



POLITECNICO DI MILANO
Facoltà di Ingegneria dell'Informazione
Dipartimento di Elettronica e Informazione

MASTER THESIS

Tutor

Tutor Politecnico di Milano: Vincenzo Caglioti
UC3M: Guillermo Carpintero

LONG EXPOSURE VIDEO-SURVEILLANCE

ASLAMIENTO DE NUEVOS OBJETOS EN LA ESCENA Y RECHAZO
ASOCIADO A LA IDENTIFICACIÓN DE MOVIMIENTO DE LOS OBJETOS DEL
FONDO.

Sandra Gómez de Dueñas

1. INTRODUCCIÓN

Cuando se utilizan imágenes de larga exposición de la misma escena, los objetos en movimiento aparecen borrosos y semi-transparentes. Además, estas características aumentan con el tiempo de exposición hasta ser totalmente invisibles cuando se utilizan tiempos muy largos. El principal contexto en el que se aplicaría esta técnica sería en Redes de Sensores Inalámbricos o similares, que requieren los mínimos cálculos y funciones de comunicación. Utilizar imágenes de larga exposición en este campo supone una importante disminución en el número de fotogramas a enviar a la estación de vigilancia y ligado al bajo esfuerzo computacional que requiere, ambos hechos llevan a ampliar la vida útil de la batería del dispositivo.

El método básico implementado con el fin de detectar la actividad en la escena vigilada es la Diferenciación de dos imágenes consecutivas, que llamaremos a partir de ahora *imagen A* (inicial) e *imagen B* (posterior). La imagen diferencia (B-A), representa el estado de movimiento que se ha llevado a cabo durante el lapso de tiempo de la toma de las fotografías A y B, es decir, en teoría, el resultado de la imagen diferencia sería una imagen negra si no se ha detectado actividad y por el contrario, si algún objeto ha entrado en escena o se ha producido algún movimiento en los objetos que forman parte de la escena vigilada, el movimiento o el “intruso” aparecerán en un color distinto al negro. Si lo extrapolamos al campo de las imágenes en blanco y negro, cuando no se produzca movimiento la imagen diferencia será totalmente negra, y cuando ha existido movimiento en alguna región de la escena, los píxeles afectados en dicha región se representarán en blanco. En consecuencia, el dispositivo de detección opera en los siguientes modos:

- 1) Modo de espera: Se desea que la mayoría de las veces el dispositivo funcione en este modo, en el cual, está a la espera de cualquier nueva actividad en la escena.
- 2) Modo de alerta: Una vez que se detecta una presencia, la red se conecta a este modo de alerta, y se debe enviar una alarma. Además se inicia un nuevo estado de detección, de mayor complejidad y detección específica.

1.1. CONDICIONES DEL ANÁLISIS

Se estudió el efecto del tiempo de exposición en la detección de actividad. Además, también se observó la influencia del tipo de escenario en el que se iba a llevar a cabo la identificación del movimiento de los objetos. Si el lugar donde se realizaba la detección de intrusos era exterior, los cambios de luminosidad temporales podrían influir considerablemente. Sin embargo, si por el contrario, se trataba de una estancia interior, la luminosidad permanecería prácticamente constante, a no ser que existieran ventanas o puertas que permitieran el paso de la luz exterior al recinto estudiado.

Por otra parte, también se analizó la influencia del tipo de movimiento del intruso con respecto al objetivo de la cámara fotográfica, es decir, si se movía de forma paralela, ortogonal o diagonalmente respecto de ella. También, trató de sacar conclusiones acerca del color de ropa (clara, oscura) que vestía el sujeto a identificar como intruso.

