Analyzing Wavelets Components to Perform Face Recognition

Pedro Isasi¹, Manuel Velasco¹, and Javier Segovia²

¹ Departamento de Informática
Universidad Carlos III
Leganes, Madrid.
isasi@gaia.uc3m.es

² Departamento de Lenguajes y Sistemas Informáticos
Facultad de Informática U.P.M.
Boadilla del Monte, Madrid.

Abstract. Face recognition is a very difficult task in real environments. In those cases a good preprocessing of the images is needed to keep the images invariant to translations, scales, luminosity, shape, aspect, rotation, noise, etc... Wavelet transformation have been probed to be a good preprocessing method for many task. However, not all the coefficients of a wavelet transform have the information needed for a classification method to be efficient. This work introduce a method to select the most appropriate coefficients for a wavelet transform to allow an unsupervised neural network to well classify a set of complex faces.

1 Introduction

Most sensitivity analysis methods using neural nets have been based almost exclusively on the use of supervised net models. In such cases, the method seek to study how a specific input affects net efficiency [3]. Having detected the inputs (X_i), which hardly affect learning, they are removed from the system and learning is repeated. After the number of inputs has been reduced, if the neural net performs similarly in terms of learning and generalisation as it did with all the inputs, the method will be said to have been successful.

This paper presents a new sensitivity analysis method based on non-supervised neural nets and more specifically on Grosberg's ART models [5, 1] (ArtSen).

In this method, some inputs are gradually removed as and when their varying influence is detected. Therefore, the process is gradual and continuous, as the inputs are received in the training phase. This means that the sudden adjustment of an input does not interfere in the sensitivity analysis process (although it does have an influence on the final result), and this can output results constantly as the relationships between inputs are detected, without having to retrain.

The following sections describe the method of non-supervised sensitivity analysis and its application for determining critical parameters in an image classification task.

2 Sensitivity Analysis using ART models (ArtSen)

The presented method is based on the *ARTMap* model, and is illustrated in figure 1.

Fig. 1. ArtSen model scheme

The model is composed of mainly three modules: a first module ($\Sigma\Pi$ connections) where weighted inputs are received, are processed (if necessary), and an output is produced, which will be the input for the second module (*ARTMap* module). In the *ARTMap* module, the signal is propagated as usual in this model. Finally, sensitivity analysis will be carried out in the Entropy Model (*EM*) producing outputs, called Influence Signals (*IS*). These outputs biunivocally determine the influence level of each input signal and are used to weight the inputs in the next time interval

$\Sigma\Pi$ Module The first module of the system consists of a set of $\Sigma\Pi$ connections [4,12] in which each cluster receives two signals, one from the input stimulus X_i and the other, the output signal, from the entropy model IS_i acting as weights. The generic equation for the $\Sigma\Pi$ cluster is:

$$X'_{tj} = \Pi_{i=1}^k W_{ji} IS_i X_i \quad (1)$$

The output produced by the $\Sigma\Pi$ cluster, the X'_i , will be input to a signal preprocessing module whether or not warranted by the problem to be solved.

Art Module The output of the first module is input to and propagated through an *ARTMap* module [2] (which can be replaced by an *ART2* model [1] if there is no reinforcement signal). The module clusters the input signals unsupervised

or by reinforcement. Once the net has been trained, the sensitivity analysis will involve determining the influence level of each input signal in the clustering process. The BottomUp connections (Z_{ij}) are used as input for the last module of the model, the entropy model.

Entropy Module A diagram of the architecture of this model is shown in figure 2. Figure 2 shows four layers: three Feed Forward layers, and one competitive layer. The main purposes of this module are:

1. To measure the influence of each system input to the on the ArtMap model.
2. To make this measure depending on the *ARTMap* learning phase.

The first task is carried out in the first three layers of the net and the second one in the fourth layer.

Fig. 2. Entropy Module

The first layer consists of n groups of cells, where n is the dimension of the X'_i input. Each group (G_i) will get m inputs, where m is related to the number of categories generated in the F2 layer of the Art model, (figure 3):

$$G_i = \{Z_{ij}, \forall j \in F_2\} \quad (2)$$

where Z_{ij} refers to the connection between the i th cell of F_1 and the j th cell of F_2 of the *ARTMap* net Art_a model (figure 3). The F_1 layer directly receives the Art net input. Therefore, each cell of this layer is related to one of the input lines of the X_i stimulus.

