MORPHOLOGICAL OPERATORS FOR VERY LOW BIT RATE VIDEO CODING

P. Salembier †, *F. Meyer* ††, *P. Brigger* ‡ *and L. Bouchard* ‡‡

† Universitat Politècnica de Catalunya, Barcelona, Spain
†† Ecole des Mines de Paris, Fontainebleau, France
‡ Ecole Polytechnique fédérale de Lausanne, Lausanne, Switzerland
‡‡ Laboratoires d'Electronique Philips, Limeil-Brevannes, France

ABSTRACT

This paper deals with the use of some morphological tools for video coding at very low bit rates. Rather than describing a complete coding algorithm, the purpose of this paper is to focus on morphological *connected operators* and segmentation tools that have recently proved to be attractive for compression.

1. INTRODUCTION

1.1. Region-based video coding

Video coding algorithms relying on a content-based representation of the sequence are becoming increasingly popular. They rely on the definition of signal-dependent partitions that are transmitted to the receiver besides the more "classical" motion and texture (gray level and color) parameters. The use of signal-dependent partitions is expected to improve the coding efficiency and also to open the door to the so-called "content-based" functionalities where objects may be processed, transmitted and manipulated individually.

In a region-based coding approach, the geometrical characteristics of the signal play an important role. For instance, the definition of the partition involves a segmentation step which should ideally extract objects which are, at least partially, characterized by some geometrical properties. Another example is the coding of the partition, where the information to transmit is purely geometrical. As a result, there may be a need for geometrical tools for image compression.

Classical linear signal processing tools are not well suited for a geometrical approach and other tools coming from nonlinear signal processing or from computer vision may be attractive for this purpose. Mathematical morphology [1] has been developed as a geometrical approach to signal processing. Its application to video coding is becoming a very active field of research in which the *Morpheco* project has played a significant role.

1.2. The *Morpheco* project

The objectives of the *Morpheco* project[1] of the European *RACE* program were to define and develop region-based coding schemes and to study the interest of Mathematical Morphology within this framework. The project has contributed to the developments of coding strategies, of video coding tools and also to the state of the art in Mathematical Morphology. A brief summary of *Morpheco*'s main contributions is given in the sequel together with a selected list of publications.

- *Coding schemes*: the project has proposed complete coding schemes for still image [2], for video coding [3, 4, 5].
- *Video coding tools*: a large set of new segmentation tools have been developed [6, 7, 8, 9, 10]. Contributions to region-based motion estimation can be found in [11]. In region-based approaches, partition coding is a major issue. Contour-oriented techniques are discussed in [12, 13, 14] whereas shape-oriented schemes can be found in [15]. Texture coding is reported in [16, 17, 18]. Finally, tools allowing the tracking of objects are presented in [19].
- *State of the art in Mathematical Morphology*: The work done within *Morpheco* has also made some contributions to the state of the art in Mathematical Morphology. Let us mention in particular the following topics: connected operators [20, 21], watersheds [22, 9], and morphological interpolation [17].

Rather than describing a complete coding / decoding scheme, this paper focuses on two topics that have been intensively studied within the *Morpheco* project: connected operators and video segmentation for coding.

2. CONNECTED OPERATORS

The systematic study and use of connected operators has been one of the major contributions of *Morpheco* to mathematical morphology. These operators are very attractive because they can simplify a signal while preserving the contour information of the non-simplified components. For region-based video coding, they are used for preprocessing, image or partition simplification and segmentation. A *connected operator*, applied on a binary image does nothing but suppressing some particles or filling some holes: hence it never creates any new contour [20, 21]. A simple criterion could be: fill all holes with an area below some threshold.

