CONNECTED OPERATORS FOR SPRITE CREATION AND LAYERED REPRESENTATION OF IMAGE SEQUENCES

P. Salembier, O. Pujol and L. Garrido
Dept. of Signal Theory and Communications,
Universitat Politècnica de Catalunya, Barcelona, Spain
Tel: (34) 9 3401 7404, Fax: (34) 9 3401 6447
Email: philippe@gps.tsc.upc.es, URL: http://gps-tsc.upc.es/imatge

ABSTRACT

This paper proposes and discusses the use of motion-oriented connected operators for sprite creation. Motion-oriented connected operators are tools allowing the simplification of frames by removing objects that do not follow a given motion. They combine features of filtering and segmentation tools. They are, however, less computationally expensive than most motion-oriented segmentation algorithms. In this paper, we show how they can be used to efficiently remove outliers with respect to the dominant motion and to create layered representation of sequences.

Keywords: Sprite representation, Mosaicing, Connected operators, Robust Motion estimation.

1 INTRODUCTION

Sprite or mosaic representations of video sequences is becoming popular for coding applications [9, 6, 5]. In particular, the MPEG-4 standard contemplates a prediction tool which is based on sprite representations [4]. Most of the time, these sprites are used to represent the background and concentrate the information that has been (or will be) visible in the sequence. They are used to improve the prediction of the background. The sprite itself can be used in different modes: it can be computed off-line and sent to the receiver at the beginning of the transmission or it can be progressively estimated on both the transmitted and the receiver sides in a causal fashion from the transmitted frames. Of course, the use of sprites is not limited to the background information and the approach can be used to deal with any components of the scene that has an homogeneous motion. This approach leads to layered representations of sequences [9]. Furthermore, sprite representation of objects is also investigated for content-based retrieval applications [10] (MPEG-7 for example). In this framework, sprite representation can be an attractive solution for representing a large number of frames in a compact and synthetic visual object that can be used for browsing and navigation purposes.

Most of the time, the sprite creation is achieved either by very simple techniques or by very complex ones. The simplest approach consists in, first, computing the dominant motion, second, compensating or wrapping all frames towards a single reference and, third, averaging the gray level or color values that are compensated at the same position. If foreground objects do not occlude the background in a large number of frames this approach gives satisfactory results. In order to improve the robustness of the scheme, the averaging operation may be replaced by a temporal median filter. On the other side of the complexity scale, motion segmentation algorithms can be used to actually segment the video sequence into a set of regions that are homogeneous in motion. The knowledge of the shape information of the visible part of the background in each frame allows us to improve the sprite creation by removing the outliers.

The goal of this paper is to propose and study the use of motion-oriented connected operators [3] to create sprite representation. Connected operators are used to remove the image components that do not follow the motion of the sprite under consideration without requiring an actual motion segmentation. The use of this approach leads to a low complexity scheme compared to real motion segmentation algorithms. The organization of this paper is as follows: section 2 reviews the basis and main features of connected operators. The application to sprite creation is presented in section 3. Finally, section 4 is devoted to the conclusions.

2 CONNECTED OPERATORS

2.1 Image processing with connected operators

Connected operators [8, 1, 7] are filtering techniques derived from mathematical morphology that eliminate part of the image content while preserving the contour information of the remaining parts of the image. They interact with the signal by means of *flat zones* [1]. A connected operator is an operator that can only merge *flat zones* of the image. As a result, it can simplify an image but without introducing any new contour.

[1] *Flat zones* are the connected components where the image is constant. Note that a *flat zone* can be reduced to a single point.

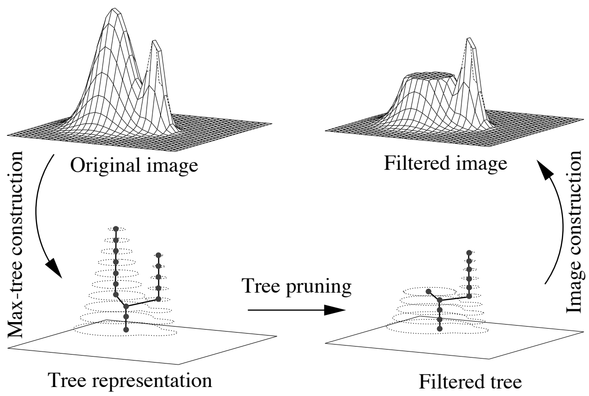


Figure 1: Connected operator: filtering strategy in three steps: 1) *Max-tree* creation, 2) pruning of the tree, 3) filtered image reconstruction

