ENTROPY AND MULTISCALE ANALYSIS: A NEW FEATURE EXTRACTION ALGORITHM FOR AERIAL IMAGES

Alexandre Winter †‡, Henri Maître ‡, Nicole Cambou † and Eric Legrand †

‡ Département IMA, ENST, 46 rue Barrault, F-75013 Paris, France

† Aérospatiale, service Vision, 1 rue Pablo Picasso, F-78114 Magny-Les-Hameaux, France

ABSTRACT

This paper presents a new, fully automatic and robust feature extraction algorithm based on the selection of a given range of scales. It compares consecutive band-pass images of a Gaussian multiscale decomposition to extract the objects that appear between given scales. The comparison is performed using original distributed entropic measures. The application to building detection in aerial images shows that scale is a robust and precise criterion for the detection of man-made objects. They also show that distributed entropic tools are relevant for the comparison of band-pass images.

From a more theoretical point of view, this method stands between Scale-space and wavelet approaches. It tries to infer the geometrical conception of scale found in the Scale-space theory into the algebraic scale of the wavelet theory.

Keywords: *Multiscale analysis, wavelets, Scale-space, feature extraction, entropy, aerial imagery, HR satellite imagery*

1. INTRODUCTION

Different notions of scale exist in the multiscale literature. Scale-space theories consider that representing an image at a larger scale means smoothing or diffusing it. In this paper, as in the wavelet theory, we consider scale as an elementary dimension. An image of a multiscale decomposition at scale t is made of objects and details whose dimension is close to t, in at least one direction. Every feature of an image should then be present in a limited range of scales: at coarse scales, the object is too small to be visible, while at fine scales, the details of the object are visible, but not the object as such. In this paper, we present an original, fully automatic and robust feature extraction method, called "entropic scalar detection", based on this parallel between an object and its scale (see figure 1). It uses original distributed entropic measures to compare consecutive band-pass images of a Gaussian multiscale decomposition, so as to extract the objects that appear between the two scales.

This work was supported by ANRT grant CIFRE #94446

E-mail: winter@ima.enst.fr

Web: http://www-ima.enst.fr/~winter/

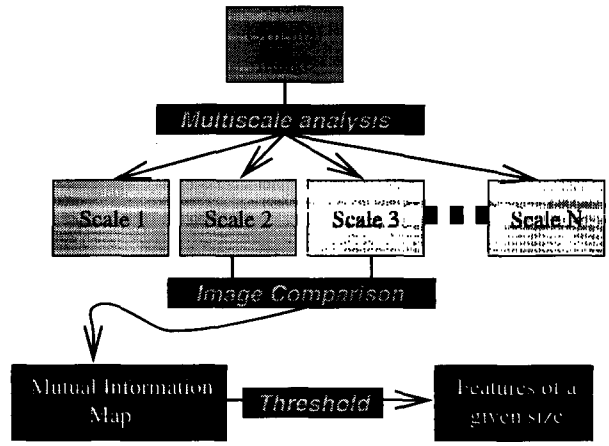


Figure 1: Entropic scalar detection

Section 2 describes the multiscale decomposition we use, while section 3 presents the magnitude we shall use to measure innovation between scales. Section 4 describes the algorithm and how it is made fully automatic. Section 5 shows some experimental results and compares them with a classical algorithm. Section 6 brings conclusions and perspectives about this peculiar approach of multiscale analysis.

2. MULTISCALE DECOMPOSITION

Let the multiscale analysis μ_N of rank N associate the detail images $X_{n,n\in\{0,\ldots,N-1\}}$ and the low-pass image X_N to an original image X:

$$\mu_N : X \mapsto \mu_N(X) = \{X_{n,n\in[0,N-1]}, X_N\} \qquad (1)$$

This work aims at comparing scale bands X_n in the same image space, i.e. the original image space. Thus, no sub-sampling will be processed. Moreover, we do not wish to favor any spatial direction in the detail images. Therefore, we use a classical Burt and Adelson algorithm (see [BA83]) with a Gaussian smoothing kernel, in order to favor neither frequency nor space localization.

Copyright 1997 IEEE

3. ENTROPY FOR IMAGE COMPARISON

This section defines original "scale-innovation maps" and the entropic measures they involve.

