

Robust motion estimation using connected operators

Philippe Salembier

Dept. of Signal Theory and Communications
Universitat Politècnica de Catalunya
Barcelona, Spain
philippe@gps.tsc.upc.es

Henri Sanson

FRANCE TELECOM - CNET/DIH
CCETT Laboratories, Department ATI
Cesson-Sevigne, France
sanson@ccett.fr

Abstract

This paper discusses the use of connected operators for robust motion estimation. The proposed strategy involves a motion estimation step, extracting the dominant motion, and a filtering step, relying on connected operators that remove objects that do not follow the dominant motion. These two steps are iterated in order to obtain an accurate motion estimation and a precise definition of the objects following this motion. This strategy can be applied on the entire frame or on individual connected components. As a result, the complete (motion-oriented) segmentation and motion estimation of the frame can be achieved.

1 Introduction

The classical approach to 2D motion estimation of image sequences generally relies on first, the definition of a partition, second, the definition of a motion model and third, the estimation of a single motion model for each region of the partition. For instance, the popular *block matching algorithm* defines a partition made of squares and assumes that the motion of each square can be modeled by a translation. Recently, a high number of contributions have been published on the use of polynomial model (affine or quadratic) to estimate and to describe the optical flow on arbitrary shaped regions (see for example [8, 2] and the references therein). This approach has the advantage of allowing the motion estimation on regions of large size which, moreover, may depend on the image content. However, it assumes that a robust motion estimation technique is available. Indeed, if the model is complex, its optimization may become a difficult issue and its sensitivity to outliers may become a serious problem. Moreover, polynomial motion models are generally used within rather large areas. As a result, outliers may likely be present.

Robust motion estimation [1] is generally achieved by modifying the weights of the values of the displaced frame difference (DFD) in such a way that the influence of outliers (generally corresponding to high values of the DFD) remain moderate. The weights are usually derived from a robust statistical formulation of the problem [4]. Note that, in practice, this approach does not take into account the spatial correlation of the outliers, that is, the DFD at each location is processed independently from the DFD of its neighbors.

In this paper, we propose the use of *motion-connected operators* to efficiently define the region on which the motion estimation has to be performed and to efficiently deal with the problem of outliers. Connected operators [6] 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. Recently, a motion connected operator has been proposed [3] and our objective in this paper is to combine it with region-based motion estimation techniques to achieve a robust estimation as well as an efficient motion segmentation. The organization of this paper is as follows: section 2 discusses the main features of connected operators. Section 3 reviews the motion estimation algorithm that has been used in this work. The combination of motion estimation and connected operators is presented in section 4. Finally, section 5 is devoted to the presentation of examples and some discussions.

2 Connected operators

2.1 Filtering strategy

Connected operators interact with the signal by means of *flat zones*. *Flat zones* are the largest connected components of the space where the image is constant (a *flat zone* can be reduced to a single point). The set of flat zones creates a partition of the image and a connected operator is an operator that can only merge *flat zones* of the image. This definition explains why they can simplify as well as preserve the contour information. In fact, they cannot introduce any new contour.

There are several ways of creating connected operators. In the sequel, we describe one relying on *Max-Tree* representations [5]. 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. Then, starting by 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). As can be seen, this is a pruning strategy and, at the end of the pruning, the filtered

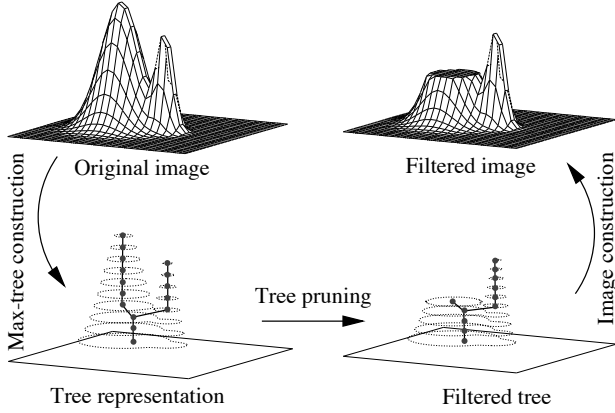


Figure 1: Connected operators: Filtering strategy

image is reconstructed by stacking the connected components corresponding to the remaining nodes.

