

# A Hybrid Based Approach for Object Tracking in Video

**Amith. R**

Dept. of Computer Science and Engineering  
Jain University  
Bengaluru, India.

**V.N. Manjunath Aradhya**

Department of MCA  
Sri Jayachamarajendra College of Engineering  
Mysuru, India.

---

DOI: <http://dx.doi.org/10.21013/jte.v4.n1.p2>

**How to cite this paper:**

**R, A., & Aradhya, V.** (2016). A Hybrid Based Approach for Object Tracking in Video. *IRA-International Journal of Technology & Engineering* (ISSN 2455-4480), 4(1).  
doi:<http://dx.doi.org/10.21013/jte.v4.n1.p2>

---

© Institute of Research Advances



This work is licensed under a [Creative Commons Attribution-Non Commercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/) subject to proper citation to the publication source of the work.

**Disclaimer:** The scholarly papers as reviewed and published by the Institute of Research Advances (IRA) are the views and opinions of their respective authors and are not the views or opinions of the IRA. The IRA disclaims of any harm or loss caused due to the published content to any party.

---

**ABSTRACT**

*Tracking of moving objects in video sequences are essential for many computer vision applications & it is considered as a challenging research issue due to dynamic changes in objects, shape, complex background, illumination changes and occlusion. Many traditional tracking algorithms fails to track the moving objects in real-time, this paper proposes a robust method to overcome the issue, based on the combination of particle filter and Principal Component Analysis (PCA), which predicts the position of the object in the image sequences using stable wavelet features, which in turn are extracted from multi scale 2-D discrete wavelet transform. Later, PCA approach is used to construct the effective subspace. Similarity degree between the object model and the prediction obtained from particle filter is used to update the feature vector to handle occlusion and complex background in video frames. Experimental results obtained from the proposed method are encouraging.*

**Keywords:** particle filter, wavelet, pca, occlusion, tracking.

**1. Introduction**

Object tracking is the significant undertaking in the field of computer vision. The conception of dynamic computers, accessibility of astounding and economical camcorders and expanding requisite for computerized video audit has created a lot of enthusiasm in tracking objects in video. Tracking can be characterized as the issue of evaluating the movement of an object in the image plane. The object detection and tracking are intently associated in light of the fact that tracking normally begins with detecting the objects, although detecting an object over and over again in consecutive image array is usually significant to benefit and verify tracking. Challenges in tracking of moving objects include partial and full occlusion handling, complex background, tracking small objects with low contrast, reflectance noise, illumination problems, and noisy videos.

Object tracking in video is been important and a challenging issue in research for long time, as it empowers few major applications [11] like: Surveillance and security to perceive individuals, to contribute exceptional sense of security using evident data. Video abstraction to obtain computerized description of videos, to produce object based outline. Traffic administration to inspect stream, to identify accidents. Video editing to eradicate unwieldy individual administrative interplay, to shape up innovative video effects. Intuitive games to offer common way of interplay with perceptive systems.

Tracking a object is simplified as template matching and updating technique. Initially, a model is developed to characterize the object and to build the template. Later, the target can be detected and tracked using template matching process. Updating the target model is essential during tracking to conform to the variations in shape of the target. Many existing target modeling approaches are available in the literature, used widely [5,8] illustrates the target using color histograms, that are prone to complex background, illumination changes and does not mirrors the spatial inter-relationship of pixels.

In the region based tracking algorithm [18, 15], the target is defined by a geometric structure and the possible location of the object in the image sequence [16, 17] is determined by calculating the similarity function or can be determined by segmentation, which focuses on certain characteristics of object like color, edges, etc., Tracking algorithms based on feature represents target using a set of different feature points. Feature based tracking algorithms are robust by tracking the target [12, 19] with the help of constant elements over numerous image sequences in video. In this type of tracking algorithms, the feature points are assigned all over the target, even though if there are any obstructions on the target, still the feature based algorithms can continue tracking by utilizing the other visible points of the target. In contour based algorithms, usually the target is initialized in the first frame and it is represented by closed silhouette which extracts the shape of target. Next, based on the image features and contour smoothness an energy function is defined [20] and active contour is approximated to local minimum energy function by solving the differential equations recursively [13]. Contour based algorithms can handle the partial

occlusion if the contour is initialized correctly and it is robust compared with region based methods because the information obtained from contour is imperceptible to the intensity variations. Tracking algorithms based on model builds a object shape model with the specification and constraints of application to carry out the tracking process by computing the similarity function among the actual frame and the shape model [14, 21].

