HAL
open science

A NEW HYBRID DIFFERENTIAL FILTER FOR MOTION DETECTION

Julien Richefeu, Antoine Manzanera

To cite this version:

Julien Richefeu, Antoine Manzanera. A NEW HYBRID DIFFERENTIAL FILTER FOR MOTION DETECTION. International Conference on Computer Vision and Graphics (ICCVG'04), Sep 2004, Varsovie, Poland. 10.1007/1-4020-4179-9_105 . hal-01222690

HAL Id: hal-01222690

https://hal.science/hal-01222690v1

Submitted on 30 Oct 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A NEW HYBRID DIFFERENTIAL FILTER FOR MOTION DETECTION

Julien Richefeu, Antoine Manzanera
Ecole Nationale Supérieure de Techniques Avancées
Unité d'Electronique et d'Informatique
32, Boulevard Victor
75739 Paris CEDEX - France

Abstract A new operator to compute time differentiation in an image sequence is presented. It is founded on hybrid filters combining morphological and linear recursive operations. It estimates recursively the amplitude of time-variation within a certain interval. It combines the change detection capability of the temporal morphological gradient, and the (exponential) smoothing effect of the linear recursive average. It is particularly suited to small and low amplitude motion. We show how to use this filter within an adaptive motion detection algorithm.

Keywords: hybrid filter, temporal morphology, motion detection.

1. Introduction and Preliminaries

In some applications such as video databases, video compression, security monitoring or medical imaging, motion information is usually a more significant cue than color or texture. More specifically, the need to detect interesting moving objects in the scene is a fundamental low-level step in many vision systems. In this paper, we will concentrate on the case of video surveillance systems using a stationary camera.

The challenge of motion detection lies in the ability of performing an accurate segmentation of the moving objects independently on their size, velocity and contrast with respect to the background of the scene. A large set of motion detection algorithms has already been proposed in the literature. We can file them into four main categories according to the type of inter-frame computations, i.e. to the way the time differentiation is performed.

The first one is based on *temporal gradient*: a motion likelihood index is measured by the instantaneous change in the image intensity computed by differentiation of consecutive frames [3]. These methods are naturally adaptive to changing environments, but are also dependent on the velocity and size of

moving objects. This drawback can be minimized using a multiple bank of spatiotemporal filters which is done at the price of an increased complexity.

The second category are the *background subtraction* techniques [11] [4] [8], that use a reference image (background), representing the stationary elements in the scene. Here the motion likelihood measure is the difference between the current frame and the background. These methods are less dependent on the velocity and size of the objects. Nevertheless, the adaptation to dynamic environment is a much more difficult task, which can penalize the detection of small amplitude motion (very slow or low contrast objects).

The third type of approach is based on the computation of the local apparent velocity (*optical flow*) [2] which is used as input of a spatial segmentation [9]. This method provides valuable information but it is in general more computationally complex and it is also sensitive to the reliability of the optical flow. Thus a trade-off has to be found between the smoothness of the flow field and the accuracy of segmentation.

More recently, *morphological filters* have been employed [5] [10] [1] for video sequence analysis. By using spatiotemporal structuring elements, a local amplitude of variation can be computed as motion likelihood index. Such measure can be useful to detect small amplitude motion, but as it is sensitive to outliers, it is usually integrated over regions using connected operators.

We propose in this paper a new differential operator based on a hybrid filter, combining morphological and linear operations. It computes a pixel-wise amplitude of time-variation over a recursively defined "temporal window". This method is designed to address the problem of small objects and slow motion while providing a certain noise immunity thanks to its linear part.

We first present the *forgetting morphological temporal gradient* in Section 2. Then we show how to use the output of this filter within a motion detection algorithm in Section 3. Results are presented and discussed in the same section.

2. The forgetting morphological temporal gradient

In this section, we introduce the forgetting morphological temporal gradient and show its interests for motion detection systems.

