Importance Sampling Based Probabilistic EigenTracker

Anup Shetty[1], Sumantra Dutta Roy[2], Subhasis Chaudhuri[1]

[1]{anupshetty, sc}@ee.iitb.ac.in, Department of Electrical Engineering, IIT Bombay
[2]sumantra@ee.iitd.ac.in, Department of Electrical Engineering, IIT Delhi

1. Introduction

Tracking objects over the video frames finds many applications in surveillance [13], human activity analysis [4] and gesture recognition [6],[12]. It becomes a very challenging task if the appearence and the dynamics of the object varies over successive frames. It becomes important to include elements of nonlinearity and non-Gaussianity in order to model accurately the underlying dynamics of a physical system [1]. We use Condensation Algorithm over the constructed eigenspace as the observation for the particle filter.

The other problem with the tracking based algorithms is loss of track after sometime and it requires reinitialization after every few frames. Hence the trackers fail to track the entity over a long duration of time. By using Importance framework [9] we are able to increase the successful tracking time by a considerable amount. We also use parameteric transformation which gives a much tighter bound on the object to be tracked.

1.1. Overview of the proposed model

Our framework for tracking moving objects consists of the following steps (i) delineating the object to be tracked, (ii) predicting best location of the entity in the next frame by observing the eigenspace,(iii) updating the eigenspace based on the error threshold and hence learning the change in appearance. The histogram can be built either real-time from first image or if the histogram is known *apriori*, the initialization can be made automatic. Importance sampling has been incorporated in the predictive framework. An importance function augments a tracker operating with one type of measurement with information from an auxiliary measurement source.

1.2. Relation to Previous Work and Contribution

Based on the conclusion of [7] we find that most existing trackers handle events of short duration with moderate changes in illumination, scene clutter and occlusion, but they fail in long run. Our approach concentrates more on making the tracker robust against the localization error accumulated over time. There has been work on other aspects of EigenTracking such as extensions for tracking of flexible objects [11], and incorporating the motion of shape in an eigenspace Active Appearance Models (AAM) [3]. Particle filtering can be incorporated with the EigenTracker framework to track objects efficiently. Particle filtering was used by [8] in tracking framework. Authors in [5], [14] introduced an index K for each particle (called the branchindex) as an auxiliary variable to improve the sample efficiecy. Development of predictive framework for EigenTracking has already been addressed [6].

We begin in Section II with basic eigen analysis and representation. Tracker initialization and occlusion handling are also discussed in section II. The predictive EigenTracker described in section III, is implemented using Condensation algorithm. Section IV describes the importance sampling mechanism used to make the system robust against clutter. The overall tracking scheme is described in section V. Finally in section VI results showing various tracking scenarios are discussed to draw conclusions in section VII.

2. Eigen Analysis

An EigenTracker [2] has an advantage over traditional feature based tracking algorithms - the ability to track objects that simultaneously undergo affine image motions and changes in view [6]. The major problem with traditional optical flow-based tracking algorithms is that it cannot distinguish between the object's motion and changes in its appearance. Other template-based trackers can only track a particular view of the object, and deviations in the appearance of the object from the template can put these algorithms off the track. Also, maintaining a large database for template based matching makes the tracker slow and tracking infeasible.

An eigenspace representation of the observations can be used to overcome these shortcomings by combining a compact, view-based representation of the object and an effective, affine parametric image transformation, both in a single framework. It incorporates an affine image transformation to map the observed object view to the canonical form. The error criteria can be modified to minimize the effects of the outliers.The modified error norm is known as robust error norm. It is more resistant to the adverse effects of structured background noise. We can define the EigenTracker as an algorithm which, given initial seed value, simultaneously gives the best estimate of the object's view and its motion.