3.1. Entropy

Considering an image X of N_c columns and N_l rows, with a probability function $P_X(k), k \in \{0,\ldots,255\}$, Fisher's information quantity q is defined as [Shan78]:

$$\forall k \in \{0,\ldots,255\}, \quad q_X(k) = -\log P_X(k). \quad (2)$$

q_X is the amount of information brought by a pixel $X_{i,j}$ whose grey level is k. The information brought by the image X is the average of q, called the entropy $H(X)$:

$$H(X) = \frac{1}{N_c \times N_l} \sum_{i=0}^{N_l-1} \sum_{j=0}^{N_c-1} q_X(X_{i,j}). \quad (3)$$

Considering another image Y and the joint probability function $P_{X,Y}(k,l), (k,l) \in \{0,\ldots,255\}^2$, the joint information quantity $q_{X,Y}$ and the joint entropy $H(X,Y)$ are defined by replacing $P_X(k)$ with $P_{X,Y}(k,l)$ in equations 2 and 3. $H(X,Y)$ represents the entropy of X and Y seen as a 2D random source.

3.2. Mutual information

Mutual information — also known as Shannon's information — is defined as the amount of information brought by the "overlapping" of the two sources:

$$I(X,Y) = H(X) + H(Y) - H(X,Y). \quad (4)$$

A simple calculus leads to:

$$I(X,Y) = \frac{1}{N_c \times N_l} \sum_{i=1}^{N_l-1} \sum_{j=1}^{N_c-1} r_{X,Y}(X_{i,j}, Y_{i,j}), \quad (5)$$

where r is the mutual information quantity of two grey level occurrences $(k,l) \in \{0,\ldots,255\}^2$:

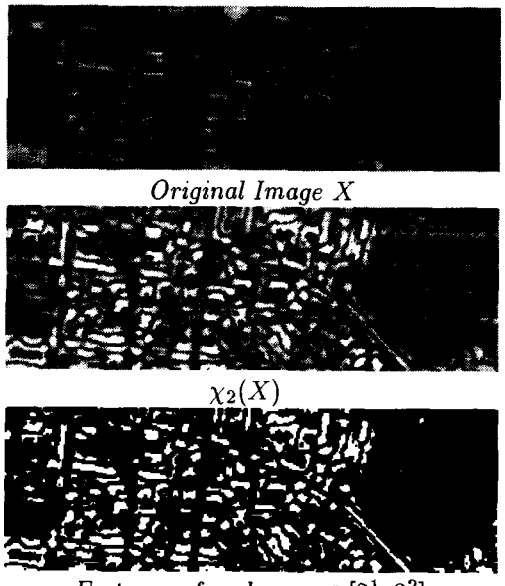
$$r_{X,Y}(k,l) = -\log \frac{P_X(k)P_Y(l)}{P_{X,Y}(k,l)}. \quad (6)$$

r is a measure of the information quantity brought by the coupling of k and l.

Let us notice that the contribution of any subset A of X and Y to $I(X,Y)$ can be computed as:

$$I_A(X,Y) = \frac{1}{N_c \times N_l} \sum_{(i,j)\in A} r(X_{i,j}, Y_{i,j}). \quad (7)$$

I_A is called the partial mutual information of A in (X,Y).



Original Image X

$\chi_2(X)$

Features of scale s, $s \in [2^1, 2^2]$

Figure 2: Original image, CIMENE and Entropic Scalar Detection images

3.3. A distributed entropic measure

Entropic measures have already been applied to image processing as global state variables of an image in [TU96, CVSM95, Dja93]. From an image analysis point of view, entropy and mutual information lack spatial localization. Thus, we propose a distribution process that enables the production of mutual information maps. A similar process was first used for self information in [Low84, Low83]. Consider two pixels $X_{i,j}$ and $Y_{i,j}$ of X and Y at the same position (i,j). Let every pixel of $\xi(X,Y)$ be defined as follows:

$$\xi(X,Y)_{i,j} = k \times r_{X,Y}(X_{i,j}, Y_{i,j}) \quad (8)$$

k being a normalization factor that enables ξ to be displayed using 256 grey-levels.

$\xi(X,Y)$ is called a CIM image. It gives every pixel (i,j) a grey tone proportional to the amount of information brought by the coupling of the two images at that point.

4. ENTROPIC SCALAR DETECTION

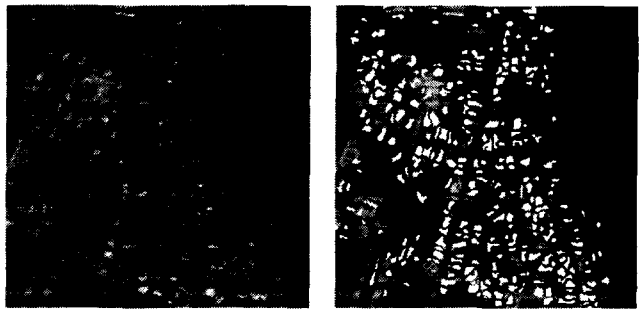
Now, let us use spatially distributed entropic measures on consecutive detail images of a multiscale analysis to perform Entropic Scalar Detection (see figure 1).

