ELSEVIER

Signal Processing: *Image Communication* 11 (1997) 83–93

SIGNAL PROCESSING: IMAGE COMMUNICATION

Empirical multiresolution models applicable to gray-level image processing[1]

María José García-Salinas[a], Juan Francisco Gómez-Lopera[a], Pedro Luis Luque-Escamilla[b], José Martínez-Aroza[c], Ramón Román-Roldán[b,*], J. Carlos Segura-Luna[d]

[a] *Departamento de Física Aplicada, Universidad de Almería, Spain*
[b] *Departamento de Física Aplicada, Facultad de Ciencias, Universidad de Granada, Av. de Fuente Nueva, Spain*
[c] *Departamento de Matemática Aplicada, Universidad de Granada, Spain*
[d] *Departamento de Electrónica, Universidad de Granada, Spain*

Received 15 November 1995

Abstract

This paper deals with empirical multiresolution linear models intended for image processing. Such models contain information about gray-level composition of regions in the image. First, a general method for building these models from samples of selected images is described. Then, a measure of their quality, based on the Jensen–Shannon divergence, is introduced. This divergence is also used as a distance measure for classifying images. Applications in non-linear image filtering are provided, giving better result than classical median filtering. © 1997 Published by Elsevier Science B.V.

Keywords: Gray-level image; Multiresolution histogram relationships; Probabilistic linear empirical models; Model-based filtering; Image classification; Jensen–Shannon divergence

1. Introduction

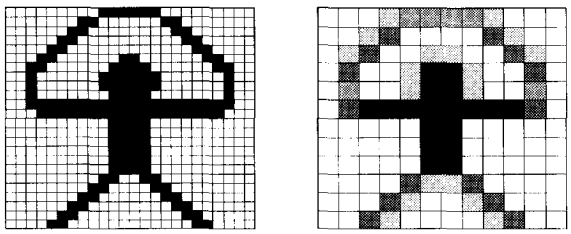
In the frame of Image Processing, a theoretical study of interrelations between gray levels of an image at different resolutions was developed by the authors [7]. As a result, certain image characteristics can be collected into a *multiresolution linear model*, which is just a matrix whose elements are probabilities of finding a particular gray level in a region of a certain size. Each model identifies an image class with a common set of characteristics, and so it can be used as prior information in order to either estimate histograms across resolutions or develope a model-based filter. In this line, this paper analyzes the empirical models, as well as their quality and filtering applications.

We deal [10] with digital gray-level images, mainly with gray-level histograms at different resolutions. Discrete gray scales (binary at the finest resolution) and additivity between gray levels are assumed. The aim is to relate gray levels and histograms at a finer resolution to gray levels and histograms at a coarser one.

[1] This work was partially supported by grant TIC91-646 from the DGCYT of the Spanish Government.

*Corresponding author.

PII S0923-5965(96)00036-7



a b

Fig. 1. An image observed at two resolutions: (a) $m = 0$, $k = \{0, 1\}$ 1 pixel = 1 dot; (b) $m' = 2$, $k' = \{0, 1, 2, 3, 4\}$ and 1 region = four pixels.

1.1. Interdependence of histograms at different resolutions

Let us consider an image observed at two different resolutions $m < m'$ in such a way that every m'-region is a set of adjacent m-pixels (see Fig. 1). Because of the assumed additivity, $k' = \sum k$ for all m-pixels included. The image histograms at both resolutions are connected by the linear relationship

$$p_{mk} = \sum_{k' \in K_{m'}} q(m, k \mid m', k') p_{m'k'}, \quad k \in K_m, \quad \frac{R_{m'}}{R_m} \in N. \tag{1}$$

Coefficients $q(m, k \mid m', k')$ stand for the average proportion of (m, k)-pixels inside (m', k')-regions and constitute the so-called "*composition matrix*" $\boldsymbol{Q}_{m,m'}$, with an opposite sense to the *aura matrix* defined by Picard [8], both in the general context of the pixel-neighbourhood relationships. Columns of $\boldsymbol{Q}_{m,m'}$ are the m-histograms averaged for all k' regions.

For our particular interest, each matrix $\boldsymbol{Q}_{m,m'} = \{q(m, k \mid m', k')\}$ (m and m' fixed), will be considered an *image model* and defines the concept of *image class*: that set of images with an internal structure obeying the coefficients $q(m, k \mid m', k')$ and therefore the relationship (1).

Not any arbitrary matrix is a composition matrix. The necessary and sufficient condition for being a composition matrix [6] is that every column $\boldsymbol{Q}_{m,m',k'}$ be a linear convex combination of possible, real m-histogram of (m', k')-regions.

An image model $\boldsymbol{Q}_{m,m'}$ can be obtained either theoretically by formulating hypotheses about the internal structure of the regions [7], or empirically, by counting regions and pixels from a preselected training set of images and by averaging their partial m-histograms for each k'. The latter is the subject of this paper.

