Adaptive Feature Selection for Object Tracking with Particle Filter

Darshan Venkatrayappa[1](✉), Désiré Sidibé[2],
Fabrice Meriaudeau[2], and Philippe Montesinos[1]

[1] Ecole des Mines d'Ales, LGI2P, Parc Scientifique Georges Besses, 30035 Nimes, France
{darshan.venkatrayappa,philippe.montesinos}@mines-ales.fr
[2] Laboratoire Le2i, 12 rue de la Fonderie, 71200 Le Creusot, France
{dro-desire.sidibe,fabrice.meriaudeau}@u-bourgogne.fr

Abstract. Object tracking is an important topic in the field of computer vision. Commonly used color-based trackers are based on a fixed set of color features such as RGB or HSV and, as a result, fail to adapt to changing illumination conditions and background clutter. These drawbacks can be overcome to an extent by using an adaptive framework which selects for each frame of a sequence the features that best discriminate the object from the background. In this paper, we use such an adaptive feature selection method embedded into a particle filter mechanism and show that our tracking method is robust to lighting changes and background distractions. Different experiments also show that the proposed method outperform other approaches.

Keywords: Tracking · Particle filter · Mean-shift filter · Feature selection

1 Introduction

Object tracking is a basic requirement in many applications related to video surveillance, robotics and interactive video games. Object trackers can be used to improve our understanding of large video sequences from medical and security applications. Tracking algorithms should be able to cope with the variation of the size of the target, orientation and pose of the target, reflectance of the target, illumination changes and background clutter. The success or failure of a tracking algorithm depends on how distinguishable an object is from the background. If the object is distinguishable from its background, then mean shift[1] or particle filters[10] can track the object successfully. On the other hand, if object and background are not distinguishable, then the tracker needs to handle the targets appearance changes as well as illumination changes and background clutter.

Many researchers have proposed different approaches to cope up with appearance change, illumination change and background clutter. In [2] visual appearance variations at a short time scale are represented as linear subspace of the

A. Campilho and M. Kamel (Eds.): ICIAR 2014, Part II, LNCS 8815, pp. 395–402, 2014.
DOI: 10.1007/978-3-319-11755-3_44

image space. Tracking algorithm updates this subspace on-line by finding a linear subspace that best approximates the observation made in the previous frames. Some methods use fixed features which are determined a priori depending on the application. A good example is the case of head tracking using skin color. Different color spaces are first evaluated and the one for which pixels values of skin clusters the most is used for tracking [3]. The authors of [13] use a new approach for face detection. This approach adaptively switches between a number of color space models as a function of the state of the environment, as well as, dynamically updates the corresponding color distribution model. In [8] the authors assemble multiple heterogeneous features then likelihood images are constructed for the various subspaces of the combined feature space, then the most discriminative feature is extracted by Principal Component Analysis (PCA) based on those likelihood images. They embed this feature selection mechanism in an mean shift tracker. In [9] the authors come up with an approach for evaluating multiple color histograms during object tracking. The method adaptively selects histograms that well distinguish foreground from background. The variance ratio is utilized to measure the separability of object and background and to extract top-ranked discriminative histograms. In [4], feature values from background patches and object observations are sampled during tracking and Fisher discriminant is used to rank the features based on sampled values. [5] combine saliency information with color features to make tracking more robust to changing illumination. [6] use mutual information to track multi-view objects in real time. They use the variance of mutual information to acquire reliable features for tracking by making use of the images of the tracked object in previous frames to refine the target model.

We address the problem of adaptive feature selection for real time tracking. As shown by Collins and Liu [7], the best features for tracking are those able to distinguish the object to be tracked from the background. The authors propose a framework for the on-line selection of the best combination of color values in each frame of a sequence. Following the same approach, we propose a particle filter based tracking algorithm which can cope with difficult conditions such as lighting variations and background distractions.

This paper is organised as follows. A brief overview of particle filtering based tracking is given in Section 2. In Section 3, the proposed method is described, explaining the adaptive feature selection method. Experimental results and discussion are shown in Section 4. Eventually, section 5 concludes this paper.