If one considers the gray-level of a pixel as an altitude, the gray-tone image may be considered as a topographic surface. With this image, most concepts of morphology become extremely familiar: flat-zones (maximal connected components

[1]The project was funded by the European Commission within the framework of the *RACE* program. It started in Jan. 1992 and ended in Dec. 1995. The consortium was formed by the Universitat Politècnica de Catalunya (Spain), Ecole des Mines de Paris (France), Ecole Polytechnique fédérale de Lausanne (Switzerland), Laboratoires d'Electronique Philips (France), Ibermática (Spain) and Zenon (Greece) among which the first four institutions were in charge of the algorithms developments.

with the same altitude), regional maxima (flat-zones without higher neighbors), regional minima, saddle points etc. An algebraic study of gray-tone connected operators uses indeed the notion of flat-zones [20, 21]. Imagine now that the topographic surface is completely covered by the see. As the level of the see decreases, lakes will appear; at altitude h we will have n lakes $(L_h^1, L_h^2, ..., L_h^n)$; at the highest altitude H, the see L_H^1 is unique. The interconnections of all lakes as they evolve when the see level goes down may be represented as a tree. The root L_H^1 of the tree represents the see covering the whole area. As the level goes down one step, the extension of the see may remain unchanged, or some flat zones at the previous altitude get uncovered; due to these dried areas, the see may become disconnected in one or several connected components. A new node is created for each lake at the new altitude and is connected to the root node representing the see one level above: a single branch will be created in the case where the see remains one connected component and several branches building a fork when there is a disconnection. As the level of the water goes down, the same recursive procedure expanding the tree is applied to each lake L_h^k still present at the current altitude and creates the subtree tree $\left(L_h^k\right)$. The leaves of the tree represent the deepest level of the lakes before they dry out one level below: they are the regional minima of the topographical surface (see Fig. 1). The forks of the tree represent saddle zones and will play an important role in the segmentation.

The effect of connected operators on the tree can be simply explained as follows: a criterion is assessed for each node $\mathcal{M}(L_h^k)$. Based on this value, the node is either preserved or removed. In this last case, the subtree tree $\left(L_h^k\right)$is removed and the node's pixels are moved towards its father's node. At the end of the process, the output *Tree* is transformed back into a gray level image.

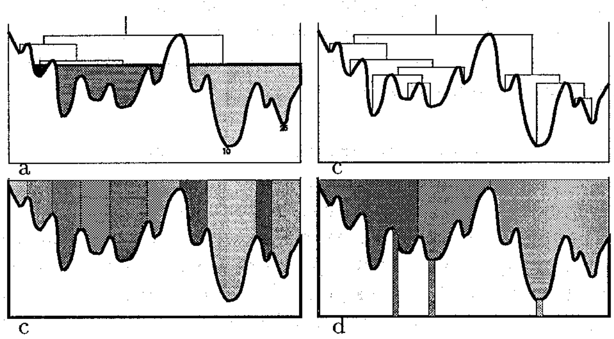

Figure 1: Top: construction of the tree of lakes; 1a) Three lakes have been created; 1b) Complete tree of lakes. Bottom: watershed lines; 1c) watershed lines associated to all minima, 1d) watershed lines associated to 3 markers. associated to all minima (left) and to a set of markers (right)

2.1. Criterion

Useful criteria for video sequence processing and coding are:

- *Dilation*: this criterion corresponds to the first known connected operator called *closing by reconstruction*. It consists in preserving a node L_h^k if the corresponding lake is not totally removed by an erosion with a structuring element. It provides a size-oriented simplification effect.
- *Area*: it is also a size-oriented operator, preserving L_h^k if the area of the corresponding lake is larger than a given threshold [23].
- *Volume:* it is size and contrast oriented operator: it preserves L_h^k if the volume below the lake is larger than a given threshold [26].
- *Contrast*: A contrast-oriented simplification can be obtained by preserving L_h^k if the depth of the corresponding lake is larger than a given threshold.
- *Complexity*: The complexity operator consists in preserving L_h^k if the ratio between the area of the lake and its perimeter is larger than a given limit [24]. This operator separates objects that have a long perimeter with respect to their size and are for this reason expensive to code within a contour-texture approach.
- *Motion*: here, the idea is to remove objects that do not undergo a specific motion. More details about this operator can be found in [25].

All these operators eliminate dark objects but if they are applied to the opposite of the image (dual operator) they also deal with bright objects. In practice, one operator is often used in cascade with its dual. Some simplification examples are shown in Fig. 2. Note the difference in the simplification effect and the very good contour preservation property.