4.1. CIMENE Images

Considering an image X and a multiscale decomposition μ_N (equation 1), we define the CIM in scales, called CIMENE image of rank n as follows:

$$\forall n \in \{1,\ldots,N\}, \chi_n(X) = \xi(X_{n-1}, X_n) \quad (9)$$

Copyright 1997 IEEE



Original *2-basis ESD*

Figure 3: Building extraction using the ESD

The higher $\chi_n(X,Y)$ is, the more information the pixel of coordinates (i,j) brings to the coupling of the scales $n-1$ and n, and the more innovative this pixel is between the two scales. CIMENE images are thus called scale innovation maps.

Three different kinds of features appear on CIMENE images (see figure 2): black non-innovative lines that correspond to stable zero-crossings of the Laplacian (i.e. contours of the objects), large dark areas that correspond to large primitives that have the same appearance in the considered scales (i.e. large non-innovative primitives) and bright innovative objects that appear between the two scales (i.e. objects of a given range of sizes). It has been shown that those bright objects have at least one dimension in the interval $[2^{n-1}, 2^n]$ in a χ_n image.

4.2. Automatic Thresholding

The next step in Entropic Scalar Detection is to select the pixels that bring information to the coupling of scales in the CIMENE image, i.e. those with bright grey levels (see figure 1). The threshold that separates scale innovative pixels from the others is determined automatically as follows.

Consider a CIMENE image $\chi_n(X)$ and a threshold σ. A_σ is the set of pixels for which $\chi_n(X)_{i,j} > \sigma$. We define the mutual information detection ratio as $\rho(\sigma) = I_{A_\sigma}/I$. Experience shows that this ratio is constantly equal to $\rho_0 = 0.5$ for thresholds visually determined to provide homogeneous detections over a set of 42 CIMENE images. Thus, the threshold can be automatically computed on any image by selecting σ such that $\rho(\sigma)$ is as close to ρ_0 as possible. This method reaches the performances of a human operator, with a mean relative error of 5%, which is definitely satisfying since the reference thresholds are visually determined with an estimated precision of 5% by a human operator.

4.3. Entropic Scalar Detection Algorithm

Now that the three steps of entropic scalar detection exposed in figure 1 are defined, we can propose the algorithm of the Entropic Scalar Detector of rank j, called ESD_j:

- perform a dyadic multiscale analysis μ_{j+1},
- compute the CIMENE image $\chi_j(X)$,
- compute the threshold σ and threshold $\chi_j(X)$.

$ESD_j(X)$ contains all the objects whose size ranges from 2^{j-1} to 2^j in at least one direction.

5. EXPERIMENTAL RESULTS

Since the only Entropic Scalar Detection (ESD) selection criterion is size, it provides interesting features in terms of detection. In this section, we use the ESD to detect man made structures in aerial images.

5.1. ESD and Feature Extraction

The ESD is fully automatic and robust with respect to photometry, unlike classical feature extraction methods like the morphological top-hat (see [Ser82]). It is also robust across various sensors — it turned out to be stable on SPOT, simulated SPOT-5, ERS-1 images and aerial photographs. Eventually, the ESD can detect a bright roof on a dark soil as well as a dark building on a bright ground (see figure 2). This is all the more crucial as the criterion of size is often more accurate than any other criterion for the low-level detection of man-made objects like roads or buildings in aerial images.

5.2. Building Detection: Performances

Assuming that a building is about 7 m wide, we use an ESD of rank 2 on 1.75 m resolution aerial images to detect them. Indeed, this algorithm extracts the objects whose smaller dimension is between 3.5 m and 7 m. An example is presented in figure 3. The results, as compared with a human operator, are presented below for 5 different areas issued from 4 different regions of France at the same resolution:

Area	Flins	Hoen.	Labou.	Miram.	Truch.
Rate	75%	76%	96%	90%	79%

Moreover, the false alarm rate is quite stable no matter how many targets the image contains: despite its statistical nature, the ESD is robust with respect to the content of the image.

5.3. Comparison with the Top-hat Operator

The ESD and the classical morphological Top-hat operator provide the same kind of results when tuned on the same size intervals. In [Pap96], the results of the ESD have been compared to those of the Top-hat in order to assess the accuracy of the methods and the relevance of the criteria of both algorithms.