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: $\psi^*(I(\vec{p})) = -\psi(-I(\vec{p}))$.

An attractive feature of this filtering approach is that it offers a large number of simplification criteria. Operators have been defined for instance to eliminate the image components either of small size, or of low contrast, or of low complexity, etc. The basic filtering strategy is to assess a specific criterion $\mathcal{C}(Z_f)$ (area, contrast, complexity, etc.) for each node Z_f of the *Max-tree* and to decide whether or not Z_f has to be pruned.

The most popular simplification criteria are increasing. In the framework of connected operators, increasingness means that: if $Z_j^1 \subseteq Z_j^2 \Rightarrow \mathcal{C}(Z_j^1) \leq \mathcal{C}(Z_j^2)$. A typical criterion is the size measured as the number of pixels. In this case, if a tree node is preserved (removed), all its ancestors (child) nodes correspond to connected components of larger (smaller) size and, as a result, they are also preserved (removed). Therefore, it is very easy to define the pruning limit for each tree branch. When the criterion is not increasing, this property does not hold anymore, and it is more difficult to define where to stop the pruning. In the case of non increasing criteria, an efficient solution based on dynamic programming techniques (Viterbi algorithm) has been proposed in [3, 5]. This solution is assumed to be used in the following.

2.2 Motion-oriented connected operators

Recently a motion-oriented connected operator has been defined [3, 5]. This operator allows the elimination of the image components that do not undergo a given motion. The filtering parameter is the motion model 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 opposite¹ 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 size 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, 5], the robustness of this criterion can be improved if a memory effect is introduced in the computation:

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

Let us illustrate the use of this operator with a simple example shown in Fig. 2. In this example, the goal of the operator is to remove all moving objects. The motion model is defined by $\vec{d}(\vec{p}) = \vec{0}$. In the part of the sequence shown in Fig. 2.a, all objects are still except the ballerina behind the two speakers and the speakers themselves. In this example, the motion connected operator is applied on the original image to remove bright moving components and the dual operator is used to remove dark moving components. As can be seen on Fig. 2.b, the operator has very precisely extracted the ballerina and the (moving) details of the speakers. Note that the contour information of the remaining objects is precisely preserved. Fig. 2.c shows the difference between the filtered and the original images. A constant value of 128 has been added to this residue so that the 0 difference appears as a mean gray level value of 128. Finally, let us mention that there exist fast algorithms to construct and to analyze the *Max-Tree* (see [5]) and connected operators can be very efficiently implemented (order of 1s of CPU for 256*256 images on a Pentium 200).

3 Region-based motion estimation

Motion estimation of regions of arbitrary shape R can be viewed as the estimation of the dense motion field $\vec{d}(\vec{p})$ that minimizes the DFD [2, 7]:

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

¹the opposite of the DFD is used in order to preserve nodes that correspond to high values of the criterion

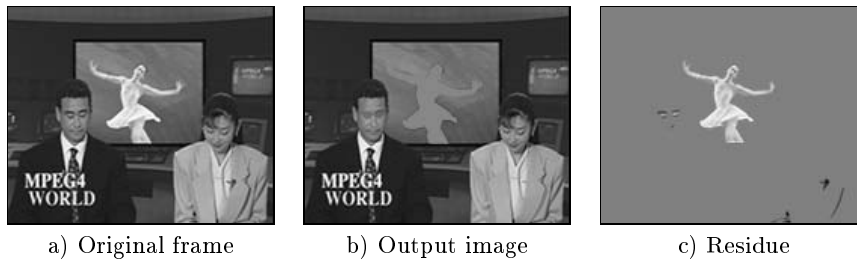


Figure 2: Example of motion oriented simplification with connected operators

The motion field is assumed to satisfy a polynomial representation:

$$\vec{d}(\vec{p}) = \sum_{k=0}^{D-1} a_k f_k(\vec{p}) \quad (4)$$

where $f_k(\vec{p})$ is a set of functions creating a basis of the space of interest. In our case, we have used first order polynomials to model affine transformations. This choice has been done to be able to deal with motion such as translation, zoom, and rotation.

