

Non-Rigid Object Tracker Based On a Robust Combination of Parametric Active Contour and Point Distribution Model

Joanna Isabelle Olszewska^a, Tom Mathes^b, Christophe De Vleeschouwer^a, Justus Piater^b and Benoit Macq^a

^aUniversité catholique de Louvain (UCL), Communications Lab., B-1348 Louvain-la-Neuve

^bUniversité de Liège (ULg), Institut Montefiore, B28, B-4000 Liège

ABSTRACT

Our study considers the development of a reliable tracker for non-rigid objects evolving on cluttered background in crowded scenes captured by moving cameras. For this purpose, we propose an original method that combines two approaches, respectively based on parametric active contours (PAC) and on point distribution model (PDM). The PAC tracker relies on an effective and efficient implementation of contour convergence mechanism to bring a smooth contour to the edges of the target in real-time. The PDM approach collects feature points in the region delineated by the PAC tracker to build and update a model of the target in term of a feature point distribution. Formally, when a novel frame is considered, its feature points are matched with the PDM model. The matching information is used to initialize the novel PAC, whose convergence identify the points that are relevant to update the PDM for the next frame. Hence, the two approaches complement each others. The a priori information provided by the PDM makes the system robust towards occlusions, while the deformation of the PAC increases its robustness towards target appearance changes. Simulations on real-world video sequences demonstrate the performance of our approach.

Keywords: Tracking, Parametric Active Contours, B-Splines, Gradient Vector Flow, Point Distribution Model

1. INTRODUCTION

In the recent years, interest on object tracking has shown up in applications like medical imaging, video surveillance or sport event analysis. In such context, the main difficulties to face with are (i) real-time tracking, (ii) non-rigid targets, (iii) occluded situations, (iv) no assumptions on the background, (v) non-static cameras, (vi) scale changes with perspective distortion, zoom effects or blurring.

For this purpose, conventional tracking methods based on a single primitive like contour or interest points, do not often meet the entire set of requirements. For example, the active contour based method presented in section 2, can deal with most of our specifications except in case of wide target occlusions. In contrast, the interest point approach presented in section 3, supports occlusions but fails in front of target appearance changes.

Hence in difficult situations, combining several types of primitive like active contours and feature points could fill in the lack of the individual methods, as shown in recent papers^{1,2}. These works predict the target trajectory, based on the matching of feature points between consecutive frames.

As two important contributions, our PAC/PDM combination, presented in section 4, proposes both an original way to drive the initialization of the PAC based on feature points matching, and a dynamic and adaptive update of the underlying target PDM. On the one hand, when a novel frame is considered, PAC initialization relies on the matching of the feature points extracted in the novel frame with the points handled by PDM. On the other hand, the target PDM is updated based on the feedback provided by the PAC convergence mechanism, i.e. only points that are inside the contour are considered to update the PDM.

Tracking results of the proposed system are presented in section 5.

Further author information: (Send correspondence to J.O.)

J.O., C.D.V. and B.M.: E-mail: {olszewska, devlees, macq}@tele.ucl.ac.be, Telephone: +32 10 47 22 71

T.M, J.P.: E-mail: Justus.Piater@ULg.ac.be, Telephone: +32 4 366 2279

2. PARAMETRIC ACTIVE CONTOUR METHOD

Active contours are deformable models whose shape evolution is controlled by internal forces and external forces using energy minimization.³

For a parameterized active contour $\mathbf{r}(s) = [x(s), y(s)]$, the energy functional can be written as

$$E_{PAC} = \int_0^1 E_{int}(\mathbf{r}(s))ds + \int_0^1 E_{ext}(\mathbf{r}(s))ds \quad (1)$$

The internal force defines its physical properties like elasticity (α) or rigidity (β) and acts as a smoothness constraint.

$$E_{int} = \frac{(\alpha |\mathbf{r}_s(s)|^2 + \beta |\mathbf{r}_{ss}(s)|^2)}{2} \quad (2)$$

The external force deforms the contour from the initial position to the feature of interest such as edges, in an image.⁴ Since they are deformable and intrinsically smooth, the active contours offer a highly capable tracker for non-rigid objects.

The user's interaction is limited to the initialization of the target in the first frame of a sequence, illustrated in Figure 1 (a). Thus for the next frames, rather than explicitly incorporating a temporal energy term in Eq. 1 like in Bouaynaya and Schonfeld,⁵ the contour is defined as follows. When the PAC tracker is used alone, without PDM, the initial contour is set to be equal to the contour computed on the previous frame. Once combined with the PDM, the PAC tracker is initialized on the PDM-based matching or on the previous PAC, according to the obtained PDM match confidence (see Section 4).

