A methodology for image segmentation based on the spectral content of regions

Marc Van Droogenbroeck
Belgacom, New Developments,
177, Bd E. Jacqmain
B-1210 Brussels, Belgium
Fax : +32 2 202 84 52
E-mail : Marc.Van.Droogenbroeck@is.belgacom.be

ABSTRACT — **As they describe images as collections of regions, object-based analysis methods start with the division of an image into several meaningful areas during a stage called segmentation. Several segmentation tools are suggested in the research literature including texture discrimination based on Gabor filters, split-and-merge techniques, quadtree approaches, watershed line,... But despite the proliferation of tools, segmentation is still a difficult task to accomplish and semi-automatic procedures often give the "best" results. Moreover, many techniques generate over-segmentation artifacts and the result depends highly on the image content.**

This paper proposes a general segmentation methodology, directly linked to an image model that uses the local spectral information. The model states that each object inside is regular, active or flat.

The segmentation process has been implemented and an illustration of different segmentation steps is provided.

1 Introduction

In all the practical cases, the segmentation process is driven by the objectives of the considered application. As a consequence, the segmentation algorithm is especially designed for a specific application and can only be applied to a reduced set of images. The situation is quite different in the field of image coding where the algorithm is required to be as universal as possible. In other words, no "a priori" information can be used during segmentation, except for low bitrate communications.

The context of object-oriented image coding is particularly relevant to illustrate some of the difficulties encountered in segmentation (see for example [1]). First, no clear link can be established between most existing segmentation techniques and the coding process itself, leading to non-optimized coding methods. Furthermore, it is difficult to detect textures with no prior information. In object-oriented image coding, research has led to methods in which details of decreasing size are progressively added to reach an acceptable level of restitution. But this hierarchical approach is not valid for textures. Finally, there is a misunderstanding of what a segmentation process should produce : *it is not necessary that detected objects coincide with real objects*, although it would be better, if possible. For example in image coding, the segmentation has to detect objects that can be coded efficiently whatever the human interpretation is.

1.1 Methodology

Segmentation algorithms usually treat the image as a whole, referring to an assumed stationnarity property. However, this property does not hold for many real images and, therefore, it is preferable for segmentation to consider an image as a collection of areas with homogeneous attributes. In our scheme, we propose to combine and cascade existing tools specific to each type of area.

The resulting methodology consists of the following steps :

(i) Define several non-redundant types of region such that the union could model a large range of images.

(ii) Select a segmentation tool for each type of region.

(iii) Determine an order for the detection of the different types of regions.

During the whole process, only areas that are not selected in a previous step may still be treated. The detection order plays an important role ; ideally the best tool takes the first place.

2 Image model based on the local spectral content

Object-based analysis methods describe images in terms of basic properties such as *contours* and

textures. In spite of the intuitive perception of contours and textures, practical implementations collide with the vagueness of the distinction between both notions. For instance, a sample of text can been seen as a set of contours or as a particular texture. To clarify our perception, we propose the following definition for textures (an example of texture is provided in figure 1) :
Definition *A texture is a signal that can naturally be extended outside the definition domain of the original object.*

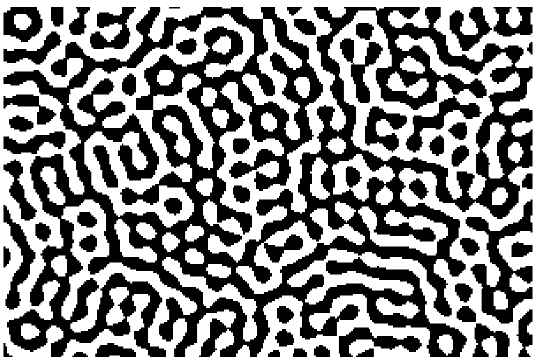

Figure 1: Example of a texture.

As expressed in the definition, *each texture is tied to an extrapolation notion.* This means that an extrapolation procedure has to be defined when talking about textures. Consequently also, a segmentation process that integrates an extrapolation algorithm will only be able to detect some specific families of textures.

In addition to textures, there exist areas, called active regions, where segmentation is almost impossible. It is better to leave them unsegmented.

In conclusion, our *model* states that regions of any image belong to one of the following three categories (illustrated in figure 2), depending on their local spectral content :

1. *Regular periodic textures*, called *textures* hereafter, are regions with an homogeneous content. The associated spectral extrapolation is described further in the text.
2. *Active regions*, where it is hard to distinguish between a contour and a texture.
3. *Flat regions* which contain no local irregularity.