4 Robust motion estimation with connected operators

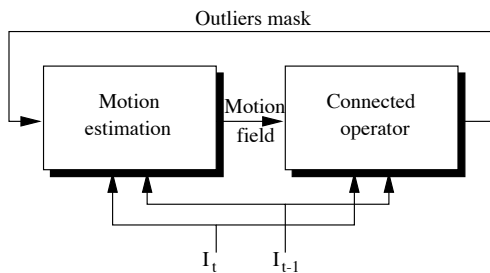


Figure 3: Motion estimation and filtering scheme

The robust motion estimation scheme proposed in this paper is described in Fig. 3. The first step is the estimation of the motion between frames I_t and I_{t-1} . The estimation depends on an outliers mask which defines which pixels have to be discarded for the estimation and which pixels have to be considered. At the beginning of the process, the outliers mask is set to 255. This means that the estimation is done on the entire frame and the resulting motion field corresponds to the dominant motion. The motion estimation output is used by connected operators to remove the image components that do not follow the dominant motion. By thresholding the difference between the original and filtered images, it is possible to define a mask of outliers. This process gives a first estimation of the dominant motion and of the corresponding areas. This first estima-

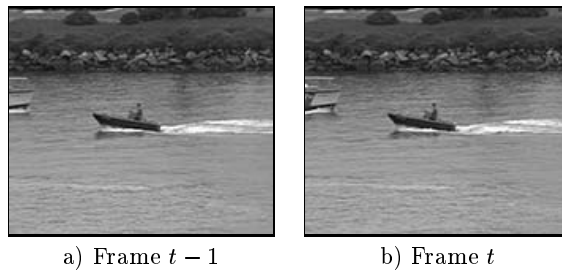


Figure 4: Original frames

tion can be improved by using the same scheme but taking into account the information about the outliers.

5 Examples and discussions

Let us illustrate the approach with an example. Fig. 4.a and 4.b show two original frames of a coast sequence involving two boats (one in the center and one appearing on the left side). The camera is tracking the black boat in the center of the image. Therefore, this boat has an apparent motion of $\vec{0}$ whereas the background is moving with an apparent horizontal translation. The dominant motion field is presented in Fig. 5.a. Following this estimation result, the upper part of the scene is moving horizontally to the left, the lower part of the scene is moving to the right and the central part is still. This result does not correspond to the actual motion of the scene. The estimation error is produced by the boats which do not follow the background motion. In particular, the boat in the center of the image forces a still motion in the center of the motion field. However, this motion field roughly describes the dominant motion and motion connected operators can be used to remove from the scene the components that do not follow this dominant motion. The operator is applied on the original frame I_t (Fig. 4.b) and the filtering result is shown in Fig. 5.b. In this result, we can see that the operator has successfully removed the two boats while preserving the background information. Fig. 5.c presents a binary image indicating where the filtered image is different from the original image. In this binary image, the black pixels define areas that do not follow the dominant motion, in other