2. Empirical models

Instead of making prior assumptions about the structure of an image class, $\boldsymbol{Q}_{m,m'}$ can be directly computed from a selected sample. Thus, the model is empirically determined, obtaining each matrix column, $\boldsymbol{Q}_{m,m',k'}$, by averaging the m-histograms of all the (m', k')-regions in the sample. Let $N_{m',k'}$ be the number of (m', k')-regions in the sample and $\mathscr{P}_m^{(i)}$, the *m-histogram* of the ith (m', k')-region $(i = 1, \ldots, N_{m',k'})$.

$$\boldsymbol{Q}_{m,m',k'} = \frac{1}{N_{m',k'}} \sum_{i=1}^{N_{m',k'}} \mathscr{P}_m^{(i)}. \tag{2}$$

By doing so for each k', a composition matrix $\boldsymbol{Q}_{m,m'}$ is obtained. The sample and the images to which the model is applied do not need an acquisition system with specific features, but the capture of these images must be done always under the same illumination conditions. That is to say, for a specific filter application, the samples for obtaining the empirical model should be processed by the same equipment and under the same conditions as in images to be filtered. These models are to be used in filtering applications, and also in image classification. It is therefore necessary to study its goodness or quality.

For each k', the m-histograms computed from all (m', k')-regions are averaged; thus, in order to know the quality of the obtained model, both the similarity (homogeneity) between the averaged m-histograms and the number of them (sample size) must be considered.

2.1. Sample homogeneity: the Jensen–Shannon divergence

A badly selected sample of images could lead to a model with very different structures, and

therefore, with a loss of validity for reaching specific objectives in image processing. In order to assess the degree to which the model corresponds to a well-defined structure, the *Jensen–Shannon (JS) divergence* measure between probability distributions has been used [5] (other measures are in [3, 4]). This measure shows the heterogeneity in the set of gray-level histograms. The (unweighted) JS-divergence is

$$\mathrm{JS}(\mathscr{P}_1, \mathscr{P}_2, \mathscr{P}_3, \ldots, \mathscr{P}_N) = H\left(\frac{1}{N}\sum_{i=1}^{N}\mathscr{P}_i\right) - \frac{1}{N}\sum_{i=1}^{N} H(\mathscr{P}_i),$$

where $H(\mathscr{P}) = -\sum p_k \log p_k$ is the Shannon entropy. Applied to model columns,

$$\mathrm{JS}(\mathscr{P}_m^{(1)}, \mathscr{P}_m^{(2)}, \ldots, \mathscr{P}^{(N_{m',k'})}) = H(\boldsymbol{Q}_{m,m',k'}) - \frac{1}{N_{m'k'}}\sum_{i=1}^{N_{m'k'}} H(\mathscr{P}_m^{(i)}). \quad (3)$$

There is a JS-divergence measure for each matrix column; therefore a divergence vector with $R_{m'} + 1$ elements stands for the image model.

2.2. An approach for classifying images

When an empirical model is constructed, the most common case is to find differences between the m-histograms of the (m', k')-regions for each k', and this can lead to a very mixed model. That is why a less restrictive class definition is proposed: an image class is now defined as that containing all images in which the m-histograms for each k' are similar, the similarity being evaluated by means of the Jenssen–Shannon divergence. For the divergence vector, the lower its elements are valued, the more unified the corresponding model is, thus representing a better-defined image class.

On the other side, this JS-divergence may also be used to compare a single image matrix ($\boldsymbol{Q}^*_{m,m'}$) with the model $\boldsymbol{Q}_{m,m'}$ of a whole class (called prototype [2]) in order to verify whether that image belongs to this class or not. Fig. 2 provides an example of classification of an image. For each gray level k', the Jenssen–Shannon divergence $\mathrm{JS}(\boldsymbol{Q}^*_{m,m'k'}, \boldsymbol{Q}_{m,m',k'})$ for several models $\boldsymbol{Q}_{m,m'}$ available has been plotted. It is clear that for nearly all gray levels the distance (as given by JS) between the image and the class prototype to which it