Such a model is well-suited for a segmentation process based on the local spectral information and is useful for coding purposes.

Existing segmentation processes usually consider flat and other regions, sometimes even only flat regions. A further distinction between textured and active regions has been introduced here to enhance the relevance of segmentation but it is also useful for coding.

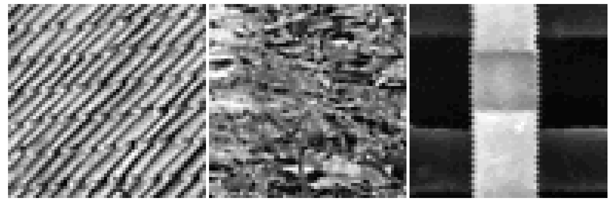

Figure 2: Examples of a (regular) texture, an active region and flat regions respectively, according to the image model.

3 Segmentation

According to the model, the segmentation process is made up of as many stages as kinds of region. Each of these stages is intended to extract one kind of region present in the image. In the following, we will focus on the guidelines but we will not enter the full details of all parts of the segmentation technique. The three stages are :

Stage 1 : *detection of textured regions.*
Gabor filters are used to detect areas where the local spectral energy is high in different directions. As detailed in [2], we use 4 oriented Gabor filters (respectively oriented at 0, 45, 90 and 135 degrees) and apply a threshold to detect areas with a high level of energy. The algorithm compares the 4 filtered images : pixel positions with only one filter output level above the threshold belong to oriented textures. This process produces a label image whose pixel values indicate the selected texture orientation.

Next, the algorithm tries to extend each connected oriented texture starting from the largest. For this, we assume that the texture segment contains all the required information to characterize the texture properly. The information content is spread in the neighborhood of the segment by means of a spectral extrapolation process developed by FRANKE [3] and extended in [4]. Pixels, not belonging to any other texture, and where the extrapolated signal is close to the original image values, are added to the considered texture. Let $f(p)$ and $g(p)$ respectively be the original and extrapolated signals (also defined inside the segment) at the location of pixel p. The initial texture sample is defined on the domain D, the set of pixel positions contained in the segment. Because only a part of the segment information is used for extrapolation, $g(p)$ may differ from $f(p)$ inside D. The measure taken for comparison outside D is related to the standard deviation σ of the reconstruction error on D :

$$\sigma = \sum_{D} \sqrt{[f(p) - g(p)]^2} \qquad (1)$$

Suppose pixel q is located in the neighborhood of D, q is added to D if

$$|f(q) - g(q)| \leq \alpha\sigma, \qquad (2)$$

where α depends upon the gray level distribution ($\alpha = 1.5$ was chosen in the examples). Note that the growing measure is adaptive because if $g(p)$ is close to $f(p)$, as would be the case for computer-generated texture, then σ is small and the growing acute. In addition, σ is fixed to the minimum of all previous values to avoid a degradation during the growing process. The entire process is iterated until no point could be added.

The complete process results in a collection of textured regions. The separation between active and flat regions is still required.

Stage 2 : *detection of active regions.*
Many pixels not selected as textures, may also have a high local spectral energy value In that case, they will be considered as parts of active regions. Unlike textures, active regions are not characterized by a direction and, consequently, there is only one type of active region. In brief, the algorithm detects active areas and the number of active regions present in an image is equal to the number of connected active regions.

DAUGMAN [5] has proved that the product of the spatial and spectral resolutions is lower bounded. Therefore, the extraction of textures and active regions has introduced some imprecision on the border. Before stage 3, the region borders are adjusted in the spatial domain by matching the border to the local morphological gradient.

Stage 3 : *detection of flat regions.*
A technique similar to the one used at stage 1 separates flat regions.

The watershed (see MEYER and BEUCHER [6]) of the morphological gradient is tuned to produce an over-segmentation. Each region, from the largest to the smallest, is extrapolated over the neighboring regions. When there is a good correspondence, which means that σ over the initial region is close to the standard deviation of the reconstruction error on a neighboring region, the two regions are merged. The process ends when no further merger is possible.

Figure 3 shows different segmentation results. The detection of textures is especially hard in figure 3(a) due to the important content of high frequency energy in a large area. Another example of texture detection is given in figure 4 where the "shirt area" is made up of textures. The remaining area was segmented as described in stage 3. More detailed examples can be found in [4].

