TEXTURES DISCRIMINATION ENHANCEMENT BY FUSION WITH SECOND AND FOURTH ORDER STATISTICS

Carlos Aviles-Cruz, Anne Guérin-Dugué
INPG-TIRF , 46, Avenue Félix Viallet, F-38031 Grenoble cedex
email : aviles@tirf.inpg.fr

ABSTRACT

In this paper, second and fourth order statistical moments are used to segment fine grain textures. The fusion of the moments is made through different implementations (serial or parallel strategies) and different formalism (Bayes and Evidence theory). A comparative study of the classification performance is presented and interpreted.

I. INTRODUCTION

The recognition of textured images is a complex task due to difficulties to well characterise the textured patterns, specially fine grain textures. For these ones, statistical features are often used. The textures are characterised by the correlation between pixels into a small neighbourhood. A number of statistical features has been proposed based on second order moments [5] and on fourth order moments [4,7,8].

The main objective of this article is to show how the fusion of second and fourth order moments can enhance the performance of textures discrimination. Two different fusion strategies are explored according to:

- A simple serial "Master-Slave" architecture.
- A parallel scheme fusion formalism:
  - a) Fusion of post probabilities,
  - b) Fusion of decisions through the Bayes theory,
  - c) Fusion of decisions through Dempster-Shafer theory (DS), i.e. the Evidence theory.

With the serial scheme, the "Slave" classifier is only active when the "Master" can not discriminate. This scheme provides also an enhancement of the discrimination associated with a computational simplicity.

With the parallel fusion scheme, all the classifiers are actives and their outputs are then combined. So this system is computationally expensive. It is widely used in character recognition, speech recognition, remote sensing and medical applications [3,6].

For all the strategies, we will use two classifiers, one is associated to the second order moments "M2" and the second to the fourth order moments "M4".

II. FEATURE EXTRACTION

The features "M2" and "M4" are evaluated inside an estimation window, and are function of the current pixel (X_n), its first neighbour (X_{n+1}) and its second neighbour (X_{n+2}). These estimations are made in four orientations (0°, 45°, 90° and 135°). Before the statistical estimations, the images are pre-processed by an equalisation of the grey level histogram in order to be, as much as possible, independent of the acquisition parameters (contrast and brightness).

II-1. Cooccurence matrices

The "M2" features are derived from the cooccurence matrices. In order to reduce the number of features, we have chosen, among all the parameters deduced from these matrices, the contrast and the angular second moment [5]. Preliminary experimental tests have shown an accurate discrimination based on these two parameters which are configured with two distances between the current pixel and its neighbour (d=1 and d=2). In this way, a texture is characterised by 16 components (2 parameters x 2 distances x 4 orientations). The size of the estimation window is 16 x 16 pixels.

II-2. Fourth order moments

With this set of features, the textures are characterised by the correlation between three pixels (X_n, X_{n+1}, X_{n+2}) through the estimation of fourth order moments. In order to reduce the number of combinations between the pixels inside the neighbourhood, and nevertheless to take into account the orientation of the textures, four neighbourhood patterns have been chosen : the three pixels are aligned according to the four selected orientations. With these limitations, for each orientation, it remains 10 "M4" moments. The general equation to estimate the fourth order moments is given by:

$$\tilde{M}4(i,j,k) = \frac{1}{N}\sum_{n=1}^{N} x_n x_{n+i} x_{n+j} x_{n+k}$$

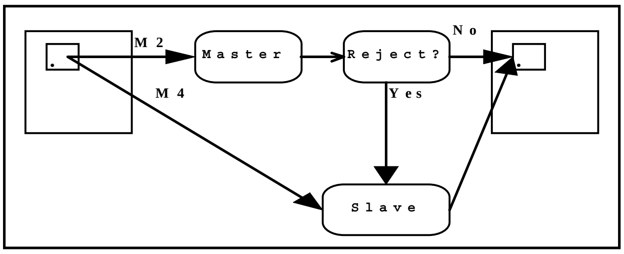
with N the size of the estimation window, and $0 \le i, j, k \le 2$.