Aquí se muestra un ejemplo de los casos estudiados:

ESCENARIO	DIRECCIÓN DEL MOVIMIENTO PARA ROPA CLARA		DIRECCIÓN DEL MOVIMIENTO PARA ROPA OSCURA	
	ORTOGONAL	PARALELO	ORTOGONAL	PARALELO
EXTERIOR 1				
	DIRECCIÓN DEL MOVIMIENTO PARA ROPA CLARA		DIRECCIÓN DEL MOVIMIENTO PARA ROPA OSCURA	
EXTERIOR 2	ORTOGONAL	DIAGONAL	ORTOGONAL	DIAGONAL
				
INTERIOR 1	DIRECCIÓN DEL MOVIMIENTO PARA ROPA CLARA		DIRECCIÓN DEL MOVIMIENTO PARA ROPA OSCURA	
	ORTOGONAL	PARALELO	ORTOGONAL	PARALELO
INTERIOR 1				
	DIRECCIÓN DEL MOVIMIENTO			
INTERIOR 2	El intruso lleva un sombrero blanco			
	ORTOGONAL		PARALELO	
INTERIOR 2				

2. OBJETIVOS

El modo predominante de funcionamiento es el modo de espera. Por ello, para que la detección se realice de forma adecuada los cambios de luminancia deberían ser corregidos. Los cambios de luminancia pueden conducir a una falsa alarma a cuando no hay actividad producida dentro de la escena. Por ello, uno de los objetivos perseguidos

en este proyecto fue la mejora de la robustez de las imágenes diferencia para intentar hacerlas menos sensibles a los cambios de luminancia.

Por otro lado, la detección de actividad general no siempre es conveniente si se habla en términos de vigilancia y a que sólo se acepta el envío de alarma cuando la actividad es de tipo humana. Por ello, en este trabajo, se trataría de limitar la detección de actividad a la producida únicamente por presencia humana.

3. AISLAMIENTO DEL MOVIMIENTO DE “NUEVOS” OBJETOS Y RECHAZO DE LA DETECCIÓN DE MOVIMIENTO ASOCIADO A OBJETOS “ANTIGUOS” EN LA ESCENA

Como ya se comentó previamente, un problema común en este tipo de vigilancia es el movimiento accidental de objetos que ya estaban presentes en la escena y que producían una falsa alarma a pesar de que ningún “nuevo” objeto entrase en la zona de vigilancia. Un ejemplo, es el movimiento de árboles y arbustos debidos a ráfagas de viento.

3.1 EXPLICACIÓN DE ALGORITMOS

Para intentar corregir dicho error se propusieron varios algoritmos, que se explicarán a continuación.

3.1.1 Imágenes Media y Desviación Estándar de un conjunto de Imágenes de Fondo

Ambos métodos calculan la imagen diferencia.

En el método de cálculo de la imagen media, en primer lugar se introduce un conjunto de imágenes del fondo. La imagen media resulta de calcular separadamente para cada canal la media aritmética, a través del conjunto total de imágenes de fondo, de las intensidades de cada pixel. Esta imagen media se utilizará como minuendo para el cálculo de la imagen diferencia y en otros procesos relacionados con la aplicación de umbrales (tresholding).

Similarmente, la imagen desviación estándar se calcula a través del conjunto de imágenes del fondo. En esta imagen resultante, independientemente para cada canal, cada pixel representa la raíz cuadrada de su valor varianza calculado a lo largo del conjunto de imágenes. El valor varianza, es la suma de los cuadrados de las diferencias entre el valor intensidad de cada pixel y su valor correspondiente en la imagen diferencia. Posteriormente, para cada canal, se calcula un valor umbral como la media de los valores de intensidad de cada pixel de la imagen desviación estándar. Se obtienen

consecuentemente 3 thresholds (uno para cada canal) que se utilizarán para distinguir el nivel de magnitud de los movimientos que se han producido en la escena y también para el proceso de distinción entre objetos “nuevos”/”antiguos” en la zona vigilada.

3.1.2 Filtro Mediano/Híbrido Mediano a la imagen diferencia

Una vez calculada la imagen media, se le aplica un filtro mediano/híbrido mediano.

La *imagen B* también es filtrada por este método para reducir su ruido. Dependiendo del escenario, el tamaño de la máscara del filtro varía, pero siempre se escoge de acuerdo al tamaño de la cabeza de una persona en la escena.

Otra opción, sería calcular la imagen diferencia y posteriormente filtrarla aunque este método se rechazó ya que el ruido incluido en ambas imágenes podría afectar negativamente a la imagen diferencia.

