J. Comp. Neurol.*, vol. 191, no. 1, pp. 1–51, 1980.
- [60] W. M. Jenkins, M. M. Merzenich, M. T. Ochs, T. Allard, and E. Guic-Robles, “Functional reorganization of primary somatosensory cortex in adult owl monkeys

- after behaviorally controlled tactile stimulation,” *J. Neurophysiol.*, vol. 63, no. 1, pp. 82–104, 1990.
- [61] L. Gyllensten, T. Malmfors, and M.-L. Norrlin, “Growth alteration in the auditory cortex of visually deprived mice,” *J. Comp. Neurol.*, vol. 126, no. 3, pp. 463–469, 1966.
- [62] M. Cynader, “Competitive Interactions in Postnatal Development of the Kitten’s Visual System,” in *Developmental Neurobiology of Vision*, Springer, 1979, pp. 109–120.
- [63] N. J. Wade and N. Wade, *A natural history of vision*. MIT Press, 2000.
- [64] G. L. Shulman, “Attentional modulation of size contrast,” *Q. J. Exp. Psychol.*, vol. 45, no. 4, pp. 529–546, 1992.
- [65] S. Coren and J. T. Enns, “Size contrast as a function of conceptual similarity between test and inducers,” *Percept. Psychophys.*, vol. 54, no. 5, pp. 579–588, 1993.
- [66] A. Hollingworth, “Does consistent scene context facilitate object perception?,” *J. Exp. Psychol. Gen.*, vol. 127, no. 4, p. 398, 1998.
- [67] I. Biederman, R. J. Mezzanotte, and J. C. Rabinowitz, “Scene perception: Detecting and judging objects undergoing relational violations,” *Cognit. Psychol.*, vol. 14, no. 2, pp. 143–177, 1982.
- [68] G. M. Reicher, “Perceptual recognition as a function of meaningfulness of stimulus material,” *J. Exp. Psychol.*, vol. 81, no. 2, p. 275, 1969.
- [69] D. D. Wheeler, “Processes in word recognition,” *Cognit. Psychol.*, vol. 1, no. 1, pp. 59–85, 1970.
- [70] J. Baron and I. Thurston, “An analysis of the word-superiority effect,” *Cognit. Psychol.*, vol. 4, no. 2, pp. 207–228, 1973.
- [71] E. E. Smith and S. M. Kosslyn, “Cognitive Psychology: Mind and Brain,” 2006. [Online]. Available: <http://www.pearsonhighered.com/educator/product/Cognitive-Psychology-Mind-and-Brain/9780131825086.page>. [Accessed: 14-Apr-2015].
- [72] G. W. Humphreys and M. J. Riddoch, *To see but not to see: A case study of visual agnosia*. Psychology Press, 1987.
- [73] O. G. Selfridge, “Pattern recognition and modern computers,” in *Proceedings of the March 1-3, 1955, western joint computer conference*, 1955, pp. 91–93.
- [74] D. Ferster and K. D. Miller, “Neural mechanisms of orientation selectivity in the visual cortex,” *Annu. Rev. Neurosci.*, vol. 23, no. 1, pp. 441–471, 2000.
- [75] D. H. Hubel and T. N. Wiesel, “Receptive fields of single neurones in the cat’s striate cortex,” *J. Physiol.*, vol. 148, no. 3, pp. 574–591, 1959.
- [76] R. Desimone, T. D. Albright, C. G. Gross, and C. Bruce, “Stimulus-selective properties of inferior temporal neurons in the macaque,” *J. Neurosci.*, vol. 4, no. 8, pp. 2051–2062, 1984.
- [77] K. Tanaka, H. Saito, Y. Fukada, and M. Moriya, “Coding visual images of objects in the inferotemporal cortex of the macaque monkey,” *J. Neurophysiol.*, vol. 66, no. 1, pp. 170–189, 1991.

-
- [78] D. Perrett, E. Rolls, and W. Caan, "Visual neurones responsive to faces in the monkey temporal cortex," *Exp. Brain Res.*, vol. 47, no. 3, pp. 329–342, 1982.
- [79] D. J. Freedman, M. Riesenhuber, T. Poggio, and E. K. Miller, "Categorical representation of visual stimuli in the primate prefrontal cortex," *Science*, vol. 291, no. 5502, pp. 312–316, 2001.
- [80] S. M. Kosslyn and C. F. Chabris, "Naming pictures," *J. Vis. Lang. Comput.*, vol. 1, no. 1, pp. 77–95, 1990.
- [81] P. S. Penev and J. J. Atick, "Local feature analysis: A general statistical theory for object representation," *Netw. Comput. Neural Syst.*, vol. 7, no. 3, pp. 477–500, 1996.
- [82] S. Pankanti and M. M. Yeung, "Verification watermarks on fingerprint recognition and retrieval," in *Electronic Imaging '99*, 1999, pp. 66–78.
- [83] Y. Jie, Z. Renjie, S. Qifa, and others, "Fingerprint minutiae matching algorithm for real time system," *Pattern Recognit.*, vol. 39, no. 1, pp. 143–146, 2006.
- [84] O. Gallo and R. Manduchi, "Reading 1d barcodes with mobile phones using deformable templates," *Pattern Anal. Mach. Intell. IEEE Trans. On*, vol. 33, no. 9, pp. 1834–1843, 2011.
- [85] K. W. Bowyer, K. Hollingsworth, and P. J. Flynn, "Image understanding for iris biometrics: A survey," *Comput. Vis. Image Underst.*, vol. 110, no. 2, pp. 281–307, 2008.
- [86] C. B. Cave and S. M. Kosslyn, "The role of parts and spatial relations in object identification," *Percept.-Lond.-*, vol. 22, pp. 229–229, 1993.
- [87] B. Laeng, J. Shah, and S. Kosslyn, "Identifying objects in conventional and contorted poses: Contributions of hemisphere-specific mechanisms," *Cognition*, vol. 70, no. 1, pp. 53–85, 1999.
- [88] I. Biederman, "Recognition-by-components: a theory of human image understanding.," *Psychol. Rev.*, vol. 94, no. 2, p. 115, 1987.
- [89] J. E. Hummel and I. Biederman, "Dynamic binding in a neural network for shape recognition.," *Psychol. Rev.*, vol. 99, no. 3, p. 480, 1992.
- [90] I. Biederman and E. E. Cooper, "Priming contour-deleted images: Evidence for intermediate representations in visual object recognition," *Cognit. Psychol.*, vol. 23, no. 3, pp. 393–419, 1991.
- [91] I. Biederman, *Visual object recognition*, vol. 2. MIT press, 1995.
- [92] E. Wachsmuth, M. Oram, and D. Perrett, "Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque," *Cereb. Cortex*, vol. 4, no. 5, pp. 509–522, 1994.
- [93] N. K. Logothetis, J. Pauls, and T. Poggio, "Shape representation in the inferior temporal cortex of monkeys," *Curr. Biol.*, vol. 5, no. 5, pp. 552–563, 1995.
- [94] R. Diamond and S. Carey, "Why faces are and are not special: an effect of expertise.," *J. Exp. Psychol. Gen.*, vol. 115, no. 2, p. 107, 1986.
- [95] G. Rhodes, S. Brennan, and S. Carey, "Identification and ratings of caricatures: Implications for mental representations of faces," *Cognit. Psychol.*, vol. 19, no. 4, pp. 473–497, 1987.

- [96] M. P. Young and S. Yamane, "Sparse population coding of faces in the inferotemporal cortex," *Science*, vol. 256, no. 5061, pp. 1327–1331, 1992.
- [97] I. Gauthier, P. Skudlarski, J. C. Gore, and A. W. Anderson, "Expertise for cars and birds recruits brain areas involved in face recognition," *Nat. Neurosci.*, vol. 3, no. 2, pp. 191–197, 2000.
- [98] D. T. Kemp, "Stimulated acoustic emissions from within the human auditory system," *J. Acoust. Soc. Am.*, vol. 64, no. 5, pp. 1386–1391, 1978.
- [99] D. DeBonis and C. Donohue, "Survey of audiology [M]," *Bosten Pearson Educ.*, pp. 77–106, 2004.
- [100] J. Groen, "The semicircular canal system of the organs of equilibrium-I," *Phys. Med. Biol.*, vol. 1, no. 2, p. 103, 1956.
- [101] R. N. Shepard, "Circularity in judgments of relative pitch," *J. Acoust. Soc. Am.*, vol. 36, no. 12, pp. 2346–2353, 1964.
- [102] I. Pollack, "The information of elementary auditory displays," *J. Acoust. Soc. Am.*, vol. 24, no. 6, pp. 745–749, 1952.
- [103] ANSI, *American National Standard Acoustical Terminology*. American National Standards Institute New York, NY, 1994.
- [104] R. Plomp, "Timbre as a multidimensional attribute of complex tones," *Freq. Anal. Period. Detect. Hear.*, pp. 397–414, 1970.
- [105] R. Plomp, *Aspects of tone sensation: a psychophysical study*. Academic Press, 1976.
- [106] R. Plomp, "Pitch of complex tones," *J. Acoust. Soc. Am.*, vol. 41, no. 6, pp. 1526–1533, 1967.
- [107] L. C. Pols, L. T. Van der Kamp, and R. Plomp, "Perceptual and physical space of vowel sounds," *J. Acoust. Soc. Am.*, vol. 46, no. 2B, pp. 458–467, 1969.
- [108] R. Plomp and H. Steeneken, "Effect of phase on the timbre of complex tones," *J. Acoust. Soc. Am.*, vol. 46, no. 2B, pp. 409–421, 1969.
- [109] C. Darwin and R. Gardner, "Mistuning a harmonic of a vowel: Grouping and phase effects on vowel quality," *J. Acoust. Soc. Am.*, vol. 79, no. 3, pp. 838–845, 1986.
- [110] R. D. Patterson, "A pulse ribbon model of monaural phase perception," *J. Acoust. Soc. Am.*, vol. 82, no. 5, pp. 1560–1586, 1987.
- [111] J. F. Schouten, "The perception of timbre," in *Reports of the 6th International Congress on Acoustics*, 1968, vol. 76, p. 10.
- [112] J.-C. Risset and D. L. Wessel, "Exploration of timbre by analysis and synthesis," *Psychol. Music*, vol. 2, p. 151, 1999.
- [113] R. D. Patterson, "The sound of a sinusoid: Spectral models," *J. Acoust. Soc. Am.*, vol. 96, no. 3, pp. 1409–1418, 1994.
- [114] R. D. Patterson, "The sound of a sinusoid: Time-interval models," *J. Acoust. Soc. Am.*, vol. 96, no. 3, pp. 1419–1428, 1994.

