ENTROPY POWER FOR THRESHOLDING TECHNIQUE IN IMAGE PROCESSING

Franck Luthon, Marc Liévin
Computer Science Dept., University of Pau & Adour Province,
IUT, Château Neuf, Place Paul Bert
64100 Bayonne, France
Tel: +33 5 59463211; fax: +33 5 59463229
e-mail: Franck.Luthon@univ-pau.fr

ABSTRACT

This paper deals with an entropic approach as unsupervised thresholding technique for image processing, in order to extract a relevant binary information from noisy data. The method is based on the computation of the entropy power of the information source, as defined by Shannon. The threshold used for binarization is proportional to the entropic deviation of the observation source. The performance of the approach is illustrated by two classical image preprocessing tasks, namely motion detection and edge detection. The evaluation set contains both synthetic data and real-world image sequences.

1 INTRODUCTION

The thresholding technique is of common use in image processing and video analysis, in order to binarize noisy observations that are coded with n bits (typ. $n = 8$), either in the spatial or temporal domain. The key point is then the choice of the threshold, in order to get rid of the noise that corrupts the data, without cutting significant information.

We propose here an unsupervised method in order to automatically and adaptively estimate the threshold based on the computation of the entropy power of the observations, under the assumption of an additive Gaussian noise that is spread all-over the grid, whereas the signal is localized. The optimal threshold (in the sense of information theory) is proportional to the square root of the entropy power.

The method is tested with real-world images both for the detection of significant temporal changes, which is a classical preprocessing step before estimation, segmentation or compression of motion, and for the detection of the maxima of spatial gradients, which is the first step towards edge detection.

2 ENTROPY POWER

As defined by Shannon [1], the entropy of an information source is a measure of its mean information. In the case of a discrete source X of observations taking values in the set $\{i = 0 \cdots M-1\}$ (typ. $M = 256$), the entropy H is given by:

$$H = -\sum_{i=0}^{M-1} p_i \log p_i \qquad (1)$$

where p_i is the probablitity that the observation at any site $s = (x, y) \in S$ takes the value: $o(s) = i$ (S being the grid supporting the observations, typically an image of size $L \times C$).

Note that the choice of the logarithmic base is arbitrary and corresponds to the choice of the unit of measure: natural unit (*nat*) with base e, binary unit (*bit*) with base 2 ; and we have: $H_{nat} = H_{bit} \cdot \log 2$.

Shannon proved that the entropy of a Gaussian source $G(0, \sigma)$ with zero mean and standard deviation σ is given by:

$$H(G) = \log(\sigma \cdot \sqrt{2\pi e}) \qquad (2)$$

It is well known that this corresponds to the maximum value of the entropy for a distribution subject to the condition that the standard deviation is fixed at σ.

For an arbitrary source with a given entropy H, Shannon also defines the **entropy power** N, which represents the power of the white noise equivalent to source X, in the sense it has the same entropy and is limited to the same band:

$$N = \frac{1}{2\pi e} \exp(2H) \qquad (3)$$

The important properties of the entropy power are:

1. the entropy power of any source is always less than or equal to its actual power,
2. the entropy power of a Gaussian source $G(0, \sigma)$ equals its power: $N(G) = \sigma^2$.
3. the entropy power of the sum of two signals is lower-bounded by the sum of their respective entropy powers, and upper-bounded by the sum of their actual average powers.
4. white Gaussian noise has the peculiar property that it can absorb any other signal which may be added

to it, provided the signal power is small, in a certain sense, compared to noise. In that case, the resultant entropy power is approximately equal to the sum of the white noise power and the signal power [1].

