MORPHOLOGICAL SEGMENTATION OF IMAGE SEQUENCES

B. MARCOTEGUI and F. MEYER
Ecole des Mines de Paris, Centre de Morphologie Mathématique,
35, rue Saint-Honoré, F–77305 Fontainebleau Cedex, France

Abstract. In image compression, object-based approaches are adapted to high compression rates, since they take into account the geometry of the objects and the human eye characteristics. Mathematical Morphology, dealing with geometrical features is a well suited technique for segmentation purposes. This paper presents a method to segment image sequences, first step of an object-oriented compression system, based on Mathematical Morphology.

Key words: Image Sequences, Segmentation, Mathematical Morphology, Image Compression.

1. Introduction

Classical methods of image compression use spectral decomposition inside spatial unities called blocks, which are not directly related with the image contents. These blocks produce artifacts in the coded image that prevent from reaching high compression rates with an acceptable quality. As an alternative, object-oriented methods have been developed [1]. In order to avoid block effects, these techniques split the image into semantic regions and then encode separately the shape and the contents of each region. Higher compression rates are expected, because the geometry of the objects and the human eye characteristics are taken into account.

The first step of an object-oriented approach is segmentation, on which is based the whole scheme. A good segmentation, with the visual contours and only them, maximizes the compression rate. In image sequences a good time continuity is the basis for a high compression rate.

Mathematical Morphology [2] [3], dealing with geometrical features is a well suited technique for segmentation purposes. This paper presents a method to segment image sequences, based on Mathematical Morphology. The first image of the sequence is segmented in an intra-frame mode, presented in section 2. In section 3 a method to reach maximum stability in image sequence segmentation is presented, as well as the possibility to include new regions, in the only cases when it is really necessary. Finally the problem of regulating the segmentation is discussed, in order to produce a bitstream compatible with the desired compression.

2. 2D segmentation

Dealing with image sequences, a 2D segmentation algorithm is also needed, in order to process intra frames (first frame of the sequence or initializing stage).

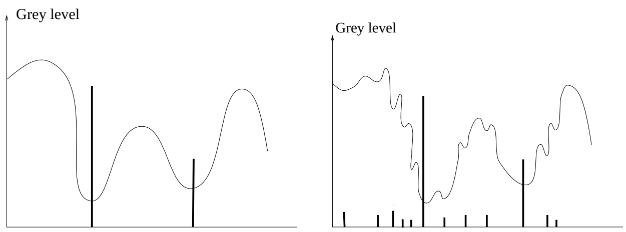
The segmentation algorithm may be split up into the following steps [4]:

– <u>filtering step</u>: the goal is to remove non perceptual features in order to simplify the detection of homogeneous regions. Morphological filters, (area opening and closing [5], and reconstruction filters) are used. They remove objects smaller than a given size preserving the contour of the remaining ones.

- feature extraction: the goal is to detect the existence of homogeneous regions.
- decision step: the goal is to place the contour of the objects that have been selected by the feature extraction. This decision is made by a watershed algorithm to a gradient image [6] [3] [7], which places the contour on the line of highest gradient between two objects.

A watershed algorithm applied directly on a gradient image produces a severe over-segmentation. Each gradient minimum becomes a region, and most of them do not match with visual regions. Gradient minima have to be classified according to their visual significance and the most significant ones have to be chosen in a number the system can afford. This is the aim of the feature extraction, the main difficulty of the segmentation.

The dynamic [8] applied to the gradient, classifies its minima according to their relative contrast. Fig. 1 illustrates this ranking for a smooth function before and after adding noise. The two main structures have a high dynamics in both functions, while the noisy minima have a comparatively lower dynamics. This measure provides for feature extraction a selection criterion based on contrast. But pure contrast does not fit exactly with visual significance. Between two regions of same contrast, the larger one, would be more visible. For this reason, in order to get a more visual criterion, we have weighted the dynamics by a function of the region size.



