Morphological Segmentation Using Texture and Coding Cost.

Beatriz Marcotegui † & Jose Crespo † ‡ & Fernand Meyer †

†Centre de Morphologie Mathématique
École Nationale Supérieure des Mines de Paris
35, rue Saint-Honoré, 77305 Fontainebleau Cedex, France
‡ E.T.S. Ingenieros de Telecomunicación
Universidad Alfonso X el Sabio, 28691 Madrid, Spain

ABSTRACT — **This paper presents a morphological segmentation algorithm oriented to coding purposes. Its characteristics are: (1) texture and coding cost information are introduced for computing the partition of an input image, and (2) feature shapes are preserved (block-type artifacts are avoided). The output of a connected filter is simplified by iteratively merging adjacent regions in order to minimize the coding cost for a given quality of representation. We use two different strategies in which flat zones (piecewise-constant regions) are considered as indivisible units. In the first one, a set of flat zones has been previously identified as significant (markers), and merges start from these markers. In the second one, no markers are used and all possible merges are considered by building the adjacency graph of the flat zones.**

1 Introduction

In the framework of an object oriented image compression system, segmentation is the first step and the basis of the whole system. The second step consists in coding contours and contents of each segmented region.

Usually segmentation and coding step are two independent stages in the system: segmentation produces an output image (with its own criteria) and afterwards a coding algorithm is applied. In that way, segmentation does not take into account the characteristics of the following coding operations and does not attempt to optimize the segmentation in terms of coding cost.

Widely used segmentation criteria are based on the homogeneity of the segmented regions (gradient minima, intensity level homogeneity...). These criteria fail when a textured region is present in the image. In a textured area a criterion of homogeneity produces an over-segmentation, which is heavy in terms of coding cost and visually not very significant.

Other segmentation approaches use texture information as criterion to locate the different regions of an image. But they do not locate precisely the contour of significant regions (comparatively with morphological methods).

This paper presents a morphological segmentation algorithm that introduces information about the texture of the regions and their coding cost. The aim is to produce a segmentation intended to be used for coding purposes so that a satisfactory quality/coding-efficiency balance is achieved.

The method presented in this paper is based on the connected operators: filters by reconstruction, area filters [1], and the flat zone approach [2, 3] which is itself the application of the connected operator concept to the image segmentation problem. Connected operators ensure the preservation of the feature shapes. The fact that flat zones (or piecewise-constant regions) of an input image are considered as entities of interest and solely fusions of them are performed, ensures the preservation of the feature shapes.

Section 2 gives a fast overview of the morphological segmentation tools. The measures we use for evaluating the quality and coding cost of a given region (or set of regions) are commented in section 3. Section 4 and 5 present the two different strategies we employ for introducing texture and coding cost in the segmentation stage.

2 Morphological Segmentation

The traditional morphological tool for segmenting is the watershed algorithm [4, 5] applied to a gradient image. The introduction of markers [6] (indicators of significant regions) avoids the typical over-segmentation problem. The watershed line is placed on the highest crest line of the gradient between two markers, which corresponds to the contour line between two significant regions. This method has been used with success for many applications(i.e. [5, 6]). However, its inability to segment thin objects makes it not very suitable for coding purposes. This effect is specially important working with image sequences:

a fast moving object is seen as very thin in the time dimension.

The flat-zone approach [2] provides a solution to the resolution problem of the traditional watershed algorithm. The flat-zone approach is not based on a gradient image but on the original image. The piecewise constant regions (flat-zones) are considered as indivisible units. The algorithm can be split up in two steps:

- Selection of a subset of flat-zones as significant regions (markers).
- The rest of the flat-zones are assigned (according to a similarity measure) to a selected one.

The similarity measure used in [2] is the grey level difference. In this paper other measures are investigated.

The aim is to produce a segmentation oriented to coding purposes. It often occurs that some sets of adjacent regions can be grouped (into one region) and that the resulting region can be satisfactorily represented by only one set of texture parameters. Therefore, merging these regions reduces the coding cost whereas the visual quality is not significantly affected. The coding cost gain arises because (1) only one set of texture parameters is employed to describe the resulting region, and (2) contour pixels (which must be coded) "disappear" in the merging process.