Wavelet analysis [22] is popularly used in noise reduction, image compression, fusion and smoothing, which primarily focus on the explicit features of the context through transformation. It has drawn attention of many research activities [9–4] in image processing and computer vision. Multi-scale 2-D Discrete wavelets provide a high dimensional vector comprised of the wavelet coefficients. When the frame is decomposed by wavelets it requires large computation cost, if vector of multi-scale 2-D wavelet coefficients are used directly in target matching it will results in poor real-time performance. In order to overcome the above said, Principal component analysis (PCA) is an efficient linear transformation approach to reduce the dimensionality; usually dimensionality of the data is high so it is projected into low dimensional space and PCA can also extract the features from the frame. Accurately estimating the object location just based on the target model is challenging, therefore it is better to integrate the tracking process with property of target motion, which results in boosting the reliability of tracking. Tracking the object is considered as probability of inference and the movements of object are depicted by the statistical models. Particle filtering [7] based on Bayesian theory and Monte Carlo method [6, 10] can be used to solve the posterior probability, where the weighted particles describes the all possible distributions of object's states.

Inspired from the above facts, particle filtering is used to predict the position of the object in the video frame. Later, eigen vector is built by extracting wavelet features from the object. Also, PCA approach is used for dimensionality reduction and stable feature extraction. Then, based on the similarity function; the object is detected and while tracking the object, eigen vectors will be updated considering challenges like complex background, illumination changes and occlusion.

## 2. Proposed Method

### 2.1 Particle Filter:

Particle filter is the generic algorithm for a function optimization where the solutions search space is searched using particles. Each particle integrates the tests whether how and where it is possibly that the object is at the position. The weights for particle are assigned accordingly how good particles are, based on particles evaluation. After assigning weights, multiply the good particles and remove the bad particles through re-sampling technique. Then predict the object where it might be using the next particle generation by evaluating, and it repeats. Compared to Kalman filter, particle filter works with the general probability density function or values. The Particle filter predicts the next state from the provided state transition (example, the motion model), then if applicable, the noisy measurement information is integrated in the correction phase.

To simulate the recursive Bayesian estimation, it is important to use Monte Carlo simulation, because it is difficult to obtain solution of priori probability density and posterior density.

Construct the posterior probability density function recursively using a set of weighted random samples  $\{x_t^{(i)}, w_t^{(i)}; i = 1, \dots, K\}$ , where K is the total number of particles. At each instance t, particle filtering algorithm repeats in two stages, they are: prediction and update.

Prediction: Every particle  $x_t^{(i)}$  is emerged individually as per the state model (1), and also to simulate unknown disturbances, an additional random noise is included. This step gives an approximation of the  $p(x_t)$ , i.e., the prior probability density function:

$$p(x_t) \approx \frac{1}{N} \sum_{i=1}^N \delta(x_t - x_t^{(i)}) \quad (1)$$

Update: Based on the latest measurement of likelihood model, weight's of the each particle has been calculated. The posterior probability density function at an instance t, in the form of a discrete approximation can be written as:

$$p(x_t | z_{1:t}) \approx \sum_{i=1}^N w_t^{(i)} \delta(x_t - x_t^{(i)}) \quad (2)$$

Where

$$w_t^{(i)} = \frac{\mathcal{L}(z_t | x_t^{(i)}) p(x_t^{(i)} | x_{t-1}^{(i)})}{q(x_t^{(i)} | x_{t-1}, z_t)} \quad (3)$$

and weight set satisfies function:

$$\sum_{i=1}^N w_t^{(i)} = 1 \quad (4)$$

## 2.2 Feature extraction

In order to differentiate the object from the background it is necessary to extract the good features and while tracking, it is vital to deal with the differences of the object and its background. Wavelet transform technique is a reversible transformation process as it preserves the information and is retained in the wavelet coefficients. The multi-scale property of wavelet transform delivers detail parameters of each frame decomposed, that consists of low frequency and high frequency components. Then, the feature vector X is constructed and the dimensionality of X is usually high. If X is used directly to track the object, it fails to perform in real-time because of high computational cost. In object tracking, the environmental changes may lead to changes in wavelet features. To obtain stable features and dimensionality reduction in feature vector, there is need for component analysis.