(a) Function's dynamics (b) Noisy function's dynamics

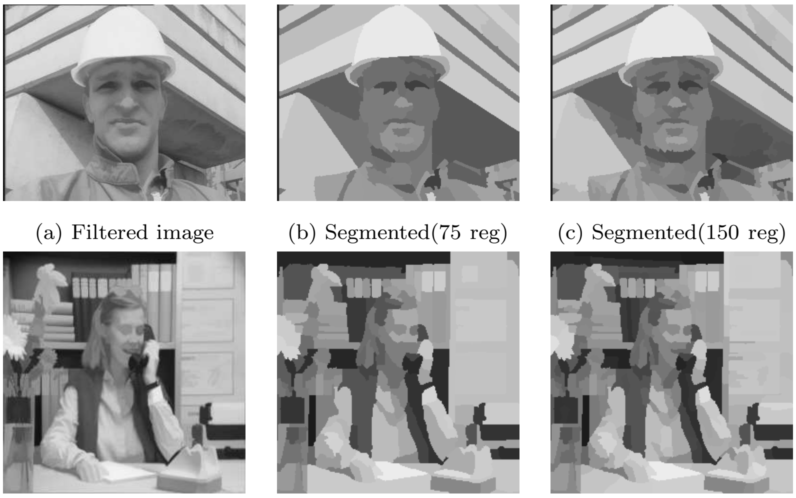
Fig. 1. Dynamics of function minima.

Some segmentations at different resolutions, based on a contrast-size criterion are shown in fig. 2.

3. 3D segmentation

In order to reach a time stability, the successive frames of a sequence are segmented as a 3D volume [9]. In order to avoid unacceptable time delays, the time depth of the 3D volume has been reduced to 2 frames(previous and current frame); this means, a sliding window through the sequence [10].

A 3D segmentation algorithm transforms a sequence into a partition of the 3D space. Contours and texture of each region are subsequently coded. Coding cost of a region is minimum if it can be predicted (only a motion vector and prediction



(a) Filtered image (b) Segmented(75 reg) (c) Segmented(150 reg)

(d) Filtered image (e) Segmented(150 reg) (f) Segmented(250 reg)

Fig. 2. 2D segmentations.

errors have to be coded). This is the case for any region keeping its label from the previous frame. On the other hand, contours and texture of new regions have to be completely coded, which is much more expensive. Therefore, to reach a high compression rate, segmentation algorithms have to favor time continuity accepting new regions only when they are really important.

3.1. Maximum stability

We first treat the ideal case, in which all regions are already present in the first frame and have been properly segmented in the intra frame mode. Thus, in inter frame mode, no further feature extraction is required. The previous segmentation works as a set of markers (put in the first frame of the sliding window) which invade the current frame.

The decision step, performed by the watershed algorithm, works on a gradient image in order to locate the contours of the selected markers. The gradient of thin objects does not contain contour information (see fig. 3). In still images the filtering step removes the objects under the gradient resolution. In video sequences this problem worsens. The uncertainty of spatial contours is proportional to the object motion (fig. 4). For this reason we perform the watershed on a time upsampled gradient (fig. 5) which allows us to locate correctly the spatial contour.

Taking the previous segmentation as a set of markers, a watershed on a 3D time upsampled gradient, produces the contours we are looking for.

The image of foreman with a smooth motion (fig. 6) is segmented in fig. 7.

If motion is faster, some objects may not be connected with their markers. Those objects disappear in the segmentation. (The algorithm processes a disconnected object as two different objects). In fig. 8 we have an example of this situation. The ping-pong ball disappears when the motion becomes faster. A high time dam in

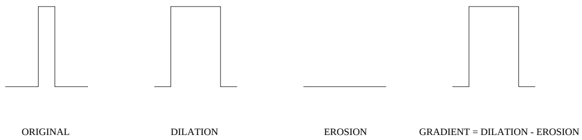


Fig. 3. Gradient of a thin object.