2 Particle Filter Based Tracking

A particle filter is a sequential Bayesian estimation technique, which recursively approximates the a posteriori distribution using a finite set of weighted samples $\{x_t^i, w_t^i\}_{i=1,...,N}$. Each sample x_t^i represents a hypothetical state of the target with a corresponding importance weight w_t^i. For tracking purpose, the state is defined as $\mathbf{x} = [x, y, s_x, s_y]^T$, where (x, y) is the center of the target and s_x and s_y are the scale of target window in the x and y directions. The particles, or samples, $\{x_t^i\}_{i=1,...,N}$ are propagated from frame t to frame $t+1$ using a dynamic model:

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + \mathbf{v}_t, \tag{1}$$

where $\mathbf{v}_t$ is a multivariate Gaussian random noise and A defines the deterministic system model. A constant velocity model is usually used for the dynamic model.

The weights are computed based on the similarity between each particle and a reference model. Finally, $E[x]$ the estimated state of the target in frame $t+1$ is obtained as the mean state of the system:

$$E[x] = \sum_{(i=1)}^{N} w_t^i x_t^i, \tag{2}$$

where x_t^i is the state of the i^{th} particle and w_t^i its weight.

3 Adaptive Feature Selection

The good performance of particle filter based tracking mainly depends on the features used to describe the target, i.e. the features used to compute the similarity measure and, hence, the weights of the particles. The basic color-based particle filter tracking algorithm [10] uses color information to represent the appearance of the target. This tracking method makes use of fixed color space such as RGB or HSV to represent the color histogram. The features used to construct the appearance model are fixed regardless of the tracking conditions. The lack of adaptation in color models leads to performance degradation when handling situations such as illumination changes or background distraction. This can be seen in the top row of figure 1 where the tracker fails to adapt to the changing illumination conditions in the sequence.

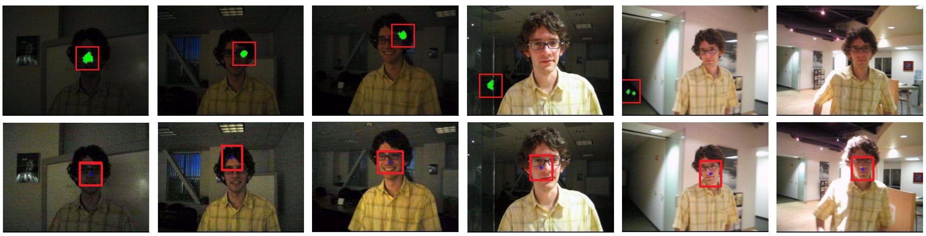

Fig. 1. Tracking results with the sequence David. Top Row: particle filter using a fixed RGB color model fails due to varying lighting conditions. Bottom Row: a particle filter with adaptive feature selection can accurately track the face over the entire sequence. Please note that this figure is best viewed in color.

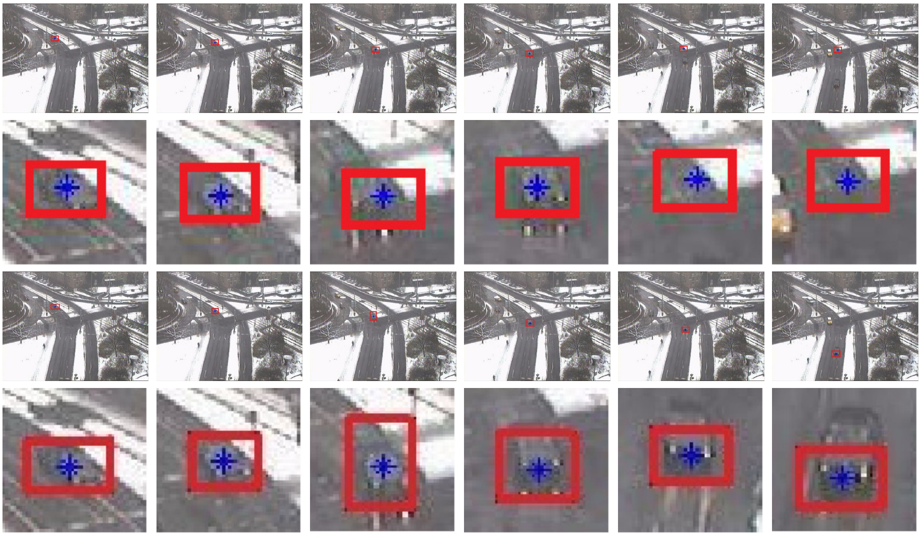

Fig. 2. Tracking results with the Winter sequence. Row 1: tracking failure with PF. Row 2: Enlarged version of the tracker window from Row 1. Row 3: tracking with PFFS. Row 4: Enlarged version of the tracker window from Row 3. Please note that this figure is best viewed in color.