By this way, the textures can be characterised with 4 x 10 components. Among these ones, for each orientation, we have selected 4 moments which are in average, the most discriminating ones. So the texture is characterised by 16 components (4 moments x 4 orientations). This selection has been experimentally done on a large dataset of textures. These features are estimated from a large estimation window (32x32 pixels). With a large window, the bias and the variance estimates are reduced but the localisation of the boundaries between the textures are not accurate; it is the contrary behaviour with a smaller window. To achieve a trade-off between accurate estimations and boundaries, a filtered method was used to estimate the moments inside a smaller window[1]. These features (called "M4F", 22x22 window) are viewed here as a downgraded version of "M4".

III. SERIAL FUSION ("MASTER-SLAVE")

A region of interest (ROI) described by the "M2" features (respectively "M4" or "M4F") are first classified by the classifier called the "Master". The reject threshold is set to a high value. If the ROI is rejected, then it is described by "M4" or "M4F" (respectively "M2") and is classified by the second classifier ("Slave"). For the "Slave", the reject threshold is smaller. The final decision is taken with the "Slave" and it can also be rejected.

The choice of the "Master" and the "Slave" depends on the performances of each group of features, on the processing time constraints, on the surface of the textured regions, etc.



Example of "Master-Slave" architecture with M2 and M4

IV. PARALLEL FUSION

Among all the possible parallel fusion schemes [3,6], we have selected 3 strategies which are better adapted to our studies with only two classifiers. Complementary between the two groups of features must enhance the global performance.

In the first scheme, the fusion is made at the measure level (here the post-probabilities). In the two other schemes, the fusion is made at the decision level, according to the Bayes theory and to the Evidence theory.

For the second and third strategies, *a priori* knowledge is required and it is given by the confusion matrix information.

The classifiers used here are Gaussian Quadratic Classifiers. For each classifier *k* (*k*=1...K), the class decision is made according to the maximum of the post-probabilities $P_k(x \in c_i / x)$ with i=1 to M, the number of classes and *x* the features vector to be classified. Here, K=2 and M=4 plus a reject class.

IV.1 Post-probabilities fusion

The fusion is done according to the post-probability generated by each classifier. With these probabilities, two types of classifiers combination are used: the "averaged classifier" and the "maximum classifier".

IV.1.1 Averaged classifier

For each class c_i, the K post-probabilities are then combined by the mean value as :

$$P_E(x \in c_i / x) = \frac{1}{K} \sum_{k=1}^{K} P_k(x \in c_i / x).$$

The final decision (the winner class) is the maximum value of these averaged post-probabilities P_E.

IV.1.2 Maximum classifier

The winner class is given by the maximum of P_k of each class, over all classifiers :

$P_E\left(x \in c_j / x\right) = \mathbf{max}\left(P_k\left(x \in c_i / x\right)\right)$ with i,j=1...M and k=1...K.

IV.2 Decision fusion

Decision fusion is weighted by the knowledge of the distribution of the errors for each classifier ($P_k(x \in c_i / e_k(x) = j_k)$ given by the confusion matrix). So, the winner class is estimated by a Belief function, meaning the proposition that *x* belongs to class c_i :

$$\mathrm{Bel}(i) = P\left[x \in c_i / e_1 = j_1, e_2 = j_2 \ldots e_K = j_K\right]$$

The winner class is defined by:

$$\mathbf{arg}\left\{\mathbf{max}\left[\mathrm{Bel}(i),\ i = 1 \ldots M\right]\right\}.$$

If the classifiers are independent then

$$\mathrm{Bel}(i) = \eta \prod_{k=1}^{K} P(x \in c_i / e_k(x) = j_k)$$

where η is the normalisation factor, computed as:

$$\frac{1}{\eta} = \sum_{i=1}^{M} \prod_{k=1}^{K} P\left[x \in c_i / e_k(x) = j_k\right].$$