In a novel incoming frame, given an initial contour, the PAC tracker evolves in three steps.

- First, B-splines⁶ are used to define a smoothed contour based on a limited number of sampling points.
- Second, an image gradient is computed⁷ and diffused to define the external force applied to the contour, minimizing the following energy functional⁸

$$E_{ext} = \int \int \mu(u_x^2 + u_y^2 + v_x^2 + v_y^2) + |\nabla f|^2 |\mathbf{v} - \nabla f|^2 dx dy \quad (3)$$

where $\mathbf{v}(x, y) = [u(x, y), v(x, y)]$ is the Gradient Vector Flow (GVF) vector field, f the edge map and μ the diffusion parameter, controlling the tradeoff between the first and the second term.

- Third, the active contour is evolving, based on the given internal (rigidity, elasticity) and external forces (Gradient Vector Flow).

Our methodology and implementation choices are meeting the real-time constraint.

Regarding the PAC tracker, our main contribution consists in combining smoothed contour description (B-splines) with a bidirectional contour convergence force (GVF) in a coherent tracking framework. In this way, the coupling of these two components ensures a precise contour tracking without introducing any shape energy term⁵ into Eq. 1.

Moreover, our PAC tracker enables a real-time implementation by decreasing evolution process computational complexity, while surrounding the target object accurately. In contrast, Park et al.² achieves accurate description at heavy computational cost using watershed segmentation, while Gouet and Lameyre¹ uses a unidirectional contour evolution and coarse contour representation to reach real-time performances.

The designed PAC tracker deals efficiently with non-rigid object tracking, as shown in Figure 1 (b), even in presence of unexpected target behaviors, illumination changes and camera movements resulting in non-static background or scaling effects. However, active contour suffers in presence of severe occlusions. In Section 4, we



(a) Contour initialization



(b) Contour convergence

Figure 1. Soccer player tracking using only the PAC tracker in a frame

explain how to circumvent that limitation based on the use of a learned appearance model of the object to track. We now explain how to formalize that a priori knowledge based on point distribution models.

3. POINT DISTRIBUTION MODEL

The point distribution model (PDM) considered in this work, uses simple feature vectors to describe object appearance as described in Mathes et Piater,⁹ rather than raw texture information like in Cootes et al.¹⁰ Local features are extracted by a Harris corner detector applied to an initial region-of-interest as in Figure 2 (a). The sparse sets of features are well-suited for non-rigid objects and tend to yield methods particularly robust to partial occlusions.

In the first frame of a sequence, the PDM tracking algorithm is initialized manually. Thus for the next frames, the region-of-interest is defined as follows. When the PDM tracker is running without PAC, the initial region-of-interest is defined by the smallest rectangle enclosing the model shape from the current frame plus a small border, considering target displacement. Once combined with the PAC, the PDM tracker is initialized based on the PAC obtained in the previous frame (see Section 4).

For each detected interest point in the region-of-interest as illustrated in Figure 2 (a), a feature vector \mathbf{a} describes its local appearance¹¹ and corresponds to the first-order local jet, which approximates the point neighborhood by a set of image derivatives, enhanced by the interest point position.

$$\mathbf{a} = (x, y, r, g, b, r_x, r_y, g_x, g_y, b_x, b_y)^T \quad (4)$$

where $\mathbf{a} \in \mathbf{A}$ with \mathbf{A} called feature space and $\mathbf{A} \subset \mathbf{R}^{11}$.

The feature vectors of a particular frame are concatenated into a shape vector.

$$\mathbf{x} = (\mathbf{a}_x^1{}^T, \mathbf{a}_x^2{}^T, \dots, \mathbf{a}_x^M{}^T)^T \quad (5)$$

A collection of these shapes generates a point distribution model, which represents the current model of the object, as shown in Figure 2 (b). The PDM combines thus local appearance information with global shape information.

The points from these learned shapes are matched to current image points using the *Hungarian method*.¹² These matchings are then used to attract the shape model to the new image measurement.



(a) Current interest points (in white)



(b) Model points (in red)

Figure 2. Soccer player tracking using only the PDM in a frame

In practice, point features tend to flicker in noisy image sequences or disappear due to out-of-plane rotations or occlusions, but as long as a reasonable subset of all the points is visible in each frame, tracking can be performed reliably. The missing points are predicted by the model.

To enable accommodation to object changes, the model has to be learnt incrementally and continuously. The adopted method has the capability to dynamically add novel relevant points and remove obsolete ones. However, deciding about the relevance or non-relevance of a point extracted in the current frame is not obvious. In Section 4, we explain how the information provided by the PAC tracker supports that decision.