2.1. Eigen Space Representation

An EigenTracker can sucessfully track moving objects, which undergo changes in appearance as well. We learn the eigenspace of appearances of the object to track, and pose the problem as estimating 2-D affine transformation coefficients **a** and the eigenspace reconstruction coefficients **c**, to minimize a robust error function between the parameterized image **I** (indexed by its pixel location **x**) and the reconstructed one **Uc** (where **U** is the matrix of the most significant eigenvectors):

$$arg \min_{\forall \mathbf{x}} \rho(I(\mathbf{x} + f(\mathbf{x}, \mathbf{a})) - [Uc](\mathbf{x}), \sigma) \tag{1}$$

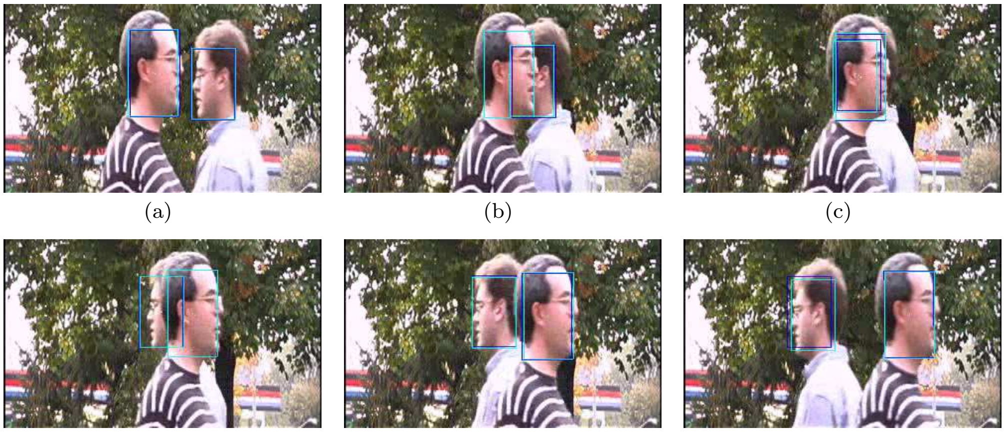
where $\rho(y, \sigma) = y^2/(y^2 + \sigma^2)$ is the robust error function, and σ is a scale parameter. The 2-D affine transformation is given by

$$f(x, a) = \begin{bmatrix} a_0 \\ a_3 \end{bmatrix} + \begin{bmatrix} a_1 & a_2 \\ a_4 & a_5 \end{bmatrix} \mathbf{x} \tag{2}$$

where **x** is the position vector of the point. A parallelogram offers tighter fit to the object being tracked as compared to a rectangular bounding box. This is an important consideration for an appearance-based method, since we do not want much background to be learnt as part of the eigenspace representation of the object.

2.2. Tracker Initialization and Occlusion Handling

Initializing a tracker automatically is a challenging problem because of clutter, other moving objects, and the possibility of misclassifying the region of interest. Our tracker performs fully automatic initialization of tracking the the object given its histogram apriori. The initialization method combines colour and motion cues. For a particular application, one may use other cues to the advantage. We augment motion cues with colour cues[10], to segment out the moving region of interest. Tracking an object in an occlusion scenario is very difficult. In our experiments we use a constant velocity model for the motion dynamics of the object. This model may not work if the object is moving with varying velocity when it is occluded. But in most real life scenario and when the entity is occluded for only few frames we can safely use this model.



(a) (b) (c)

(d) (e) (f)

FIG 1. (a-f) Results of our approach on video frames showing tracked objects in occluded scenario