3 THRESHOLD SETTING

Consider an observation source $O = \{o(s), s \in S\}$ consisting of a useful signal X corrupted by an additive Gaussian noise: $O = X + G$. Assuming that the noise level is low but present all-over the grid, whereas the signal is of high amplitude but remains localized on the grid, property 4 holds and one can estimate the equivalent **entropic deviation** σ_e by computing the entropy H of the observation source:

$$\sigma_e = \sqrt{N} = \frac{\exp(H)}{\sqrt{2\pi e}} \tag{4}$$

The threshold value θ may then be fixed as a quantity proportional to σ_e:

$$\theta = \kappa \cdot \sigma_e \tag{5}$$

The choice of the multiplicative parameter κ is based on the table of probability of the normal distribution which gives the correspondance between κ values and the percentage of thresholded distribution (Tab. 1). For $\kappa \approx 4$, 100% of the Gaussian distribution is taken into account. Therefore, in order to properly cancel the noise contribution, and also for simplicity, we take $\kappa = 4$ for all video sequences shown here.

Table 1: Percentage of standard distribution vs. κ.

κ	.43	.67	.97	1.28	1.65	1.96	2.57	3.9
%	33	50	66	80	90	95	99	100

Note the connection with a bit-plane slicing technique if we take the bit as entropy unit (cf. $\sqrt{2\pi e} = 4.13$):

$$\kappa \approx 4 \Rightarrow \theta = \kappa \frac{2^{H_{bit}}}{\sqrt{2\pi e}} \approx 2^{H_{bit}} \tag{6}$$

To illustrate the principle of the method, let consider the sample image of Fig. 1 made of 7 grey levels $\{i = 0 \cdots 6\}$ with respective probabilities: $\{p_i = \frac{1}{2}, \frac{1}{4}, \frac{1}{16}, \frac{1}{32}, \frac{1}{8}, \frac{1}{32}, 0\}$. This image simulates a noisy map with a rectangular pattern. Tab. 2 gives the parameters computed from this image. The estimated threshold θ is well positionned on the histogram. Note that σ_e is significantly different from the usual computation of the standard deviation σ.

Table 2: Estimated parameters

parameter	H_{bit}	$2^{H_{bit}}$	σ_e	$\theta = 4\sigma_e$	σ
value	1.94	3.83	0.93	3.70	1.52

Fig. 2 shows the result of the proposed entropic binarization with $\kappa = 4$ for change detection on a synthetic sequence.

0	1	0	1	1	0	0	0
0	0	3	2	2	1	2	0
0	1	4	4	4	5	3	0
0	0	4	1	0	4	1	0
0	1	4	4	4	5	0	0
0	1	0	2	1	0	1	0
0	0	1	1	0	1	0	0
0	1	0	0	1	0	0	0

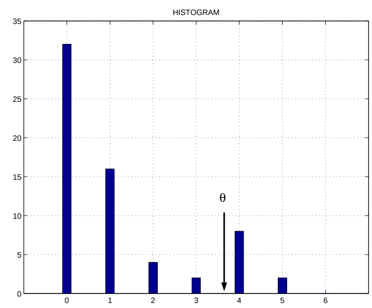

Figure 1: Sample image of size 8×8 with corresponding histogram and estimated threshold for $\kappa = 4$.

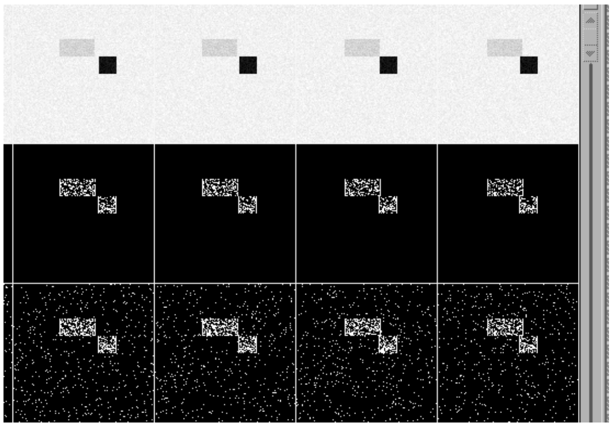

Figure 2: From top to bottom: **Synthetic** sequence with two mobiles; Automatic binarization ($\kappa = 4 \Rightarrow \theta \approx 7.8$); Binarization with $\theta = 7$ for comparison purpose.