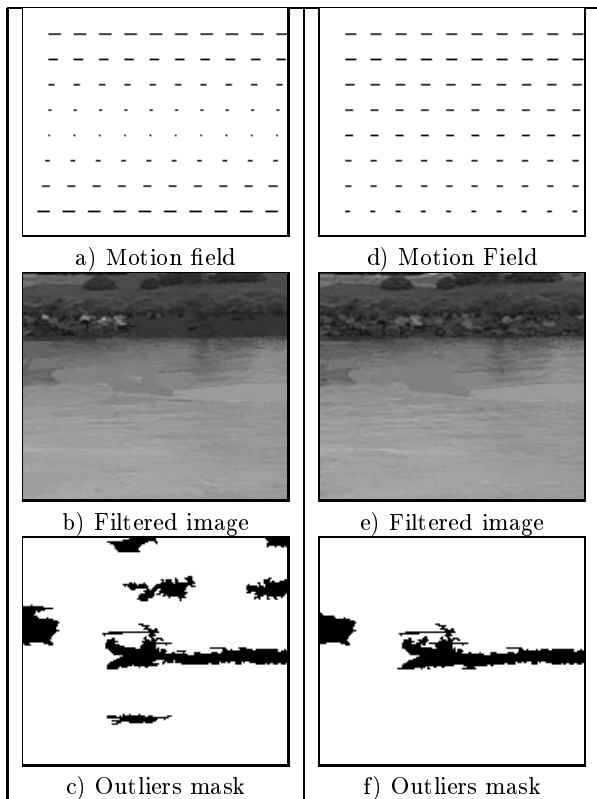


Figure 5: Example of robust motion estimation with connected operators. First column: first estimation and filtering; Second column: second estimation and filtering

words outliers.

Now, using the outliers mask, a second motion estimation is performed but only on the areas without outliers. If I_t^0 denote the area of dominant motion without outliers, the estimation algorithm minimizes $\mathcal{D}(I_t^0)$ as defined by equation 3. The resulting motion field is shown in Fig. 5.d. The dominant motion now corresponds to an horizontal translation towards the left. Note that that the magnitude of the translation is function of the objects distance from the camera. Based on this estimation, a new motion simplification can be performed (Fig. 5.e). The resulting outliers mask is shown in Fig. 5.f. In this mask, the two boats that do not follow the dominant motion are clearly identified whereas the background area is properly defined. Note that parts of the background which were lost during the first estimation are now well classified.

This procedure can be iterated in a hierarchical way in order to analyze the whole frame. That is, the same algorithm can be applied on each connected component of the outliers (dark areas of Fig. 5.f). In this case, the motion estimator defines the dominant motion within each connected component, as shown in Fig. 6.a. Then, for

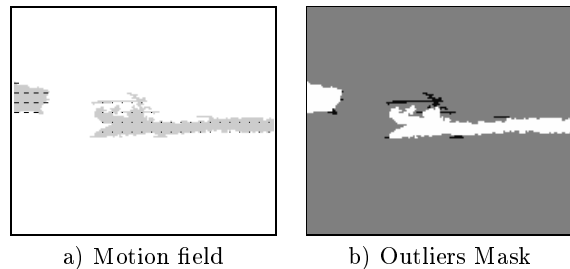


Figure 6: Non-dominant motion estimation (in b) the outliers are in black)

each connected component, the motion connected operator detects the outliers with respect to its motion. Fig. 6.b shows the final mask which defines the background area in gray, the two boats in white and some area of outliers.

This motion estimation approach is at the same time robust and efficient in terms of computational complexity. Finally, note that the outliers mask is a very good starting point for motion-oriented segmentation or online creation of mosaic representation.

References

- [1] M. J. Black and P. Anandan. The robust estimation of multiple motions: Parametric and peicewise-smooth flow fields. *Computer Vision and Image Understanding*, 63(1):75–104, January 1996.
- [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] P.J. Huber. *Robust statistics*. Wiley, New York, 1981.
- [5] P. Salembier, A. Oliveras, and L. Garrido. Anti-extensive connected operators for image and sequence processing. *IEEE Transactions on Image Processing*, To be published.
- [6] 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.
- [7] H. Sanson. Toward a robust parametric identification of motion on regions of arbitrary shape by non-linear optimization. In *Proceedings of IEEE Internatioanl Conference on Image Processing, ICIP'95*, volume I, pages 203–206, October 1995.
- [8] G. Tziritas and C. Labit. *Motion analysis for image sequence coding*, volume 4 of *Advances in Image Communication*. Elsevier, 1994.