3.1 Selecting the Best Features

The limitation of using fixed color features is the main motivation for an adaptive feature selection mechanism. The main idea of feature selection is to select for each frame of a sequence, the set of features that best discriminates the background and the object [7].

Different features can be used for tracking including color, shape or texture. Color distribution [10] is robust against noise and partial occlusions, but becomes ineffective in the presence of illumination changes, or when the background and the target have similar colors. Edges or contour features [11] are more robust to illumination variations, but are sensitive to clutter and are computationally expensive.

In this work, we focus on color features and use RGB color space, HSV color space and the transformed RGB space. The latter color space is based on the normalization of each channel independently [12]:

$$\begin{pmatrix} R' & G' & B' \end{pmatrix} = \begin{pmatrix} \frac{R-\mu_R}{\sigma_R} & \frac{G-\mu_G}{\sigma_G} & \frac{B-\mu_B}{\sigma_B} \end{pmatrix} \tag{3}$$

where σ is the standard deviation of the color channel and μ is its mean value.

Given the targets position in the frame, we generate two color distributions p_f and q_f for the object and the background, respectively, for each color feature f. The background area can be defined as the region surrounding the target location. The separability between the background and the foreground for feature f is given by the log-likelihood ratio computed as $L_f = \log(p_f/q_f)$. The log-likelihood ratio provides natural separability between the object and the background. Thus, thresholding L_f at zero is equivalent to classifying the object and the background using maximum likelihood rule.

In order to rank the different features, Collins and Liu [7] suggest the use of the two-class variance ratio of the likelihood function:

$$var(L; p, q) \equiv \frac{var(L; (p+q)/2)}{var(L; p) + var(L; q)}, \tag{4}$$

The variance ratio in equation 4 is large for the features f which clearly separate the object from the background. We can then use this measure to rank all our features and use to one with highest variance ratio value to track the object in the next frame. The main advantage here is that, depending on the conditions, different features will be selected in different frames of the sequence.

4 Experiments and Discussion

In this section, we evaluate the performance of the proposed tracking method using four video sequences acquired in different conditions. The conditions include indoor and outdoor scenes, moving objects and persons, difficult illumination changes, occlusion and background distraction (background similar in color with the target).

For a quantitative evaluation, we have manually generated the ground truth for all sequences and use the distance between the centres of the ground-truth window and the output of the tracker as a measure of performance. In all experiments the target is manually initialized in the first frame. We compare the performance of the proposed particle filter with feature selection method (PFFS) against the conventional particle filter (PF) and mean-shift (MS) using a fixed set of color features, and a mean-shift with feature selection method (MSFS).

All the 4 filters MS, MSFS, PF and PFFS are our own implementation. Both PF and PFFS use variable size windows. The number of particles used in both the type of particle filters is 250. Increasing the number of particles reduces the frame rate producing almost the same performance. We have experimented with correlation of histograms, intersection of histograms , χ^2 and Bhattacharyya coefficient as the similarity measure. In our experiments Bhattacharyya coefficient performed best for most of the sequences. All the implementation were carried out on Matlab platform.

The first sequence under consideration is the 'David' [14]. In this video sequence, we see a person slowly emerging out of a dark room to a brightly illuminated room with varying scale and out-of-plane pose changes. As he emerges out, he performs certain actions with his hand which partially occludes his face. The results obtained with the David sequence are shown in figure 1. As can be seen in Top row of figure 1, a simple particle filter (PF) using a fixed set of color features, in this case fixed RGB features, fail to correctly track the face when the illumination conditions are varying. This is the same in the case of mean shift filter (MS). On the other hand, the proposed method using an adaptive feature selection technique with particle filter (PFFS) performs extremely well in this difficult situation. The same observations apply for the results with the Winter sequence presented in figure 2.

A comparison of the different tracking algorithms for the David sequence is shown in figure 3. Both the particle filter (PF) and the mean-shift (MS) trackers lose the target before the end of the sequence. On the contrary, an adaptive feature selection approach makes the trackers to robustly follow the target despite illumination variations. We can also observe that the particle filter with feature selection (PFFS) performs better than mean-shift with feature selection (MSFS) as shown by the error in figure 3. Note that the target is said to be lost by a tracker if the error, i.e. the distance between the centres of the ground-truth window and the output of the tracker, is greater or equal to the size of the ground-truth window.