-
- [115] M. A. Akeroyd and R. D. Patterson, "Discrimination of wideband noises modulated by a temporally asymmetric function," *J. Acoust. Soc. Am.*, vol. 98, no. 5, pp. 2466–2474, 1995.
- [116] S. Handel, "Timbre perception and auditory object identification," *Hearing*, pp. 425–461, 1995.
- [117] A. S. Bregman, "The formation of auditory streams," *Atten. Perform.*, vol. 7, pp. 63–76, 1978.
- [118] A. S. Bregman and S. Pinker, "Auditory streaming and the building of timbre.," *Can. J. Psychol. Can. Psychol.*, vol. 32, no. 1, p. 19, 1978.
- [119] A. S. Bregman, *Auditory scene analysis: The perceptual organization of sound*. MIT press, 1994.
- [120] C. J. Darwin, "Auditory grouping," *Trends Cogn. Sci.*, vol. 1, no. 9, pp. 327–333, 1997.
- [121] D. Broadbent and P. Ladefoged, "On the fusion of sounds reaching different sense organs," *J. Acoust. Soc. Am.*, vol. 29, no. 6, pp. 708–710, 1957.
- [122] J. F. Culling and C. Darwin, "Perceptual separation of simultaneous vowels: Within and across-formant grouping by F0," *J. Acoust. Soc. Am.*, vol. 93, no. 6, pp. 3454–3467, 1993.
- [123] J. F. Culling and C. Darwin, "Perceptual and computational separation of simultaneous vowels: Cues arising from low-frequency beating," *J. Acoust. Soc. Am.*, vol. 95, no. 3, pp. 1559–1569, 1994.
- [124] P. F. Assmann and Q. Summerfield, "Modeling the perception of concurrent vowels: Vowels with different fundamental frequencies," *J. Acoust. Soc. Am.*, vol. 88, no. 2, pp. 680–697, 1990.
- [125] R. Meddis and M. J. Hewitt, "Modeling the identification of concurrent vowels with different fundamental frequencies," *J. Acoust. Soc. Am.*, vol. 91, no. 1, pp. 233–245, 1992.
- [126] A. De Cheveigné, "Concurrent vowel identification. III. A neural model of harmonic interference cancellation," *J. Acoust. Soc. Am.*, vol. 101, no. 5, pp. 2857–2865, 1997.
- [127] H. E. Gockel, E. R. Hafter, and B. C. Moore, "Pitch discrimination interference: The role of ear of entry and of octave similarity," *J. Acoust. Soc. Am.*, vol. 125, no. 1, p. 324, 2009.
- [128] R. A. Rasch, "The perception of simultaneous notes such as in polyphonic music," *Acta Acust. United Acust.*, vol. 40, no. 1, pp. 21–33, 1978.
- [129] T. Houtgast, "Psychophysical evidence for lateral inhibition in hearing," *J. Acoust. Soc. Am.*, vol. 51, no. 6B, pp. 1885–1894, 1972.
- [130] C. Darwin and N. Sutherland, "Grouping frequency components of vowels: When is a harmonic not a harmonic?," *Q. J. Exp. Psychol.*, vol. 36, no. 2, pp. 193–208, 1984.
- [131] B. C. Moore, "Temporal integration and context effects in hearing," *J. Phon.*, vol. 31, no. 3, pp. 563–574, 2003.

- [132] K. R. Kluender, J. A. Coady, and M. Kiefte, “Sensitivity to change in perception of speech,” *Speech Commun.*, vol. 41, no. 1, pp. 59–69, 2003.
- [133] E. Zwicker, “‘Negative afterimage’ in hearing,” *J. Acoust. Soc. Am.*, vol. 36, no. 12, pp. 2413–2415, 1964.
- [134] Q. Summerfield, A. Sidwell, and T. Nelson, “Auditory enhancement of changes in spectral amplitude,” *J. Acoust. Soc. Am.*, vol. 81, no. 3, pp. 700–708, 1987.
- [135] S. McAdams, “Spectral fusion and the creation of auditory images,” in *Music, mind, and brain*, Springer, 1982, pp. 279–298.
- [136] S. McAdams, “Segregation of concurrent sounds. I: Effects of frequency modulation coherence,” *J. Acoust. Soc. Am.*, vol. 86, no. 6, pp. 2148–2159, 1989.
- [137] Q. Summerfield, J. F. Culling, and A. Fourcin, “Auditory Segregation of Competing Voices: Absence of Effects of FM or AM Coherence [and Discussion],” *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 336, no. 1278, pp. 357–366, 1992.
- [138] P. Joris, C. Schreiner, and A. Rees, “Neural processing of amplitude-modulated sounds,” *Physiol. Rev.*, vol. 84, no. 2, pp. 541–577, 2004.
- [139] B. C. Moore, B. R. Glasberg, and A. J. Oxenham, “Effects of pulsing of a target tone on the ability to hear it out in different types of complex sounds,” *J. Acoust. Soc. Am.*, vol. 131, no. 4, pp. 2927–2937, 2012.
- [140] A. S. Bregman, J. Abramson, P. Doehring, and C. J. Darwin, “Spectral integration based on common amplitude modulation,” *Percept. Psychophys.*, vol. 37, no. 5, pp. 483–493, 1985.
- [141] J. W. Hall III and J. H. Grose, “Comodulation masking release and auditory grouping,” *J. Acoust. Soc. Am.*, vol. 88, no. 1, pp. 119–125, 1990.
- [142] B. C. Moore and M. J. Shailer, “Comodulation masking release as a function of level,” *J. Acoust. Soc. Am.*, vol. 90, no. 2, pp. 829–835, 1991.
- [143] N. I. Durlach and H. S. Colburn, “Binaural phenomena,” *Handb. Percept.*, vol. 4, pp. 365–466, 1978.
- [144] M. Kubovy, J. E. Cutting, and R. M. McGuire, “Hearing with the third ear: dichotic perception of a melody without monaural familiarity cues,” *Science*, vol. 186, no. 4160, pp. 272–274, 1974.
- [145] G. A. Miller and G. A. Heise, “The trill threshold,” *J. Acoust. Soc. Am.*, vol. 22, no. 5, pp. 637–638, 1950.
- [146] A. S. Bregman and J. Campbell, “Primary auditory stream segregation and perception of order in rapid sequences of tones,” *J. Exp. Psychol.*, vol. 89, no. 2, p. 244, 1971.
- [147] van L. Noorden, “Rhythmic fission as a function of tone rate,” *IPO Annu. Prog. Rep.*, vol. 6, p. 9, 1971.
- [148] M. A. Akeroyd, R. P. Carlyon, and J. M. Deeks, “Can dichotic pitches form two streams?,” *J. Acoust. Soc. Am.*, vol. 118, no. 2, pp. 977–981, 2005.
- [149] W. J. Dowling, “The perception of interleaved melodies,” *Cognit. Psychol.*, vol. 5, no. 3, pp. 322–337, 1973.

-
- [150] A. S. Bregman and G. L. Dannenbring, "The effect of continuity on auditory stream segregation," *Percept. Psychophys.*, vol. 13, no. 2, pp. 308–312, 1973.
- [151] C. Darwin and R. Hukin, "Auditory objects of attention: the role of interaural time differences.," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 25, no. 3, p. 617, 1999.
- [152] K. Koffka, *Principles of Gestalt psychology*. Routledge, 2013.
- [153] C. J. Darwin and C. E. Bethell-Fox, "Pitch continuity and speech source attribution.," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 3, no. 4, p. 665, 1977.
- [154] A. S. Bregman and A. I. Rudnicki, "Auditory segregation: Stream or streams?," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 1, no. 3, p. 263, 1975.
- [155] A. S. Bregman, "The meaning of duplex perception: Sounds as transparent objects," in *The psychophysics of speech perception*, Springer, 1987, pp. 95–111.
- [156] V. Ciocca and A. S. Bregman, "Perceived continuity of gliding and steady-state tones through interrupting noise," *Percept. Psychophys.*, vol. 42, no. 5, pp. 476–484, 1987.
- [157] J. Lyzenga, R. Carlyon, and B. Moore, "Dynamic aspects of the continuity illusion: perception of level and of the depth, rate, and phase of modulation," *Hear. Res.*, vol. 210, no. 1, pp. 30–41, 2005.
- [158] G. A. Miller and J. Licklider, "The intelligibility of interrupted speech," *J. Acoust. Soc. Am.*, vol. 22, no. 2, pp. 167–173, 1950.
- [159] U. Neisser, *Cognitive Psychology: Classic Edition*. Psychology Press, 2014.
- [160] A. Ihlefeld and B. Shinn-Cunningham, "Disentangling the effects of spatial cues on selection and formation of auditory objects," *J. Acoust. Soc. Am.*, vol. 124, no. 4, pp. 2224–2235, 2008.
- [161] L. P. A. S. van Noorden, *Temporal coherence in the perception of tone sequences*. Institute for Perceptual Research, 1975.
- [162] D. Pressnitzer and J.-M. Hupé, "Temporal dynamics of auditory and visual bistability reveal common principles of perceptual organization," *Curr. Biol.*, vol. 16, no. 13, pp. 1351–1357, 2006.
- [163] G. Rodríguez-Gil, "El Poderoso Sentido del Olfato," *Resour. San Franc. State Univ.*, vol. 11, pp. 5–11, 2010.
- [164] M. O. Hutchins, "Chemical Senses: Olfaction and Gustation," in *Neuroscience Online*, vol. 1, 1 vols., University of Texas. Medical School, 2014.
- [165] C.-Y. Su, K. Menz, and J. R. Carlson, "Olfactory perception: receptors, cells, and circuits," *Cell*, vol. 139, no. 1, pp. 45–59, 2009.
- [166] V. M. Alcaraz, *Estructura y función del sistema nervioso*. UNAM, 2000.
- [167] B. G. Green and D. Nachtigal, "Temperature Affects Human Sweet Taste via At Least Two Mechanisms," *Chem. Senses*, p. bju021, May 2015.
- [168] B. Kolb and I. Q. Whishaw, *Neuropsicología humana*. Ed. Médica Panamericana, 2006.