The second and third layers are each composed of a cell, connected as shown in figure 2. The output of this last layer has the same dimension as the F_1 layer of Art_a , which also corresponds to the dimension of the X_i input. This output directly addresses the next competitive layer. The equations that govern the behavior of the module are:

$$\tau_s \frac{\partial S_{ij}^1}{\partial t} = -S_{ij}^1 + f_1(Z_{ij}) \quad (3)$$

Fig. 3. Connection between the input of the Entropy Model and the BottomUp connections of the ArtMap net

$$\tau_s \frac{\partial S_i^2}{\partial t} = -S_i^2 + f_2(\sum_{j=1}^M W_{ji} S_{ji}^1) \quad (4)$$

$$\tau_s \frac{\partial S_i^3}{\partial t} = -S_i^3 + f_3(\sum_{j=1}^M W_{ji} (S_{ji}^1 - S_i^2)^2) \quad (5)$$

Where:

S_{ij}^1 is the output of the i th cell of the j th group of the first layer of the entropy module, f_1 , f_2 y f_3 are sigmoidal functions, Z_{ij} is the j th input of the i th group of the entropy module, M is the number of categories created so far in Art_a (F_2 dimension, coinciding with the number of neural groups), S_i^2 is the output of the i th cell of the second layer of the entropy module, W_{ji} is the weight of the connection between the j th cell of the first layer and the i th cell of the second layer. A constant value $W_{ji} = \frac{1}{M}$ can be selected, S_i^3 is the output of the i th cell of the third layer of the entropy module, W_{ji} is the weight of the connection between the j th cell of the first layer and the i th cell of the third layer. A constant value $W_{ji} = \frac{1}{M}$ can be selected, s is a temporal constant

Equation 2 shows how S_i^2 works as a threshold to compare the signals of the input groups of the entropy module, in absolute terms, thus outputting the variances of these signals, grouped by the input of F_1 . This variance, S_i^3 , acts as the input for the competitive layer of the entropy module, where the dynamics of the model is defined by:

$$\tau_c \frac{\partial IS_i}{\partial t} = -IS_i + f_c(S_i^3 + IS_i - \sum_{j=1, j \neq i}^N W_{ij} IS_j \delta(t)) \quad (6)$$

Where:

IS_i is the activation value of the units, W_{ij} are the connection weight of the i th cell of the third layer with the j th cell of the fourth layer, n is the dimension of the fourth layer, $\delta(t)$ is a positive value lower than 1, cell competitiveness rating of the cells, f_c is a sigmoidal threshold function, and τ_c is a temporal constant.

Values for $\tau_c \gg \tau_s$ should be used that assure that the short-term variations of S_i^3 do not strongly affect IS_i .

$\delta(t)$ introduces what has been called **competitive vigilance factor (cvf)** in the competitive layer, where the higher the value, the higher the competition. Its variation in time will be governed by the equation:

$$\delta(t) = 1 - \frac{1}{e^{\frac{t}{\tau}}} \quad (7)$$

3 Experimental validation

The experiment chosen to validate the system was automatic face recognition. This problem has some special features that makes it suitable for use in system validation [9, 8]. It is a very complicated task involving a wide variety of factors, including translations, scales, luminosity, shape, aspect, etc. [11, 10],

The experiment involves classifying a set of 300 images of the faces of 20 different people.

Two kind of experiments are presented, those provided by the Pattern Recognition System without performing Sensitivity Analysis (Type A), and those provided by the system with the Sensitivity Analysis Process previous to the pattern selections (Type B).

The preprocessing is the same for both experiments and it is very simple. First, the images are compressed to 64x64 pixels ([3]), where each pixel represents a grey level between 0 and 255, totalling 4096 points. These images are filtered by means of a wavelet transformation into two dimensions, using the Daubechie 4 wavelet, outputting a 64x64 coefficient matrix. This matrix will be the input of the neural net, which will classify these images according to their similarity.