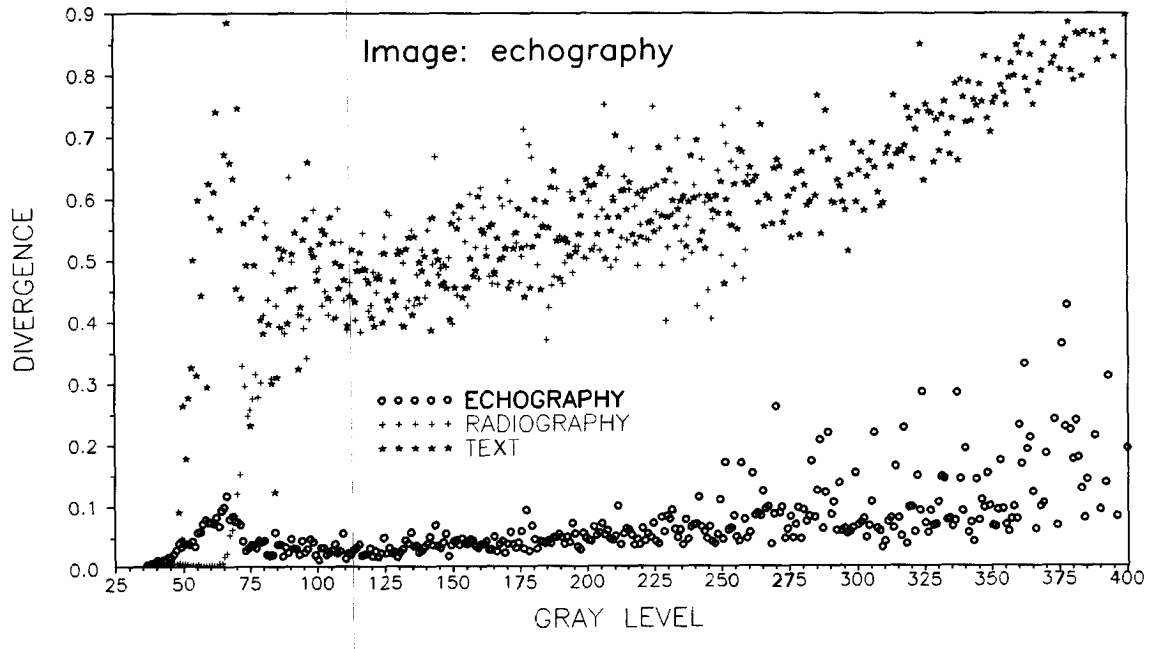


Fig. 2. Divergence values for the classification of an echograph. The gray scale $\{0, \ldots, 63\}$ and a region 3×3 size has been used.

Table 1
Discriminating indices for the examples of Section 2.2

Image to classify (i)	DI	DI	DI
	DI(i, RX)	DI(i, TX)	DI(i, EC)
rx	1	76,16	29,66
tx	3,18	1	2,80
ec	6,41	8,78	1

belongs is much less than the distances to the other prototypes.

Now, let us introduce the following notation: A, RX ... stand for image classes, while a, rx ... stand for single images ($a \in A$, rx $\in$ Rx, ...); $\boldsymbol{Q}(A)$, $\boldsymbol{Q}$(RX) denote class prototypes (models of classes A, RX) and $\boldsymbol{Q}^*(a)$, $\boldsymbol{Q}^*$(rx) stand for composition matrices of single images a, rx ... A *discriminating index* (*DI*) in the classification of an element $a \in A$ with respect to B is defined as the quotient between the following column-averaged divergences:

$$\mathrm{DI}(a, B) = \frac{\overline{\mathrm{JS}}(\boldsymbol{Q}^*(a), \boldsymbol{Q}(B))}{\overline{\mathrm{JS}}(\boldsymbol{Q}^*(a), \boldsymbol{Q}(A))}. \tag{4}$$

Some classification examples done have the values for DI given in Table 1. The DI is higher for radiographs, confirming that it is the most homogeneous class. The used DI, based on the JS-divergence measure, has proved to be better than others based on different measures (Kullback, etc.).

2.3. Training sample size

For a given resolution pair (m, m'), the size of the sample is actually realized by a size vector whose elements, one for each $k' \in K_m$, are the numbers of k'-regions in the training sample, which is a visually selected set of images.

In the training process the sample size may be small or even null – that is, for one or more k' levels there may be no k'-regions. In this case, the matrix would have one or more undetermined columns. This difficulty is overcome by initializing the empirical models with an appropriate theoretical one having unit weight. The initialization proceeds by considering the theoretical base model as coming from a hypothetical sample. This means that the first-scanned samples are assumed as if they were regions corresponding to a theoretical model. Thus, a theoretical *base* model is added to every empirical one, in order to prevent undetermined columns. Since one theoretical m-histogram is included in every column, its influence depends on the sample size.

The most reliable procedure to select a suitable base model is a rather empirical one, by adding different bases to the empirical model with undetermined columns. This leads to several complete models which will have the same sample size but different divergence values. The best of the added bases can then be determined as that of the minimum increase in the original divergence.

Let us now consider the divergence of the uncompleted model and the divergence of the complete one. The two values differ because in the second model one more m-histogram has been averaged. The difference is

$$\Delta\mathrm{Div}(X) = \mathrm{JS}(\boldsymbol{Q}^+_{m,m',k'}) - \mathrm{JS}(\boldsymbol{Q}_{m,m',k'}),$$

where the letter X stands for the theoretical base added ($X = B, H, Y, Z, \ldots$) and $\boldsymbol{Q}^+_{m,m',k'}$ is the matrix column of the complete model.

In the following example, Fig. 3, $\Delta\mathrm{Div}(X)$ is plotted versus k' for $X = B, H, Y, Z, \ldots$. As it can be seen, a curve corresponding to a certain base remains clearly below the others: model B is the best to be added as a base for an echograph model. This behaviour is general for echographs.

2.4. Examples

Several models with different m and m' values have been done. The best way of visualizing them is by means of a 3-D representation of the matrices $\boldsymbol{Q}_{m,m'}$: gray levels k at the finer resolution m are on the X-axis; gray levels at resolution $m' > m$ are on the Y-axis. Probabilities (matrix's elements) are on the Z-axis.

Three examples of empirical models (for echographs, radiographs and texts) are shown in Figs 4–6. The gray scale at the finer resolution ($m = 6$) is $K_6 = \{0, \ldots, 63\}$. The region at the coarser resolution m' is nine pixels in size (for a 3×3 observation