3. Predictive Eigen Tracker

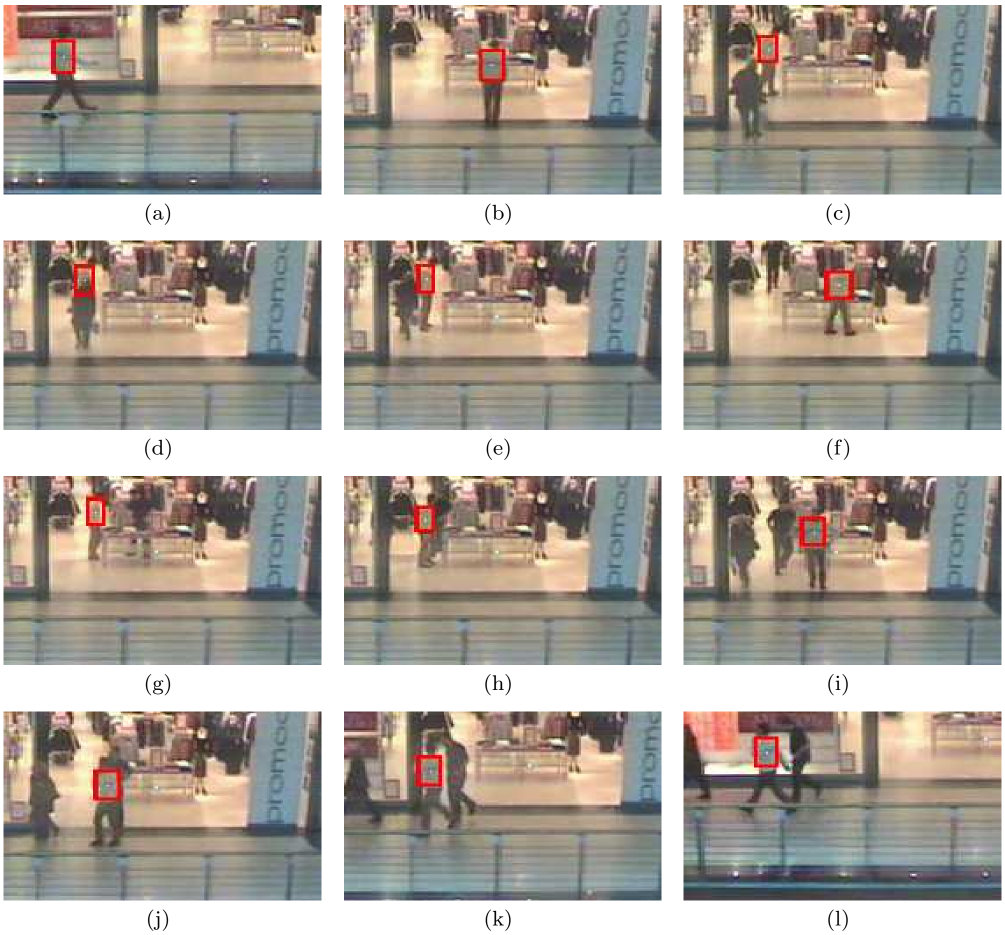
One of the main reasons for the inefficiecy of the eigentracking algorithm is the absence of predictive framework. After each frame is processed, it simply updates the eigenspace and affine coefficients. In each case it requires a good seed value for non-linear optimization. A predictive framework helps in speeding up the tracking process. Sequential Monte Carlo techniques for filtering time series [1] and their use in the context of visual tracking [8] have been described at length in the literature. The starting point is a standard state space model, where a Markovian prior on the hidden states is coupled with a conditionally independent observation process. Denoting by x_t and y_t the state and the data respectively, at time t, and fixing the order of the dynamics to one, the sequence of filtering distributions $p(x_t|y_{0:t})$ to be tracked obeys the recursion.

$$p(x_{t+1}|y_{0:t+1}) \;\propto\; p(y_{t+1}|x_{t+1}) \int_{x_t} p(x_{t+1}|x_t)p(x_t|y_{0:t})\, dx_t, \tag{3}$$