3.1.3. Imagen Diferencia

La imagen diferencia se obtiene aplicando el valor absoluto a la imagen resultante de calcular la diferencia pixel a pixel entre la *imagen B* y la imagen media previamente filtradas.

3.1.4. Identificación y Distinción entre el movimiento de objetos “nuevos y antiguos en la escena

3.1.4.1. Métodos basados en la comparación de las medias aritméticas de las intensidades de dos imágenes consecutivas

Para distinguir el movimiento de objetos “antiguos” en la escena vigilada, tales como el movimiento de una planta, del movimiento relativo a un intruso (objeto nuevo) en la escena, los dos primeros métodos se basan en la comparación entre la media de intensidades en la imagen fondo (en nuestro caso la imagen fondo usada es la imagen media) con la media de intensidades en la *imagen B*, la cual contiene información sobre el desplazamiento de objetos o/ y entrada de intrusos respecto de la imagen inicial o la imagen del fondo. Esta comparación de medias de intensidades se lleva a cabo en cada área de la imagen diferencia ya que éstas son las áreas en las que se ha detectado actividad.

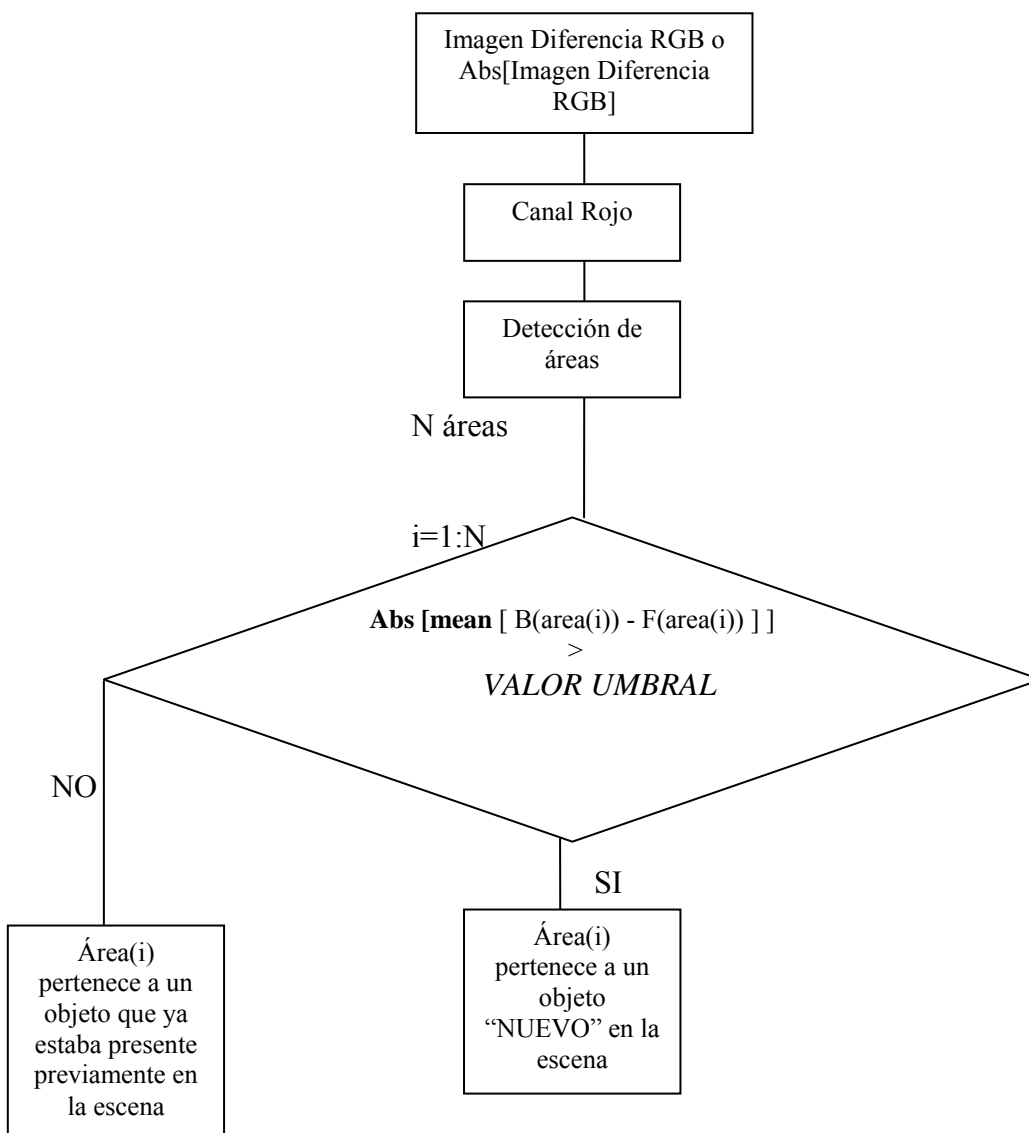
Teóricamente, si ningún objeto ha entrado en la zona vigilada durante varios frames consecutivos, la media de las intensidades de una imagen debería permanecer constante a pesar de que los objetos incluidos en ella se hayan desplazado. Pero, si algún individuo u objeto nuevo entrase en la escena controlada, alterándola consiguientemente, la media aritmética de las intensidades en la imagen diferencia (imagen de fondo sustraída a la imagen alterada) se verá considerablemente afectada por las intensidades del nuevo objeto. Esto significa que si ningún nuevo objeto entrase en la escena, la diferencia entre las medias aritméticas de las intensidades de dos imágenes consecutivas debería ser cero. Sin embargo, esto no es realmente cierto, ya que entre dos imágenes consecutivas de la misma escena siempre hay cambios de