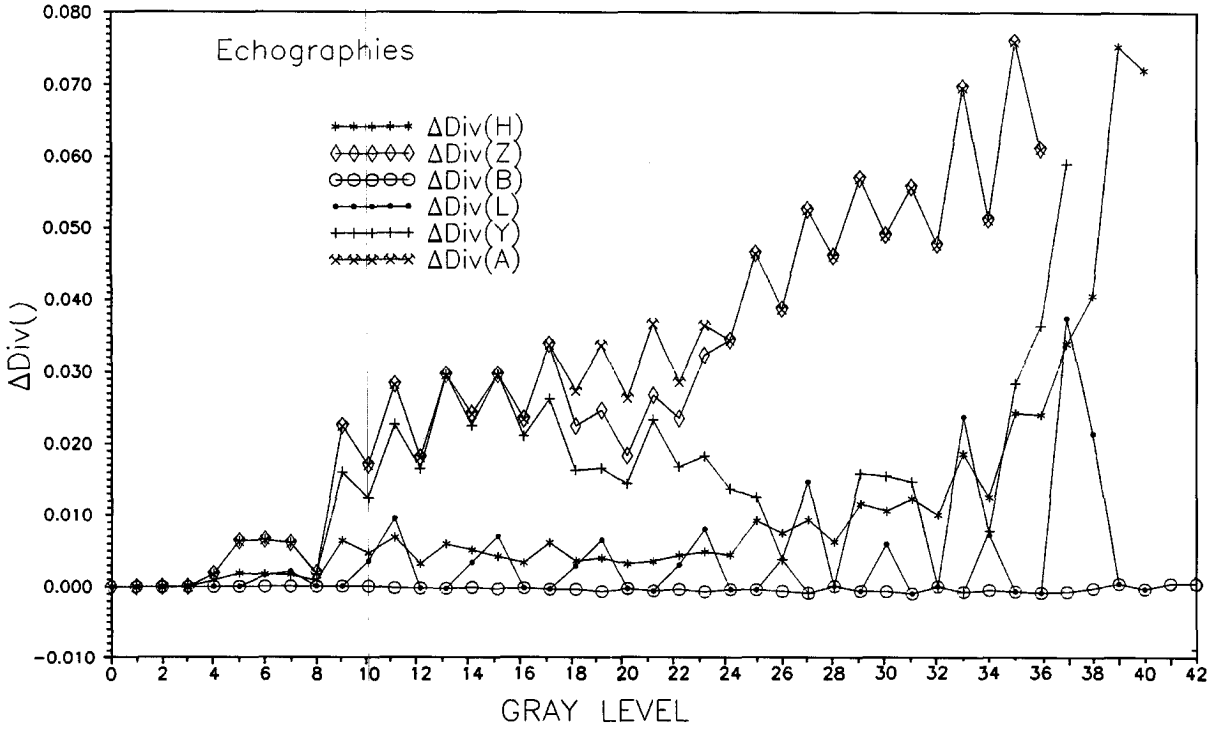


Fig. 3. Determination of a theoretical base model for an echograph one ($k = 0, \ldots, 15$ and region size 2×2).

window). Below each matrix, the divergence values and the sample size are represented for each gray level k'. U and W stand for unweighted and weighted (with sample size) averaged divergence, respectively.

3. Filtering applications

In previous works with theoretical models, some useful applications in histogram estimation and image processing [10] were developed. Recently, filters based on empirical models have been designed. In these, each time a pixel is to be processed, the gray level k' of a region containing it is observed, and the information given by the column $\boldsymbol{Q}_{m,m',k'}$ of the appropriate empirical model is used to find the filtered gray level k_f of the pixel. Different ways of using this information are the following:

(1) k_f is a central value (mean, median, mode) of the distribution $\boldsymbol{Q}_{m,m',k'}$.
(2) k_f is the result of an empirically weighted median filtering (EWMF). Median filters [1,9] sort the gray levels k of the pixels contained in a region and assign the middle value to the pixel centred in the region. The EWMF puts a weight to every k given by its probability in the corresponding $\boldsymbol{Q}_{m,m',k'}$ column of the empirical model.
(3) One of the previous procedures is performed for every region containing the pixel, and then again, a central value of these filtered gray levels is chosen.

The examples in Fig. 7 show the results of a noisy-image filtering using the EWMF. The source image (b) is the original one (a) corrupted by 30% of salt-and-pepper noise. First, the classical median filtering with region size 3×3 is shown in (c) for comparison. Part (d) is the result of an EWMF using a text model with no theoretical base