The reason for using 2D wavelet data processing before inputting data to the neural net is that most of the useful information in the images is located in certain transformation coefficients, particularly the low frequency coefficients. This is easily demonstrable by rebuilding the images using the inverse transformation from these coefficients only. Rebuilding is nearly perfect, albeit less perfect when fewer coefficients are selected. However, if the images are rebuilt by randomly selecting the same number of coefficients or using any other criterion, then the rebuilding process is not possible.

3.1 Type A experiments

In a first stage the patterns results from the wavelet transformations are directly introduced to the net. Each pattern has 4096 coefficients (64x64), i. e., the experiment is developed with all the information available. Table 1 presents the results from this experiment. First column shows the vigilance factor. Second column shows the number of categories generated by the net. Third column presents the number of categories classified perfectly. Fourth column indicates the percentage of global success. Finally, last column shows the typical deviation of the success by category.

Vigilance Factor	Number of Categories Generated	Number of Categories with 100% Success Rate	Global Percentage of Success	Typical Deviation of the Success Rate by Category
0.999900	19	17	68%	20.69
0.999910	22	19	67%	20.42
0.999920	27	25	68%	18.64
0.999930	35	32	76%	15.68
0.999940	45	40	84%	11.76
0.999950	64	62	97%	6.22
0.999960	97	95	98%	4.10
0.999970	133	130	99%	4.94

Table 1. Experiment with 4096 coefficients

In the next stages, the low frequency properties of the wavelet transformation are applied, obtaining sub-patterns with 1024 coefficients (32x32), removing 75% of the total information.

In a third stage 256 sub-patterns (16x16) are used, and so on, with a fourth stage (64 coefficients), and a fifth stage (16 coefficients). In this last stage only 0.39% of the information is used.

Table 2 shows partial information of these stages. A new column, with the number of coefficients, is included in the table.

Number of Coefficients	Vigilance Factor	Number of Categories Generated	Number of Categories with 100% Success Rate	Global Percentage of Success	Typical Deviation of the Success Rate by Category
1024	0,999950	52	48	92%	8,09
1024	0,999960	73	71	97%	5,56
256	0,999960	45	41	94%	6,99
256	0,999970	75	72	99%	2,33
64	0,999970	42	38	88%	10,62
64	0,999980	67	64	99%	1,47
16	0,999990	54	48	93%	8,15
16	0,999995	102	99	99%	4,02

Table 2. Experiment with rest of coefficients

To increase the vigilance factor implies that the number of categories and the success rate are higher. We consider a categorization as valid when there is a 95% of success at least and the number of categories generated is lower or equal than 40. This value is two times the number of different people.

In the first experiment (4096 coefficients) to get a 97% success 64 categories are generated, value very high for the purpose of the method. Same occurs for

experiment 2, 4, and 5. However, for experiment 3 (256 coefficients), we can observe an acceptable categorization, with a 94% success and 45 categories.

3.2 Type B experiments

These experiments add the Sensitivity Analysis to make the process faster and more efficient. The patterns which are result from the wavelet transformation are processed by a module. This module has two main processes. First process weights the significance of each component. Second process transforms again the patterns building another sub-patterns only with the relevant components. The relevant components chosen by the method depends on the value for the competitive vigilance factor. In this case increasing this facto the number of coefficients selected by the method are: 171, 124, 72, 45, 37, 21, 14, 13 and 10.

Therefore the experiment is developed again with the components calculated by the process. Table 3 indicates partial results of the Type B Experiments, in the same way as was shown in type A experiments.