luminosidad que afectan a los valores de intensidad de cada pixel, a pesar de que ningún nuevo objeto aparezca en la escena.

Por lo que, en general, cuando se consideran dos imágenes consecutivas, se analiza cada área particular de su imagen diferencia. Para cada una de dichas áreas, se calculan los valores de las medias aritméticas regionales de sus intensidades en la imagen del fondo y en la imagen alterada (*imagen B*). Más tarde, se calcula la diferencia entre ambos valores medios y el área correspondiente estudiada será sólo considerada como un área de “nuevo” objeto si dicha diferencia supera un cierto valor umbral. En este aspecto del valor umbral es donde los dos primeros métodos estudiados, que se basan en la media, difieren.

Hay que notar, que trabajando con imágenes RGB la comparación de medias debe realizarse para cada canal.

Ahora, un diagrama muestra conceptualmente los procesos seguidos en los dos métodos de comparación de medias de intensidades, aunque sólo se representa el proceso para el primer canal, es decir, el rojo. La misma secuencia se implementará para el segundo y el tercer canal.



Nota: En este diagrama, *B* se refiere a la imagen tomada tras un lapso de tiempo desde que se tomó la imagen *A* y *F* se refiere a la imagen de fondo (imagen de la zona vigilada sin ningún intruso). En general, la imagen *A* se refiere a la imagen *F*.

3.1.4.1.1. Cálculo automático del umbral por región basado en la media de la desviación estándar regional de intensidades en la imagen de fondo

En primer lugar, se debe recordar que nuestra imagen de fondo es la media de un conjunto de imágenes de fondo.

Como se explicó previamente, a partir de dos frames consecutivos, *A* y *B*, donde *A* representa una imagen del fondo (equivalente a la imagen *F* del diagrama anterior), la imagen diferencia *D* se calcula siguiendo la siguiente expresión $D = \text{abs}(B - A)$ –con la excepción de un caso particular que se explica en el proyecto-. Posteriormente, para cada canal, se estudia cada una de las áreas de la imagen *D* y se calculan los siguientes valores

- la media aritmética de sus correspondientes intensidades regionales (*referidas a cada área*) intensidades en la imagen de fondo *A*
- desviación estándar de sus correspondientes intensidades regionales en la imagen de fondo *A*.
- la media aritmética de sus correspondientes intensidades regionales en el frame *B*.