An efficient way of creating and implementing connected operators relies on *Max-Tree* representations [7]. The filtering strategy is illustrated by Fig. 1. The image is considered as a 3D relief and the first step is to construct a *Max-Tree* representation of the image. The nodes of the tree represent the binary connected components resulting from the thresholding of the original image at all possible gray level values. The links between the nodes describe how the *flat zones* may be merged. Note that there exist fast algorithms to construct the tree (see [7]). Then, starting from the leaves of the tree, each node is studied and a particular criterion is assessed for each node. If the criterion value is above (below) a given threshold, the node is preserved (removed). If the criterion is increasing, that is if the criterion value of a node is always smaller or equal to the criterion value of its father, then the algorithm defines a tree pruning strategy and, at the end of the pruning, the filtered image is reconstructed by stacking the connected components corresponding to the remaining nodes. If the criterion is not increasing, the definition of the pruning strategy is less straightforward. As discussed in [7], the non-increasingness of a criterion is most of the time a drawback that implies a lack of robustness of the operator (similar images may give different results). In [7], a solution relying on dynamic programming technique (Viterbi algorithm) was proposed. In the sequel, since the motion criterion is not increasing, this solution is assumed to be used and the interested reader is referred to [7] for more details on this issue.

The operator is said to be *anti-extensive* because the filtered image is, for each pixel, smaller than the original image. In practice, this means that the operator simplifies the image by removing its bright components that do not fulfill a given criterion. To simplify dark components, the dual operator should be used. If $\psi(I(\vec{p}))$ is a connected operator applied on image $I(\vec{p})$, its dual operator is: $\psi^*(I(\vec{p})) = -\psi(-I(\vec{p}))$.

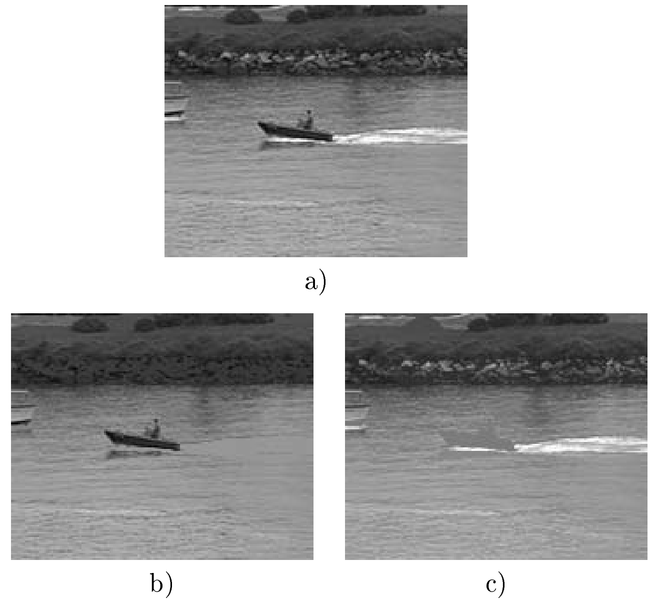


Figure 2: Example of simplification with motion-oriented connected operator. The operator eliminates the image components that do not follow the global motion. a) Original frame, b) Elimination of bright components with motion-oriented connected operators, c) elimination of dark components with the dual operator.

2.2 Motion-oriented connected operators

Recently a motion-oriented connected operator has been defined [3, 7]. This operator allows the elimination of the image components that do not undergo a given motion. The filtering parameter is the motion field itself. Let us denote by $\vec{d}(\vec{p})$ the motion field corresponding to an image $I_t(\vec{p})$ at time t. The basic filtering strategy is to measure for each node Z_f of the *Max-tree* of image I_t, the opposite [2] of the energy of the mean displaced frame difference (DFD):

$$\mathcal{C}_{I_t}^{I_{t-1}}(Z_f) = -\sum_{\vec{p} \in Z_f} [I_t(\vec{p}) - I_{t-1}(\vec{p} - \vec{d}(\vec{p}))]^2 / N \quad (1)$$

where N is the number of pixels of Z_f. If the criterion value is high, the connected component corresponding to Z_f actually follows the motion $\vec{d}$ and should be preserved. However, if the criterion value is low, the connected component can be considered as an outlier and should be removed. As explained in [3, 7], the criterion robustness can be improved if a memory term is

[2] the opposite of the DFD is used in order to preserve nodes that correspond to high values of the criterion