Copyright 1997 IEEE

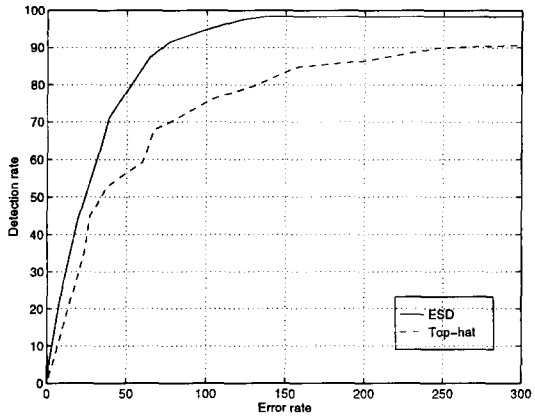


Figure 4: Comparison of the error versus detection rate plots of the *ESD* and the morphological Top-hat.

Our reference database is made of manually detected buildings in aerial image. We compared the error rate versus detection rate plots provided by Top-hat and CIMENE images thresholded at various levels (figure 4). It turned out that the *ESD* behaves better than the Top-hat on all the tested images: for the same error rate, the thresholded CIMENE image contains more buildings than the Top-hat. Eventually, the Top-hat is neither automatic nor robust towards the sensor and the statistics of the image.

6. CONCLUSION AND PERSPECTIVES

In this work, we have presented an original and robust scale-based detection method that offers interesting perspectives.

First, the good results of the entropic scalar detector show that scale is a relevant detection parameter for some objects at least, provided a genuine scale-based method is used. Moreover, they show that, even though they are limited for usual image analysis [Low84, Low83] entropic magnitudes are well suited to multiscale image analysis. Entropic tools are originally measures of a deviance to the equilibrium, therefore, they are useful when such an equilibrium exists. Unlike raw images, detail images are strongly centered around an "equilibrium" value, i.e. zero. They can be analyzed by distributed global entropic tools, while raw images can only be analyzed this way on small stationary windows [Mai96].

The success of this approach also offers interesting theoretical perspectives. Two major trends exist in the field of multiscale image analysis. On the one hand, the wavelet theory provides a variety of orthogonal scale-based decompositions [Mal89], but remains mainly algebraic and lacks the notion of scale and space dependence. On the other hand, Scale-spaces provide an interesting tracking of the behavior of structures along scales [LT94, Koe88], but lack scale-orthogonality and efficient implementations. Entropic multiscale detection stands between the two former theories: it uses wavelet techniques to perform discrete orthogonal multiscale decompositions while, like Scale-spaces, it is based on the behavior of features through scales. The generalization of the *ESD* (resp. *CIMENE*) to a large range of scales and the merging of its results in a "feature Scale-space" (resp. an "entropic Scale-space") could introduce a Scale-space geometry into the very algebraic wavelet theory, and efficiently merge some advantages of the two approaches.

7. REFERENCES

[BA83] P.J. Burt and A.E. Adelson. The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, 31:532–540, Apr. 1983.

[Mal89] S. Mallat. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on PAMI*, 11(7):674–693, Jul. 1989.

[Shan78] C.E. Shannon. *The Mathematical Theory of Communication.* The Illinois University Press, Denver, 1978.

[TU96] P. Thévenaz and M. Unser. A pyramid approach to sub-pixel image fusion based on mutual information. In *Proceedings of the ICIP*, volume 1, pages 265–269, Lausanne, Sep. 1996.

[CVSM95] A. Collignon, D. Vandermeulen, P. Suetens, and G. Marchal. Automated multi-modality image registration based on information theory. In Y. Bizais, editor, *Information processing in medical imaging*, pages 263–274. 1995.

[Dja93] J-P. Djamdji. *Analyse en ondelettes et mise en correspondance en télédétection.* PhD thesis, Université de Nice-Sophia Antipolis, Dec. 1993.

[Low83] G. E. Lowitz. Can a local histogram really map texture information? *Pattern Recognition*, 16(2):141–147, 1983.

[Low84] G. E. Lowitz. Mapping the local information content of a spatial image. *Pattern Recognition*, 17(5):545–550, 1984.

[Ser82] J. Serra. *Image analysis and mathematical morphology.* Academic Press, London, 1992.

[Pap96] C. Papin. Méthodes morphologiques et multiéchelles pour la détection de bâtiments. *DEA report University Paris VI, ENST*, Aug. 1996.

[Mai96] H. Maître. Entropy, information and image. *Progress in Picture Processing*, H. Maître and J. Zinn-Justin eds., p. 83–117, North Holland, Amsterdam, 1996.

[LT94] T. Lindeberg and B. Ter Haar Romeny. Linear Scale-space. In Bart Ter Haar Romeny, editor, *Geometry Driven Diffusion in Computer Vision.* Kluwer Academy, 1994.

[Koe88] J.J. Koenderink. Scale-time. *Biological Cybernetics*, 58:159–162, 1988.

Copyright 1997 IEEE