- [169] L. M. Bartoshuk, "Comparing Sensory Experiences Across Individuals: Recent Psychophysical Advances Illuminate Genetic Variation in Taste Perception," *Chem. Senses*, vol. 25, no. 4, pp. 447–460, Aug. 2000.
- [170] L. M. Bartoshuk, V. B. Duffy, and I. J. Miller, "PTC/PROP tasting: Anatomy, psychophysics, and sex effects," *Physiol. Behav.*, vol. 56, no. 6, pp. 1165–1171, Dec. 1994.
- [171] D. U. Silverthorn, *Fisiología Humana. Un enfoque integrado 4a edición*. Ed. Médica Panamericana, 2008.
- [172] A. L. Benton, *Brain and Behavior: Research in Clinical Neuropsychology*. Transaction Publishers, 2009.
- [173] J. A. Boulant and J. B. Dean, "Temperature Receptors in the Central Nervous System," *Annu. Rev. Physiol.*, vol. 48, no. 1, pp. 639–654, 1986.
- [174] I. Kojo and A. Pertovaara, "The Effects of Stimulus Area and Adaptation Temperature on Warm and Heat Pain Thresholds in Man," *Int. J. Neurosci.*, vol. 32, no. 3–4, pp. 875–880, Jan. 1987.
- [175] T. Engen, *Odor Sensation and Memory*. Greenwood Publishing Group, 1991.
- [176] C. Assaiante and B. Amblard, "Peripheral vision and age-related differences in dynamic balance," *Hum. Mov. Sci.*, vol. 11, no. 5, pp. 533–548, 1992.
- [177] R. Collings, J. Paton, S. Glasser, and J. Marsden, "The effect of vision impairment on dynamic balance," *J. Foot Ankle Res.*, vol. 8, no. Suppl 1, p. A6, 2015.
- [178] H. Wallach, "The role of head movements and vestibular and visual cues in sound localization.," *J. Exp. Psychol.*, vol. 27, no. 4, p. 339, 1940.
- [179] G. Corfas and Y. Dudai, "Adaptation and fatigue of a mechanosensory neuron in wild-type *Drosophila* and in memory mutants," *J. Neurosci.*, vol. 10, no. 2, pp. 491–499, Feb. 1990.
- [180] A. Stocker and E. P. Simoncelli, "Sensory adaptation within a Bayesian framework for perception," *Adv. Neural Inf. Process. Syst.*, vol. 18, p. 1289, 2006.
- [181] R. H and G. SINGER, "SENSORY ADAPTATION AND BEHAVIORAL COMPENSATION WITH SPATIALLY TRANSFORMED VISION AND HEARING," *Psychol. Bull.*, vol. 67, no. 5, pp. 307–322, 1967.
- [182] L. A. Westerman and R. L. Smith, "Rapid and short-term adaptation in auditory nerve responses," *Hear. Res.*, vol. 15, no. 3, pp. 249–260, Sep. 1984.
- [183] S. B. Laughlin, "The role of sensory adaptation in the retina," *J. Exp. Biol.*, vol. 146, no. 1, pp. 39–62, Sep. 1989.
- [184] J. V. M. Hanson, J. Heron, and D. Whitaker, "Recalibration of perceived time across sensory modalities," *Exp. Brain Res.*, vol. 185, no. 2, pp. 347–352, Jan. 2008.
- [185] P. Carruthers, "Human knowledge and human nature: a new introduction to an ancient debate," 1992.
- [186] K. Lehrer, "Theory of knowledge," 2000.

-
- [187] E. Dietrich, A. B. Markman, and others, *Cognitive dynamics: Conceptual and representational change in humans and machines*. Psychology Press, 2014.
- [188] T. S. Hyde and J. J. Jenkins, “Differential effects of incidental tasks on the organization of recall of a list of highly associated words,” *J. Exp. Psychol.*, vol. 82, no. 3, p. 472, 1969.
- [189] L. Hasher and R. T. Zacks, “Automatic and effortful processes in memory,” *J. Exp. Psychol. Gen.*, vol. 108, no. 3, p. 356, 1979.
- [190] R. B. Tootell, M. S. Silverman, E. Switkes, and R. L. De Valois, “Deoxyglucose analysis of retinotopic organization in primate striate cortex,” *Science*, vol. 218, no. 4575, pp. 902–904, 1982.
- [191] M. J. Farah, M. J. Soso, and R. M. Dasheiff, “Visual angle of the mind’s eye before and after unilateral occipital lobectomy,” *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 18, no. 1, p. 241, 1992.
- [192] S. M. Kosslyn, *Image and mind*. Harvard University Press, 1980.
- [193] M. J. Farah, “The neurological basis of mental imagery: A componential analysis,” *Cognition*, vol. 18, no. 1, pp. 245–272, 1984.
- [194] R. N. Shepard and L. A. Cooper, *Mental images and their transformations*. The MIT Press, 1986.
- [195] R. A. Finke, *Principles of mental imagery*. The MIT Press, 1989.
- [196] S. M. Kosslyn, *Image and brain: The resolution of the imagery debate*. MIT press, 1996.
- [197] S. M. Kosslyn, “Mental images and the Brain,” *Cogn. Neuropsychol.*, vol. 22, no. 3–4, pp. 333–347, May 2005.
- [198] S. M. Kosslyn, W. L. Thompson, and G. Ganis, *The case for mental imagery*. Oxford University Press, 2006.
- [199] J. M. Henderson and A. Hollingworth, “Eye movements and visual memory: Detecting changes to saccade targets in scenes,” *Percept. Psychophys.*, vol. 65, no. 1, pp. 58–71, 2003.
- [200] J. Hochberg, “Gestalt theory and its legacy,” *Percept. Cogn. Century’s End*, pp. 253–306, 1998.
- [201] J. F. Kihlstrom, “Joseph Jastrow and his duck—or is it a rabbit,” *Trends Cogn. Sci. TICS Last Visit.*, vol. 31, no. 7, p. 2007, 2004.
- [202] J. Y. Lettvin, H. R. Maturana, W. S. McCulloch, and W. H. Pitts, “What the frog’s eye tells the frog’s brain,” *Proc. IRE*, vol. 47, no. 11, pp. 1940–1951, 1959.
- [203] R. E. Crist, W. Li, and C. D. Gilbert, “Learning to see: experience and attention in primary visual cortex,” *Nat. Neurosci.*, vol. 4, no. 5, pp. 519–525, 2001.
- [204] M. F. Bear, B. W. Connors, and M. A. Paradiso, *Neuroscience*, vol. 2. Lippincott Williams & Wilkins, 2007.
- [205] P. Smolensky, “On the proper treatment of connectionism,” *Behav. Brain Sci.*, vol. 11, no. 1, pp. 1–23, 1988.
- [206] J. S. Bruner, “Going beyond the information given,” *Contemp. Approaches Cogn.*, vol. 1, pp. 119–160, 1957.

- [207] D. L. Hintzman, "MINERVA 2: A simulation model of human memory," *Behav. Res. Methods Instrum. Comput.*, vol. 16, no. 2, pp. 96–101, 1984.
- [208] D. L. Hintzman, "Judgments of frequency and recognition memory in a multiple-trace memory model," *Psychol. Rev.*, vol. 95, no. 4, p. 528, 1988.
- [209] G. S. Cree and K. McRae, "Analyzing the factors underlying the structure and computation of the meaning of chipmunk, cherry, chisel, cheese, and cello (and many other such concrete nouns)," *J. Exp. Psychol. Gen.*, vol. 132, no. 2, p. 163, 2003.
- [210] A. R. Damasio, "Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition," *Cognition*, vol. 33, no. 1, pp. 25–62, 1989.
- [211] W. K. Simmons and L. W. Barsalou, "The similarity-in-topography principle: Reconciling theories of conceptual deficits," *Cogn. Neuropsychol.*, vol. 20, no. 3–6, pp. 451–486, 2003.
- [212] J. Tanaka and I. Gauthier, "Expertise in object and face recognition," *Psychol. Learn. Motiv.*, vol. 36, pp. 83–125, 1997.
- [213] K. Tanaka, "Inferotemporal cortex and object recognition," *Adv. Psychol.*, vol. 121, pp. 160–188, 1997.
- [214] L. R. Brooks, "Nonanalytic Concept Formation and Memory for Instances," in *Cognition and Categorization*, E. Rosch and B. Lloyd, Eds. Lawrence Erlbaum Associates, 1978, pp. 3–170.
- [215] R. M. Nosofsky, "Choice, similarity, and the context theory of classification," *J. Exp. Psychol. Learn. Mem. Cogn.*, vol. 10, no. 1, p. 104, 1984.
- [216] S. W. Allen and L. R. Brooks, "Specializing the operation of an explicit rule," *J. Exp. Psychol. Gen.*, vol. 120, no. 1, p. 3, 1991.
- [217] A. L. Patalano, E. E. Smith, J. Jonides, and R. A. Koeppel, "PET evidence for multiple strategies of categorization," *Cogn. Affect. Behav. Neurosci.*, vol. 1, no. 4, pp. 360–370, 2001.
- [218] F. G. Ashby and S. W. Ell, "The neurobiology of human category learning," *Trends Cogn. Sci.*, vol. 5, no. 5, pp. 204–210, 2001.
- [219] E. E. Smith and D. L. Medin, *Categories and concepts*. Harvard University Press Cambridge, MA, 1981.
- [220] L. W. Barsalou, "On the indistinguishability of exemplar memory and abstraction in category representation," *Content Process Specif. Eff. Prior Exp. Adv. Soc. Cogn.*, vol. 3, pp. 61–88, 1990.
- [221] J. D. Smith and J. P. Minda, "Distinguishing prototype-based and exemplar-based processes in dot-pattern category learning," *J. Exp. Psychol. Learn. Mem. Cogn.*, vol. 28, no. 4, p. 800, 2002.
- [222] E. H. Rosch, "On the internal structure of perceptual and semantic categories," in *Cognitive development and the acquisition of language*, Oxford, England: Academic Press, 1973, p. xii, 308.
- [223] L. W. Barsalou, "The instability of graded structure: Implications for the nature of concepts," 1987.