with the notation $\mathbf{x_{0:t}} = (\mathbf{x_0}, \ldots, \mathbf{x_t})$ and similarly for $\mathbf{y}$ [15]. The recursion can however be used within a sequential Monte Carlo framework where the posterior probability $\mathbf{p(x_t|y_{0:t})}$ is approximated by a finite set $\mathbf{x_t^m}, \mathbf{m} = \mathbf{1}...\mathbf{M}$ of M samples, the particles. A commonly used state dynamics model is the first order autoregressive (AR) process: $\mathbf{X_t} = \mathbf{A}_1\mathbf{X}_{t-1} + \mathbf{W}_t$, where t denotes time. The particular form of the model will depend on the application. We use a Conditional Density Propagation (CONDENSATION) based framework[8] for propagation of state densities across frames. We model the measurement as: $\mathbf{Z}_t = \mathbf{B}\mathbf{X}_t + \mathbf{F}_t$ where $\mathbf{X}_t$ is the state vector at time t and $\mathbf{Z}_t$ is the observation vector. $\mathbf{A}_1$, $\mathbf{B}$ are coefficient matrices and $\mathbf{W}_t$ and $\mathbf{F}_t$ are assumed to be zero mean, white Gaussian noise vectors. We show the results (Fig. 2) of tracking a long sequence with changing appearance in a cluttered backgroung, having multiple moving objects and having scenes with partial occlusion.

4. Importance Sampling Mechanism

An Importance function augments a tracker operating with one type of measurement with information from an auxiliary measurement source. This additional knowledge makes the system robust to failures as compared to *single measurement* systems. For Importance Sampling, one samples from an Importance function $g(X)$, rather than the state density $P(X)$. We use a new importance sampling mechanism, a general technique for estimating the properties of a particular distribution, while only having samples



(a) (b) (c)

(d) (e) (f)

(g) (h) (i)

(j) (k) (l)

Fig 2. Caviar Sequence: The object to be tracked is very small and has same colour as the background. There are multiple moving objects and occlusion is also handled well.