4 Conclusions

This paper has been concerned with the difficult task of segmentation for which a new methodological framework has been proposed. We have discussed and defined the concept of texture, restricted to regular periodic signals in our implementation.

The segmentation scheme we have presented stresses the distinction between several types of regions (regular textures, active or flat regions) characterized by local spectral attributes and applies an appropriate segmentation tool to each type of region. The guidelines of the complete process have been given and illustrated by several images.

References

[1] M. Biggar, O. Morris, and A. Constantinides, "Segmented-image coding: performance comparison with the discrete cosine transform," *IEE Proceedings F, Radar and Signal Processing*, vol. 135, pp. 121–132, April 1988.

[2] A. Bovik, M. Clark, and W. Geisler, "Multichannel texture analysis using localized spatial filters," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, pp. 55–73, January 1990.

[3] U. Franke, "Selective deconvolution: A new approach to extrapolation and spectral analysis of discrete signals," in *Int. Conf. on Acoustics, Speech and Signal Processing*, pp. 30.3.1–30.3.4, IEEE, May 1987.

[4] M. Van Droogenbroeck, *Traitement d'images numériques au moyen d'algorithmes utilisant la morphologie mathématique et la notion d'objet : application au codage.* PhD thesis, Catholic University of Louvain, May 1994.

[5] J. Daugman, "Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters," *J. Opt. Soc. Am. A*, vol. 2, pp. 1160–1169, July 1985.

[6] F. Meyer and S. Beucher, "Morphological segmentation," *Journal of Visual Communication and Image Representation*, vol. 1, pp. 21–46, September 1990.

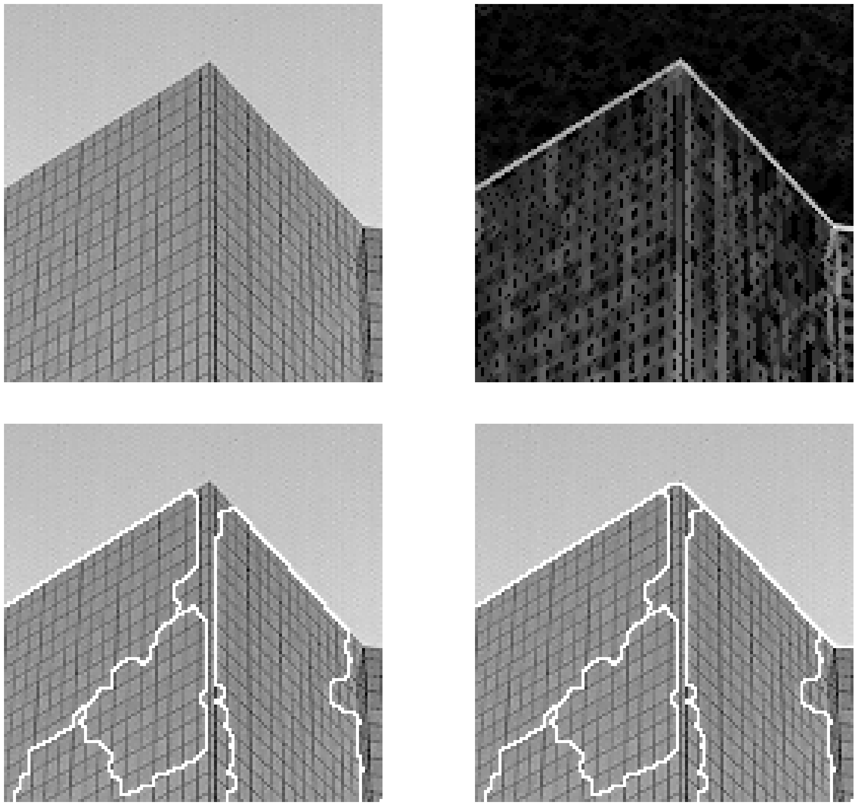

Figure 3: Illustration of the segmentation process.

(a) Image representing the top of a Belgacom tower.

(b) The morphological gradient (dilation minus erosion by a lozenge structuring element of 5 pixels) of the original image (a). The gradient shows how noisy the image is.

(c) Textures detected at the end of stage 1.

(d) Final segmentation. Only the upper region was detected as a flat region; other regions are textures or active regions.

Figure 4: Final result of the segmentation process. All the shown parts of the shirt in the right-hand image have been detected as textures. The algorithm considers all the other regions as flat regions.