Considering an image sequence $I_t(x)$ where t is a time index and x a (bi-dimensional) space index, morphological temporal filters are defined using temporal structuring element $\tau = [t_1, t_2]$. The temporal erosion (resp. dilation) is defined by $\varepsilon_\tau(I_t)(x) = \min_{z \in \tau}\{I_{t+z}(x)\}$ (resp. $\delta_\tau(I_t)(x) = \max_{z \in \tau}\{I_{t+z}(x)\}$). The temporal (morphological) gradient γ is then defined by $\gamma_\tau(I_t) = \delta_\tau(I_t) - \varepsilon_\tau(I_t)$.

τ represents the temporal interval of interaction, which can be causal (e.g. $[-3, 0]$), anti-causal (e.g. $[0, +5]$) or both (e.g. $[-1, +1]$). Thus the temporal gradient corresponds to the amplitude of variation within this interval.

The use of this operator suffers from two major drawbacks: (1) it implies the use of a buffer with size corresponding to the diameter of the structuring element, which can be very memory consuming; (2) it is very sensitive to sudden large variations (like impulse noise or slight oscillations of the sensor).

To cope with these two problems, we use hybrid filters, which can be viewed as a recursive estimation of the values of the temporal erosion and dilation. Using parameter α, which is a real number between 0 and 1, the *forgetting temporal dilation* M_t (resp. erosion m_t) is defined as shown in Figure 1. As in the classical running average (or exponential smoothing) defined by $A_t(x) = \alpha I_t(x) + (1 - \alpha)A_{t-1}(x)$, the inverse of α has the dimension of time. The semantics of $M_t(x)$ (resp. $m_t(x)$) is then the (estimated) maximal (resp. minimal) value observed at pixel x within the $1/\alpha$ last frames. So as α tends to unity, M_t (resp. m_t) tends to I_t, and as α tends to zero, M_t (resp. m_t) tends to the maximal (resp. minimal) value observed during the whole sequence. The use of the term "forgetting" is justified by the fact that these operators attach more importance to the near past than to the far past.

Initialization	
for each pixel x:	$M_0(x) = m_0(x) = I_0(x)$
For each frame t	
for each pixel x:	$M_t(x) = \alpha I_t(x) + (1 - \alpha) max\{I_t(x), M_{t-1}(x)\}$
	$m_t(x) = \alpha I_t(x) + (1 - \alpha) min\{I_t(x), m_{t-1}(x)\}$
For each frame t	
for each pixel x:	$\Gamma_t(x) = M_t(x) - m_t(x)$

Figure 1. The forgetting morphological temporal operators: M_t is the forgetting dilation, m_t the forgetting erosion and Γ_t the forgetting morphological gradient.

Γ_t, the forgetting morphological temporal gradient, is used in the following as a differentiation filter because of its interesting properties: (1) it has the dimension of an amplitude of time-variation, so it is able to integrate motion over a long period depending on $1/\alpha$, and then to detect small or slow moving objects; (2) it is less sensitive to impulse noise because of its forgetting term, corresponding to the exponentially decreasing weights attached to the past values; (3) it only requires the use of two buffers to compute the forgetting erosion and dilation. Figure 2 displays the forgetting morphological operations compared with their morphological counterparts.

The forgetting morphological gradient thus represents a relevant motion likelihood measure. We show on the next section how it can be used within a complete moving objects detection algorithm.