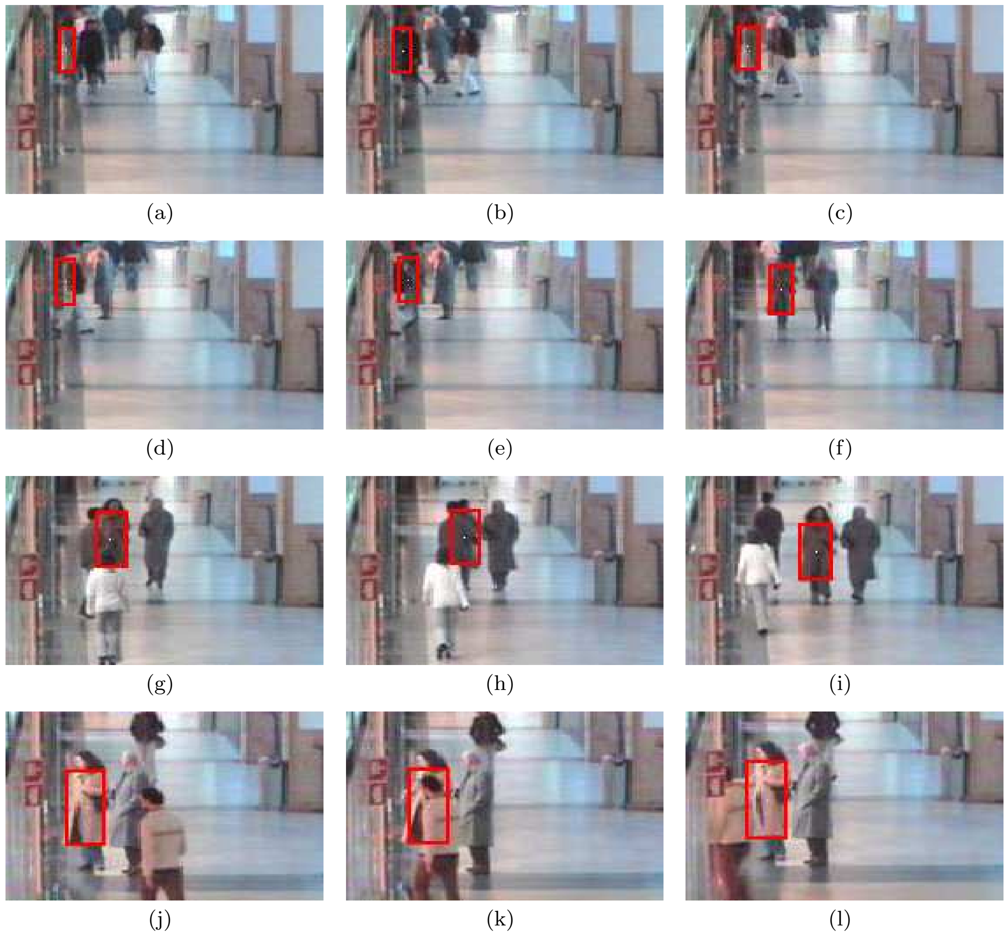


Fig 3. Another similar Caviar Sequence: The object to be tracked is very small and is scale varying with time.

generated from a different distribution than the distribution of interest. This approach segments out the region of interest using another source of measurement.

Our framework of using an auxiliary independent measurement augments the original trcker, and enhances its reliability by making it more robust. For example, let us consider a non-convex shape enclosed in a bounding parallelogram. A plain Eigentracker will have problems with changing backgrounds in the bounding parallelogram. This is especially important in the context of incrementally constructing the eigenspace, on the fly. Using only texture or colour-based properties, for example may give an improper segmentation. A combination of the two in an importance sampling framework can result in more reliable tracking. We use a uniformity predicate to form a new Importance Eigenspace representing a view of the object of interest in a frame, with the background eliminated. We optimize the new affine image parameters $\mathbf{a}'$ and the reconstruction coefficients $\mathbf{c}'$ of the Importance eigenspace to obtain the Importance measurement.

5. Overall Tracking Scheme

For tracking the desired object, first we have to delineate the object of interest in the first frame. For all the frames sequentially obtain the image measurement optimizing image parameters $\mathbf{a}$ and reconstruction coefficients $\mathbf{c}$. For importance sampling optimize $\mathbf{a}'$ and $\mathbf{c}'$ parameters in the *importance*