-
- [224] S. L. Armstrong, L. R. Gleitman, and H. Gleitman, "What some concepts might not be," *Cognition*, vol. 13, no. 3, pp. 263–308, 1983.
- [225] E. Rosch, "Cognitive representations of semantic categories.," *J. Exp. Psychol. Gen.*, vol. 104, no. 3, p. 192, 1975.
- [226] M. I. Posner and S. W. Keele, "On the genesis of abstract ideas.," *J. Exp. Psychol.*, vol. 77, no. 3p1, p. 353, 1968.
- [227] C. B. Mervis and J. R. Pani, "Acquisition of basic object categories," *Cognit. Psychol.*, vol. 12, no. 4, pp. 496–522, 1980.
- [228] D. N. Osherson, E. E. Smith, O. Wilkie, A. Lopez, and E. Shafir, "Category-based induction.," *Psychol. Rev.*, vol. 97, no. 2, p. 185, 1990.
- [229] N. Goodman, *Fact, fiction, and forecast*. Harvard University Press, 1983.
- [230] G. L. Murphy and D. L. Medin, "The role of theories in conceptual coherence.," *Psychol. Rev.*, vol. 92, no. 3, p. 289, 1985.
- [231] W. Ahn and C. C. Luhmann, "Demystifying theory-based categorization," *Build. Object Categ. Dev. Time*, pp. 277–300, 2005.
- [232] L. W. Barsalou, "Ad hoc categories," *Mem. Cognit.*, vol. 11, no. 3, pp. 211–227, 1983.
- [233] D. E. Rumelhart and D. A. Norman, "Representation in Memory.," Jun. 1983.
- [234] L. W. Barsalou, "Frames, concepts, and conceptual fields.," 1992.
- [235] L. Barsalou and C. Hale, "Components of conceptual representation. From feature lists to recursive frames," 1993.
- [236] S. L. Greenspan, "Semantic flexibility and referential specificity of concrete nouns," *J. Mem. Lang.*, vol. 25, no. 5, pp. 539–557, 1986.
- [237] P. Tabossi, "Effects of context on the immediate interpretation of unambiguous nouns.," *J. Exp. Psychol. Learn. Mem. Cogn.*, vol. 14, no. 1, p. 153, 1988.
- [238] L. Squire, "Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory," *Cogn. Neurosci. J. Of*, vol. 4, no. 3, pp. 232–243, 1992.
- [239] E. R. Kandel, J. H. Schwartz, T. M. Jessell, and others, *Principles of neural science*, vol. 4. McGraw-Hill New York, 2000.
- [240] E. Tulving, "Episodic and semantic memory," *Organ. Mem. Lond. Acad.*, vol. 381, no. e402, p. 4, 1972.
- [241] R. L. Klatzky, *Human memory: Structures and processes*. Oxford, England: W. H. Freeman, 1975.
- [242] P. Robinson, "Attention, Memory, and the 'Noticing' Hypothesis," *Lang. Learn.*, vol. 45, no. 2, pp. 283–331, Jun. 1995.
- [243] F. I. Craik, R. Govoni, M. Naveh-Benjamin, and N. D. Anderson, "The effects of divided attention on encoding and retrieval processes in human memory," *J. Exp. Psychol. Gen.*, vol. 125, no. 2, pp. 159–180, 1996.
- [244] N. B. Turk-Browne, D.-J. Yi, and M. M. Chun, "Linking Implicit and Explicit Memory: Common Encoding Factors and Shared Representations," *Neuron*, vol. 49, no. 6, pp. 917–927, Mar. 2006.
-

- [245] N. Cowan, *Attention and Memory*. Oxford University Press, 1997.
- [246] T. Shallice, P. Fletcher, C. Frith, P. Grasby, R. Frackowiak, and R. Dolan, "Brain regions associated with acquisition and retrieval of verbal episodic memory," *Nature*, vol. 368, no. 6472, pp. 633–635, 1994.
- [247] M. R. Uncapher and M. D. Rugg, "Effects of divided attention on fMRI correlates of memory encoding," *J. Cogn. Neurosci.*, vol. 17, no. 12, pp. 1923–1935, 2005.
- [248] F. I. Craik and R. S. Lockhart, "Levels of processing: A framework for memory research," *J. Verbal Learn. Verbal Behav.*, vol. 11, no. 6, pp. 671–684, 1972.
- [249] N. J. Slamecka and P. Graf, "The generation effect: Delineation of a phenomenon," *J. Exp. Psychol. [Hum. Learn.]*, vol. 4, no. 6, p. 592, 1978.
- [250] H. Ebbinghaus, *Memory: A contribution to experimental psychology*. Teachers college, Columbia university, 1913.
- [251] S. E. Petersen, P. T. Fox, M. I. Posner, M. Mintun, and M. E. Raichle, "Positron emission tomographic studies of the cortical anatomy of single-word processing," *Nature*, vol. 331, no. 6157, pp. 585–589, 1988.
- [252] E. Tulving, "Elements of episodic memory," *Oxf. Univ. Press*, 1985.
- [253] L. R. Squire, C. E. L. Stark, and R. E. Clark, "The Medial Temporal Lobe," *Annu. Rev. Neurosci.*, vol. 27, no. 1, pp. 279–306, 2004.
- [254] M. A. Wilson and B. L. McNaughton, "Reactivation of hippocampal ensemble memories during sleep," *Science*, vol. 265, no. 5172, pp. 676–679, Jul. 1994.
- [255] J. L. McClelland, B. L. McNaughton, and R. C. O'Reilly, "Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionist models of learning and memory," *Psychol. Rev.*, vol. 102, no. 3, pp. 419–457, 1995.
- [256] J. L. McGaugh, "Memory--a Century of Consolidation," *Science*, vol. 287, no. 5451, pp. 248–251, Jan. 2000.
- [257] K. Nakazawa *et al.*, "Requirement for Hippocampal CA3 NMDA Receptors in Associative Memory Recall," *Science*, vol. 297, no. 5579, pp. 211–218, Jul. 2002.
- [258] D. R. Godden and A. D. Baddeley, "Context-Dependent Memory in Two Natural Environments: On Land and Underwater," *Br. J. Psychol.*, vol. 66, no. 3, pp. 325–331, Aug. 1975.
- [259] J. E. Eich, H. Weingartner, R. C. Stillman, and J. C. Gillin, "State-dependent accessibility of retrieval cues in the retention of a categorized list," *J. Verbal Learn. Verbal Behav.*, vol. 14, no. 4, pp. 408–417, Aug. 1975.
- [260] J. E. Eich, "The cue-dependent nature of state-dependent retrieval," *Mem. Cognit.*, vol. 8, no. 2, pp. 157–173, Mar. 1980.
- [261] A. P. Yonelinas, "The Nature of Recollection and Familiarity: A Review of 30 Years of Research," *J. Mem. Lang.*, vol. 46, no. 3, pp. 441–517, Apr. 2002.
- [262] D. L. Hintzman and T. Curran, "Retrieval Dynamics of Recognition and Frequency Judgments: Evidence for Separate Processes of Familiarity and Recall," *J. Mem. Lang.*, vol. 33, no. 1, pp. 1–18, Feb. 1994.

-
- [263] A. P. Yonelinas and L. L. Jacoby, "Dissociations of processes in recognition memory: Effects of interference and of response speed," *Can. J. Exp. Psychol. Can. Psychol. Expérimentale*, vol. 48, no. 4, pp. 516–535, 1994.
- [264] C. S. Dodson and M. K. Johnson, "Some problems with the process-dissociation approach to memory," *J. Exp. Psychol. Gen.*, vol. 125, no. 2, pp. 181–194, 1996.
- [265] L. L. Jacoby, J. P. Toth, and A. P. Yonelinas, "Separating conscious and unconscious influences of memory: Measuring recollection," *J. Exp. Psychol. Gen.*, vol. 122, no. 2, pp. 139–154, 1993.
- [266] D. L. Schacter, *The Seven Sins of Memory: How the Mind Forgets and Remembers*. Houghton Mifflin Harcourt, 2002.
- [267] R. Buckner and D. L. Schacter, *The Cognitive Neuroscience, Third Edition*. Cambridge: MIT Press, 2004.
- [268] "perceptual computing Definition in the Cambridge English Dictionary." [Online]. Available: <http://dictionary.cambridge.org/us/dictionary/english/perceptual-computing>. [Accessed: 30-Nov-2016].
- [269] D. Vernon, G. Metta, and G. Sandini, "A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents," *IEEE Trans. Evol. Comput.*, vol. 11, no. 2, p. 151, 2007.
- [270] D. E. Kieras and D. E. Meyer, "An overview of the EPIC architecture for cognition and performance with application to human-computer interaction," *Hum.-Comput. Interact.*, vol. 12, no. 4, pp. 391–438, 1997.
- [271] J. R. Anderson, D. Bothell, M. D. Byrne, S. Douglass, C. Lebiere, and Y. Qin, "An integrated theory of the mind.," *Psychol. Rev.*, vol. 111, no. 4, p. 1036, 2004.
- [272] J. R. Anderson, "ACT: A simple theory of complex cognition.," *Am. Psychol.*, vol. 51, no. 4, p. 355, 1996.
- [273] P. Langley, "An cognitive architectures and the construction of intelligent agents," in *Proceedings of the AAAI-2004 Workshop on Intelligent Agent Architectures*, 2004, p. 82.
- [274] D. Choi, M. Kaufman, P. Langley, N. Nejati, and D. Shapiro, "An architecture for persistent reactive behavior," in *Proceedings of the Third International Joint Conference on Autonomous Agents and Multiagent Systems-Volume 2*, 2004, pp. 988–995.
- [275] P. Langley, "Cognitive architectures and general intelligent systems," *AI Mag.*, vol. 27, no. 2, p. 33, 2006.
- [276] D. P. Benjamin, D. M. Lyons, and D. W. Lonsdale, "ADAPT: A Cognitive Architecture for Robotics.," in *ICCM*, 2004, pp. 337–338.
- [277] M. Shanahan and B. Baars, "Applying global workspace theory to the frame problem," *Cognition*, vol. 98, no. 2, pp. 157–176, 2005.
- [278] M. Shanahan, "Consciousness, Emotion, and Imagination A Brain-Inspired Architecture for Cognitive Robotics. Aisb 2005, 26-35," in *Proceedings of the AISB'05 Workshop: Next Generation Approaches to Machine Consciousness*. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary>, 2005.