Una vez calculados dichos valores, para cada área se realiza el siguiente análisis. Se toma el valor absoluto a la diferencia entre los valores medios previos. En caso de que dicho valor absoluto sea mayor que el valor devuelto al aplicar una función de la media de la desviación estándar de las intensidades regionales en *A*, entonces el área estudiada se considera como perteneciente a un “nuevo objeto en el escenario” (intruso). Si por el contrario, es inferior que la desviación estándar, la región estudiada corresponde a un movimiento de un objeto que ya estaba presente en la zona bajo vigilancia. A éste método se le dio el nombre de ***Corrección por Media1***.

3.1.4.1.2. Cálculo automático del umbral por región basado en la media de la desviación estándar de las intensidades inherentes a cada pixel a través de un conjunto de imágenes. Aplicación al método de la media de intensidades regionales.

En este segundo método, para cada canal, se compara el valor absoluto de la diferencia entre las medias regionales de intensidades de las imágenes *A* y *B*, con el umbral correspondiente obtenido de la media de sus intensidades respectivas (para cada canal) en la imagen desviación estándar. Si ese valor absoluto supera el valor umbral, el área seleccionada pertenece a un área de “nuevo objeto”, mientras que en caso contrario, el área corresponde a un movimiento de uno de los objetos ya presentes en la escena (“objeto antiguo”). Este método se llamó ***Corrección por Media2***.

3.1.4.2. Métodos basados en la comparación de medianas de intensidades en dos imágenes consecutivas

De forma similar al proceso previo relacionado con la identificación de movimiento basado en la media de intensidades en dos frames consecutivos se propuso el estudio basado en el cálculo de la mediana ya que este parámetro es un mejor

indicador de la tendencia central y también un método más robusto en presencia de valores atípicos que la media aritmética. Por lo que, en teoría, este método estaría menos influenciado por los cambios de luminosidad eventuales, que aumentan de forma considerable la intensidad de los píxeles y afectan negativamente al proceso de detección.

3.1.4.2.1. Cálculo automático del umbral por región basado en la mediana de la desviación estándar regional de intensidades en la imagen de fondo

Este proceso es análogo al método explicado en el punto 3.1.4.1.1. pero con la diferencia de que en vez de calcular para cada canal y área las medias de las intensidades de los píxeles del área correspondiente, se calcula su mediana. Además, el umbral de comparación se determina mediante una función de la mediana de la desviación estándar de las intensidades en el área considerada. Este método fue llamado ***Corrección por Mediana1.***

3.1.4.2.2. Cálculo automático del umbral por región basado en la media de la desviación estándar de las intensidades inherentes a cada pixel a través de un conjunto de imágenes. Aplicación al método de la media de intensidades regionales.

En esta cuarta propuesta, para cada canal, el valor absoluto de la diferencia entre las medianas regionales de las intensidades de las imágenes A y B, se compara con el valor umbral correspondiente, obtenido a partir de la media de las intensidades respectivas por canal en la imagen desviación estándar, de la forma que se explica en el punto 3.1.1. .

En caso de que el valor absoluto de la diferencia de medianas sea mayor que este valor umbral, el área estudiada pertenecerá a un objeto que no estaba antes en la imagen de fondo, es decir, se tratará de un “nuevo objeto”. En caso contrario el área se corresponde con un “objeto antiguo”. El nombre asignado a este método fue ***Corrección por Mediana2.***

3.1.5. Apertura de pequeñas áreas

Después de la eliminación de áreas que no pertenecen, en teoría, a un “nuevo objeto”, un proceso de apertura se llevó a cabo con el fin de reconocer únicamente a **intrusos humanos**. La aplicación de este método está aún más justificada en el caso de video-vigilancia de larga exposición.

3.1.6. Distinción de la magnitud de la actividad

Las áreas detectadas en la imagen *abierta* son sometidas a una clasificación basada en una comparación de intensidades para cada uno de los píxeles que constituyen el área, con varios múltiplos de cada uno de los umbrales definidos en el punto 3.1.4.1.2.