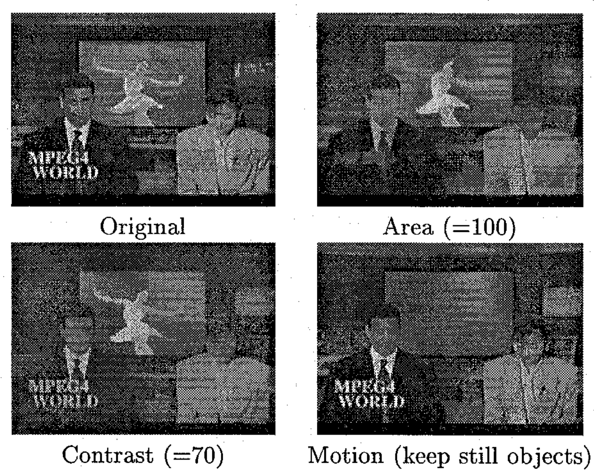

Figure 2: Examples of connected operator

3. SEGMENTATION

3.1. From connected filters to segmentation

In order to present the watershed transform, we will let the water invade again the topographic surface, starting from the minima. Imagine that we pierce each minimum of the topographic surface S and that we plunge this surface into the see with a constant vertical speed. The water entering through the holes floods the surface S. During the flooding, two or more floods coming from different minima may merge. We want to avoid this event and we build a dam on the points of the surface S where the floods would merge. At the end of the process, only the dams emerge, defining the watershed of the function; the watershed lines separate the various catchment basins; in fig.1c we have represented each catchment basin associated with a different grey-tone. For segmentation purpose, the watershed transform is usually applied to the gradient and not to the original image. Fig.3b shows the gradient of the cameraman image. In this image however there are far too many minima and the watershed line would produce far too many regions. But remember that the connected filters have been specially designed for filling valleys and hence suppressing regional minima. If we apply a series of connected filters of increasing size to the gradient image, the weakest minima will progressively vanish. If we construct the watershed line associated to each of theses images we produce a series of partitions with less and less regions with the following property: each contour present in a coarse partition is also present in any of the finer partitions. Fig.3c,d,f present a segmentation with the same number of regions associated to respectively a contrast filter, an area filter and a volumic filter; the volumic filter seems to have the best psycho-visual characteristics [26].

Fig.1d presents the last possibility to control the segmentation resulting from the watershed transform: pierce holes in the topographic surface only inside the regions one desires to segment. In the case of fig.1d, we have pierced 3 holes, yielding a partition of three regions. This procedure is at the heart of the recursive segmentation of sequences presented below: the previous segmentation serves as markers for the segmentation of the current image.

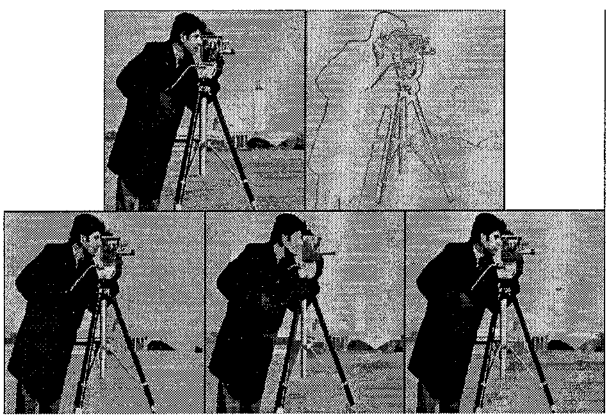

Figure 3: a) Original image; b) gradient; c) contrast; d) area; e) volume

3.2. Segmentation for coding

Segmenting for coding means finding a partition yielding the best trade-off between cost and quality of representation. Obviously, this result cannot rely on geometrical criteria alone: the coding capabilities and the coding cost have to be taken into account. If two adjacent regions can be represented with a unique texture model or can be predicted from a past image by a unique motion model, it will be cheaper to group them than to let them separate. In other words, there is a feedback between the coding capabilities and the segmentation itself.