- [279] M. Shanahan, “A cognitive architecture that combines internal simulation with a global workspace,” *Conscious. Cogn.*, vol. 15, no. 2, pp. 433–449, 2006.
- [280] W. D. Christensen and C. A. Hooker, “An interactivist-constructivist approach to intelligence: self-directed anticipative learning,” *Philos. Psychol.*, vol. 13, no. 1, pp. 5–45, 2000.
- [281] J. Weng, W. S. Hwang, Y. Zhang, C. Yang, and R. Smith, “Developmental humanoids: Humanoids that develop skills automatically,” in *Proc. The First IEEE-RAS International Conference on Humanoid Robots*, 2000, vol. 78.
- [282] J. Weng, “Developmental robotics: Theory and experiments,” *Int. J. Humanoid Robot.*, vol. 1, no. 2, pp. 199–236, 2004.
- [283] J. Weng, “A theory of developmental architecture,” in *Proceedings of the 3rd International Conference on Development and Learning (ICDL 2004)*, 2004, pp. 357–362.
- [284] J. L. Krichmar, D. A. Nitz, J. A. Gally, and G. M. Edelman, “Characterizing functional hippocampal pathways in a brain-based device as it solves a spatial memory task,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 6, pp. 2111–2116, 2005.
- [285] A. K. Seth, J. L. McKinstry, G. M. Edelman, and J. L. Krichmar, “Active sensing of visual and tactile stimuli by brain-based devices,” *Int. J. Robot. Autom.*, vol. 19, no. 4, pp. 222–238, 2004.
- [286] J. Krichmar and G. Reeke, “The darwin brain-based automata: Synthetic neural models and real-world devices,” *Model. Neurosci. Biol. Syst. Neuromimetic Robot.*, pp. 613–638, 2005.
- [287] J. L. Krichmar and G. M. Edelman, “Brain-based devices for the study of nervous systems and the development of intelligent machines,” *Artif. Life*, vol. 11, no. 1–2, pp. 63–77, 2005.
- [288] J. L. Krichmar, A. K. Seth, D. A. Nitz, J. G. Fleischer, and G. M. Edelman, “Spatial navigation and causal analysis in a brain-based device modeling cortical-hippocampal interactions,” *Neuroinformatics*, vol. 3, no. 3, pp. 197–221, 2005.
- [289] C. Burghart *et al.*, “A cognitive architecture for a humanoid robot: A first approach,” in *5th IEEE-RAS International Conference on Humanoid Robots, 2005.*, 2005, pp. 357–362.
- [290] I. Horswill, “Cerebus: A higher-order behavior-based system,” *AI Mag.*, vol. 23, no. 1, pp. 27–27, 2002.
- [291] R. A. Brooks, C. Breazeal, M. Marjanović, B. Scassellati, and M. M. Williamson, “The Cog project: Building a humanoid robot,” in *Computation for metaphors, analogy, and agents*, Springer, 1999, pp. 52–87.
- [292] A. M. Leslie, “ToMM, ToBy, and Agency: Core architecture and domain specificity,” *Mapp. Mind Domain Specif. Cogn. Cult.*, pp. 119–148, 1994.
- [293] S. Baron-Cohen, *Mindblindness: An essay on autism and theory of mind*. MIT press, 1997.
- [294] C. L. Breazeal, “Sociable machines: Expressive social exchange between humans and robots,” Massachusetts Institute of Technology, 2000.

-
- [295] C. Breazeal, “Emotion and sociable humanoid robots,” *Int. J. Hum.-Comput. Stud.*, vol. 59, no. 1, pp. 119–155, 2003.
- [296] Y. LeCun *et al.*, “Backpropagation applied to handwritten zip code recognition,” *Neural Comput.*, vol. 1, no. 4, pp. 541–551, 1989.
- [297] S. Lawrence, C. L. Giles, A. C. Tsoi, and A. D. Back, “Face recognition: A convolutional neural-network approach,” *Neural Netw. IEEE Trans. On*, vol. 8, no. 1, pp. 98–113, 1997.
- [298] M. Matsugu, K. Mori, Y. Mitari, and Y. Kaneda, “Subject independent facial expression recognition with robust face detection using a convolutional neural network,” *Neural Netw.*, vol. 16, no. 5, pp. 555–559, 2003.
- [299] P. Le Callet, C. Viard-Gaudin, and D. Barba, “A convolutional neural network approach for objective video quality assessment,” *Neural Netw. IEEE Trans. On*, vol. 17, no. 5, pp. 1316–1327, 2006.
- [300] S. Marsella and J. Gratch, “Computationally Modeling Human Emotion,” *Commun ACM*, vol. 57, no. 12, pp. 56–67, Nov. 2014.
- [301] Y. Bengio, “Learning deep architectures for AI,” *Found. Trends® Mach. Learn.*, vol. 2, no. 1, pp. 1–127, 2009.
- [302] J. Schmidhuber, “Deep Learning in Neural Networks: An Overview,” *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.
- [303] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
- [304] M. Zeiler *et al.*, *Clarifai*. New York, NY, USA: Clarifai, 2015.
- [305] E. Berry, S. Zacharia, and A. Lewine, *TripleLift*. New York, NY, USA: TripleLift, 2015.
- [306] C. F. Cadieu *et al.*, “Deep Neural Networks Rival the Representation of Primate IT Cortex for Core Visual Object Recognition,” *PLoS Comput. Biol.*, vol. 10, no. 12, p. 35, 2014.
- [307] A. Wang, “The Shazam Music Recognition Service,” *Commun ACM*, vol. 49, no. 8, pp. 44–48, Aug. 2006.
- [308] A. S. Master, T. P. Stonehocker, B. J. Levitt, J. Huang, and K. Mohajer, *Systems and methods for sound recognition*. Google Patents, 2016.
- [309] C. Brinkman, M. Fragkiadakis, and X. Bos, “Online music recognition: the Echoprint system,” 2016.
- [310] K. Noda, Y. Yamaguchi, K. Nakadai, H. G. Okuno, and T. Ogata, “Audio-visual speech recognition using deep learning,” *Appl. Intell.*, vol. 42, no. 4, pp. 722–737, Jun. 2015.
- [311] R. Collobert and J. Weston, “A unified architecture for natural language processing: Deep neural networks with multitask learning,” in *Proceedings of the 25th international conference on Machine learning*, 2008, pp. 160–167.
- [312] A. Kumar *et al.*, “Ask me anything: Dynamic memory networks for natural language processing,” *ArXiv Prepr. ArXiv150607285*, 2015.

- [313] J. Hirschberg and C. D. Manning, “Advances in natural language processing,” *Science*, vol. 349, no. 6245, pp. 261–266, 2015.
- [314] C. Clavel and Z. Callejas, “Sentiment analysis: from opinion mining to human-agent interaction,” *IEEE Trans. Affect. Comput.*, vol. 7, no. 1, pp. 74–93, 2016.
- [315] K. Ravi and V. Ravi, “A survey on opinion mining and sentiment analysis: Tasks, approaches and applications,” *Knowl.-Based Syst.*, vol. 89, pp. 14–46, 2015.
- [316] “Gartner’s 2016 Hype Cycle for Emerging Technologies Identifies Three Key Trends That Organizations Must Track to Gain Competitive Advantage.” [Online]. Available: <http://www.gartner.com/newsroom/id/3412017>. [Accessed: 01-Dec-2016].
- [317] J. Fisher, “Who’s set to make money from the coming artificial intelligence boom?,” *Business Insider*, 23-Feb-2015. [Online]. Available: <http://www.businessinsider.com/artificial-intelligence-how-to-invest-2015-2>. [Accessed: 01-Dec-2016].
- [318] K. Kelly, “The Three Breakthroughs That Have Finally Unleashed AI on the World,” *WIRED*, 27-Oct-2014. [Online]. Available: <https://www.wired.com/2014/10/future-of-artificial-intelligence/>. [Accessed: 01-Dec-2016].
- [319] E. Jhonsa, “Nvidia’s deals with IBM, Microsoft and Google strengthen its artificial intelligence lead,” *TheStreet*, 16-Nov-2016. [Online]. Available: <https://www.thestreet.com/story/13896777/1/nvidia-s-deals-with-ibm-microsoft-and-google-strengthen-its-artificial-intelligence-lead.html>. [Accessed: 01-Dec-2016].
- [320] A. Hern, “‘Partnership on AI’ formed by Google, Facebook, Amazon, IBM and Microsoft,” *The Guardian*, 28-Sep-2016.
- [321] “IBM Watson,” *IBM Watson*. [Online]. Available: <https://www.ibm.com/watson/>. [Accessed: 01-Dec-2016].
- [322] “IBM - DeepQA Project,” 22-Apr-2009. [Online]. Available: <https://www.research.ibm.com/deepqa/deepqa.shtml>. [Accessed: 01-Dec-2016].
- [323] D. A. Ferrucci, “Introduction to ‘this is watson,’” *IBM J. Res. Dev.*, vol. 56, no. 3.4, pp. 1–1, 2012.
- [324] “Products & Services | Google Cloud Platform,” *Google Cloud Platform*, Nov-2016. [Online]. Available: <https://cloud.google.com/products/>. [Accessed: 01-Dec-2016].
- [325] “Microsoft Cognitive Services - APIs.” [Online]. Available: <https://www.microsoft.com/cognitive-services/en-us/apis>. [Accessed: 01-Dec-2016].
- [326] “TensorFlow — an Open Source Software Library for Machine Intelligence.” [Online]. Available: <https://www.tensorflow.org/>. [Accessed: 01-Dec-2016].
- [327] “Cloud Dataflow - Batch & Stream Data Processing,” *Google Cloud Platform*. [Online]. Available: <https://cloud.google.com/dataflow/>. [Accessed: 01-Dec-2016].
- [328] “Cloud Storage - Online Data Storage,” *Google Cloud Platform*. [Online]. Available: <https://cloud.google.com/storage/>. [Accessed: 01-Dec-2016].