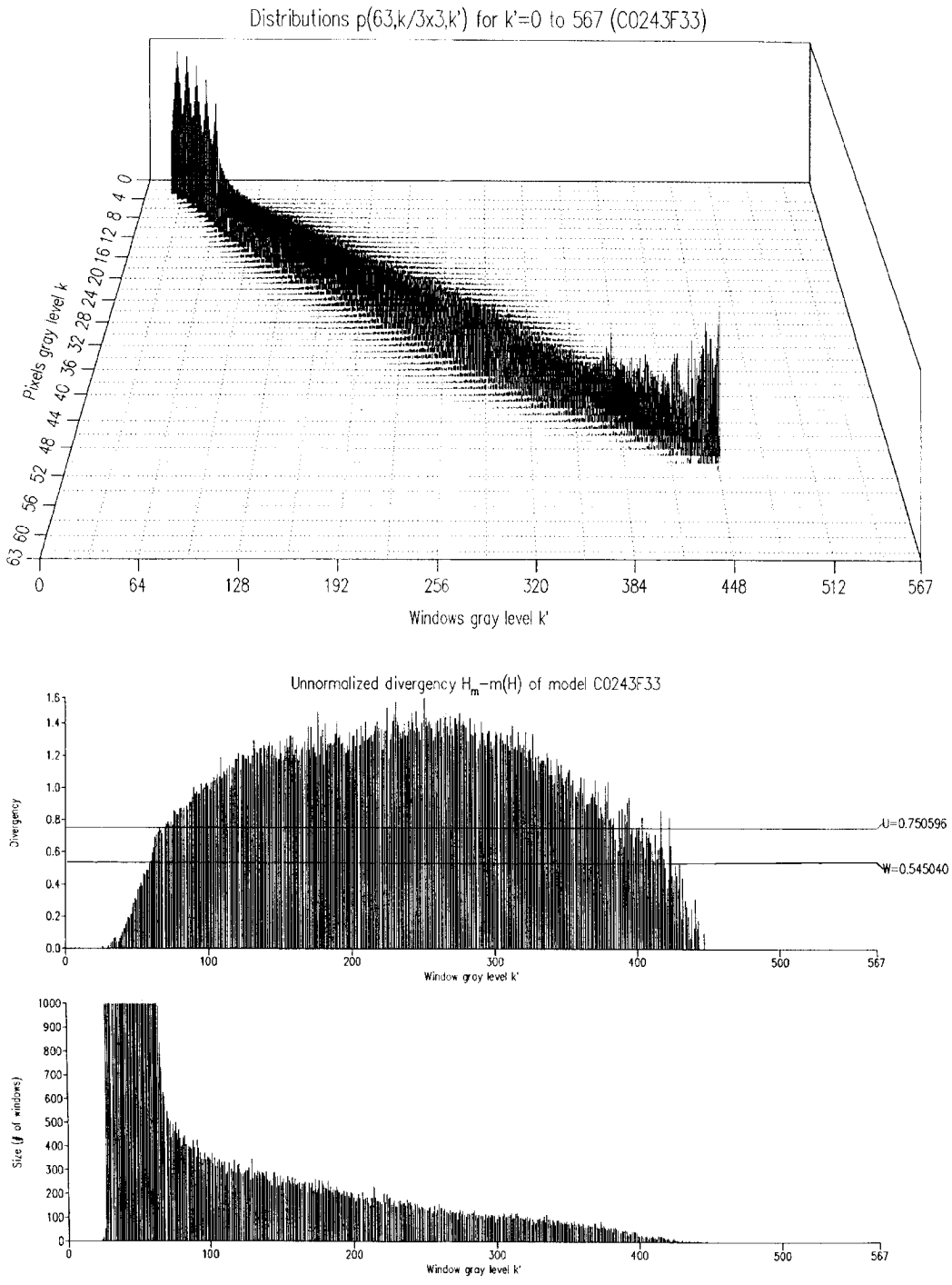


Fig. 4. Representation of the matrix of an echograph model. The divergence and histogram versus region gray-level are shown below.