4. COMBINED TRACKER

The combined tracker is composed of two trackers running in parallel and a control level on top of them. The control level enables an exchange of information between these two individual trackers. The essence of the combination is to take basically the strong points of the PAC and PDM trackers, relying on their complementary advantages. The resulting new tracker is more robust, as it survives even if one of the trackers fails.

Principle After feature vector computation and shape vector generation as explained in Section 3, the PDM is obtained like in Figure 3 (a). Next, the PAC is initialized based on the PDM as shown in Figure 3 (b). After convergence, the PAC system delineates accurately the region of interest wherein the target object is lying (see Figure 3 (c)). Then, at the control level, the result provided by the PAC tracker feeds the PDM tracker by providing it with an information about the confidence to give to the feature points extracted in the current image, based on whether or not these points belong to the region defined by the contour. Hence, PAC convergence process supports the PDM updates as illustrated in Figure 3 (d).

Algorithm Information exchange between the two trackers is performed by applying the following algorithm to each frame i and given an initial region-of-interest (ROI) \mathcal{A}_{INIT} .

1. Extraction of interest points that lie into the region-of-interest, defined by
 - \mathcal{A}_{INIT} , for the first frame;
 - an enhanced rectangle enclosing the model shape from the previous frame \mathcal{A}_{PDM}^{i-1} , in occluded situations;



(a) PDM computation



(b) PDM-based-on PAC initialization



(c) PAC convergence



(d) PAC-based-on PDM update

Figure 3. Illustration of the PAC-PDM combination for one frame

- the active contour from the previous frame \mathcal{A}_{PAC}^{i-1} , in all other cases.
2. Matching every model point with the best nearby interest point extracted in the step 1 and obtaining only the matched feature vectors.
 3. Computing the nearest neighbours of these features vectors and obtaining new feature vectors.
 4. First update of the PDM, by adding new relevant feature vectors found in step 3, leading to the most probable target object shape.
 5.
 - If $M \geq 3$ then initialization of the PAC based on the point model. The region-of-interest \mathcal{A}_{PDM}^i is defined by an enhanced rectangle enclosing the model shape from the current frame.
 - Else initialization of the PAC based on the contour founded in the previous frame $\mathcal{A}_{PAC}^{(i-1)}$.
 6. Convergence of the PAC and determination of the precise and accurate region-of-interest \mathcal{A}_{PAC}^i
 7. Second update of the PDM, by removing the feature vectors corresponding to the interest points not lying inside the \mathcal{A}_{PAC}^i

Occlusions In case of partial occlusions or cluttered background, the PAC tracker could be unstable. In that case, the a priori model defined by the PDM tracker is helpful. Indeed, after occlusion or within the cluttered background, matching the feature points extracted on the current frame with the PDM features provides

information about feature point displacement. Thus, at the level control, the a priori information provided by the PDM tracker is used to keep the whole system on track and to initialize the current frame contour, in case of occlusions or cluttered background.

In this work, occlusions are detected when the region-of-interests of two different objects are intersecting each other. This approach is more robust to appearance changes rather than considering as an occluded situation, a case of poor matches between the interest points extracted in consecutive frames like Gouet and Lameyre work.¹

Moreover to increase the robustness of the tracker, the PDM update is disabled during occlusions. Indeed, no feature vectors are added into the model or removed. On the other hand, the multiple possibilities of tracker initialization in the algorithm prevent the whole tracking system drifting. In fact, the combined tracker can still operate even if one of the tracker has failed, and the latter one can be properly reinitialized.

Hence, PAC and PDM complements each others. Indeed, the PAC tracker does not include any prior knowledge, to preserve flexibility in cases of unexpected target deformation. In contrast, the PDM tracker relies on a priori models of the object, to support occlusions and cluttered backgrounds. Moreover, for sports applications like soccer, PDM gives only the player’s center of gravity while the PAC provides additional information like player feet’s position.

5. RESULTS

We have tested all the presented trackers on soccer and video-surveillance sequences. In Figure 5, we present results of our combined tracker on a 1000 frame soccer sequence.

The PAC tracker works successfully in situations where target appearance changes (Figure 5, frames 202, 506, 655, 891, 926), for example due to player fall, out-of-plane rotation or severe deformations.

On the other hand, the PDM approach is robust to partial occlusions because only a subset of the object points need to be visible in order to constrain the model parameters. Thus the combined tracker deal well with such situations as illustrated in Figure 5 frame 702, where a second player is crossing the target player.

Figure 4 illustrates for some particular frames of the sequence, tracking before, during and after a full occlusion.