In the Browse While Waiting sequence from CAVIAR test data set, the object of interest is a person. In this sequence a person standing in the sunlight gradually moves towards a shadowed area and again moves back to the region with sunlight. For this sequence the three filters MS, PF and MSFS fail as the person moves from sunlight to the shadowed area. Where as the PFFS overcomes all the difficulties. The results are shown in figure 4.

The Fog sequence is a challenging one even for a human observer due to the foggy condition. For this sequence, the target car is hardly distinguishable from the background and all tracking algorithms fail since none of the color features used can discriminate the car from the background. This is illustrated in figure 5. The results for four sequences are summarized in table 1. For three

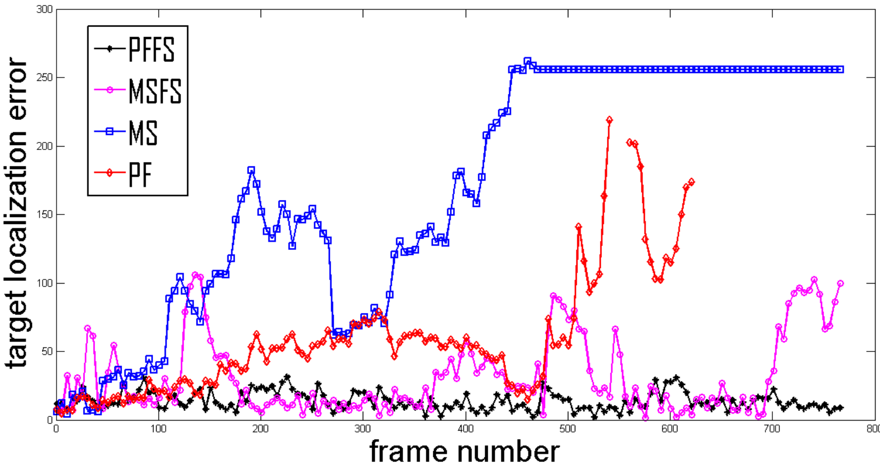

Fig. 3. Performances comparison for the sequence David

Table 1. Mean Error in pixels for different sequences using all the four filters. PF = particle filter; MS = mean-shift; MSFS = mean-shift with feature selection; PFFS = particle filter with feature selection; TF = tracking failure.

Sequence	MS	PF	MSFS	PFFS
David	TF	TF	31.41	13.46
Browse While Waiting	TF	TF	TF	14.48
Winter	TF	TF	16.64	6.78
Fog	TF	TF	TF	TF

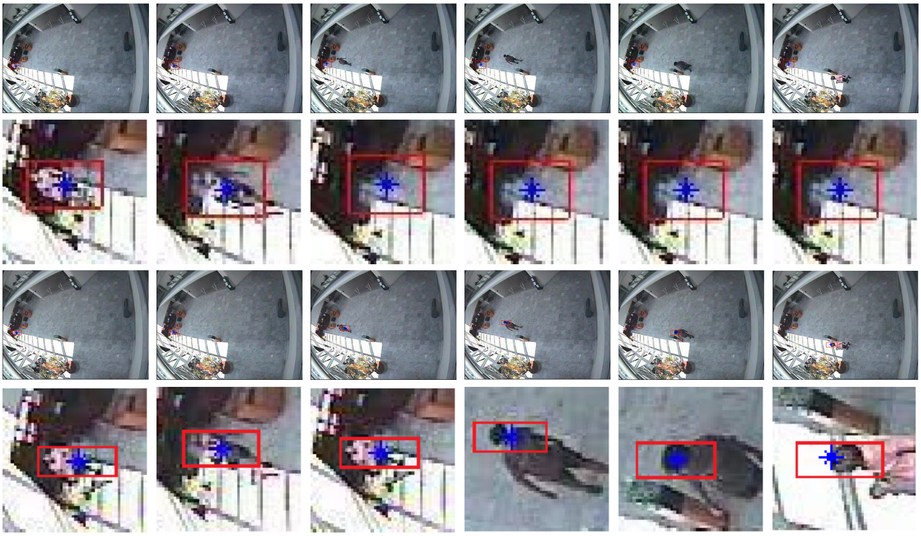

Fig. 4. Tracking results with the Browse While Waiting sequence. Row 1: tracking failure with PF. Row 2: Enlarged version of the tracker window from Row 1. Row 3: tracking with PFFS. Row 4: Enlarged version of the tracker window from Row 3. Please note that this figure is best viewed in color.