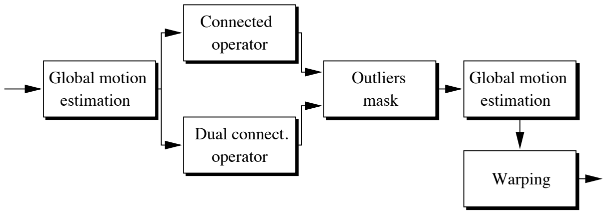


Figure 3: Diagram flow for sprite creation

introduced:

$$\mathcal{C}(Z_f) = \alpha \mathcal{C}_{I_t}^{I_{t-1}}(Z_f) + (1-\alpha)\mathcal{C}_{I_t}^{\psi(I_{t-1})}(Z_f) \quad (2)$$

This equation states that the criterion is a weighted combination of 1) the DFD between the original frames at time t and $t-1$ and 2) the DFD between the original frame at time t and the filtered frame at time $t-1$.

Let us illustrate the effect of the operator. Fig. 2.a shows an original frame. Assume that the dominant motion has been estimated. This motion is forwarded to the operator in order to eliminate the image components that do not follow this motion. Fig. 2.b shows the output of the operator. As can be seen, bright objects that do not follow the background motion have been removed. However, the remaining components have been well preserved. Fig. 2.c shows the output of the dual operator applied on the original frame. Here, dark objects not following the global motion have been removed.

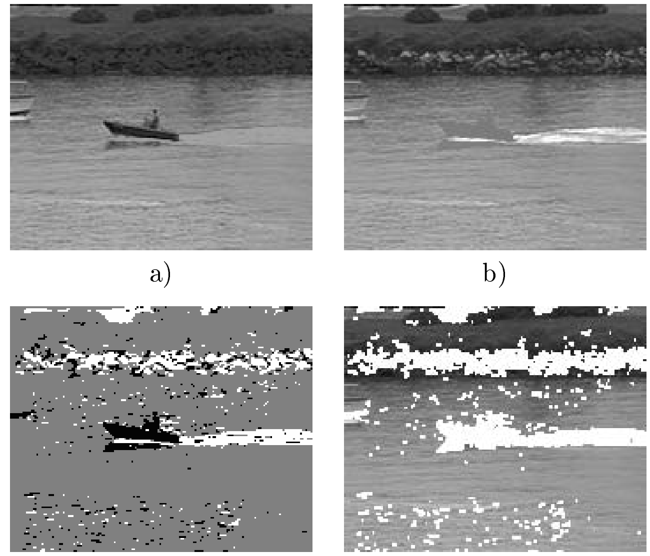
3 SPRITE CREATION

The proposed scheme for sprite creation is illustrated in Fig. 3. Let us assume that we deal with the background (however, the approach can be generalized to any group of objects that undergo a similar motion). The first step is to estimate the dominant motion. This can viewed as the estimation of the dense motion field $\vec{d}(\vec{p})$ that minimizes the DFD [2]:

$$\mathcal{D}(R) = \sum_{\vec{p} \in R} [I_t(\vec{p}) - I_{t-1}(\vec{p} - \vec{d}(\vec{p}))]^2 \quad (3)$$

where R is the region on which the estimation is performed.

For the estimation of the dominant motion, R is the entire frame. In our experiment, a affine model of the dense motion field has been assumed. This dense motion field is forwarded to a motion-oriented connected operator that removes all bright components that do not follow the dominant motion. In parallel, the dual operator is applied to remove dark components. Examples of filtered frames is shown in Fig. 4.a and 4.b (the original frame is the one of Fig. 2.a). As can be seen, the background objects have been preserved and the two boats have been removed. By simple comparison between the filtered and the original frames, the outliers mask can be extracted. In Fig. 4.c, the pixels in white (dark) correspond to locations where the output values of the connected operators (dual connected operators) significantly differ from the original image. This outliers mask identifies the pixel that have to be warped towards the sprite (see Fig. 4.d).



a) b) c) d)

Figure 4: Example of sprite creation: Processing of the first frame (original frame can be seen in Fig 2.a) a) output of the connected operator, b) output of the dual connected operator, c) outliers mask: white and black pixels are outliers respectively removed by the connected operator and its dual, d) first image warped in the sprite

Before actually warping the frame, the motion can be re-estimated. Indeed the presence of outliers may have corrupted the initial estimation. Once detected, these outliers can be removed from the frame and the motion can be re-estimated on a region of support R that contains no outliers. As usual, the entire sprite is progressively constructed by averaging the pixels that are compensated at the same position. Fig 5 shows the progressive construction of the sprite. Finally, the interest of outliers removal is illustrated in Fig. 6. In this case, the sprite has been created without removing the outliers. As can be seen, the presence of these outliers has a strong negative influence in the final result.