IV.3 Theory of evidence

The theory of evidence was developed by Dempster and Shafer [3,6]. Its goal is to combine the propositions or evidences following the *a priori* knowledge given by the matrix confusion. For the K classifiers: εr_k is the recognition rate (mean value of the diagonal components) and εs_k substitution rate (mean value of the no diagonal components). Given a number of exhaustive and mutually exclusive propositions $A_i = x \in c_i,\ i = 1 \ldots M$, which form a universal set $\Theta = \left\{A_1 \ldots A_M\right\}$. "Bel($A_i$)" indicates the degree to which the evidences support the proposition A_i. It is calculated from a function called the basic probability assignment (*m, mass function*), which represents the individual impact of each evidence on the subset of Θ. Therefore, the combination rule is designed by $m(A) = m_1(A) \oplus \ldots \oplus m_K(A)$. Let us notice here only the illustrative formulas for this architecture with two classifiers. Then, in this simple case, the mass function is calculated as:

i) if the classifiers have the same decision A_i ,

$$m(A_i) = \eta\left[m_1(A_i).m_2(\Theta) + m_1(A_i).m_2(A_i) + m_1(\Theta).m_2(A_i)\right]$$

with $\eta^{-1} = 1 - m_1(A_i).m_2\left(\overline{A_i}\right) - m_1\left(\overline{A_i}\right).m_2(A_i)$

i) if the classifiers have two different decisions A_1 and A_2:

$$m(A_1) = \eta\left[m_1(A_1).m_2(\Theta) + m_1(A_1).m_2\left(\overline{A_2}\right)\right]$$

$$m(A_2) = \eta\left[m_1(\Theta).m_2(A_2) + m_1\left(\overline{A_1}\right).m_2(A_2)\right]$$

with $\eta^{-1} = 1 - m_1(A_1).m_2(A_2)$

The elementary mass function are $m_k(A_k) = \varepsilon r_k$, $m_k\left(\overline{A_k}\right) = \varepsilon s_k$ and $m_k(\Theta) = 1 - \varepsilon r_k - \varepsilon s_k$.

The Belief function is determined by $Bel(A_i) = m(A_i)$.

VI. PERFORMANCE OF THE FUSION SCHEMES

Eight textures were selected from the Brodatz album[2], organised into two patchwork images (256x256 pixels) of 4 textures (see figure 1).

For each separated features set ("M2", "M4" or "M4F") and for their fusion ("M2M4", "M2M4F"), the recognition rate is estimated by a cross validation method. The learn set is extracted from the isolated textures (100 samples per class) and the test set comes from a sequential scanning of the patchwork image.

The figure 2 shows the different recognition rates for the isolated features and for their fusion.

- ♦ In all the cases, the recognition rates are increased with the fusion (up to 12 % for the second dataset and up to 5 % for the first one) comparing to the isolated features ("M2", "M4" and "M4F").
- ♦ Concerning the fusion of "M4F" and "M2" ("M2M4F"), the resulting performance is a downgraded version of the fusion "M2M4". Comparated to this combination, the deterioration is down to 8 % for the second datasets and down to 3 % for the first one. But the "M2M4F" fusion is less time consuming than "M2M4.
- ♦ The decision fusion through the Bayes theory or the Evidence theory both use *a priori* information about each features set by the confusion matrix. For the former, all the estimated error probabilities are used to computed the function "Belief". For the later, only the recognition and the substitution rates are used. Then in this case, a greater robustness (due to the mean operation to compute εr_k and εs_k) on the fusion recognition is obtained relatively to the variability of the estimation of the confusion matrix. Table 1 shows the best stability for the DS method. Here the matrix confusions (*a priori* information) are estimated through an averaged "Hold Out" error counting method (the mean, maximum and minimum values are obtained over 20 estimations).

	Dataset 1			**Dataset 2**		
	Mean	**Min**	**Max**	**Mean**	**Min**	**Max**
DS (M2M4)	90,56	+0	-0	77,04	-0	+0
Bayes (M2M4)	86,3	-1,58	+3,65	72,55	-4,29	+5,83
DS (M2M4F)	88,37	-1,79	+2,7	68,98	-0	+0
Bayes (M2M4F)	86,71	-3,43	+2,28	73,7	-4,02	+2,7

Table 1 : Robustness of the Bayes and Evidence theories to the matrix confusion estimation.