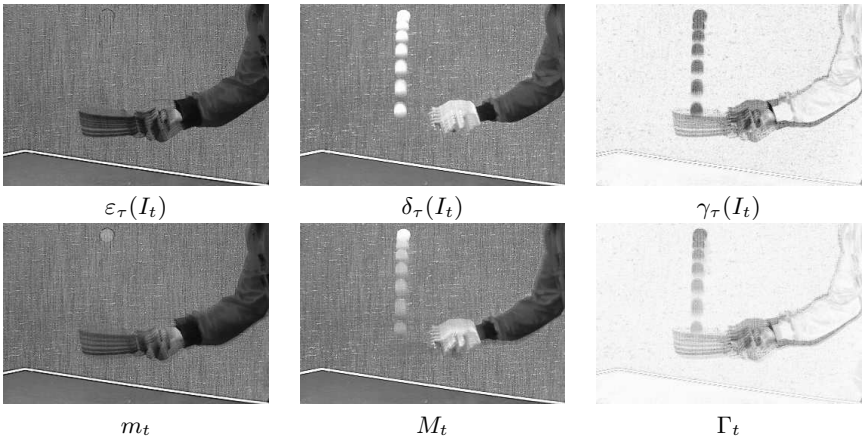

Figure 2. Application of the forgetting morphological operators (bottom), compared with the classical morphological temporal operators (top), computed on the frame $t = 19$ of the classical "Tennis" sequence (Berkeley). For comparison purposes the structuring element is $\tau = [-8, 0]$, and the forgetting term is $\alpha = 1/9$. The gradients are displayed in reverse video mode. Note that we use the symmetrical gradient in order to treat dark and light objects the same way.

3. Motion detection algorithm

The filter presented above makes possible a good level of detection for motion whose amplitude is below the spatiotemporal discretization. Nevertheless, the forgetting term α needs to be adapted to the velocity of the moving object. As there are several objects with different sizes and velocities within the scene, it is necessary to adjust locally the value of α to the observations.

In addition to this, we need a decision rule to discriminate the moving objects from the background. Because the scene is constantly evolving - typically under illumination or weather condition changes - the decision criterion has to be temporally adaptive. Furthermore the temporal variation, possibly due to moving objects, but also to noise or irrelevant motion, is not uniformly distributed through the scene. So the decision must be locally differentiated.

In our algorithm, we compute, from the forgetting morphological temporal filter, a local estimation of the spatiotemporal activity. This estimation is used both for deciding the pixel-level motion label and for adjusting the value of the forgetting term α.

Following [6], we use the Σ-Δ filter to compute a second order statistics on the sequence: The Σ-Δ filter S_t of a time series X_t is a recursive approximation of the median defined by: $S_t = S_{t-1} + 1$ if $S_t < X_t$ and $S_t = S_{t-1} - 1$ if $S_t > X_t$.

Now, what we compute exactly is V_t, the Σ-Δ filter of N times the non-zero values of the forgetting morphological temporal filter. Then, the local

estimation of the spatiotemporal activity is defined by $\Theta_t = G_\sigma(V_t)$, where G_σ is the bidimensionnal Gaussian filter with standard deviation σ.

Θ_t is used as the first step of decision for the moving label of each pixel: $D_t = 1$ if $\Gamma_t > \Theta_t$ and $D_t = 0$ elsewhere.

Θ_t is also used to update the forgetting term for every pixel. Inspired by [7] who used a locally adaptive learning rate for recursive background estimation, we compute the local α for pixel x by a similar formula: $\alpha_t(x) = 2^{-\frac{\overline{\Theta_t(x)}^2}{k^2}}$ where $\overline{\Theta_t}$ is the complementary version of Θ_t (e.g. $\overline{\Theta_t} = 255 - \Theta_t$ for images coded on 255 gray level) and k is a constant used to set the range value of α. α is then locally adjusted in such a way that the forgetting filters use long term memory in poor spatiotemporal activity areas, and short term memory in areas with high spatiotemporal activity. Thus, the detection of slow, small, or low-contrast moving objects will be enhanced, while large moving objects with high contrast will be better segmented.