-
- [329] “BigQuery - Analytics Data Warehouse,” *Google Cloud Platform*. [Online]. Available: <https://cloud.google.com/bigquery/>. [Accessed: 01-Dec-2016].
- [330] M. Masse, *REST API design rulebook*. O’Reilly Media, Inc., 2011.
- [331] L. Richardson and S. Ruby, *RESTful web services*. O’Reilly Media, Inc., 2008.
- [332] J. Webber, S. Parastatidis, and I. Robinson, *REST in practice: Hypermedia and systems architecture*. O’Reilly Media, Inc., 2010.
- [333] R. DeMaria and J. L. Wilson, *High score!: the illustrated history of electronic games*. McGraw-Hill Osborne Media, 2003.
- [334] S. Okada, *Video target control and sensing circuit for photosensitive gun*. Google Patents, 1989.
- [335] C. A. Wingrave *et al.*, “The wiimote and beyond: Spatially convenient devices for 3d user interfaces,” *IEEE Comput. Graph. Appl.*, vol. 30, no. 2, pp. 71–85, 2010.
- [336] “PS4™ Technical Specifications - PlayStation®4 System,” *us.playstation.com*. [Online]. Available: <http://us.playstation.com/ps4/features/techspecs/index.htm>. [Accessed: 01-Dec-2016].
- [337] “Kinect for Xbox One,” *Xbox.com*. [Online]. Available: <http://www.xbox.com/en-US/xbox-one/accessories/kinect>. [Accessed: 01-Dec-2016].
- [338] “Kinect for Windows SDK 2.0,” *Microsoft Download Center*. [Online]. Available: <https://www.microsoft.com/en-us/download/details.aspx?id=44561>. [Accessed: 01-Dec-2016].
- [339] Microsoft, “Microsoft HoloLens,” *Microsoft HoloLens*. [Online]. Available: <https://www.microsoft.com/microsoft-hololens/en-us/hardware>. [Accessed: 01-Dec-2016].
- [340] “VIVE™ | Buy Vive Hardware.” [Online]. Available: <https://www.vive.com/us/product/>. [Accessed: 01-Dec-2016].
- [341] “Intel® RealSense™ Technology,” *Intel*. [Online]. Available: <http://www.intel.com/content/www/us/en/architecture-and-technology/realsense-overview.html>. [Accessed: 01-Dec-2016].
- [342] P. Mell and T. Grance, “The NIST definition of cloud computing,” *Natl. Inst. Stand. Technol.*, vol. 53, no. 6, p. 50, 2009.
- [343] “Virtualization Technology & Virtual Machine Software.” [Online]. Available: <http://www.vmware.com/solutions/virtualization.html>. [Accessed: 02-Dec-2016].
- [344] S. Angeles, B. S. W. J. 20, and 2014 06:34 am EST, “Virtualization vs. Cloud Computing: What’s the Difference?,” *Business News Daily*, 20-Jan-2014. [Online]. Available: <http://www.businessnewsdaily.com/5791-virtualization-vs-cloud-computing.html>. [Accessed: 02-Dec-2016].
- [345] “Oracle VM VirtualBox.” [Online]. Available: <https://www.virtualbox.org/>. [Accessed: 02-Dec-2016].
- [346] “Parallels Virtualization.” [Online]. Available: <http://www.parallels.com/es/all-products/>. [Accessed: 02-Dec-2016].

- [347] “VMware Products - Workstation.” [Online]. Available: <http://www.vmware.com/products/workstation.html>. [Accessed: 02-Dec-2016].
- [348] “XenServer – Server Virtualization and Consolidation - Citrix,” *Citrix.com*. [Online]. Available: <https://www.citrix.com/products/xenserver/>. [Accessed: 02-Dec-2016].
- [349] “Server Virtualization with VMware vSphere.” [Online]. Available: <http://www.vmware.com/products/vsphere.html>. [Accessed: 02-Dec-2016].
- [350] “Server virtualization management with Proxmox VE.” [Online]. Available: <http://www.proxmox.com/en/proxmox-ve>. [Accessed: 02-Dec-2016].
- [351] “Microsoft Azure: Cloud Computing Platform & Services.” [Online]. Available: <https://azure.microsoft.com/en-us/>. [Accessed: 02-Dec-2016].
- [352] “Amazon Web Services - Elastic compute cloud (EC2),” *Amazon Web Services, Inc.* [Online]. Available: <http://aws.amazon.com/es/ec2/>. [Accessed: 02-Dec-2016].
- [353] M. J. Scheepers, “Virtualization and containerization of application infrastructure: A comparison,” *Univ. Twente Enschede*, 2014.
- [354] “Docker,” *Docker*. [Online]. Available: <https://www.docker.com/>. [Accessed: 02-Dec-2016].
- [355] J. R. Okajima, “Aufs4 - Advanced Multi Layered Unification Filesystem version 4.x.” [Online]. Available: <http://aufs.sourceforge.net/>. [Accessed: 03-Dec-2016].
- [356] “Raspberry Pi DockerCon Challenge,” *Docker Blog*, 15-Sep-2015. [Online]. Available: <http://dockr.ly/2f3Jyvz>. [Accessed: 03-Dec-2016].
- [357] “HypriotOS - About Us. Docker pirates.” [Online]. Available: <http://blog.hypriot.com/about/>. [Accessed: 03-Dec-2016].
- [358] S. F. Piraghaj, A. V. Dastjerdi, R. N. Calheiros, and R. Buyya, “Efficient virtual machine sizing for hosting containers as a service (services 2015),” in *2015 IEEE World Congress on Services*, 2015, pp. 31–38.
- [359] “Docker Swarm,” *Docker*, 03-Dec-2016. [Online]. Available: <https://docs.docker.com/swarm/>. [Accessed: 03-Dec-2016].
- [360] “Kubernetes - Production-Grade Container Orchestration.” [Online]. Available: <http://kubernetes.io/>. [Accessed: 03-Dec-2016].
- [361] “Apache Mesos,” *Apache Mesos*. [Online]. Available: <http://mesos.apache.org/>. [Accessed: 03-Dec-2016].
- [362] “openstack/nova-docker - Git repository,” *GitHub*. [Online]. Available: <https://github.com/openstack/nova-docker>. [Accessed: 03-Dec-2016].
- [363] “Carina by Rackspace - High-performance, instant-on Docker containers.” [Online]. Available: <https://getcarina.com/>. [Accessed: 03-Dec-2016].
- [364] “Container Engine,” *Google Cloud Platform*. [Online]. Available: <https://cloud.google.com/container-engine/>. [Accessed: 03-Dec-2016].
- [365] “AWS - Amazon EC2 Container Service,” *Amazon Web Services, Inc.* [Online]. Available: <http://aws.amazon.com/es/ecs/>. [Accessed: 03-Dec-2016].
- [366] “IBM Bluemix - Containers.” [Online]. Available: <https://www.ibm.com/cloud-computing/bluemix/containers>. [Accessed: 03-Dec-2016].

-
- [367] “Rackspace: Managed Dedicated & Cloud Computing Services,” *Rackspace Hosting*. [Online]. Available: <https://www.rackspace.com/en-gb>. [Accessed: 03-Dec-2016].
- [368] L. Prieto-González, G. Tamm, and V. Stantchev, “Governance of cloud computing: semantic aspects and cloud service brokers,” *Int. J. Web Grid Serv.*, vol. 11, no. 4, pp. 377–389, 2015.
- [369] G. Stoneburner, A. Goguen, and A. Feringa, “Risk management guide for information technology systems,” *Nist Spec. Publ.*, vol. 800, no. 30, pp. 800–30, 2002.
- [370] J. T. FORCE and T. INITIATIVE, “Guide for Applying the Risk Management Framework to Federal Information Systems,” *NIST Spec. Publ.*, vol. 800, p. 37, 2010.
- [371] D. M. Parrilli, “Legal Issues in Grid and cloud computing,” in *Grid and Cloud Computing*, Springer, 2010, pp. 97–118.
- [372] B. R. Kandukuri, V. R. Paturi, and A. Rakshit, “Cloud security issues,” in *Services Computing, 2009. SCC’09. IEEE International Conference on*, 2009, pp. 517–520.
- [373] S. Dzombeta, V. Stantchev, R. Colomo-Palacios, K. Brandis, and K. Haufe, “Governance of Cloud Computing Services for the Life Sciences,” *IT Prof.*, vol. 16, no. 4, pp. 30–37, 2014.
- [374] T. Mather, S. Kumaraswamy, and S. Latif, *Cloud security and privacy: an enterprise perspective on risks and compliance*. O’Reilly, 2009.
- [375] W. Jansen, T. Grance, and others, “Guidelines on security and privacy in public cloud computing,” *NIST Spec. Publ.*, vol. 800, p. 144, 2011.
- [376] S. De Haes and W. Van Grembergen, “IT governance and its mechanisms,” *Inf. Syst. Control J.*, vol. 1, pp. 27–33, 2004.
- [377] B. Nguyen, *A comparison of the business and technical drivers for ISO 27001, ISO 27002, CobiT and ITIL*. Retrieved from trongbang86. blogspot. com: <http://trongbang86.blogspot.com/2010/11/comparison-of-business-and-technical.html>, 2010.
- [378] C. CobiT, “Control Objectives for Information and related Technology,” *IT Gov. Inst. Www Isaca Org*, 2002.
- [379] “ITIL - IT Service Management Books.” [Online]. Available: <http://www.itil.org.uk/>. [Accessed: 03-Dec-2016].
- [380] C. Horwath, W. Chan, E. Leung, and H. Pili, “Enterprise risk management for Cloud Computing,” *COSO*, 2012.
- [381] “COSO - ERM Integrated Framework Update.” [Online]. Available: <http://www.coso.org/ermupdate.html>. [Accessed: 03-Dec-2016].
- [382] ISO, “ISO/IEC DIS 27017.” 22-Apr-2015.
- [383] “ISO 9000:2015 - Quality management systems -- Fundamentals and vocabulary,” *ISO*. [Online]. Available: http://www.iso.org/iso/home/store/catalogue_tc/catalogue_detail.htm?csnumber=45481. [Accessed: 03-Dec-2016].