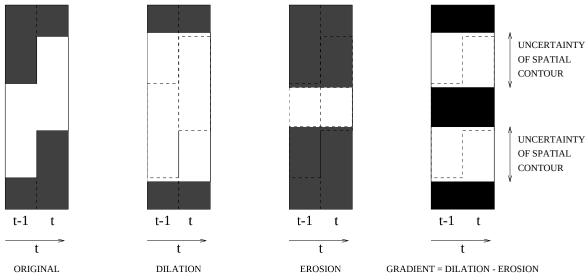


Fig. 4. 3D gradient of a moving object.

the gradient prevents the marker of the disappeared region from reaching the next frame.

This algorithm, intrinsically fails when a new object appears; the only markers we have are those of the previous frame. On the other hand, the algorithm is perfectly able to remove from the segmentation an object that has disappeared. Therefore the next problem to solve is the inclusion of new regions.

3.2. Inclusion of new regions

The algorithm presented in the previous section offers a maximum stability; it maintains permanent objects in the scene as long as possible, until they eventually vanish. Unable to let new regions appear, it has to be completed by a procedure for introducing new regions. Performing a feature extraction for each frame would introduce highly unstable segmentations, since the ranking of regions according to their visual significance would fluctuate one image to another. Stability will be introduced by a three step procedure:

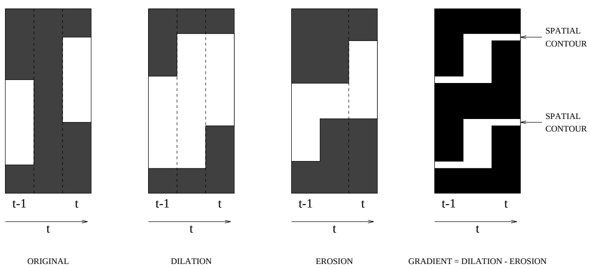


Fig. 5. 3D time upsampled gradient.

Fig. 6. Filtered image sequence

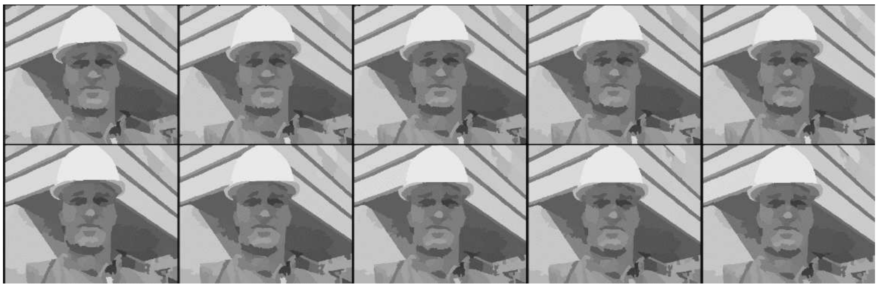
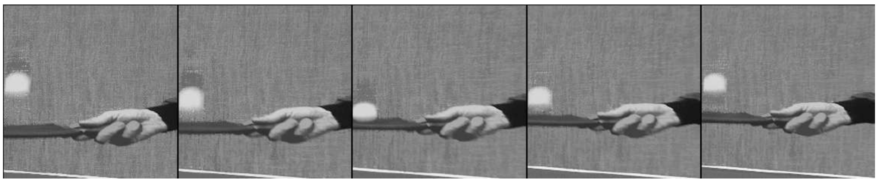


Fig. 7. Segmented image sequence

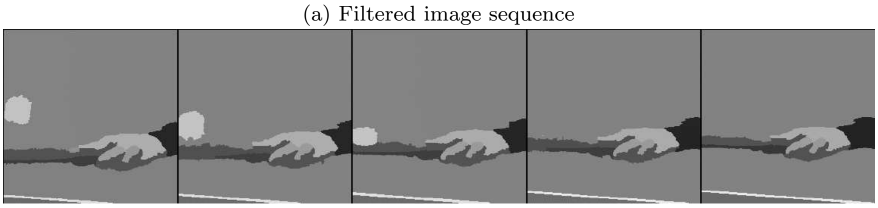
1. independent feature extraction.
2. validation of new markers.
3. merging new and old markers.

This procedure is hereafter explained in detail.