When an occlusion is occurring like in Figure 4 (a), the PAC tracker starts to be initialized based on the PDM as explained in the algorithm of Section 4, without updating the model. Moreover, no points are added into the PDM. For the case of full occlusion as shown in Figure 4 (b), the combined tracker could successfully track the target object by incorporating into the PDM tracker, a Kalman filter to predict the target object trajectory. After the full occlusion, only matches between the current extracted points and the model points initialize the PAC tracker, until the occluded situation is entirely finished like in Figure 4 (c). In this case, the PAC provides again information to update the PDM and new shapes could be learned.

The combined tracker could deal with situations such like occlusions, out-of-plane rotations, illumination changes.

6. CONCLUSIONS

The paper proposes a new way to combine the active contour and interest point tracking approaches. The information exchange like PAC convergence results or points matching information, between the trackers increases robustness of the combined tracker. Thus, the combined tracker can still operate even if one of the tracker has failed, and the latter one can be properly reinitialized. The method has been validated for usual situations

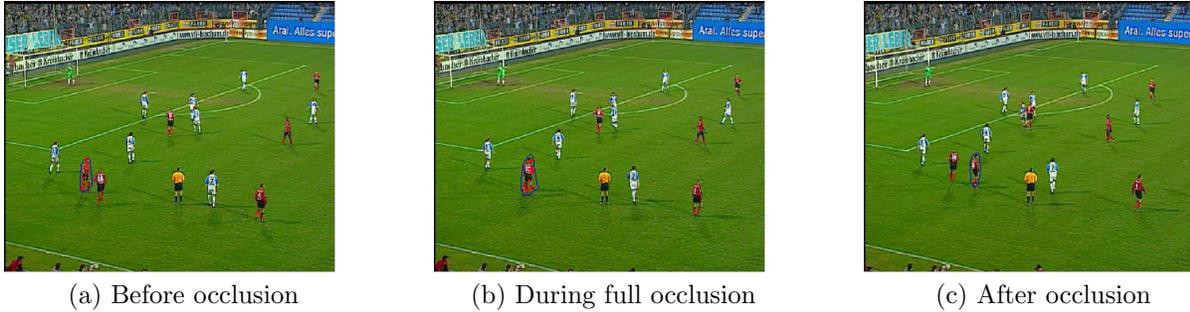


Figure 4. Soccer player tracking in a full occlusion case

occurring in non-rigid object tracking as well as challenging ones like occlusions. The developed algorithms are robust and amenable to real-time implementation.

ACKNOWLEDGMENTS

The authors are grateful for the support of the Région Wallonne (DGTRE, Belgium) and the Belgian NSF.

REFERENCES

1. V. Gouet and B. Lameyre, "SAP: A robust approach to track objects in video streams with snakes and points," in *Proceedings of the British Machine Vision Conference (BMVC'04)*, pp. 737–746.
2. S.-C. Park, S.-H. Lim, B.-K. Sin, and S.-W. Lee, "Tracking non-rigid objects using probabilistic Hausdorff distance matching," *Pattern Recognition* **38**, pp. 2373–2384, December 2005.
3. M. Kass, A. Witkin, and D. Terzopoulos., "Snakes: Active contour models," *International Journal of Computer Vision* **1**, pp. 321–331, January 1988.
4. A. Blake and M. Isard, *Active Contours*, Springer-Verlag, 1998.
5. N. Bouaynaya and D. Schonfeld, "Active surfaces for video tracking and 3-D segmentation based on a new method for multidimensional optimization," in *Proceedings of the IS&T/SPIE Conference on Visual Communications and Image Processing (VCIP'06)*, **6077**.
6. M. Unser, "Splines: A perfect fit for signal and image processing," *IEEE Signal Processing Magazine* **16**, pp. 22–38, Nov. 1999.
7. D. Forsyth and J. Ponce, *Computer Vision, A Modern Approach*, Prentice Hall, 2003.
8. C. Xu and J. Prince, "Snakes, shapes and Gradient Vector Flow," *IEEE Transactions on Image Processing* **7**, pp. 359–369, March 1998.
9. T. Mathes and J. Piater, "Robust non-rigid object tracking using point distribution models," in *Proceedings of British Machine Vision Conference (BMVC'05)*, pp. 849–858.
10. T. Cootes, C. Taylor, D. Cooper, and J. Graham, "Active shape models - Their training and application," *Computer Vision and Image Understanding* **61**, pp. 38–59, January 1995.
11. J. Koenderink and A. van Doorn, "Representation of local geometry in the visual system," *Biological Cybernetics* **55**(6), pp. 367–375, 1987.
12. H. Kuhn, "The Hungarian method for solving the assignment problem," *Naval Research Logistics Quarterly* **2**, pp. 83–97, 1955.



frame202



frame 506



frame 655



frame 702



frame 891



frame 926

Figure 5. Soccer player tracking during 1000 frames