- [384] D. Niyato, E. Hossain, and S. Camorlinga, “Remote patient monitoring service using heterogeneous wireless access networks: architecture and optimization,” *Sel. Areas Commun. IEEE J. On*, vol. 27, no. 4, pp. 412–423, 2009.
- [385] V. Suryani, A. Rizal, A. Herutomo, M. Abdurrohman, T. Magedanz, and A. Elmangoush, “Electrocardiogram monitoring on OpenMTC platform,” in *Local Computer Networks Workshops (LCN Workshops), 2013 IEEE 38th Conference on*, 2013, pp. 843–847.
- [386] C. Buckl *et al.*, “Services to the field: An approach for resource constrained sensor/actor networks,” in *Advanced Information Networking and Applications Workshops, 2009. WAINA’09. International Conference on*, 2009, pp. 476–481.
- [387] J. Joseph, *Inventory and Anti-Theft Alarm System*. 2015.
- [388] L. Simon and R. Anderson, “Security analysis of consumer-grade anti-theft solutions provided by android mobile anti-virus apps,” in *4th Mobile Security Technologies Workshop (MoST)*, 2015.
- [389] “Internet of Things By The Numbers: Market Estimates And Forecasts,” *Forbes*. [Online]. Available: <http://www.forbes.com/sites/gilpress/2014/08/22/internet-of-things-by-the-numbers-market-estimates-and-forecasts/>. [Accessed: 06-Jun-2015].
- [390] D. Evans, “The internet of things: How the next evolution of the internet is changing everything,” *CISCO White Pap.*, vol. 1, 2011.
- [391] H. Abie and I. Balasingham, “Risk-based Adaptive Security for Smart IoT in eHealth,” in *Proceedings of the 7th International Conference on Body Area Networks*, ICST, Brussels, Belgium, Belgium, 2012, pp. 269–275.
- [392] M. Feldhofer, S. Dominikus, and J. Wolkerstorfer, “Strong authentication for RFID systems using the AES algorithm,” in *Cryptographic Hardware and Embedded Systems-CHES 2004*, Springer, 2004, pp. 357–370.
- [393] B. Calmels, S. Canard, M. Girault, and H. Sibert, “Low-cost cryptography for privacy in RFID systems,” in *Smart Card Research and Advanced Applications*, Springer, 2006, pp. 237–251.
- [394] M. Feldhofer and J. Wolkerstorfer, “Strong crypto for RFID tags-a comparison of low-power hardware implementations,” in *Circuits and Systems, 2007. ISCAS 2007. IEEE International Symposium on*, 2007, pp. 1839–1842.
- [395] L. Eschenauer and V. D. Gligor, “A key-management scheme for distributed sensor networks,” in *Proceedings of the 9th ACM conference on Computer and communications security*, 2002, pp. 41–47.
- [396] A. Perrig, J. Stankovic, and D. Wagner, “Security in wireless sensor networks,” *Commun. ACM*, vol. 47, no. 6, pp. 53–57, 2004.
- [397] D. Macedonio and M. Merro, “A semantic analysis of key management protocols for wireless sensor networks,” *Sci. Comput. Program.*, vol. 81, pp. 53–78, 2014.
- [398] Arduino, “Arduino boards - Hardware index,” 2015. .
- [399] Raspberry Pi Foundation, “Raspberry Pi Products,” *Raspberry Pi*, 2015. .
- [400] Intel Corp., “Intel® Edison Development Platform.” 2015.

-
- [401] J. Horvath and R. Cameron, “Arduino, Raspberry Pi, and Programming Physical Things,” in *The New Shop Class*, Springer, 2015, pp. 17–30.
- [402] W. Anwaar and M. A. Shah, “Energy Efficient Computing: A Comparison of Raspberry PI with Modern Devices,” *Energy*, vol. 4, no. 2, 2015.
- [403] Symantec Corp., “Internet Security Threat Report 2016 - Appendices,” Symantec Corporation, Volume 21, Apr. 2016.
- [404] Symantec Corp., “Internet Security Threat Report 2016,” Symantec Corporation, Volume 21, Apr. 2016.
- [405] Norton, “Norton Cybercrime Index,” 2015. [Online]. Available: <http://us.norton.com/cybercrimeindex>.
- [406] A. Shabtai, D. Mimran, and Y. Elovici, “Evaluation of Security Solutions for Android Systems,” *ArXiv Prepr. ArXiv150204870*, 2015.
- [407] D. Maiorca, D. Ariu, I. Corona, M. Aresu, and G. Giacinto, “Stealth attacks: An extended insight into the obfuscation effects on Android malware,” *Comput. Secur.*, vol. 51, pp. 16–31, Jun. 2015.
- [408] W. Meng, W. H. Lee, S. R. Murali, and S. P. T. Krishnan, “Charging Me and I Know Your Secrets!: Towards Juice Filming Attacks on Smartphones,” in *Proceedings of the 1st ACM Workshop on Cyber-Physical System Security*, New York, NY, USA, 2015, pp. 89–98.
- [409] S. Hwang, S. Lee, Y. Kim, and S. Ryu, “Bittersweet ADB: Attacks and Defenses,” in *Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security*, New York, NY, USA, 2015, pp. 579–584.
- [410] Z. Zhang, P. Liu, J. Xiang, J. Jing, and L. Lei, “How Your Phone Camera Can Be Used to Stealthily Spy on You: Transplantation Attacks Against Android Camera Service,” in *Proceedings of the 5th ACM Conference on Data and Application Security and Privacy*, New York, NY, USA, 2015, pp. 99–110.
- [411] T. Pott, “Ransomware attack hits Synology’s NAS boxen,” 05-Aug-2014. [Online]. Available: http://www.theregister.co.uk/2014/08/05/synologys_synolocker_crisis_its_as_bad_as_you_think/. [Accessed: 08-Jun-2015].
- [412] D. Eric, “Is The Future of Content Protection Cloud(y)?,” *SMPTE Conf.*, vol. 2014, no. 10, pp. 1–17, Oct. 2014.
- [413] D. O’Brien, “The Internet of Things: New Threats Emerge in a Connected World,” *Symantec Security Response*, 20-Jan-2014. [Online]. Available: <http://www.symantec.com/connect/blogs/internet-things-new-threats-emerge-connected-world>. [Accessed: 08-Jun-2015].
- [414] K. Hayashi, “IoT Worm Used to Mine Cryptocurrency,” *Symantec Security Response*, 19-Mar-2014. [Online]. Available: <http://www.symantec.com/connect/blogs/iot-worm-used-mine-cryptocurrency>. [Accessed: 08-Jun-2015].
- [415] Shodan, “Shodan,” 2015. [Online]. Available: <https://www.shodan.io/>.
- [416] Offensive Security, “Kali Linux,” 2015. [Online]. Available: <https://www.kali.org/>.

- [417] B. J. Feder, “A Heart Device Is Found Vulnerable to Hacker Attacks,” *N. Y. Times Available Httpwww Nytimes Com2008031*, vol. 2, 2008.
- [418] U.S. Department of Homeland Security, “Industrial Control System Cyber Emergency Response Team (ICS-CERT) - ALERT-13-164-0,” 29-Oct-2013. [Online]. Available: <https://ics-cert.us-cert.gov/alerts/ICS-ALERT-13-164-01>.
- [419] U.S. Food and Drug Administration (FDA), “Safety Communications - Cybersecurity for Medical Devices and Hospital Networks: FDA Safety Communication,” 13-Jun-2013. [Online]. Available: <http://www.fda.gov/MedicalDevices/Safety/AlertsandNotices/ucm356423.htm>. [Accessed: 08-Jun-2015].
- [420] J. Radcliffe, “Hacking medical devices for fun and insulin: Breaking the human SCADA system,” in *Black Hat Conference presentation slides*, 2011, vol. 2011.
- [421] U.S Department of Homeland Security, “The Industrial Control Systems Cyber Emergency Response Team (ICS-CERT),” 2015. [Online]. Available: <https://ics-cert.us-cert.gov/>.
- [422] Juan Carlos I, *LOPD*, vol. Ley Orgánica. 1999.
- [423] Bundesministerium der Justiz und für Verbraucherschutz, *Bundesdatenschutzgesetz (Federal Data Protection Act)*. 1990, p. 38.
- [424] ISACA, “COBIT 5.” 2014.
- [425] Council on CyberSecurity, “The Critical Security Controls for Effective Cyber Defense v5.1.” 10-Jul-2014.
- [426] Symantec Corp., “An Internet of Things Reference Architecture,” p. 21, Jun. 2015.
- [427] Online Trust Alliance, “Internet of Things,” 28-Jul-2015. [Online]. Available: <https://otalliance.org/initiatives/internet-things>. [Accessed: 02-Dec-2016].
- [428] US Department of Homeland Security, “Security Tenets for Life Critical Embedded Systems,” Mar-2016. [Online]. Available: <https://www.dhs.gov/publication/security-tenets-lces>. [Accessed: 02-Dec-2016].
- [429] T. Ring, “Connected cars—the next target for hackers,” *Netw. Secur.*, vol. 2015, no. 11, pp. 11–16, 2015.
- [430] M. Hazewinkel, *Encyclopaedia of Mathematics*. Springer, 2002.
- [431] J. Łukasiewicz and L. Borkowski, *Selected works*, vol. 59. North-Holland, 1970.
- [432] E. N. Zalta, *Stanford encyclopedia of philosophy*, 2009th ed. Stanford, Calif. : Metaphysics Research Lab, Center for the Study of Language and Information, Stanford University, 1995.
- [433] L. A. Zadeh, “Fuzzy sets,” *Inf. Control*, vol. 8, no. 3, pp. 338–353, 1965.
- [434] B. Smith, “Lógica y ontología formal,” *Univ. Estatal Nueva York En Buffalo Dpto Filos.*, Nov. 2011.
- [435] F. Lehmann, “Semantic networks,” *Comput. Math. Appl.*, vol. 23, no. 2, pp. 1–50, 1992.
- [436] L. K. Schubert, “Extending the expressive power of semantic networks,” *Artif. Intell.*, vol. 7, no. 2, pp. 163–198, 1976.

-
- [437] M. Minsky, “A framework for representing knowledge,” 1974.
- [438] E. Rich and K. Knight, *Inteligencia artificial*. McGraw-Hill, 1994.
- [439] M. R. Quillian, “Word concepts: A theory and simulation of some basic semantic capabilities,” *Behav. Sci.*, vol. 12, no. 5, pp. 410–430, 1967.
- [440] O. R. E. Motlagh, S. H. Tang, and S. M. Homayouni, “A New Strategy for Relationship Modelling of Complex Systems Using Self-Evolving Semantic Networks,” *FEMTEC 2013*, p. 58, 2013.
- [441] W. A. Woods, “What’s in a link: Foundations for semantic networks,” DTIC Document, 1975.
- [442] F. Sebastiani, “Machine learning in automated text categorization,” *ACM Comput. Surv. CSUR*, vol. 34, no. 1, pp. 1–47, 2002.
- [443] P. C. Costa, *Bayesian semantics for the Semantic Web*. George Mason University, 2005.
- [444] G. Weidl, A. Madsen, and S. Israelson, “Applications of object-oriented Bayesian networks for condition monitoring, root cause analysis and decision support on operation of complex continuous processes,” *Comput. Chem. Eng.*, vol. 29, no. 9, pp. 1996–2009, 2005.
- [445] T. Gruber and al., “A translation approach to portable ontology specifications,” *Knowl. Acquis.*, vol. 5, no. 2, pp. 199–220, 1993.
- [446] M. Uschold and M. Gruninger, “Ontologies: Principles, methods and applications,” *Knowl. Eng. Rev.*, vol. 11, no. 2, pp. 93–136, 1996.
- [447] B. Swartout, R. Patil, K. Knight, and T. Russ, “Toward distributed use of large-scale ontologies,” in *Proc. of the Tenth Workshop on Knowledge Acquisition for Knowledge-Based Systems*, 1996.
- [448] N. F. Noy and D. L. McGuinness, “Ontology Development 101,” *Knowl. Syst. Lab. Stanf. Univ.*, vol. 15, no. 2, 2001.
- [449] F. T. Fonseca, M. J. Egenhofer, P. Agouris, and G. Câmara, “Using ontologies for integrated geographic information systems,” *Trans. GIS*, vol. 6, no. 3, pp. 231–257, 2002.
- [450] J. F. Sowa and others, *Knowledge representation: logical, philosophical, and computational foundations*, vol. 13. MIT Press, 2000.
- [451] R. Colomo-Palacios, V. Stantchev, and A. Rodríguez-González, “Special issue on exploiting semantic technologies with particularization on linked data over grid and cloud architectures,” *Future Gener. Comput. Syst.*, vol. 32, pp. 260–262, 2014.
- [452] V. Stantchev, R. Colomo-Palacios, and M. Niedermayer, “Cloud Computing Based Systems for Healthcare,” *Sci. World J.*, vol. 2014, 2014.
- [453] V. Stantchev, K. Petrush, and G. Tamm, “Assessing and governing IT-staff behavior by performance-based simulation,” *Comput Hum Behav*, vol. 29, no. 2, pp. 473–485, Mar. 2013.
- [454] A. Farquhar, R. Fikes, and J. Rice, “The ontolingua server: A tool for collaborative ontology construction,” *Int. J. Hum.-Comput. Stud.*, vol. 46, no. 6, pp. 707–727, 1997.