- ♦ Relatively to the serial fusion ("M2M4" for both databases), the recognition rate can be enhanced by parallel fusion (up to 2 % for the second dataset and up to 1.5 % for the first one). Here, the enhancement by the parallel fusion is not too advantageous compared to the serial fusion.

An accurate analysis of the localisation of the misclassified pixels in the patchwork images shows that "M2" and "M4" are complementary in almost all the cases, inside the textured regions. This is not the case for the boundaries where the misclassification errors are randomly distributed. This property suggests that the very simple fusion scheme "Master-Slave" is better adapted to the texture features. And also, the computation complexity required for the parallel fusion is greater than for the serial scheme.

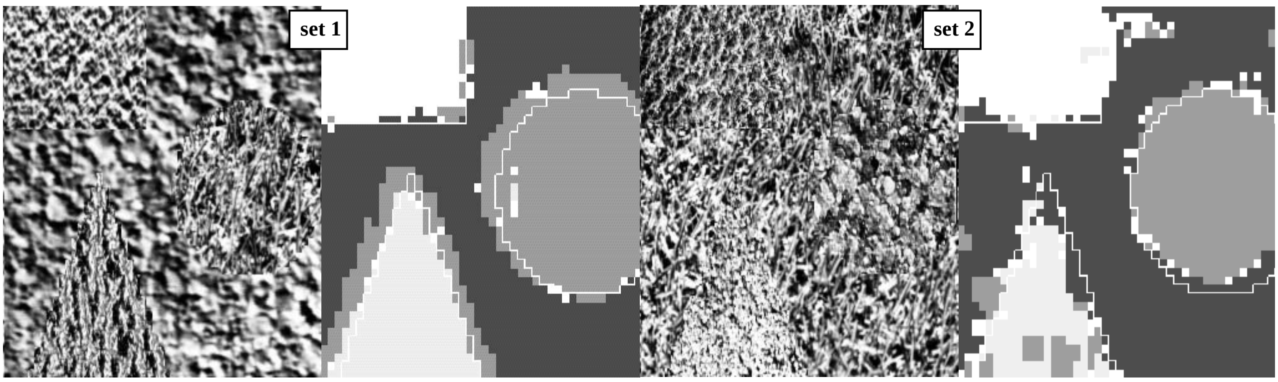


figure 1: Segmentation according to "Master-Slave" architecture ("M2" as master, "M4" as slave)