4 APLICATION TO MOTION DETECTION

Motion detection consists in labelling each pixel or site s of image at time t to get a binary map of temporal changes. Assuming a static camera and a constant ligthning of the scene, the basic observation is the absolute value of the temporal intensity difference between two consecutive images (frame difference):

$$o(s) = |I_t(s) - I_{t-1}(s)|. \tag{7}$$

This observation is noisy since the frame difference is sensitive not only to actual motion in the scene, but also to lighting variation (slight illumination changes) and to acquisition noise (due to camera and quantization). Therefore, an adequate thresholding technique is required to detect significant temporal changes.

The hypothesis of an additive Gaussian noise is commonly assumed. Various techniques based on maximum likelihood tests have been proposed for the estimation of the noise level and the automatic setting of the threshold value [2, 3]. Here, we show that the entropic approach described in section 3 is a proper solution in order to automatically estimate the motion threshold.

Since the frame difference O is classically modeled as the sum of the relevant motion signal X plus an additive

Gaussian noise $G(0, \sigma)$, property 3 yields:

$$N(X) + N(G) \leq N(O) \leq P_X + \sigma^2. \qquad (8)$$

where P_X is the average power of the motion signal.

Considering that the useful signal remains localized in the image (relevant temporal changes arise for a limited amount of pixels) whereas the random noise is present everywhere in the frame, the entropic contribution due to the useful signal is actually small in the sense understood by Shannon. Indeed P_X is small compared to the noise power σ^2 (i.e., the average SNR over the entire frame difference is low, although the local SNR may be high). Therefore, property 4 holds and we have: $\sigma_e^2 = N(O) \approx P_X + \sigma^2$. Whereas the noise power σ^2 is almost constant over time, the signal power P_X varies over time, since it depends on the actual motion present in the scene. Hence, $N(O)$ is sensitive both to the noise level σ and to the amplitude of actual motion in the scene. Its measure provides a means to set the motion threshold as given by Eq. 5.

One can see (Fig. 3) that the threshold adapts over time, depending on the amount of motion present in the scene. For sequence **Street 1**, the entropic deviation grows when a car enters the camera field of view, since more pixels undergo a temporal variation in intensity. On the opposite, when there is little motion in the scene, σ_e falls since there are very few temporal changes.

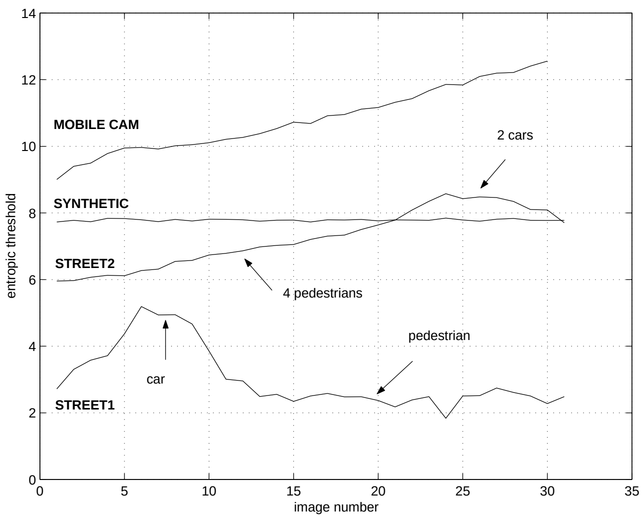

Figure 3: Evolution with time of the threshold $\theta = \kappa.\sigma_e$, on four sequences for $\kappa = 4$.

Fig. 4 shows the influence of κ for the detection of mobile pixels in the case of a static camera.

The sequence **Street 1** (Fig. 5) corresponds to a scene acquired also with a static camera, but with low-pass filtering. Therefore the noise power is filtered, yielding a lower value for σ_e and hence a lower threshold compared to the other sequences (Fig. 3).