Let, the dimensionality of X is N, we need to obtain less than N dimensional feature vector S. The feature components of S should have the minimum variance, as shown in Eq. (5)

$$E(\|x - \hat{x}\|^2) \quad (5)$$

To find linear transformation, it is necessary to discard the components with lower variance and to retain the higher variance components. PCA's intent is to find orthogonal transformation matrix  $W^T = [w_1 w_2 \dots w_M]$  such that the N dimensional vector X makes the components of the transform results S, such that all components  $C_i$  ( $i = 1, 2, \dots, M$ ) to be uncorrelated. When all the features X along the  $w_1$  projection,  $C_1$  will have the high variance. Then  $C_1$  is called the first principal component, the next principal component can be found in the same way by magnifying variance in  $w_2$  subject to the orthogonal condition of  $w_1$ . Likewise N number of components can be obtained from N dimensional vector. Feature vector and dimensionality reduction is achieved using PCA by calculating the mean and variance of both low frequency and high frequency components to build the subspace.

To find the object's location, similarity measure function  $\Delta s$  is used to find the similarity degree between the object model in initial frame and predicted object model in current frame, and in most of the cases the probability distribution of both the models are same. If  $s = [s_1, s_2, \dots, s_n]^T$  is the subspace built,  $w = [w_1, w_2, \dots, w_n]$  is the weight respectively and  $s' = [s'_1, s'_2, \dots, s'_n]^T$  is the eigen spaces of prediction done by particle filter. Threshold is determined based on experiments.

$$\Delta s = \sum_{i=1}^n w_i |s_i - s'_i| < Threshold \quad (6)$$

Match the feature vector with respect to eigen space  $s'$  and measure the similarity. In order to handle the changes in wavelet features, caused from intervention of external environmental factors, it is essential to update the feature vectors while tracking the object. If the object does not match in the current, wavelet features can be extracted from the previous frame, because it was matched successfully in earlier frame to construct the subspace. Then, PCA will extract the feature vectors and sort according to their respective weight from the constructed subspace. Hence it is effective to update the model consists of the features with greater weights meets the formula (7)

$$\sum_{i=1}^N w_i > Threshold \quad (7)$$

### 3. Experimental Results

To assess the performance of proposed method, we make use of videos, 500 frames (Average) each, consists of various issues like tracking people, traffic management, circumstantial variations, complex background etc. Openly accessible standard PETS, ViSOR, iLIDS and other datasets are used for this purpose. The frequency of complete occlusion among the objects in video is often and the challenging feature of these datasets is tracking the objects with less contrast, complex background and occlusion.

In Figure 3.1 we can see the target objects are tracked successfully with various real-time video sequences, which consists of sunny day sequences, low contrast sequences, complex background, occlusion, etc., both in indoor and outdoor environment. Stable wavelet based features are extracted using PCA and the object model is updated accordingly. Due to this, any environmental changes, complex background and occlusion will not affect the tracking process. The proposed method outperforms the wavelet and traditional PCA based method and Mean-Shift tracking method as demonstrated in Table I. The object to be tracked is randomly initialized in the first frame. As shown the figure 3.1 (a) Image sequences from PETS dataset are tested with the proposed method by tracking car in outdoor environment handling full occlusion, (b) tiny remote car is fully occluded in the video frame, but the proposed method tracks the car successfully after recovering from the occlusion in the upcoming frames, (c) a man walking in a subway is tracked successfully by overcoming the occlusion problem with the help of effective model updating mechanism, (d) and (e) tracking people both in indoor and outdoor environment with complex background and different illuminations. Figure 3.2 shows the results obtained from wavelet and traditional PCA based approach and figure 3.3 shows results achieved from mean-shift tracking. In most of the cases both the methods fail to track the object with complex background when occlusion occurs as shown below. False positive rate is relatively high with respect to wavelet based PCA approach and while tracking object when occlusion occurs, after occlusion it assigns new label to the same old object treating it as a new object. Tracker Detection Rate, False Alarm Rate, Detection Rate, Accuracy, Positive Prediction, Negative Prediction, False Positive Rate and False Negative Rate achieved from the proposed method, wavelet and PCA based method without using particle filter and Mean-Shift tracking methods are demonstrated in Table I.