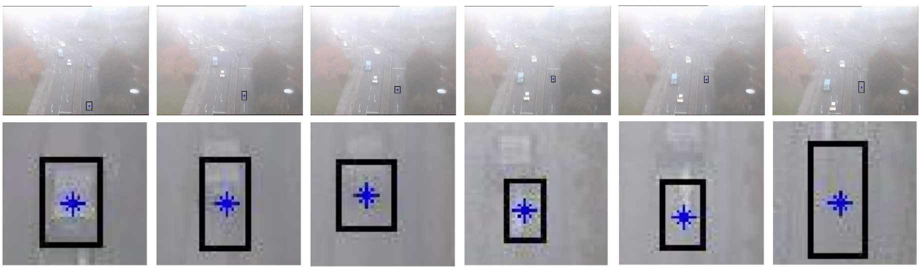

Fig. 5. 1st Row : Tracking results for the FOG sequence using PFFS. 2nd Row: Enlarged version of the tracker window. Note that this figure is best viewed in color.

of the four sequences, the PFFS algorithm outperforms other methods resulting in less average tracking error.

5 Conclusion

In this paper a robust tracking method is proposed. It is based on an adaptive feature selection mechanism which makes the tracker robust against occlusion, confusing background color and large illumination variation. Experiments with different sequences show that a particle filter based tracker with adaptive feature selection outperforms other established color based tracker in difficult tracking environments. A direction of future work would be the integration of shape and texture features and an extension for multi-objects tracking.

References

1. Comaniciu, D., Ramesh, V., Meer, P.: Real-Time Tracking of Non-Rigid Objects using Mean Shift. In: Computer Vision and Pattern Recognition (CVPR), pp. 142–149. IEEE Computer Society (2000)
2. Ho, J., Lee, K.C., Yang, M.H., Kriegman, D.J.: Visual Tracking Using Learned Linear Subspaces. In: Computer Vision and Pattern Recognition (CVPR), pp. 782–789. IEEE Computer Society (2004)
3. Zarit, B.D., Super, B.J., Quek, F.K.H.: Comparison of Five Colour Models in Skin Pixel Classification. In: International Conference in Computer Vision (ICCV), pp. 58–63 (1999)
4. Grabner, H., Grabner, M., Bischof, H.: Realtime Tracking via On-line Boosting. In: British Machine Vision Conference (BMVC), pp. 47–56 (2006)
5. Sidibe, D., Fofi, D., Meriaudeau, F.: Using Visual Saliency For Object Tracking with Particle Filter. In: European Signal Processing Conference (EUSIPCO), pp. 1776–1780 (2010)
6. Leung, A.P., Gong, S.: Online Feature Selection using Mutual Information for Real-time Multi-view Object Tracking. In: Zhao, W., Gong, S., Tang, X. (eds.) AMFG 2005. LNCS, vol. 3723, pp. 184–197. Springer, Heidelberg (2005)
7. Collins, R.T., Liu, Y.: On-line Selection of Discriminative Tracking Features. In: International Conference in Computer Vision (ICCV), pp. 346–352 (2003)
8. Han. B., Davis. L: Object Tracking by Adaptive Feature Extraction. In: International Conference in Computer Vision (ICCV), pp. 1501–1504 (2004)
9. Kwolek, B.: Object Tracking using Discriminative Feature Selection. In: Blanc-Talon, J., Philips, W., Popescu, D., Scheunders, P. (eds.) ACIVS 2006. LNCS, vol. 4179, pp. 287–298. Springer, Heidelberg (2006)
10. Nummiaro, K., Koller-Meier, E., Van Gool, L.J.: An Adaptive Color-based Particle Filter. In: Image and Vision Computing, pp. 99–110 (2003)
11. Isard, M., Blake, A.: Condensation - Conditional Density Propagation for Visual Tracking. Int. Journal of Computer Vision (IJCV) **28**, 99–110 (1998)
12. Van de Sande, K.E.A., Gevers, T., Snoek, C.G.M.: Evaluating Color Descriptors for Object and Scene Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) **32**, 1582–1596 (2010)
13. Stern, H., Efros, B.: Adaptive Color Space Switching for Tracking under Varying Illumination. Image and Vision Computing **23**, 3–346 (2005)
14. Ross, D., Lim, J., Lin, R.S., Yang, M.H.: Incremental Learning for Robust Visual Tracking. Int. Journal of Computer Vision (IJCV) **77**, 125–141 (2008)