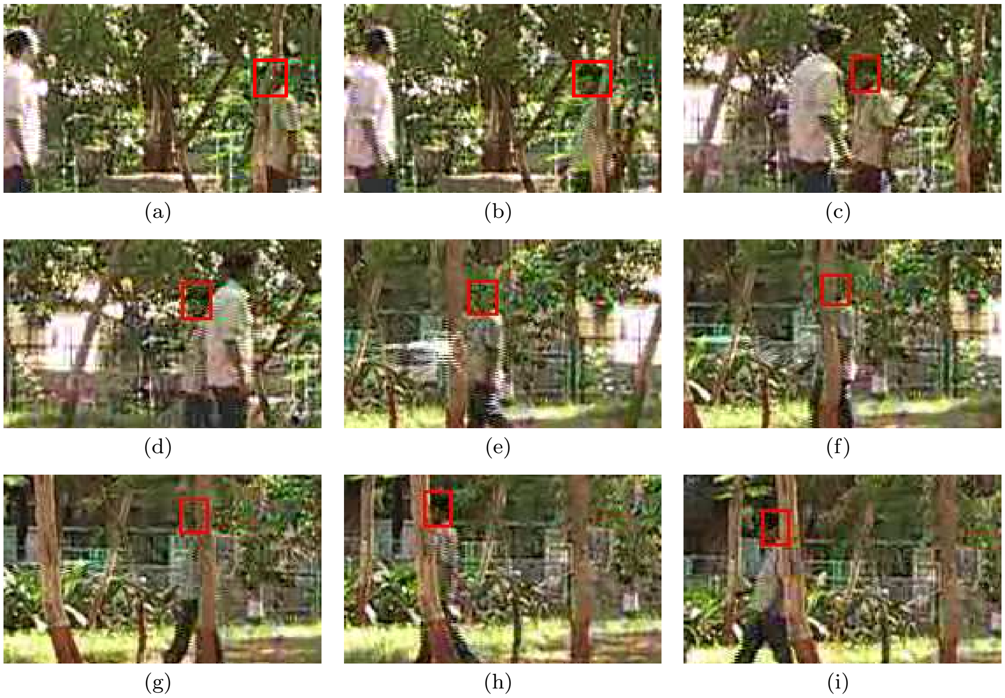


FIG 4. Scene with moving backgroud, occlusion and multiple moving objects

eigenspace to compute importance measurement. Estimate new affine parameters using previous output. Now, for each eigenspace check for reconstruction error, and based on that update the eigenspace. Repeat the process by predicting the affine parameters for the next frame.

6. Results and Discussions

We have performed extensive experiments and analysis of our algorithm. First, we show results on standard data, and analyze the performance of the proposed method. Next, we show results on motion in an out door environment, including significant occlusions, changing background and multiple interacting people.

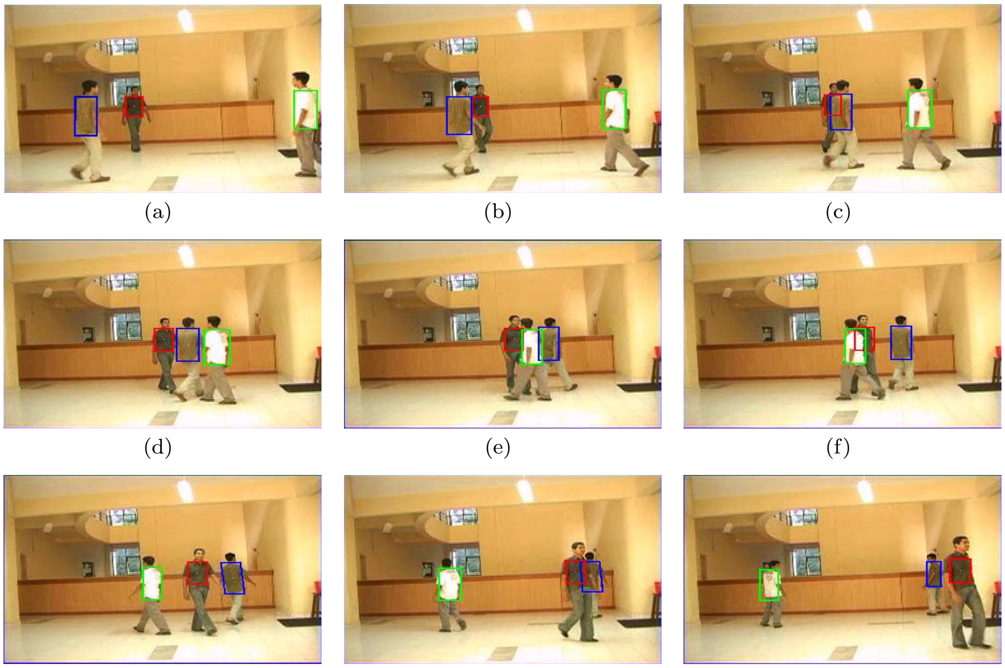
6.1. Tracking under occlusion scenario

In Fig. 1 two person cross each other resulting in complete occlusion of one. The colour of both the entities are similar. Constant velocity model is used to handle this scenario.

6.2. Tracking in large clutter

The sequences (Fig. 2 and 3) used is from the Caviar data set . Here the tracker sucessfully tracks the person for a long time. The tracker also takes into accout the scale change with time. Also the tracker handles occlusion very efficiently.

In Fig. 4 two persons cross each other. The colour of the stem and the branches are very similar to the colour of the face, also the leaves are moving. In the background, people and vehicles are also in motion.The tracker loses the person when occluded by the branch but regains after few frames.



