IMPROVED OBJECT TRACKING WITH CAMSHIFT ALGORITHM

OULD-DRIS Nouar[*], *GANOUN Ali*[*,**] *and CANALS Raphaël*[*]

[*]Laboratory of Electronics, Signals and Images (LESI), University of Orléans, FRANCE
[**] Faculty of Engineering, EE Department, University of GARYOUNIS, LIBYA

Nouar.Ould-Driss@univ-orleans.fr, Ali.Ganoun @univ-orleans.fr, Raphael.Canals @univ-orleans.fr

ABSTRACT

In this paper, we present a new object tracking approach based on the analysis of a two-dimensional image distribution histogram calculated from two colorimetric channels automatically selected on criteria of representativeness. Among the essential contributions of this work, we can quote a better modelling of the object to track and the management of the target appearance changes during the sequence. Our approach is a prolongation of the CamShift algorithm applications (Continuously Adaptive MeanShift) in order to track object presenting strong modifications of shape and luminosity.

1. INTRODUCTION

In the literature, many methods have been developed to solve the object tracking problem. Numerous approaches are based on the visual primitives detected and tracked in images by employing correlation. Other techniques process the edge or region data, or the movement of the object in order to track it in the images sequence [01], [02], [03].

Another robust and nonparametric technique is proposed in the library of computer vision "Intel Corporation, 2001" [04]: it implements the CamShift algorithm which uses a one-dimensional histogram to track an object with known hue in colour images sequences. The difficulty emerges when one wishes to employ this algorithm to track objects without a priori knowledge nor training phase. In [05], it was then stated that the use of a three-dimensional histogram solves the problem and leads to a target localization improvement. The histogram back-projection permits to obtain a probability distribution image which is processed by the iterative CamShift algorithm in order to find the maximum of the distribution and thus the object centre, its dimensions and attitude [06], [07], [08].

The approach we present in this work is based on this algorithm. After this introduction, section 2 briefly presents the CamShift algorithm such as it was stated in the literature. Our new tracking system is described in section 3 and the obtained results are presented in section 4. We finish by a conclusion and propose some orientations for future research.

2. THE CAMSHIFT ALGORITHM

The principle of the CamShift algorithm is given in [04] and [08]. Each image of the sequence is converted into a probability distribution image relatively to the histogram of the object to be tracked. From this image, the centre and the size of the object are given thanks to the CamShift algorithm. These new centre and size are employed to place the search window in the next image. This process is then repeated for a continuous target tracking in the video sequence.

The algorithm of CamShift thus employs a 2D probability distribution image produced from a back-projection of the target histogram with the image to process.

The CamShift algorithm calls upon the MeanShift one to calculate the target centre in the probability distribution image [9]. It is a matter of finding a rectangle presenting the same moments as those measured on the probability image. These parameters are given from the first and second moments [04], [05].

3. THE TRACKING SYSTEM

The global tracking system we have implemented is primarily based on the CamShift method. The principal steps of this algorithm are stated as follows:

1. Determination of the interest region of the target R_t in current image I_t.
2. Automatic selection of two colorimetric channels.
3. In each region constituting the target, calculation and affectation of the average colour to all the corresponding pixels.
4. The two-dimensional histogram is calculated with the colorimetric channels selected in 2.
5. Back-projection of this histogram with the image I_{t+1} in order to obtain the probability distribution image.

6. Application of the MeanShift algorithm on this image to determine the new target centre in image I_{t+1}.
7. Determination of regions constituting the target R_{t+1} by using the same averages calculated in 3.
8. To take into consideration changes of the object, we may need to go to step 2, otherwise go to step 4.

3.1. Interest region of the target

In the majority of works proposed on tracking methods in general and on the CamShift method in particular [05], [10], the definition of the object to be tracked is manually carried out by the operator thanks to a rectangle or an ellipse around the target in image I_t. A convex mask is then used to reduce the spectral influence of the pixels far from the centre which are less reliable than those located near the centre.

In this work we propose that the area corresponding to the target is improved by edge detection for a better object modelling. A Canny filter [11] is used for that in the selected rectangle. Only the external boundaries are preserved and an interconnection is carried out to obtain a single closed contour. This method permits to eliminate the background often included in the object modelling in order to achieve a better model of the target.

3.2. Selection of two colorimetric channels

Because the management of a 3D histogram and its back-projection with a colour image are very time-consuming, we have decided to work with a 2D histogram. Its two automatically-chosen colorimetric channels must then constitute as well as possible our tracking criteria:

- The first one gives the best representation of the target;
- The second one gives the least satisfactory representation of the background.

We suppose that the most representative channel of the target is that for which the corresponding histogram has the smallest number of local minima. Indeed, less there are regions in the target, plus this one is represented by large and significant regions.

We seek now the least representative channel in the image without taking into account the content of the target. This choice enables us to ignore objects which can present the same properties as the target if we would select the second most representative channel of the target R_t. To do that, we calculate simply the average of the background pixel values for each colorimetric channel and we keep the channel presenting the smallest average.