3.3. Segmentation strategies

As we have seen, it is most natural to construct segmentation trees based on geometrical criteria. It is however unlikely that any threshold of the tree would yield an optimal segmentation for coding. The extraction of the optimal partition from this tree becomes a classical covering problem: finding the covering yielding the lowest distorsion for a given budget (see the presentation on *SESAME* in the same conference).

3.3.1. Intra mode

The goal of the intra mode is to produce the coding of the first image in the sequence and of all other image parts which cannot be predicted. Two strategies have been experimented: top-down and bottom-up.

Top-down approach [2, 9] The root of the tree is constructed first. After a strong simplification with connected filters, the initial image is segmented into a coarse partition. This partition is then coded by adjustment of a texture model. The coded image is subtracted from the original image: the coding residue produces a low signal in all regions which are already correctly represented and a high signal in the regions which ought to be resegmented. The tiles of the partitions where the coding residue is high are then recursively resegmented by using markers extracted after simplification of the current coding residue. As the partition gets finer, the criteria are changed: size and contrast criteria are successively applied.

Bottom-up approach [10] The bottom-up approach chooses for the leaves of the tree all flat-zones of a mildly simplified image (use of area openings and closings of small size). The most similar regions are then iteratively merged until the desired level of coarseness is reached. This last level will be the root of the tree. The similarity criteria are first very simple: colour distance between adjacent regions, integral of the gradient along the contour line. As the regions become bigger, the criteria become more sophisticated: ability of two adjacent regions to be represented by a unique texture model.

3.3.2. Inter mode: time recursive segmentation

In inter mode, the segmentation has to fulfill two goals: (a) ensure a good temporal stability of the regions in order to render the new partition as predictable as possible from the previous one (b) let important new regions appear. In order to avoid unacceptable time delays, a special spatio-temporal frame made of two images will be segmented: the first image contains the previous partition and the second image contains the current image to be segmented. The tiles of the previous partitions are treated as markers. According to the assignement rules of the flat zones of the current image, two strategies may be distinguished. Both of them operate fusions of regions in the order of decreasing similarity and two markers are never allowed to merge.

Strong projection [7, 11] A flat zone of the current image may only be assigned to a marker. The procedure stops when all flat-zones have been assigned to markers. The result of this projection stage is a partition where each tile is connected with a marker in the previous image. Since no new region could appear with this projection mechanism, a residue image is produced and used for resegmenting all tiles containing new regions. The procedure is very similar to the procedure used in the top-down approach in intra.

Weak projection [27] Flat zones of the current image may merge one with another or with a marker. Only the merging of two markers is forbidden. In this way, new regions may appear directly in the elaboration of the new partition tree.

4. CONCLUSION

With the tools of mathematical morphology presented in this paper, we were able to construct synthetic representations of sequences like mosaics and partition trees. The temporal stability of this representation makes is an ideal support for the new functionalities of *MPEG4*, like object tracking [19] and content based manipulation.

5. REFERENCES

[1] J. Serra. *Image Analysis and Mathematical Morphology.* Academic Press, 1982 (vol.1) and 1988 (vol.2)

[2] P. Salembier. Morphological multiscale segmentation for image coding. *Signal Processing*, 38(3):359–386, September 1994.

[3] P. Salembier, L. Torres, F. Meyer, and C. Gu. Region-based video coding using mathematical morphology. *Proc. of IEEE (Invited paper)*, 83(6):843–857, June 95.

[4] C. Gu and M. Kunt. Very low bit-rate video coding using multi-criterion segmentation. In *IEEE ICIP'94*, volume II, pages 418–422, Texas, U.S.A., Nov. 1994.

[5] P. Salembier and X. Ayuso. Very low bit rate video coding using active triangular mesh. In *ICASSP'96*, Atlanta (GA), USA, May 1996.

[6] F. Meyer. Morphological image segmentation for coding. In *First Workshop on Mathematical Morphology and its Applications to Signal Processing*, pages 46–51, Barcelona, Spain, May 1993.