Number of Coefficients	Vigilance Factor	Number of Categories	Number of Categories with 100%	Global Percentage	Typical Deviation of the Success Rate
171	0,999950	42	37	86%	11,49
171	0,999960	51	48	95%	7,98
171	0,999970	77	75	99%	3,48
124	0,999960	47	43	95%	6,01
124	0,999970	66	64	98%	4,95
72	0,999970	50	46	93%	8,65
72	0,999980	85	83	99%	2,07
45	0,999970	41	36	95%	6,82
45	0,999980	59	55	97%	4,79
37	0,999980	57	52	95%	6,72
37	0,999990	111	109	99%	2,90
21	0,999980	43	37	88%	12,20
21	0,999990	78	77	99%	1,13
14	0,999990	59	53	91%	10,86
14	0,999995	109	105	99%	5,32
13	0,999990	56	49	92%	11,16
13	0,999995	108	103	98%	4,68
10	0,999990	44	37	92%	11,33
10	0,999995	82	74	96%	10,18

Table 3. Results from type B experiments

Considering the same principles as in Type A Experiments (95% success rate and 40 categories) we can observe that in first experiment (171 coefficients) 51 categories are generated for a 95% success. This can be considered as acceptable, as well as in second experiment (47 categories for a 95% success). In the fourth

experiment (45 coefficients) a good categorization is observed (41 categories for a 95% success), but decreasing this number (37, 21, 14, 13, and 10 categories) no acceptable categorization is showed. Therefore, an optimum decreasing of the number of categories is demonstrated. Under this value, results begin to be worse.

Another method efficiency test has been run on a time series prediction problem [7]. The method was completely validated against a classical sensitivity analysis applied to a multilayer neural net. The efficiency of the proposed model with noisy signals was also verified in this problem.

4 Conclusions

The possibility of continuous learning, as in ART models, is highly important when dealing with some sensitivity analysis problems. In situations where the influence of the signals changes over time it is essential to have a method capable of picking up these variations, as and when they occur at a reasonable computational cost, and to be able to automatically determine when the variation took place.

The new sensitivity analysis presented here has these features. As the input changes, if the weights of the categories change or new categories are produced, these changes are automatically included in the sensitivity analysis calculations for the next input. The results vary due to the new situation.

As shown in the experimentation presented, the generality level of a system with inputs removed by the model is not reduced and, in some cases is increased due to the elimination of the noisy signals.

The analysis can be carried out in a supervised or nonsupervised manner or by reinforcement using this method by merely selecting the respective ART model.

References

1. A. Carpenter and S. Grossberg. Art2: Stable self-organization of pattern recognition code for analog input patterns. *Applied Optics*, 26:4919–4930, 1987.
2. A. Carpenter, S. Grossberg, and J. Reynolds. Artmap: Supervised real-time learning and classification of nonstationary data by a self-organizing neural network. *Neural Networks*, 4:565–588, 1991.
3. A. Carpintero, J. Castellanos, J. Rios, and J. Segovia. Automatic face recognition for access control. In *Proceedings of 1993 IJCNN*, pages 1289–1292. International Joint Conference on Neural Networks, 1993.
4. J.A. Feldman and D.H. Ballard. Connectionist models and their properties. *Cognitive science*, 6:205–254, 1982. Reprinted in Anderson and Rosenberg-1988.
5. S. Grossberg. Competitive learning: From interactive activation to adaptive resonance theory. *Cognitive science*, 11:23–63, 1987.
6. P. Isasi. *Sistema Neuronal Sensible al Entorno con Arquitectura Autoorganizada*. PhD thesis, Facultad de Inform'atica, Universidad Polit'ecnica, Madrid, 1994.

7. P. Isasi and J. Segovia. Sensitivity analysis in art models: An application in forecasting. In *XV International Symposium on Forecasting*, Toronto, Canada, June 1995.
8. Shen Jun, Shen Wei, H. J. Sun, and J. Y. Yang. Fuzzy neural nets with non-symmetric membership functions and applications in signal processing and image analysis. *Signal Processing*, 80(6):965–983, June 2000.
9. Hongbong Kim and Kwanghee Nam. Object recognition of one-dof tools by a back-propagation neural net. *IEEE Transactions on Neural Networks*, 6(2):484–487, March 1995.
10. Sheng Liu, R. Olivia, Chih-Ping Wei, and Paul Jen-Hwa Hu. Neural net learning for intelligent patient-image retrieval. *IEEE Expert*, 13(1):49–57, February 1998.
11. K. Nezis and G. Vosniakos. Recognizing 2d shape features using a neural network and heuristics. *Computer-Aided Design*, 29(7):523–539, July 1997.
12. D. Rumelhart and McClelland. *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, volume II. MIT Press., Cambridge, 1986.