Para cada canal, cada pixel de cada una de las áreas se compara con varios múltiplos del correspondiente valor umbral para dicho canal. Así, después de este tratamiento, cada pixel en el área abierta tendrá un valor de intensidad de salida tan alto como el valor definido por el rango al que el pixel input pertenece. Consecuentemente, los píxeles que correspondan a la actividad más intensa tendrán el mayor nivel de intensidad, mientras que aquellos que correspondan a niveles de actividad media tendrán una intensidad algo

menor y los que representan áreas donde casi no hubo movimiento, no se representarán en la imagen de salida final. De esta forma se obtiene una imagen con graduación de niveles de movimiento.

3.1.7. Detección de bordes

Para encontrar los bordes de la imagen obtenida después del proceso de apertura, se aplicó un detector de bordes tipo *Canny*.

3.1.8. Detección de fronteras y extracción de agujeros

3.1.9. Deblurring-enfoco de imágenes borrosas-

En este tipo de video-vigilancia el sujeto en movimiento aparece con una silueta borrosa y semitransparente, por ello, fue preciso estudiar varios métodos de deconvolución incluidos en MATLAB con el fin de intentar eliminar dichos efectos:

- Deblurring con Wiener Filter*
- Deblurring con un filtro regularizado*
- Deblurring con Blind Deconvolution*
- Deblurring con Algoritmo de Lucy-Richardson*

Finalmente, el que introdujo una mayor mejora en los resultados fue el método del *Algoritmo de Lucy-Richardson*.

4. COMPARACIÓN DE FILTROS PARA LA DETECCIÓN DE ACTIVIDAD



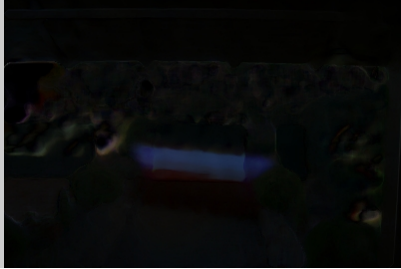



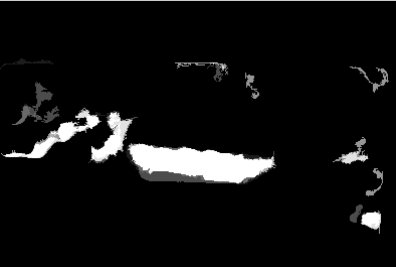

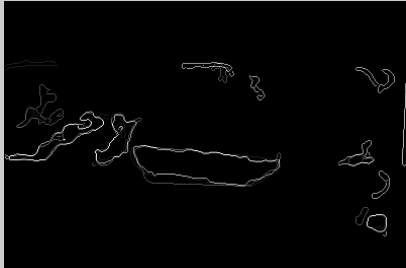
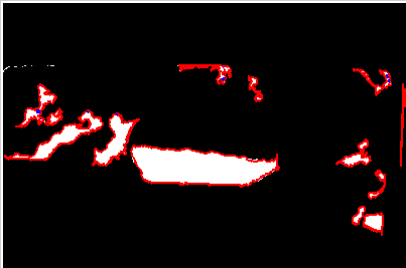
En este proyecto se estudió la efectividad de dos tipos de filtros espaciales diferentes para tratar de eliminar el ruido y el pequeño detalle no significativo de las imágenes. Los filtros no lineales estudiados fueron:

- 1) Filtro mediano: Útil cuando existe gran cantidad de ruido aleatorio.
- 2) Filtro híbrido-mediano: Es un tipo de filtro mediano más efectivo en la preservación de bordes que el filtro mediano común.

En la siguiente tabla se muestra un ejemplo paso a paso, de la mejora en los resultados obtenidos en la detección del intruso tras la aplicación del filtro mediano y algunos de los métodos que se muestran en el punto 3.1.4.

EXTERIOR 1

Mañana

	IMAGEN DEL MOVIMIENTO	IMAGEN MEDIA
		
	IMAGEN DIFERENCIA RGB	IMAGEN DIFERENCIA (ESCALA DE GRISES)
		
	IMAGEN DIFERENCIA (BLANCO Y NEGRO)	DIFERENCIA: TRAS CORRECCIÓN POR MEDIA1
		
	IMAGEN FINAL (MOVIMIENTOS EXTREMOS).(Corrección por Media1)	IMAGEN FINAL (ESCALA PROGRESIVA DE MOVIMIENTO)(Corrección por Media1)
		
	CANNY EDGE DETECTOR*	BOUNDARY EXTRACTION*
		

Comparación de los efectos de los diferentes métodos para detección de actividad de “nuevos/antiguos objetos” en el Canal 1 de la imagen diferencia inicial de un escenario exterior

Nota: las imágenes marcadas con un * se refieren a la imagen obtenida tras el tratamiento con Corrección por Media 1.