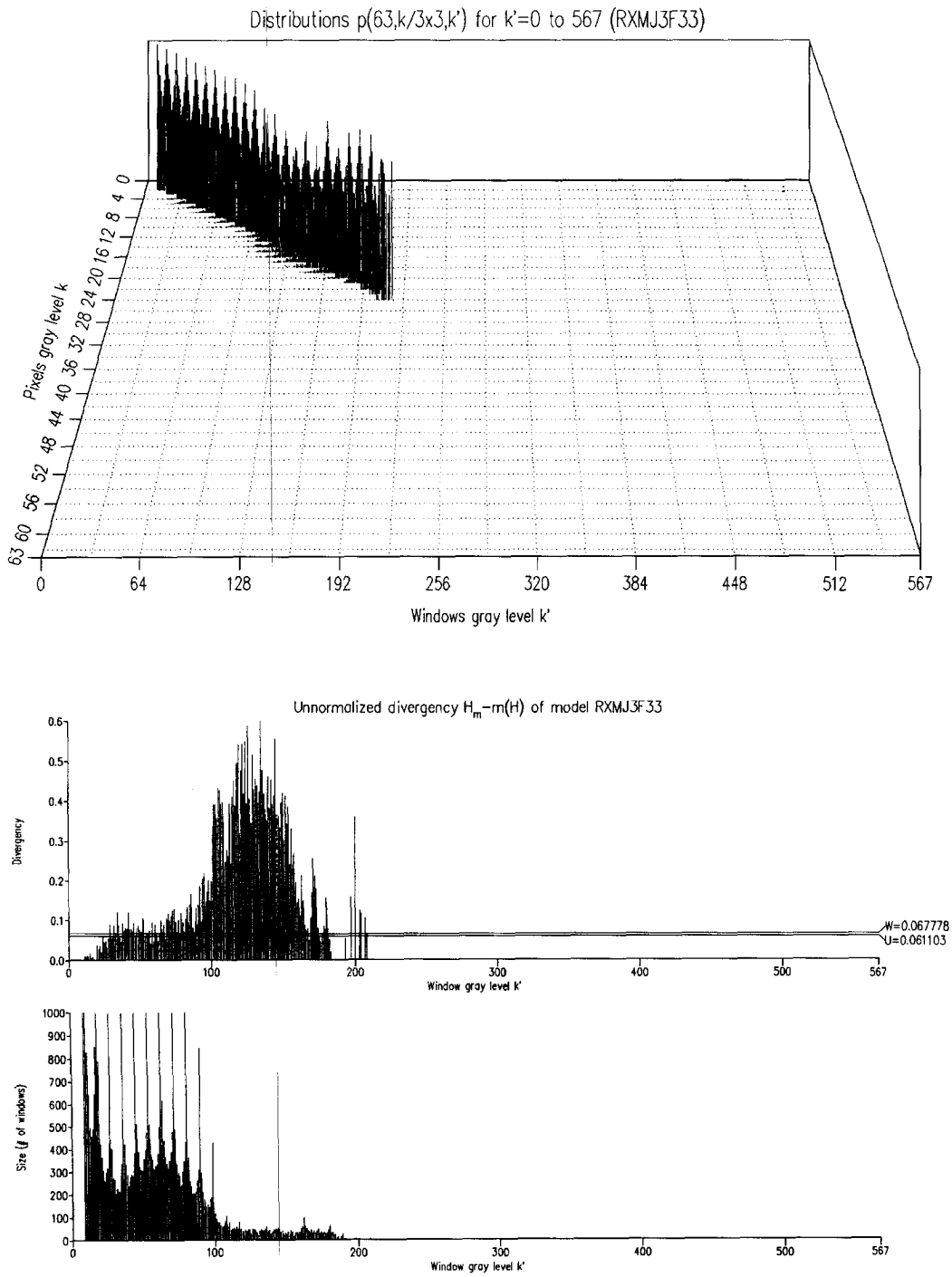


Fig. 5. Representation of the matrix of a radiograph model. The divergence and histogram versus region gray level are shown below.

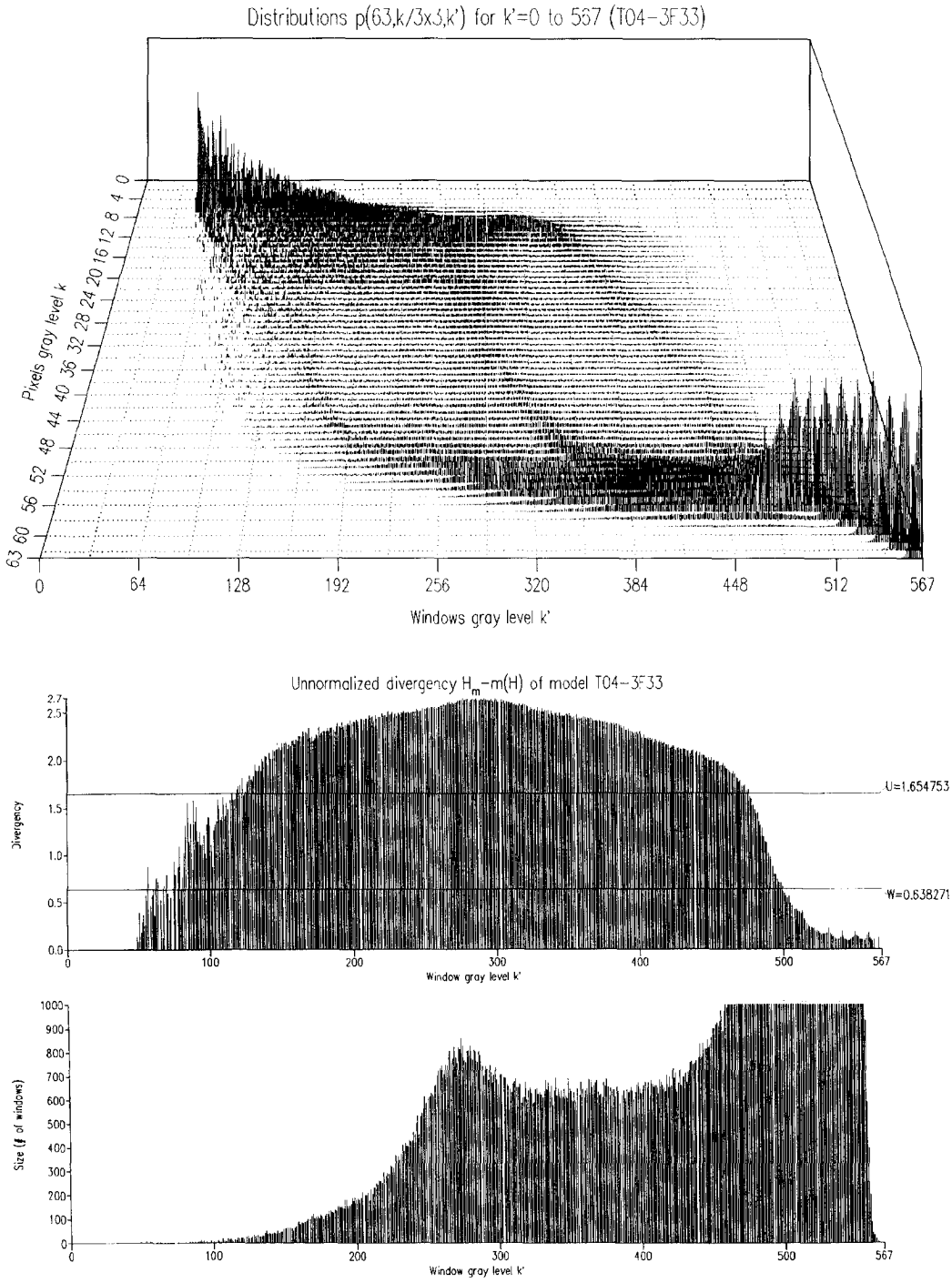


Fig. 6. Representation of the matrix of a text model. The divergence and histogram versus region gray level are shown below.