1. A feature extraction (with a contrast-size criterion, as presented in section 2) selects the objects visually important. This selection is performed without taking previous segmentations into account, and its result is taken as a provisional set of markers. Some of them already exist in the previous frame and they have not to be considered as new markers. Others are markers of new significant regions in the scene and have to be kept. The rest are ranking fluctuations and we want to reject them, in order to keep a good stability.
2. We want to introduce new regions only when they are justified by an important change in the scene. This criterion has been implemented by means of a moving mask, which removes those markers that are associated to regions that have not changed in time. The moving mask is generated as follows: the difference between two successive simplified frames is computed. Moving areas appear with a high grey level. The feature extraction step computes a contrast threshold, that corresponds to the required contrast of a region in order to be considered visually important. This threshold is applied to the difference image producing a mask of changing areas. Only markers under this moving mask will be kept as markers of new regions.



(a) Filtered image sequence



(b) Segmented image sequence

Fig. 8. No new regions.

3. Old and new markers are merged into a 3D image; in the first frame the previous segmentation (as markers of permanent objects) and in the frame corresponding to the current time, markers selected by the feature extraction that are under the moving mask (as markers of new regions).

After the feature extraction step, the decision step is performed in a 3D time upsampled image as explained in the previous section. Fig. 9 contains a block diagram of this technique.

With this technique a new region appears only when an important change justifies it. In fig. 10 a segmented image including new regions is presented.

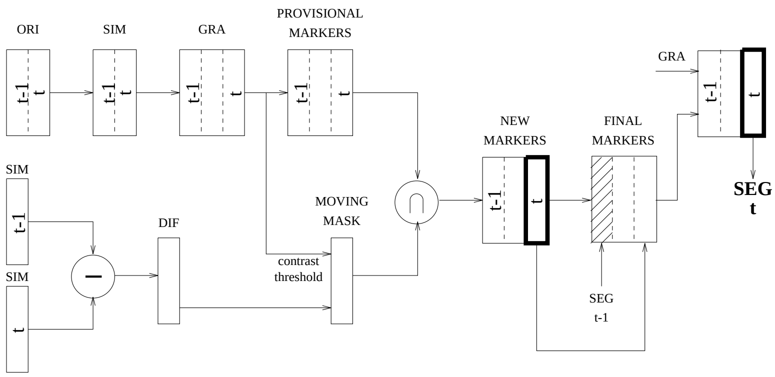


Fig. 9. Block diagram of inclusion of new regions.

3.3. Regulation of the segmentation

For a given compression rate the quality should be as high as possible. If motion in the sequence is predictable from the previous frames a fine segmentation is afforded. On the other hand if new significant regions appear all the time the accuracy of

Fig. 10. Inclusion of new regions.

the segmentation have to be reduced. Therefore a mechanism should be provided enabling to introduce more regions upon request if predictable motion allows it or to merge them in order to assign bits for new significant regions. Several segmentations followed by coding step should be necessary before an optimum is reached. Complexity and processing time forbid this option. Another solution is to produce during the segmentation a more complete representation, such as a tree structure [11]. Each node of the tree represents a region linked by an edge to an ancestor region containing it and by two or several edges to daughter nodes representing all regions contained in it. Going down in the tree means merging regions, going up means splitting regions. Producing the best segmentation then simply results in cutting the tree at the level producing the required bitstream. For generating the initial tree in intra mode, the tree of increasing dynamics is a good candidate. We are currently developing algorithms for updating it in time: this means ranking the fusions according to their cost/benefit in inter mode and providing a mechanism for introducing new regions. In order to evaluate the stability of tree representations we have made the following experience. We consider the simple situation where no new regions appear. We construct from the first intra frame a tree of fusions and derive from it two segmentations: a fine segmentation (fig. 11) included in a coarse one(fig. 12). A 3D time recursive segmentation as presented in section 3.1 is applied, one using the fine segmentation and the other the coarse one; we obtain a fine and a coarse segmentations of the whole sequence. We then apply to all frames of the fine segmentation the same fusions of labels as defined in the initial intra frame; this gives us again a coarse segmentation. The coarse segmentation obtained by both methods are the same. This shows that our tree representation is indeed stable.