3.3. Segmentation of the target

Segmentation permits to manipulate the colour average of each region instead of the values of each pixel, thus gaining in simplicity and robustness against noise and colour variations. Inspired by [12], our target segmentation method exploits the one-dimensional histograms which characterize the distribution of each colour channel in order to determine local minima defining the edges of the different regions, while avoiding an over-segmentation. Thus the 2D histogram constituting the target model is limited at some peaks.

Region data issued of this segmentation will be also applied in the search window in the next images, unless we wish to start a new segmentation to update the model.

3.4. Generation of the probability distribution image

The search window is defined around the target and is larger than the target window, increased by a distance *d* generally ranging between 10 and 20 pixels.

The back-projection of the 2D histogram with the part of the image I_{t+1} contained in the search window gives a probability distribution image.

The use of the region averages permits to filter the probability image without influencing the probability of the object itself, while eliminating the small non-significant regions in the search window.

The probability distribution image being available, we can now calculate the new position of the object thanks to the MeanShift algorithm.

3.5. Application of the MeanShift algorithm

The search window is initially centred at the position of the object in image I_t and the mass centre of the distribution is calculated. If this one is different from the window centre, then the search window moves towards this mass centre. We thus repeat the operation until the mass centre of the distribution in the window is identical to the window centre. The procedure is stopped if we obtain a minimal variation between the new window position and the preceding one; we can also specify a maximum number of iterations.

Dimensions and attitude of the target could be calculated here, permitting to adjust its size in case of appearance modifications. But this advantage becomes a problem when a near region of background presents appearance similarities, including thus this region in the target. So a fixed-size target is used, and a fixed-size search window too.

4. RESULTS

We have applied our method on various video sequences. The 576x720 images in figures 4 and 5 illustrate some obtained results.

In Fig.4, we track a woman in a video sequence. In spite of some appearance modifications, the algorithm tracks the woman as she moves from one frame to the next one in approximately 1.6 second on a 2GHz PC in Matlab.

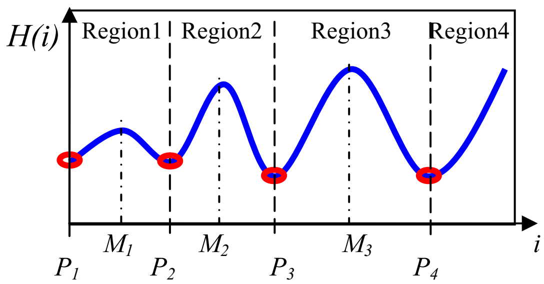

Fig.1: Histogram of one colorimetric channel and definition of the different regions.

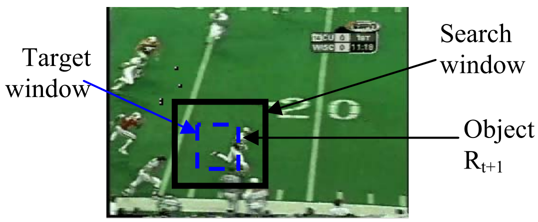

Fig.2: Search window in I_{t+1}.

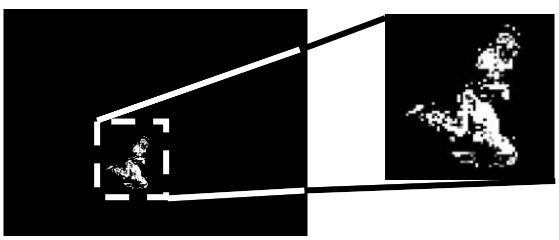

Fig.3: Probability distribution image.

In the same way, the algorithm succeeds in the tracking when the woman walks behind a road sign, although the window centre is momentarily ill-positioned because a part of the target is masked and the mass centre is then shifted. By way of comparison, tracking results are obtained with the basic method using a 3D histogram in about 4 minutes. Moreover these results are slightly degraded because the target model includes sometimes background pixels, according to the target shape.

Fig.5 demonstrates the tracking of a woman in an images sequence while managing occlusions with a cyclist, a pedestrian and two cars. The algorithm takes about 1.8 second to track correctly the woman. Although these occlusions, our algorithm successes in tracking the woman, even until the last images where the occlusion is high. To note the same phenomenon of mass centre shift in images presenting occlusions.

5. CONCLUSION

We have proposed in this paper a new object tracking approach in colour images sequences, based at the same time on the CamShift algorithm and on the modelling of the target to be tracked. In our approach, the model is built on two colour channels: the first one gives the best representation of the target and the second one the least acceptable description of the background.

The use of a 2D histogram instead of the most commonly used 3D one, coupled with a target edge detection and a target segmentation in order to improve the modelling while simplifying it, allows to gain greatly in computational time.

The choice of a fixed-size target limits the managing of appearance modifications and occlusions, but it prevents to consider some similar regions of the background as parts of the target. This problem could be solved by decomposing the object in sub-objects to be tracked separately in variable-size search windows and depending on each other under a group model. That should permit to manage easily deformations and occlusions of the target in a practical and effective manner.