The morphological segmentation approach is performed in two steps: first significant regions are marked and afterwards contours of the selected regions are located. This approach can be useful for the new functionalities of MPEG4 (as object tracking) since an interaction with the user for selecting markers is possible. Section 4 presents an extension of the flat-zone approach that uses more complex criteria for assigning non-marked regions to marked ones. Nevertheless, in a coding context the concept of a unique good segmentation does not exist: the segmented regions depend on the quality of the texture model used and on the aimed compression rate. Therefore the marker selection a priori is a very complex problem in a general case. This is why we have investigated a method that instead of starting the merging process from a set of markers, allows the merging of any two adjacent regions (as proposed in [7]). Section 5 presents a method based on neighborhood graphs, in which no set of markers is required and whose criteria includes coding information.

3 Texture Quality and Coding Cost Measures

The quality measure employed in this paper is the standard signal-to-noise ratio (SNR).

$$\text{SNR} = 20 \times log \left(\frac{S}{N} \right)$$

where

$$S = \sum_{(i,j)\epsilon R} L(i,j)$$

$$N = \sum_{(i,j)\epsilon R} |L(i,j) - L'(i,j)|$$

$L(i,j)$: pixel luminance

$L'(i,j)$: texture model approximation at pixel (i,j).

R: a segmented region.

The coding cost is the sum of the required number of bits for coding contours (using chain-code the cost is proportional to the length) and texture. For coding the texture, spectral description has been used.

4 Method Based on Markers

Markers are connected set of pixels that are located inside significant regions. After markers have been obtained, features that have not been "marked" are assigned to the marked ones (the final segmentation has as many regions as markers have been selected).

In a general case, there is not a unique characteristic that can successfully extract the most significant regions in various circumstances. Sometimes the size of a feature is important, but not always (for example the eyes in a face are small but significant regions). This is the reason why a combination of criteria is used. In [2] size criterion, contrast criterion, extrema criterion are combined. In this paper a new criterion including coding cost is introduced.

The complexity of the segmented image depends on the target compression rate and the final goal is to obtain a coded image as close as possible to the original image. The strategy undertaken to fulfill these two specifications is the following:

1. A connected filter is applied to the input image [1].
2. A preliminary rich (i.e., with a large number of regions) segmentation P is computed using the flat zone approach where markers are the extrema (maxima and minima) regions at the output of the connected filter.

3. The sets of (a) extrema regions of P, and of (b) those regions of P that can be coded with a large compression factor are selected as markers M.

4. A flat zone merging algorithm is performed such that regions belonging to M merge with non marked regions, until all regions are assigned to a member of M. When a flat zone must choose between two possible members of M, the alternative resulting in lower coding cost (or higher compression rate) is chosen when the quality requirements adopted are met.

As can be observed, extrema regions play an important role as marker candidates. The reason is that usually extrema regions are visually significant and should in general be preserved in the final segmentation.

Nevertheless, marker extraction in a coding context remains a complex problem. For this reason we investigate an algorithm (presented in next section) which does not require a set of markers.

5 Method Based on Neighborhood Graphs

In a flat-zone approach an image is seen in terms of plateaus. A plateau is defined as the largest connected component of pixels possessing the same function value. The original image contains a large amount of small flat-zones and the segmentation consists in merging some of them producing larger zones in smaller number. A region of the final segmentation is the union of some initial flat-zones. Initial flat-zones merge according to some similarity measure. The algorithm presented in the previous section needs a set of markers for beginning the process of merging. This second strategy tries to avoid the complex problem of marker selection for an arbitrary input image. We propose to evaluate each possible fusion of two adjacent regions with a value that indicates their similarity, to class them and to merge adjacent regions until reaching the required complexity of the segmented image.

The algorithm can be implemented in an efficient way with a graph structure [8, 9] whose vertices and edges are valuated. Associating a graph to a segmented image, vertices represent the segmented regions and edges the neighborhood relations among them. Edges correspond to the pairs of adjacent regions and represents all the potential fusions. Their valuation, a measure of similarity between the two adjacent regions, provides the order of fusion. The edge of lowest valuation will be the first edge to be removed because it separates the two most similar regions.