Fig. 6 corresponds to the case of a mobile camera

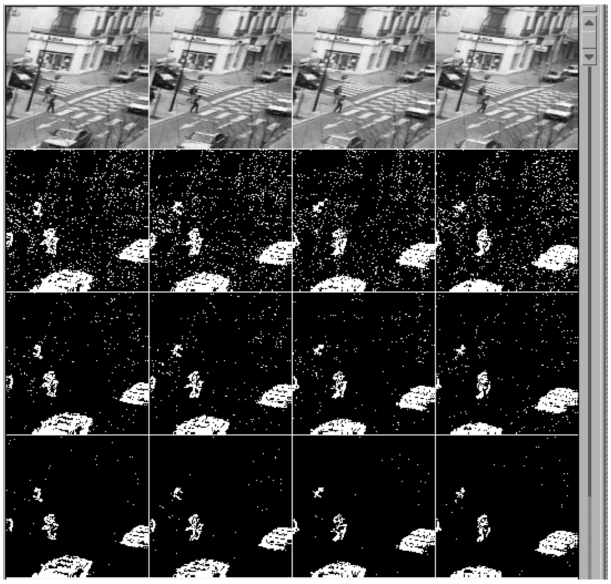

Figure 4: From top to bottom : four frames of sequence **Street 2**; binary maps obtained by entropic thresholding for $\kappa = 2; 3; 4$ respectively.

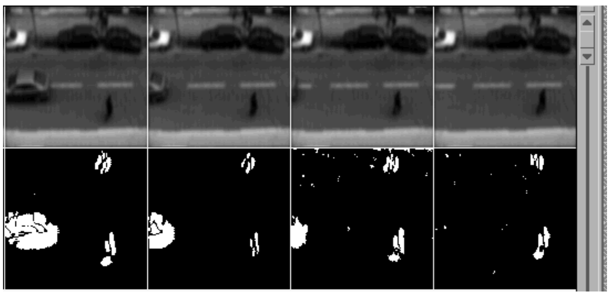

Figure 5: Sequence **Street 1** that was low-pass filtered; change detection with $\kappa = 4$.

translating from right to left. The mobile contours are detected properly.

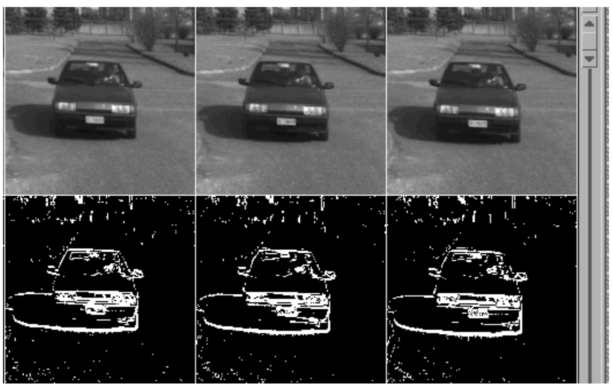

Figure 6: Sequence **Mobile cam**: car approaching towards a mobile camera ; change detection with $\kappa = 4$.

The efficiency and robustness of this threshold estimation has also been tested for face analysis application

(Fig. 7). The white points correspond to pixels detected as mobile, informative of facial feature motion.

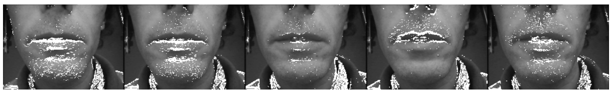

Figure 7: **Lip** sequence: change detection with $\kappa = 4$.

5 APPLICATION TO EDGE DETECTION

Edge detection is classicaly based on thresholding the modulus of the spatial gradients in the image. The observation is computed as:

$$o(s) = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \tag{9}$$

The choice of the threshold is the key-point to extract the pixels that are good candidates for edges. Hysteresis thresholding for example requires the use of two thresholds (high and low) that are usually set manually.