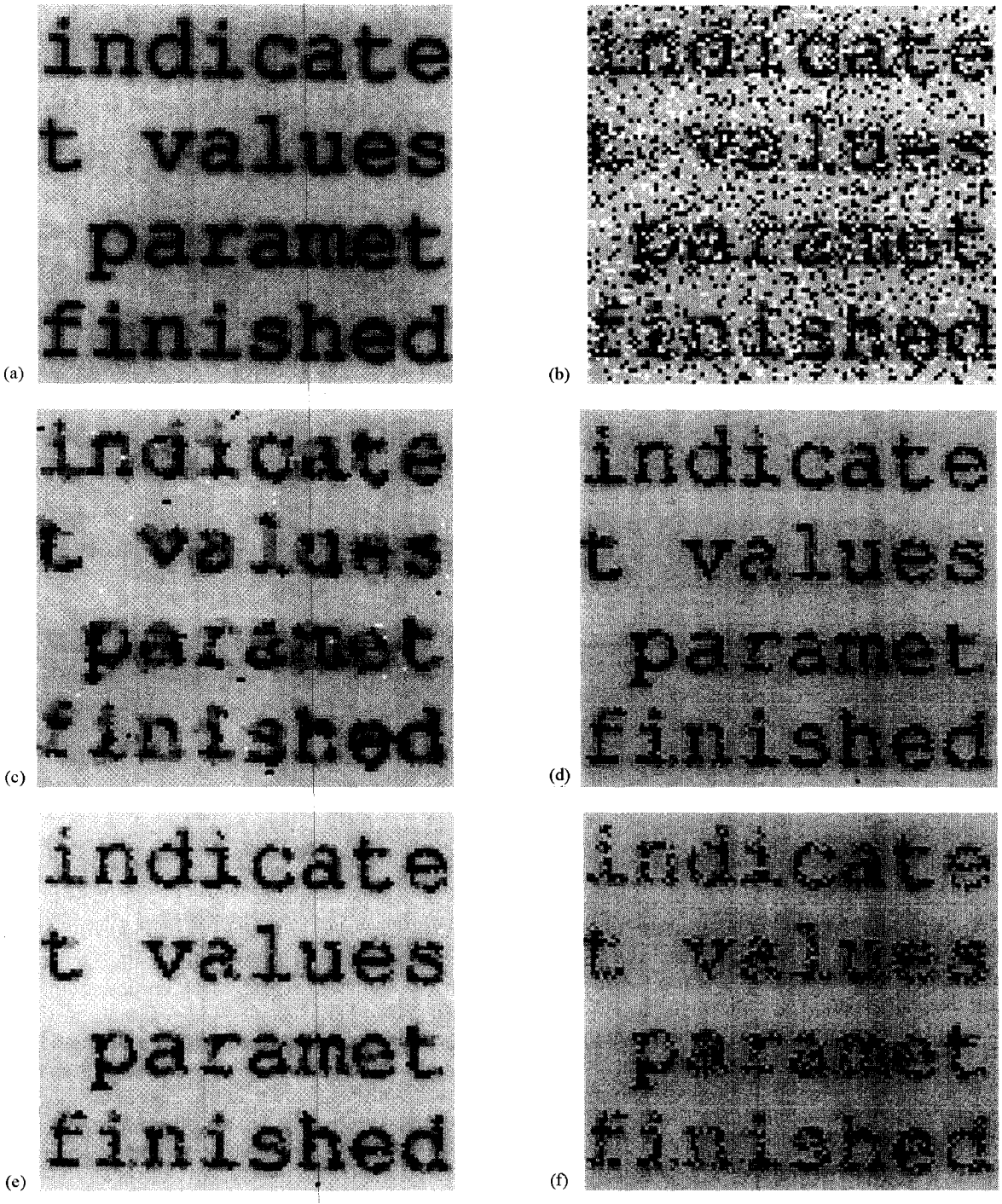
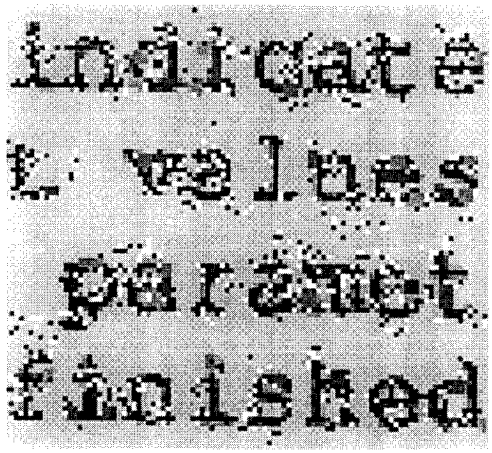
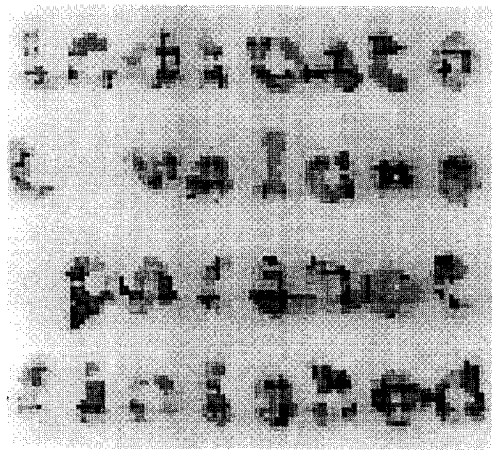