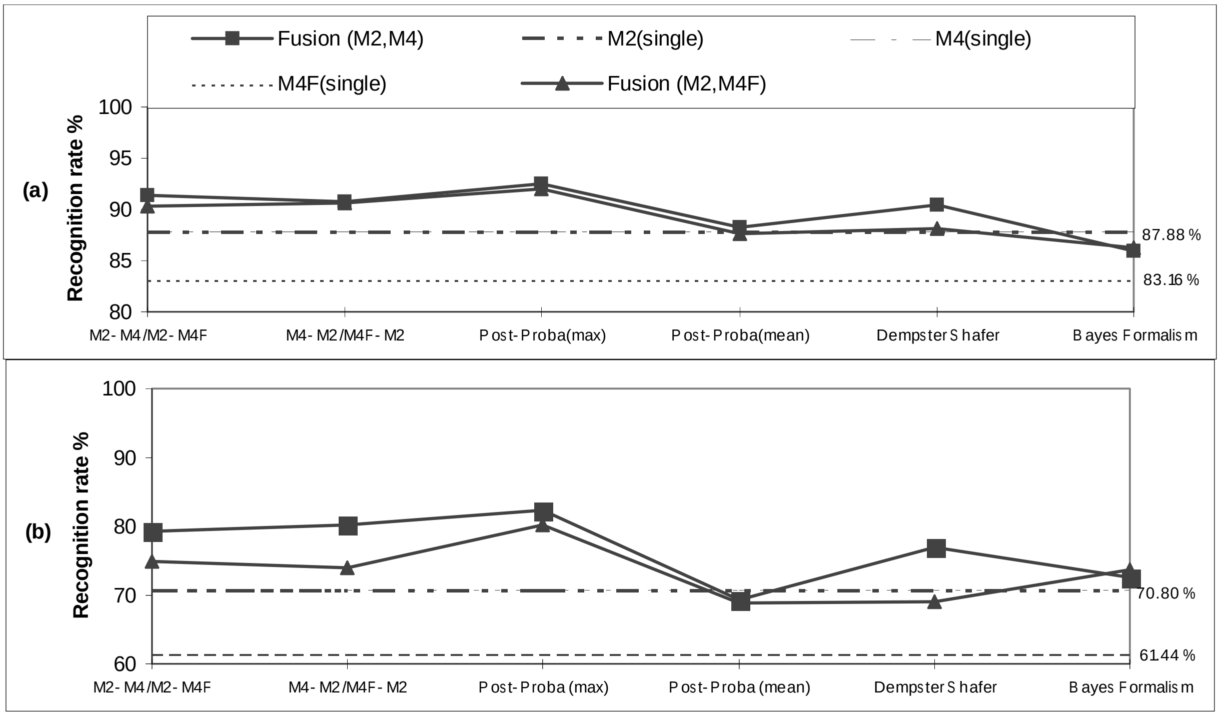


Figure 2: Recognition rate for the different fusion methods : (a) first dateset, (b) second dataset.

VII. CONCLUSION

In all the fusion cases, the recognition rates are increased (up to 12%) compared to the isolated features. "Master-Slave" architecture is justified if we want to reduce the computation time. So, the "Slave" features are not estimated in the whole image but only when the "Master" cannot discriminate. Nevertheless, when the "Master" makes a misclassification with a great a posterior probability, the "Slave" is not "called" for this pixel or region. This situation can be avoid in certain cases by taking into account the whole behaviour of the features with parallel fusion architectures. Then, it is necessary to compute for each pixel test, the two groups of features. The performances and also the processing time are obviously increased. Here, a reduction of about 40% of the processing time has been achieved with this serial architecture compared to the parallel. In parallel schemes fusion, best stability for DS method is obtained compared to Bayes method (both use a *priori* information about each features set given by the confusion matrix). Due to the mean operation to compute the recognition and the substitution rate (which are used to determine the Belief function), contrary to Bayes method where all the estimated error probabilities are used to computed the function "Belief".

Finally, the main problem in the serial and in the parallel fusion is the texture borders because the misclassification errors in both methods are randomly distributed.

REFERENCES

[1] Amblard P. O., Lacoume J. L., "Statistiques d'ordre supérieur des signaux complexes stationnaires et non-stationnaires", GRETSI, pp. 121-124, Juan-Les-Pins, France, sept 1993.

[2] Brodatz P., "Textures: A Photographic Album for Artists and Designers", Dover Publications INC., june 1966.

[3] G. Shafer., "A mathematical theory of evidence", Princeton, NJ., Princeton University Press, 1976.

[4] Guérin-Dugué A., Aviles-Cruz C., "Modified fourth order moments for texture recognition", Workshop on System Identification and High Order Statistics, Sophia Antipolis, France, sept 1993.

[5] Haralick R. M.et al., "Texture features for image classification", IEEE Trans on SMC, vol 3, n° 1, pp.610-621, 1973.

[6] L.Xu, A. Crzyzak and C. Y. Suen R., "Methods of combining multiple classifiers and their applications to handwriting recognition", IEEE Transaction Systemes, Man and Cybernetics , vol 22, n° 3, May/June 1992.

[7] Ramponi G., "High order moments for texture segmentation in presence of noise", Workshop on High Order Statistics in Signal Processing, Edinburg, UK, sept 1994.

[8] Tsatsanis M. K., Giannakis G. B., "Object and Texture Classification using High Order Statistics", IEEE Trans SMC, vol 14, n° 7, pp.733-750, 1992.