Our approach allows an automatic determination of the threshold to be applied. Typical results of entropic thresholding are shown in Fig. 8 and 9 (outdoor and indoor images). Shown in the upper row are the real-

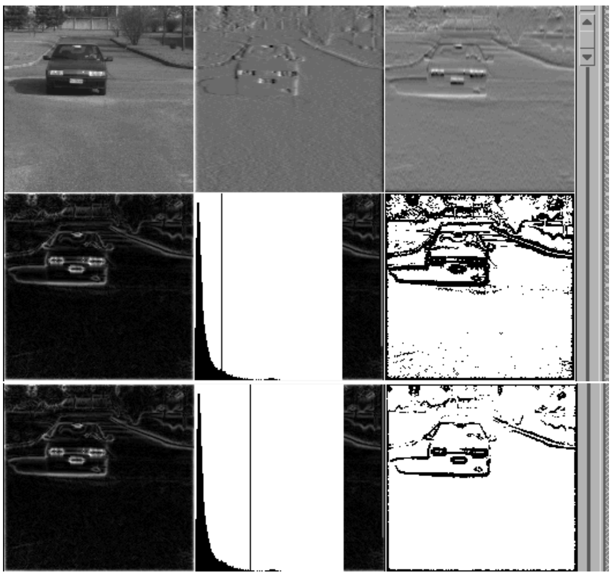

Figure 8: Contour detection on outdoor image for $\kappa = 2$ and 4 respectively ($H_{bit} = 5.61$, $\sigma_e = 11.88$).

world image, the vertical and horizontal gradients computed with an exponential derivation filter. The middle or lower rows show the modulus of gradient (which is the observation on which the entropic thresholding is applied), the histogram of observations with the position of the estimated threshold, and the binary map obtained after thresholding showing the contour-candidates.

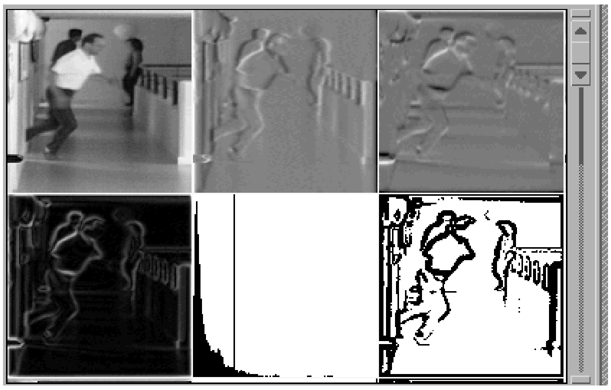

Figure 9: Contour detection on indoor image ($H_{bit} = 5.88$, $\sigma_e = 14.34$, $\kappa = 2$).

6 DISCUSSION

The entropic thresholding technique described here is both simple and efficient. It may be applied in many situations encountered in image preprocessing, where the data is corrupted by an additive noise spread all-over the observation grid. The binarization technique was used successfully for face analysis. Another potential application for 2-D spectral analysis is described in [4]. In actual applications, a post-processing is often applied after binarization (regularization by MRF, mathematical morphology like erosion or dilation...). In such cases, one should lower the threshold by taking $\kappa = 2$ typically. It should be pointed out that this technique is costless when used in the context of image compression based on an entropic coding scheme (Huffman-like coding), since the entropy of the source is computed anyway for compression purpose. Further developments we are currently working on are the computation of a local entropic deviation, instead of the global computation over the whole frame, and the use for video indexing (automatic scene cutting).

References

[1] C. E. Shannon. A mathematical theory of communication. *The Bell System Technical Journal*, 27:379–423,623–656, July, October 1948.

[2] T. Aach, A. Kaup, and R. Mester. Statistical model-based change detection in moving video. *Signal Processing*, 31(2):165–180, March 1993.

[3] Y. Z. Hsu, H. H. Nagel, and G. Reckers. New likelihood test methods for change detection in image sequences. *Computer Vision, Graphics, and Image Processing*, 26:73–106, 1984.

[4] F. Luthon, X. Navarro, and M. Liévin. Seuillage entropique en traitement d'images. In *18e Colloque sur le traitement du signal et des images (GRETSI'01)*, pages 353–356, Toulouse, France, 10-13 Sept. 2001. Vol.2.