(a) (b) (c)
(d) (e) (f)
(g) (h) (i)

FIG 5. Multiple Importance Predictive Eigen Tracker

6.3. Multiple Predictive Tracker

We have also tried tracking multiple objects simultaeously (Fig. 5).This can be done by developing eigenspace for each object individually and then predicting their next position independently. We show the results of tracking three objects simultaneously in a cluttered background with complete occlusion in few frame. This is a real life scenario where trackers usually fail in such highly occluded scenes. Though the scene is highly cluttered the objects are tracked successfully.

7. Conclusion

The Importance Sampling based Predictive Eigen Tracker is robust to background clutter and noise. It can also update slowly varying appearance which makes it a promising tracker to track long duration sequence. It can also handle slowly varying illumination and the tracker can very efficiently handle occlusion in few frames. Computing and matching eigenspace is computationally very expensive and it takes time. But the accuracy can be bit compromised for speed and hence it can be very promising for tracking in complicated scenarios.

Acknowledgement

The authors would like to thank V. Srikrishnan, Electrical Engineering Department, IIT Bombay for interesting discussions and comments on the work.

References

[1] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for on-line non-linear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–

188, Feb. 2002.

[2] M. J. Black and A. D. Jepson. Eigentracking: Robust matching and tracking of articulated objects using a view-based representation. *Technical Report: RBCV-TR- 96-50, Dept. of Computer Science, University of Toronto*, October 1996.

[3] T. Cootes, G. J. Edwards, and C. Taylor. Active appearance models. In *Proc. European Conference on Computer Vision (ECCV)*, 1998.

[4] J. Gao, A. G. Hauptmann, and H. D. Wactlar. Combining motion segmentation with tracking for activity analysis. In *International Conference on Automatic Face and Gesture Recognition (FGR'04)*, pages 699–704, 2004.

[5] F. Guo and Q. G. Sample-efficiency-optimized auxiliary particle filter. In *Proceedings of IEEE Workshop on Statistical Signal Processing*, 2005.

[6] N. Gupta, P. Mittal, K. S. Patwardhan, S. D. Roy, S. Chaudhury, and S. Banerjee. On-line predictive appearance-based tracking. In *Proc. IEEE International Conference on Image Processing (ICIP)*, pages 1041 – 1044, 2004.

[7] W. Hu, T. Tan, L.Wang, and S. Maybank. A survey on visual surveillance of object motion and behaviors. *IEEE Trans. on Systems, Man, Cybernetics - Part C: Applications and Reviews*, 34, 2004.

[8] M. Isard and A. Blake. Condensation - conditional density propagation for visual tracking. *International Journal of Computer Vision*, 28(1):5–28, 1998.

[9] M. Isard and A. Blake. Icondensation: Unifying low-level and high-level tracking in a stochastic framework. In *Proc. European Conference on Computer Vision (ECCV)*, pages 893–908, 1998.

[10] R. Kjeldsen and J. Kender. Finding skin in color images. In *Proc. Intl. Conf. on Automatic Face and Gesture Recognition*, page 312 317, 1996.

[11] F. D. la Torre, J. Vitria, P. Radeva, and J. Melenchon. Eigen filtering for flexible eigentracking (efe). In *Proc. International Conference on Pattern Recognition (ICPR),*, pages 1118–1121, 2000.

[12] J. Mammen, S. Chaudhuri, and T. Agrawal. Tracking of both hands by estimation of erroneous observations. In *Proc. British Machine Vision Conference (BMVC)*, 2001.

[13] C. P. and J. R. Multiple target tracking for surveillance: A particle filter approach. In *Intelligent Sensors, Sensor Networks and Information Processing Conference*, pages 181– 186, 2005.

[14] M. K. Pitt and N. Shephard. Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Association*, page 590599, 1999.

[15] P. Perez, J. V. C. Hue, and M. Gangnet. Color-based probabilistic tracking. In *Proc. Europ. Conf. Computer Vision(ECCV)*, pages 661–375, 2004.