- [455] S. FIPA00006, “FIPA Ontology Service Specification,” vol. XC00086D, Aug. 2001.
- [456] P. Shvaiko and J. Euzenat, “Ten challenges for ontology matching,” in *On the Move to Meaningful Internet Systems: OTM 2008*, Springer, 2008, pp. 1164–1182.
- [457] W. Gao and T. Xu, “Stability Analysis of Learning Algorithms for Ontology Similarity Computation,” in *Abstract and Applied Analysis*, 2013, vol. 2013.
- [458] E. Simperl, S. Wölger, S. Thaler, B. Norton, and T. Bürger, “Combining human and computation intelligence: the case of data interlinking tools,” *Int. J. Metadata Semant. Ontol.*, vol. 7, no. 2, pp. 77–92, 2012.
- [459] W. E. Djeddi and M. T. Khadir, “Ontology alignment using artificial neural network for large-scale ontologies,” *Int. J. Metadata Semant. Ontol.*, vol. 8, no. 1, pp. 75–92, 2013.
- [460] D. L. McCuinness, “Ontologies come of age,” *Spinn. Semantic Web Bringing World Wide Web Its Full Potential*, p. 171, 2005.
- [461] E. Van Herwijnen, *Practical sgml*. Springer, 1994.
- [462] W3C, *Owl 2 web ontology language document overview*. 2009.
- [463] T. Bray, “The JavaScript Object Notation (JSON) Data Interchange Format.” Internet Engineering Task Force (IETF), Mar-2014.
- [464] “Unicode 9.0.0.” [Online]. Available: <http://www.unicode.org/versions/Unicode9.0.0/>. [Accessed: 03-Dec-2016].
- [465] “MongoDB for GIANT Ideas,” *MongoDB*. [Online]. Available: <https://www.mongodb.com/index>. [Accessed: 03-Dec-2016].
- [466] “JSON for Linking Data Community Group.” .
- [467] “JSON-LD 1.0.” [Online]. Available: <https://www.w3.org/TR/json-ld/>. [Accessed: 03-Dec-2016].
- [468] “JSON-LD Public Repositories,” *GitHub*. [Online]. Available: <https://github.com/json-ld>. [Accessed: 03-Dec-2016].
- [469] “WordNet RDF - WordNet - WordNet RDF.” [Online]. Available: <http://wordnet-rdf.princeton.edu/>. [Accessed: 03-Dec-2016].
- [470] “Google Knowledge Graph Search API - Method entities.search,” *Google Developers*. [Online]. Available: <https://developers.google.com/knowledge-graph/reference/rest/v1/>. [Accessed: 03-Dec-2016].
- [471] V. Nebot, R. Berlanga, J. M. Pérez, M. J. Aramburu, and T. B. Pedersen, “Multidimensional Integrated Ontologies: A Framework for Designing Semantic Data Warehouses,” in *Journal on Data Semantics XIII*, S. Spaccapietra, E. Zimányi, and I.-Y. Song, Eds. Springer Berlin Heidelberg, 2009, pp. 1–36.
- [472] P. Gärdenfors, *Conceptual spaces: The geometry of thought*. MIT press, 2004.
- [473] M. Klein, “Combining and relating ontologies: an analysis of problems and solutions,” in *IJCAI-2001 Workshop on ontologies and information sharing*, 2001, pp. 53–62.
- [474] J. Euzenat and P. Shvaiko, *Ontology Matching*. Springer, 2007.

- [475] C. M. Willis *et al.*, “Olfactory detection of human bladder cancer by dogs: proof of principle study,” *Bmj*, vol. 329, no. 7468, p. 712, 2004.
- [476] B. Buszewski, J. Rudnicka, T. Ligor, M. Walczak, T. Jezierski, and A. Amann, “Analytical and unconventional methods of cancer detection using odor,” *TrAC Trends Anal. Chem.*, vol. 38, pp. 1–12, 2012.
- [477] D. L. Wells, S. W. Lawson, and A. N. Siriwardena, “Canine responses to hypoglycemia in patients with type 1 diabetes,” *J. Altern. Complement. Med.*, vol. 14, no. 10, pp. 1235–1241, 2008.
- [478] N. Maillot, M. Thonnat, and A. Boucher, “Towards ontology-based cognitive vision,” *Mach. Vis. Appl.*, vol. 16, no. 1, pp. 33–40, 2004.
- [479] S. Tongphu, B. Suntisrivaraporn, B. Uyyanonvara, and M. N. Dailey, “Ontology-based object recognition of car sides,” in *Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 2012 9th International Conference on*, 2012, pp. 1–4.
- [480] E. Griffith, “Algorithms that can ‘see’ images on the Web,” *Fortune Online Ed*, Jan. 2015.
- [481] V. Stantchev and G. Tamm, “Reducing Information Asymmetry in Cloud Marketplaces,” *Int. J. Hum. Cap. Inf. Technol. Prof. IJHCITP*, vol. 3, no. 4, pp. 1–10, 2012.
- [482] S. Dzombeta, V. Stantchev, R. Colomo-Palacios, K. Brandis, and K. Haufe, “Governance of Cloud Computing Services for the Life Sciences,” *IT Prof.*, vol. 16, no. 4, pp. 30–37, 2014.
- [483] V. Stantchev, *Enhancing Health Care Services with Mixed Reality Systems*. Berlin, New York: Springer, 2009.
- [484] C. Batini, P. Buisseret, M. Lasserre, and M. Toupet, “Does proprioception of the extrinsic eye muscles participate in equilibrium, vision and oculomotor action?,” in *Annales d’oto-laryngologie et de chirurgie cervico faciale: bulletin de la Societe d’oto-laryngologie des hopitaux de Paris*, 1984, vol. 102, pp. 7–18.
- [485] L. Prieto González, B. Puerta Hoyas, and A. de Amescua Seco, “Distributed multi-dimensional feature recognition and multi-level semantic classification model for computer vision,” *DYNA New Technol.*, vol. 3, no. 1, Dec. 2016.
- [486] X. S. Zhou and T. S. Huang, “CBIR: from low-level features to high-level semantics,” *Proc SPIE*, vol. 3974, pp. 426–431, 2000.
- [487] L. Prieto-González, V. Stantchev, and R. Colomo-Palacios, “Applications of ontologies in knowledge representation of human perception,” *Int. J. Metadata Semant. Ontol.*, vol. 9, no. 1, pp. 74–80, 2014.
- [488] J. L. López-Cuadrado, R. Colomo-Palacios, I. González-Carrasco, Á. García-Crespo, and B. Ruiz-Mezcua, “SABUMO: Towards a collaborative and semantic framework for knowledge sharing,” *Expert Syst. Appl.*, vol. 39, no. 10, pp. 8671–8680, 2012.
- [489] E. Jiménez-Domingo, J. M. Gómez-Berbis, R. Colomo-Palacios, and Á. García-Crespo, “CARL: A Complex Applications Interoperability Language based on Semantic Technologies for Platform-as-a-Service Integration and Cloud Computing,” *J. Res. Pract. Inf. Technol.*, vol. 43, no. 3, 2011.

- [490] V. Stantchev, “Effects of Replication on Web Service Performance in WebSphere,” International Computer Science Institute, Berkeley, California 94704, USA, ICSI Tech Report 2008-03, Feb. 2008.
- [491] Message Passing Interface Forum, *MPI: A Message-Passing Interface Standard Version 3.0*. 2012.
- [492] D. E. Culler, J. P. Singh, and A. Gupta, *Parallel computer architecture: a hardware/software approach*. Morgan Kaufmann, 1998.
- [493] J. J. Hopfield, “Neural networks and physical systems with emergent collective computational abilities,” *PNAS*, vol. 79, no. 8, pp. 2554–2558, 1982.
- [494] “Git.” [Online]. Available: <https://git-scm.com/>. [Accessed: 05-Dec-2016].
- [495] GitLab Inc., “GitLab Community,” *GitLab*. [Online]. Available: <https://about.gitlab.com/community/>. [Accessed: 05-Dec-2016].
- [496] “OpenCV | OpenCV.” [Online]. Available: <http://opencv.org/>. [Accessed: 05-Dec-2016].
- [497] “Medidor de acidez en cítricos Atago PAL-Easy ACID1.” [Online]. Available: http://www.infoagro.com/instrumentos_medida/medidor.asp?id=2142. [Accessed: 05-Dec-2016].
- [498] U. S.-F. D. A. Center for Food Safety and Applied Nutrition, “Bad Bug Book - BBB - pH Values of Various Foods.” [Online]. Available: <http://www.fda.gov/Food/FoodborneIllnessContaminants/CausesOfIllnessBadBugBook/ucm122561.htm>. [Accessed: 05-Dec-2016].
- [499] “NoSQL Databases Explained,” *MongoDB*. [Online]. Available: <https://www.mongodb.com/nosql-explained>. [Accessed: 05-Dec-2016].
- [500] “Redis.” [Online]. Available: <https://redis.io/>. [Accessed: 05-Dec-2016].