Computationally attractive experimental results validate our approach and show its efficiency in tracking an object quickly. This work is currently in progress on an embedded DSP board equipping a video-surveillance system.

REFERENCES

[01] V. Nguyen and Y. Tan, “Fast Block-Based Motion Estimation Using Integral Frames”, *IEEE Signal Processing Letters*, Vol. 11, N° 9, September 2004.

[02] S. Besson, M. Barlaud, and G. Aubert, “A 3-Step Algorithm Using Region-Based Active Contours For Video Objects Detection”, *EURASIP Journal on Applied Signal Processing*, pp. 572-581, Issue 6, 2002.

[03] M.Irani, and P. Anandan, “A Unified Approach to Moving Object Detection in 2D and 3D Scenes”, *IEEE Trans. on PAMI*, Vol. 20, pp. 577-589, June 1998.

[04] Intel Corporation, “Open Source Computer Vision Library Reference Manual”, 123456-001, 2001.

[05] G. John Allen, Y. D. Richard Xu and S. Jin Jesse, “Object Tracking Using CamShift Algorithm and Multiple Quantized Feature Spaces”, *Inc. Australian Computer Society*, vol.36, 2004.

[06] B. K. P. Horn, “Robot vision”, *MIT Press*, 1986.

[07] W. T., Freeman, K. Tanaka, J. Ohta, and K. Kyuma, “Computer Vision for Computer Games”, *Int. Conf. on Automatic Face and Gesture Recognition*, pp.100-105, 1996.

[08] G. R. Bradski. “Computer vision face tracking for use in a perceptual user interface”, *Intel Technology Journal*, 2nd Quarter, 1998.

[09] Y. Cheng, “MeanShift, mode seeking, and clustering”, *IEEE Trans. on PAMI*, vol.17, pp.790-799, 1995.

[10] D. Comaniciu, V. Ramesh, and P. Meer “Kernel-Based Object Tracking”, *IEEE Trans. on PAMI*, vol. 25, pp. 564-575, 2003.

[11] J. Canny, “A computational approach to edge detection”, *IEEE Trans. on PAMI*, pp. 679–698, 1986.

[12] R. Ohlander, K. Price, et D. Reddy, “Picture segmentation using a recursive region splitting method”, Computer Graphics and Image Processing, pp. 313-333, 1978.

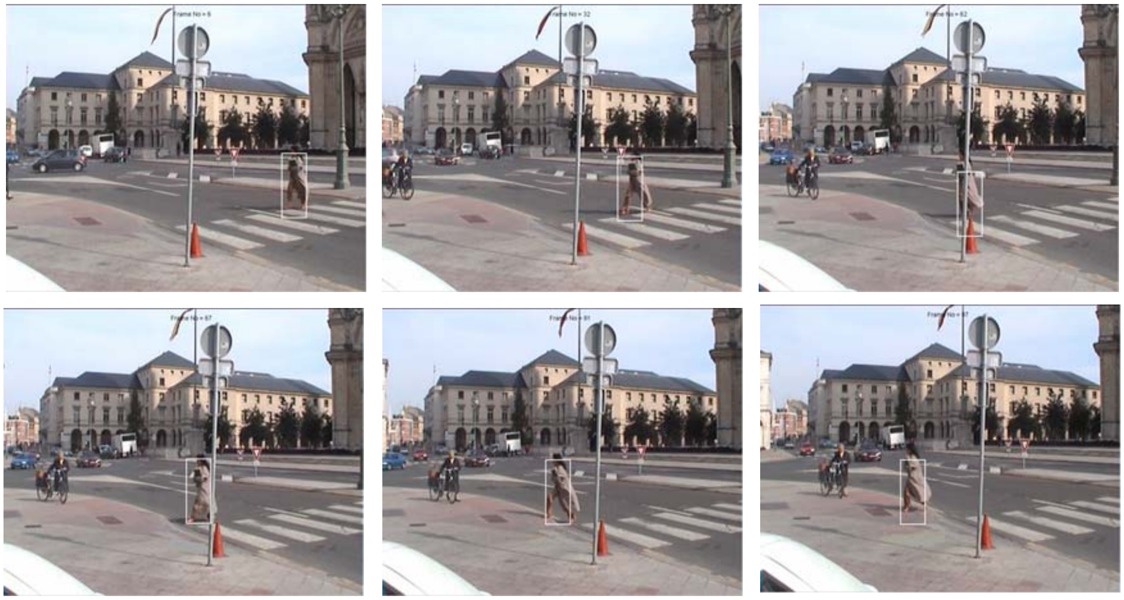

Fig.4: Tracking of a woman in presence of occlusions.
Images 6, 32, 62, 67, 81 and 97 of the sequence.

Fig.5: Tracking of a second woman in presence of several occlusions.
Images 1, 36, 38, 44, 66, 93, 104, 152 and 175 of the sequence.