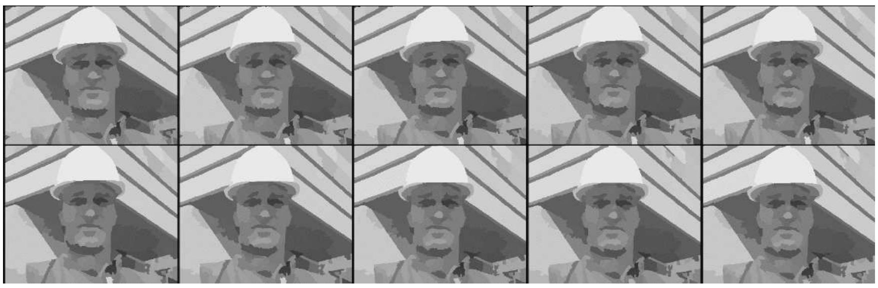


Fig. 11. Fine segmentation.

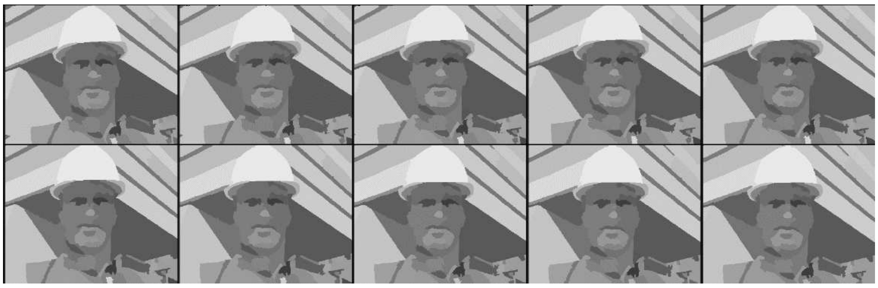


Fig. 12. Coarse segmentation.

4. Conclusion

A morphological segmentation algorithm for image sequences has been presented in this paper. Good stability is reached, including new regions only when an important change in the scene justifies it. A way to regulate the segmentation accuracy is introduced. In future work we intend to generate a tree structure with a criterion taking into account not only visual significance but also coding cost. In that way we expect to maximize the quality/cost rate of the system.

References

1. M. Kunt, A. Ikonomopoulos, and M.Kocher, "Second-generation image coding techniques," *Proc. IEEE*, vol. 73, April 1985.
2. J. Serra, *Image analysis and mathematical morphology.* Academic Press, 1982.
3. S. Beucher, *Segmentation d'images et morphologie mathématique.* PhD thesis, Ecole des Mines de Paris, Juin 1990.
4. P. Salembier and J. Serra, "Morphological multiscale image segmentation," in *Visual Communications and Image Processing*, vol. 1818, SPIE, 1992.
5. L. Vincent, "Grayscale area openings and closings, their efficient implementation and applications," in *Mathematical Morphology and its Applications to Signal Processing*, (Barcelona), Mai 1993.
6. S. Beucher and C. Lantuéjoul, "Sur l'utilisation de la ligne de partage des eaux en détection de contours," Rapport Interne N-598, Ecole des Mines de Paris, Centre de Morphologie Mathématique, Mai 1979.
7. F. Meyer and S. Beucher, "Morphological segmentation," *Journal of Visual Communication and Image Representation*, vol. 1, pp. 21–46, Septembre 1990.
8. M. Grimaud, *La géodésie numérique en morphologie mathématique. Application à la détection automatique de microcalcifications en mammographies numériques.* PhD thesis, Ecole des Mines de Paris, Décembre 1991.
9. F. Friedlander, *Le traitement morphologique d'images de cardiologie nucléaire.* PhD thesis, Ecole des Mines de Paris, Décembre 1989.
10. M. Pardas, P. Salembier, and L. Torres, "3d morphological segmentation for image sequences," in *IEEE Winter Workshop on Nonlinear Signal Processing*, (Tampere), 1993.
11. F. Meyer, "Arbre des minima et dynamique," rapport interne, Ecole des Mines de Paris, Centre de Morphologie Mathématique, mars 1990.