[7] M. Pardàs and P. Salembier. 3D morphological segmentation and motion estimation for image sequences. *Signal Processing*, 38(2):31–43, September 1994.

[8] B. Marcotegui and F. Meyer. Morphological segmentation of image sequences. In*Second workshop on Mathematical morphology and its applications to image processing*, pages 101–108. Kluwer Academic Publishers, Fontainebleau, France, 1994.

[9] P. Salembier and M. Pardàs. Hierarchical morphological segmentation for image sequence coding. *IEEE Trans. on Image Processing*, 3(5):639–651, September 1994.

[10] B. Marcotegui, J. Crespo, and F. Meyer. Morphological segmentation using texture and coding cost. In *1995 IEEE workshop on Nonlinear Signal and Image Processing*, pages 246–249, Halkidiki, Greece, June 20-22 1995.

[11] M. Pardàs, P. Salembier, and B. González. Motion and region overlapping estimation for segmentation-based video coding. In *IEEE ICIP'94*, volume II, pages 428–432, Austin, Texas, November 1994.

[12] F. Marqués, J. Sauleda, and A. Gasull. Shape and location coding for contour images. In *Picture Coding Symposium*, pages 18.6.1–18.6.2, Lausanne, Switzerland, March 1993.

[13] C. Gu and M. Kunt. Contour simplification and motion compensated coding. *Signal Processing: Image Communication*, 7(4-6):279–296, November 1995.

[14] F. Meyer and O. Ribes. Contour coding system in the hexagonal raster. In *1995 IEEE Workshop on Nonlinear Signal and Image Processing*, pages 274–277, Halkidiki, Greece, June 20-22 1995.

[15] P. Brigger and M. Kunt. Morphological shape representation for very low bit-rate video coding. *Signal Processing: Image Communication*, 7(4-6):297–312, Nov. 1995.

[16] J. R. Casas and L. Torres. Coding of details in very-low bit-rate video systems. *IEEE Trans. on Circuits and Systems for Video Technology*, 4(3):317–327, June 1994.

[17] J.R. Casas, P. Salembier, and L. Torres. Morphological interpolation for texture coding. In *IEEE ICIP'95*, volume I, pages 526–529, Washington, USA, Oct. 95.

[18] D. Gimeno, L. Torres, and Casas J.R. A new approach to texture coding using stochastic vector quantization. In *IEEE ICIP'94*, Autsin, USA, November 1994.

[19] F. Marqués, B. Marcotegui, and F. Meyer. Tracking areas of interest for content-based functionalities in segmentation-based video coding. In *IEEE ICASSP'96*, Atlanta, USA, May 1996.

[20] J. Serra and P. Salembier. Connected operators and pyramids. In SPIE *Image Algebra and Mathematical Morphology*, volume 2030, pages 65–76, San Diego (CA), USA, July 1993.

[21] P. Salembier and J. Serra. Flat zones filtering, connected operators and filters by reconstruction. *IEEE Trans. on Image Processing*, 3(8):1153–1160, August 1995.

[22] F. Meyer. Color image segmentation. In *4th Int. Conf. on Image Processing and its Applications*, pages 303–304, Maastricht, The Netherlands, May 1992.

[23] L. Vincent. Grayscale area openings and closings, their efficient implementation and applications. In *First Workshop on Mathematical Morphology and its Applications to Signal Processing*, pages 22–27, Barcelona, Spain, May 1993.

[24] A. Oliveras and P. Salembier. Generalized connected operators. In SPIE *VCIP'96*, volume 2727, pages 761–773, Orlando (FL), USA, March 1996.

[25] P. Salembier, A. Oliveras, and L. Garrido. Motion connected operators for image sequences. In *EUSIPCO'96*, Trieste, Italy, September 1996.

[26] C. Vachier Extraction de caracteristiques, segmentation d'images et morphologie mathematique In *PhD thesis*, Ecole des Mines, Paris, 1996

[27] B. Marcotegui Segmentation algorithm by multicriteria region merging. In *ISMM96*, Atlanta, May, 1996.