7. OTROS INTENTOS

Con el fin de intentar mejorar la identificación de intrusos se estudiaron diferentes algoritmos, que se citan a continuación:

7.1 INTENTO DE COMPENSACIÓN DE LOS EFECTOS DE LUMINOSIDAD

7.1.1. Calcular la media de un conjunto de imágenes de fondo y hallar la imagen diferencia con respecto a la imagen media.

7.1.2. Intento de cálculo de la imagen diferencia en el espacio HSV

7.1.3. Intento de cálculo de la imagen diferencia en el espacio NTSC

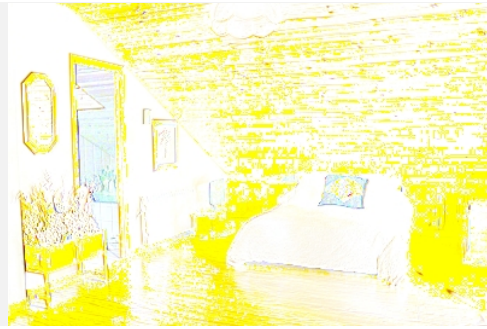
7.1.4. Mejora de imágenes multispectrales. Cociente de intensidades

Ejemplo de la mejora en la identificación del sujeto intruso tras este método:

Imagen del movimiento-Velocidad de obturación 2''- Paralelo



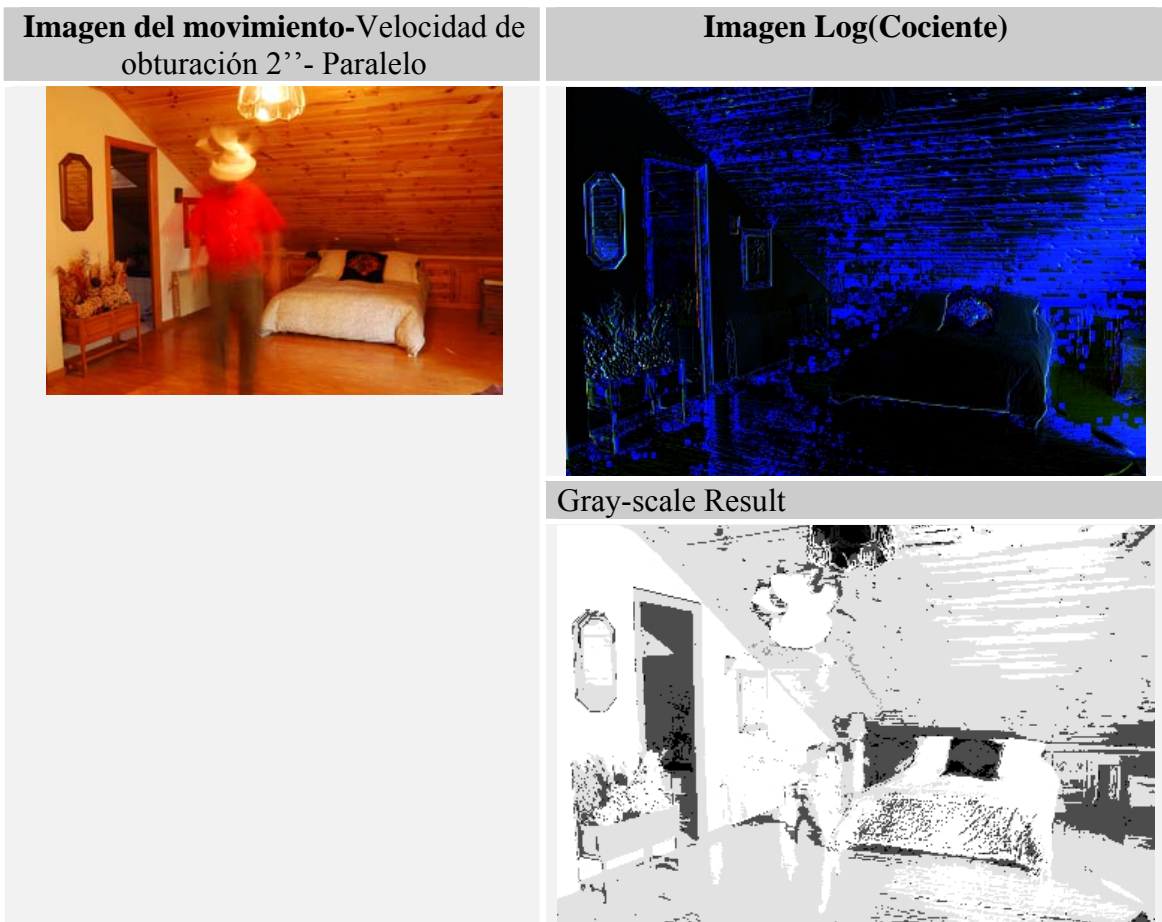
Imagen Cociente



Resultado en Escala de Grises



8.1.5. Mejora de imágenes multispectrales. Logaritmo del cociente de intensidades



8.2. COMPARACIÓN CON UMBRAL PARA LA IMAGEN DIFERENCIA POR EL MÉTODO MAD/MED

8.3. UMBRAL AUTOMÁTICO OBTENIDO REALIZANDO EL CUARTIL 0.75 DE LAS DESVIACIONES ESTÁNDAR A TRAVÉS DE UN CONJUNTO DE IMÁGENES DE FONDO

9. CONCLUSIONES

- **Todos los métodos** (Corrección por Media 1, Corrección por Media 2, Corrección por Mediana 1 y Corrección por Mediana 2) **presentaron resultados significativos** al tratar de eliminar las alarmas procedentes del movimiento de árboles y preservando la activación de la alarma por intruso. Los resultados de todos ellos fueron muy parecidos, aunque ocurren pequeñas desviaciones entre las Correcciones por Media y por Mediana 1.

- En algunos casos, la Corrección por Mediana 1, tiende a preservar mejor el contorno del intruso que la Corrección por Media 1, y en general, tiende a eliminar áreas eventuales que no pertenecen al intruso, proporcionando una identificación "más limpia". De todas formas, ambos métodos, arrojaron resultados positivos.