The developed algorithm can be split up in the following steps:

1. A connected filter (area filter of small size) is applied to the input image.

2. The output of the connected filter is simplified by merging all the regions smaller than 4 pixels (for a qcif image: 176×144 pixels). The goal is to remove non perceptual features in order to simplify the detection of regions. A modified flat-zone approach is used to compute this simplification. The contrast priority used in [2] tends to produce irregular contours that are expensive to code. The modification consists in controlling the merging process by two factors: (1) the contrast between flat-zones and (2) the length of the final contour. The obtained image is a too-rich segmentation that will be simplified in the following.

3. The goal of this point is to merge the couples of regions that produce the lowest loss of quality until the required compression rate is reached. No set of markers is used, instead any two adjacent regions can merge. The exhaustive algorithm that maximizes the final quality for a given compression rate would be:

   - find the couple of regions whose merging produces the lowest loss of quality.
   - merge them.
   - reevaluate edges around the new region.
   - iterate.

   However, the computational cost forbids this solution. As a suboptimal (and practical) solution, a non exhaustive search algorithm is implemented. Neighborhood graph edges are valuated with a simple similarity measure (i.e. contrast criterion) and considered one by one in increasing order of this valuation. Before removing an edge (merging two regions) the merging is validated by a quality criterion: if the loss of quality is not important the merge is performed; otherwise, both regions remain as two different regions in the final segmentation.

With this technique we simplify a given segmentation while controlling the loss of quality.

(a) Original image

(d) Coded image

Figure 1: Method based on markers

(a) Oversegmented image

(b) Coded image

Figure 2: Method based on neighborhood graphs

6 Results

Figure 1 illustrates the results obtained by the method based on markers. Markers have been computed automatically (extrema plus regions that can be coded with a high compression rate) but they can be introduced by the user in the framework of a content based system. Fig 1(a) shows the original image and fig 1(b) the final coded image.

Fig 2 illustrates the method based on neighborhood graphs. Fig 2(a) shows the simplification of the input image (connected filter plus elimination of small regions). It contains 735 regions and 8250 contour points. Without a set of markers regions merge until a specified complexity is reached or the elimination of an additional edge damages significantly the coded image. Fig 2(b) shows the final coded image. It contains 57 regions and 2500 contour points.

7 Conclusion

A morphological segmentation algorithm based on connected operators is presented in this paper. This algorithm differs from current morphological segmentation in that both (a) the texture model of the regions and (b) the coding cost associated with merging operations are important factors for determining whether or not adjacent regions are merged.

By introducing coding-cost criteria in the segmentation stage, the problem of the bit-stream regulation can be treated a priori and not a posteriori, after a coding stage that exceeds the target bit-stream.

We have discussed two different strategies. The first one is a direct extension of the flat zone approach and uses a set of significant flat zones as markers. The second strategy does not require a previously computed set of markers, and it can incorporate complex segmentation criteria in a flexible and efficient manner.

References

[1] L. Vincent, "Greyscale area openings and closings, their efficient implementation and applications," in *Workshop on Mathematical Morphology, Barcelona*, pp. 22–27, May 1993.

[2] J. Crespo and J. Serra, "Morphological pyramids for image coding," in *Proceedings of SPIE, Cambridge*, 1993.

[3] J. Crespo, *Morphological Connected Filters and Intra-Region Smoothing for Image Segmentation.* PhD thesis, School of Electrical Engineering, Georgia Institute of Technology, 1993.

[4] S. Beucher and C. Lantuéjoul, "Use of watersheds in contour detection," in *Proc. Int. Workshop Image Processing, Real-Time Edge and Motion Detection/Estimation*, 1979.

[5] S. Beucher, *Segmentation d'Images et Morphologie Mathématique.* PhD thesis, E.N.S. des Mines de Paris, 1990.

[6] F. Meyer and S. Beucher, "Morphological segmentation," *J. Visual Commun. Image Repres.*, vol. 1, no. 1, pp. 21–45, 1990.

[7] O. Monga, *Segmentation d'Images par Croissance Hiérarchique de Régions.* PhD thesis, Université Paris Sud. Centre d'Orsay, 1988.

[8] L. Vincent, "Graphs and mathematical morphology," *Signal Processing*, vol. 16, no. 4, 1989.

[9] L. Vincent, *Algorithmes Morphologiques à Base de Files d'Attente et de Lacets. Extension aux Graphs.* PhD thesis, E.N.S. des Mines de Paris, 1990.