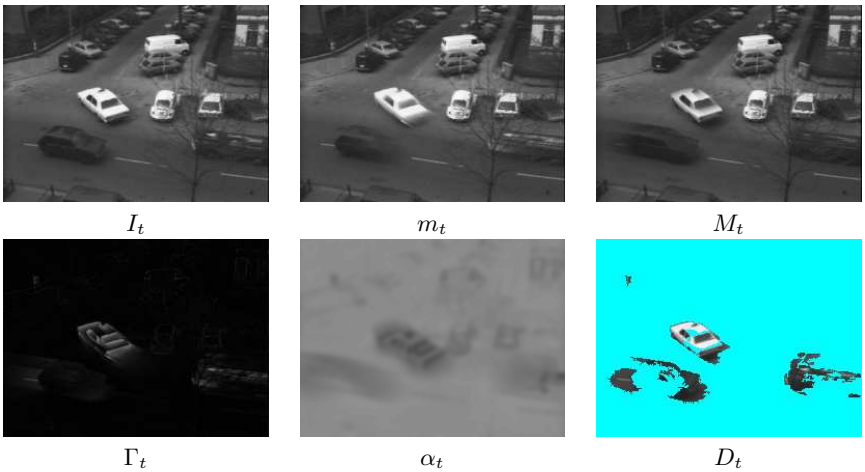

Figure 3. Results for the Hamburg Taxi sequence (frame n. 20). The parameters used are $N = 2$ (number of deviations) for S_t, $\sigma = 2.5$ for the standard deviation of the Gaussian filter used to compute Θ_t, $k = 140$ as the constant used in the computation of α_t.

Figure 3 shows the motion detection algorithm steps applied on the Hamburg Taxi sequence. The last image in the figure represents D_t, the detection result after removal of the smallest regions using an alternated filter by reconstruction with a ball of radius 1 [12]. It can be seen that the small amplitude motion like the pedestrian on the top left, as well as low-contrast moving object like the dark car on the bottom left are well detected while the high contrasted taxi at the center is better segmented. This is an effect of the adaptable memory

of the forgetting filters, in order to fit their detecting ability to the amount of motion.

4. Conclusion

We have presented a new hybrid differential filter and shown how it can be used in a motion detection algorithm with a local adaptation of the forgetting terms. Like the recursive filters, the computation time and memory consumption does not depend on the size of the temporal window. At the present time, we are investigating more sophisticated spatial interaction, in order to improve the spatiotemporal adaptivity, and quantify precisely the validity range of the algorithm.

References

[1] V. Agnus, C. Ronse, and F. Heitz. Spatio-temporal segmentation using morphological tools. In *15th ICPR*, pages 885–888, Barcelona, Spain, September 2000.

[2] S.S. Beauchemin and J.L. Barron. The computation of optical flow. In *ACM Computing Surveys*, volume 27(3), pages 434–467, September 1995.

[3] P. Bouthemy and P. Lalande. Recovery of moving objects masks in an image sequence using local spatiotemporal contextual information. *Optical Engineering*, 32(6):1205–1212, June 1993.

[4] S-C.S. Cheung and C. Kamath. Robust techniques for background subtraction in urban traffic video. In *IEEE Int. Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, Nice, France, 2003.

[5] E. Decenciere Ferrandiere, S. Marshall, and J. Serra. Application of the morphological geodesic reconstruction to image sequence analysis. *IEE Proceedings - Vision, Image and Signal Processing*, 144(6):339–344, December 1997.

[6] A. Manzanera and J. Richefeu. A robust and computationally efficient motion detection algorithm based on Σ-Δ background estimation. In *ICVGIP*, Kolkata, India, 2004.

[7] M. Pic, L. Berthouze, and T. Kurita. Active background estimation: Computing a pixel-wise learning rate from local confidence and global correlation values. *IEICE trans. Inf. & Syst.*, E87-D(1):1–7, January 2004.

[8] M. Piccardi. Background subtraction techniques: a review. In *Proc. IEEE Conference on Computer*, 2004. http://www-staff.it.uts.edu.au/∼massimo.

[9] F. Ranchin and F. Dibos. Moving objects segmentation using optical flow estimation. Technical report, UPD Ceremade, December 2003. http://www.ceremade.dauphine.fr/CMD/preprints03/0343.pdf.

[10] P. Salembier, A. Oliveras, and L. Garrido. Anti-extensive connected operators for image and sequence processing. *IEEE trans. on Image Processing*, 7(4):555–580, April 1998.

[11] K. Toyoma, J. Krumm, B. Brumitt, and B. Meyers. Wallflower: principles and practice of background maintenance. In *ICCV*, pages 255–261, Kerkyra, Greece, 1999.

[12] L. Vincent. Morphological grayscale reconstruction in image analysis: applications and efficient algorithms. *IEEE trans. on Image Analysis*, 2(2):176–201, April 1993.