- **En escenarios exteriores y para los tiempos de exposición más largos, cuando se procesan las imágenes con las Correcciones por Media/Mediana 1 (o 2), parte de la estela del intruso en movimiento desaparece, y consecuentemente el área del intruso es menor que la respectiva en la imagen diferencia inicial.**

- Cuando se trabaja en escenarios de exterior, se aconseja el uso de esta técnica de detección de intrusos con imágenes de larga exposición en condiciones nocturnas, ya que la aparición de sombras es menos probable en la oscuridad, pero teniendo en cuenta que para tiempos de exposición superiores a 10' una fuente de luz debe estar presente porque por el contrario, ningún intruso será detectado.

- En escenarios donde la sombra del intruso podría afectar negativamente en el proceso de detección de actividad es conveniente no aplicar el valor absoluto a la diferencia de los dos frames consecutivos, sino utilizar sólo esta imagen diferencia. Sin embargo, se perderá detalle en la silueta y forma del intruso

- El efecto del Filtro Híbrido Mediano es una mayor preservación de bordes con respecto al Filtro Mediano, pero una mayor cantidad de detalles en reposo del escenario fueron detectado. Consecuentemente, el filtro mediano estándar proporciona una imagen más limpia en la detección del intruso.

- **Para compensar los cambios de luminosidad eventuales que interfieren en el proceso de detección el mejor método resultó ser el cociente de imágenes multispectrales.** Dicho método **incrementa considerablemente la nitidez** de la imagen diferencia aunque algunos objetos que estaban detrás del intruso en el escenario fueron también detectados.