4 CONCLUSIONS

In this paper, the application of motion-oriented connected operators to the problem of sprite generation has been proposed. Connected operators are efficient tools to define and to extract outliers with respect to a given motion field. They are particularly efficient is terms of

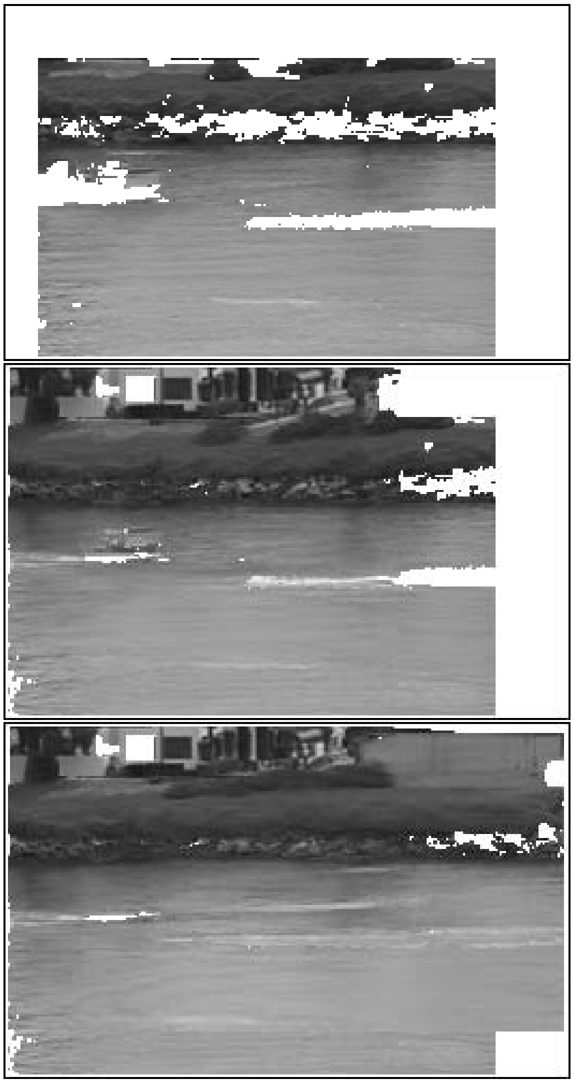


Figure 5: Progressive construction of the sprite. Top: 50 frames, Center: 100 frames, Bottom: 200 frames.

Figure 6: Sprite without outliers elimination (200 frames)

computational complexity and avoid the use of a complex motion segmentation algorithm. Furthermore, they are independent from the motion estimation and from the warping stages. As a result, they can be introduced, just before the warping, in any sprite creation scheme.

References

[1] J. Crespo. *Morphological Connected Filters and Intra-region Smoothing for Image Segmentation.* PhD thesis, Georgia Institute of Technology, 1993.

[2] J. L. Dugelay and H. Sanson. Differential methods for the identification of 2D and 3D motion models in image sequences. *Signal Processing, Image Communication*, 7:105–127, 1995.

[3] L. Garrido, A. Oliveras, and P. Salembier. Motion analysis of image sequences using connected operators. In *SPIE Visual Communications and Image Processing, VCIP'97*, volume 3024, pages 546–557, San Jose, CA, USA, February 1997.

[4] MPEG/N1648 ISO/IEC/ JTC1/SC29/WG11. Core experiment on sprites and global motion compensation, April 1997.

[5] K. Jinzenji, S. Ishibashi, and H. Kotera. Algorithm for automatically producing layered sprites by detecting camera movement. In *IEEE International Conference on Image Processing, ICIP'97*, pages I:767–770, Santa Barbara, USA, October 1997.

[6] M.C Lee, W. Chen, C.B. Lin, C. Gu, T. Markoc, S.I. Zabinsky, and R. Szeliski. A layered video object coding system using sprite and affine motion model. *IEEE Transactions on circuits and systems for video technology*, 7(1):130–145, February 1997.

[7] P. Salembier, A. Oliveras, and L. Garrido. Anti-extensive connected operators for image and sequence processing. *IEEE Transactions on Image Processing*, 7(4):555–570, April 1998.

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

[9] J.Y.A. Wang and E.H. Adelson. Representing moving images with layers. *IEEE Transactions on Image Processing*, 3(5):625–638, September 1994.

[10] H. Zhang, J.Y.A. Wang, and Y. Altunbasak. Content-based video retrieval and compression: A unified solution. In *IEEE International Conference on Image Processing, ICIP'97*, pages I:13–16, Santa Barbara, USA, October 1997.