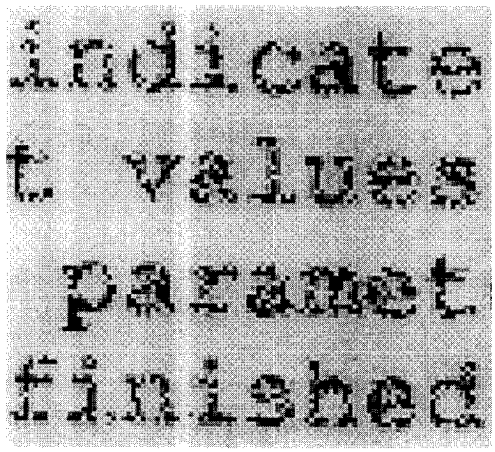
Fig. 7. Experimental results of filtering a text image: (a) original image; (b) source; (c) classical median filtering region 3×3; (d) EWMF with an uncompleted text model (without base); (e) EWMF with the same model completed with base H; (f) EWMF with the same model completed with base B; (g) theoretically weighted (B) median filtering; (h) classical median filtering region 5×5; (i) EWMF with an uncompleted text model (region 5×5).



(g)



(h)



(i)

Fig. 7. Continued.

added, and still having undetermined columns. When one of these is required, then the lack of information prevents filtering. The result was much better than the classical one, although the noise was not completely removed. If we choose a complete model, the results are shown in parts (e), with base H, and (f), with base B, where Model B is clearly the best choice: the noise was eliminated from the background and the next is easily readable. However, this is not due only to Model B, as demonstrated in (g), where the filter used only the theoretical Model B.

It might be thought that a larger region (such as 5×5) should remove the noise better in a classical median filtering, as shown in (h). The impulsive noise has in fact disappeared, but the text is now blurred and unreadable. The empirical filtering again gives a better result (i); even without any base, the noise is completely removed from the background and the text is still readable. Empirical models built from other image classes did not give any good result, even with a different region size.

EWMF made for source images belonging to other classes (radiographs and texts) exhibited the same behavior: the best result was always the one obtained with an EWMF using an allied model (of the same class) completed with the theoretical one that best suit that class.

4. Conclusions

In this paper the Empirical Multiresolution Models have been presented as a tool in image processing. A general method for building these models from samples of selected images has been described. The Jensen–Shannon divergence and the sample size are used as parameters for measuring the quality of a model. The Jensen–Shannon divergence has also been used as a distance measure to classify images. Finally, applications in image filtering have been outlined.

Notation and definition

dot	elementary component of a digital gray-level image. It can only be black or white
pixel	a rectangular maximally convex set of dots
image	array of equal-sized, equal-shaped pixels
$m = 0, 1, 2, \ldots$	resolution index (from fine to coarse)
R_m	size in dots of the pixel at the resolution m (called m-pixel)
$K_m = \{0, 1, \ldots, R_m\}$	gray scale for each resolution
$k \in K_m$	gray level of a pixel, given by the number of black dots included
$\mathscr{P}_m = \{p_{m,k}\}$	gray-level m-histogram
$\boldsymbol{Q}_{m,m'}$	image model matrix
$\boldsymbol{Q}_{m,m,k'}$	each k'-column of the above matrix ($k' = 0, 1, \ldots$)
JS	the Jensen–Shannon divergence
DI	discriminating index
EWMF	empirically weighted median filter

References

[1] M. Gabbouj, P. Haavisto and Y. Neuvo, "Recent advances in median filtering", in: E. Arikan, ed., *Communications, Control and Signal Processing*, Elsevier, Amsterdam, 1990.

[2] D.M. Gómez Allende, "Reconocimiento de formas y visión artificial", RAMA, Madrid, 1993, Chapter 2, pp. 15–38.

[3] R.W. Johnson, "Axiomatic characterization of the directed divergences and their linear combinations", *IEEE Trans. Inform. Theory*, Vol. IT-25, No. 6, November 1979, pp. 709–716.

[4] J. Kittler, "Feature selection and extraction in pattern recognition", in: R.A. Vaughan, ed., *Pattern Recognition and Image Processing in Physics*, Adam Hilger, Bristol, 1990.

[5] J. Lin, "Divergence measures based on the Shannon entropy", *IEEE Trans. Inform. Theory*. Vol. 37, No. 1, January 1991, pp. 145–150.

[6] J. Martínez Aroza, Contribución al análisis y estimación de imágenes digitales mediante la teoría de la información, Ph.D. Thesis, Universidad de Granada, 1990.

[7] J. Martínez Aroza and R. Román Roldán, "Probabilistic linear models for multiresolution estimation in gray-level images", *Multidimensional Systems and Signal Processing*, 1995.

[8] R.W. Picard and I.M. Elfadel, "Structure of aura and co-occurrence matrices for the Gibbs texture model", *J. Math. Imaging Vision*, Vol. 2, 1992, pp. 5–25.

[9] I. Pitas and A.N. Venetsanopoulos, *Nonlinear Digital Filters. Principles and Applications*, Kluwer Academic Publishers, Dordrecht, 1990.

[10] R. Román Roldán, J.J. Quesada-Molina and J. Martínez Aroza, "Multiresolution-information analysis for images", *Signal Processing*, Vol. 24, 1991, pp. 77–91.