(a)



(b)



(c)



(d)



(e)



(f)

Figure 3.1: Results obtained from proposed method; (a) Tracking car by handling occlusion. (b) Tracking small remote car under full occlusion. (c) Tracking a man walking in subway with complex background under occlusion. (d) Tracking a man crossing road with complex background. (e) & (f) Tracking people with complex background under different illumination.



(a)



(b)



(c)



(d)

Figure 3.2: Results obtained from wavelet and traditional PCA (Without particle filter); (a) Tracking car by handling occlusion but fails to preserve shape. (b) Fails to track small remote car under full occlusion. (c) Fails to track a man walking in subway under occlusion with many false positives. & (d) Tracking a man crossing road.



(a)



(b)



(c)





(d)

Figure 3.3: Results obtained from mean-shift tracking; (a) Fails to track car when occlusion occurs. (b) Fails to track small remote car under full occlusion. (c) Fails to track a man walking in subway under occlusion. & (d) Fails to track a man crossing road.

TABLE I. PERFORMANCE ANALYSIS

Metrics	Proposed Method	Wavelet and PCA based method Without using Particle Filter	Mean-Shift Tracking
Tracker Detection Rate (TP/TG)	77.32%	67.94%	48.81%
False Alarm Rate (FP/(TP+FP))	11.72%	28.28%	24.14%
Detection Rate (TP/(TP+FN))	81.31%	73.26%	58.67%
Accuracy ((TP+TN)/TF)	77.85%	69.29%	60.72%
Positive Prediction (TP/(TP+FP))	87.27%	71.70%	62.63%
Negative Prediction (TN/(FN+TN))	1.67%	26.36%	55.54%
False Negative Rate (FN/(FN+TP))	18.68%	26.73%	33.76%
False Positive Rate (FP/(FP+TN))	35.34%	74.76%	56.86%

Note: TP: True Positive; TN: True Negative; FP: False Positive; FN: False Negative; TF: Total number of frames in the video; TG: Total number of frames for the ground truth objects.

#### 4. Conclusion

The proposed method extracts the wavelet coefficients consists of low frequency and high frequency components of each video frame using the 2-D discrete wavelet transform, which also results in noise reduction. Accuracy of tracking object is enhanced by utilizing the object structure and appearance. Based on the combination of particle filter and PCA approach, the tracking can be accomplished by predicting the likely position of the object and building the subspace by feature extraction and dimensionality reduction. The proposed method is tested with publicly available standard PETS, iLIDS dataset and other real time video sequences by considering challenges like occlusion and complex background, which achieves fast and accurate tracking. Also proposed method outperforms both mean-shift tracking and wavelet based PCA method by tracking object with complex background under occlusion.

## References

- [1] Rajiv Kapoor & Rajesh Rohilla, Modified foreground segmentation for object tracking using wavelets in a tensor framework, *International Journal of Electronics*, 102:9, 2015, 1560-1582.
- [2] Kwon, J.H. and Lee, H.S. and Park, F.C. and Lee, K.M., A Geometric Particle Filter for Template-Based Visual Tracking, *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 36:4, 2014, 625-643.
- [3] Barga Deori and Dalton Meitei Thounaojam, A survey on moving object tracking in video, *International Journal on Information Theory*, 3 (3), July 2014, 31-46.
- [4] QIAN. Cheng, X.U. Shuchang, Zhang Yin, Zhang San-yuan, Object tracking method based on incrementally updating linear discriminant subspace, *Pattern Recognit. Artif. Intell.* 23 (2), 2010, 250–255.
- [5] H.Y. Zhou, Y. Yuan, Y. Zhang, et al., Non-rigid object tracking in complex scenes [J], *Pattern Recognit. Lett.* 30 (2), 2009, 98–102.
- [6] Wang Lei, Feature selection with kernel class separability, *IEEE Trans. Pattern Anal. Mach. Intell.* 30 (9), 2008, 1534–1546.
- [7] A. WANG Fasheng, New particle filter for nonlinear filtering problems [J], *Chin. J. Comput.* 31 (2), 2008, 346–352.
- [8] Takala V., Pietikainen M. Multi-object tracking using color, texture and motion[C], *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Piscataway, NJ, USA: IEEE, 2007, pp. 1–7.
- [9] Chai Yuhua, Gao Liqun, Lu Shun. Wavelet-based watershed for image segmentation algorithm[C], in: *Proceedings of the Sixth World Congress on Intelligent Control and Automation*, Dalian, China, 2006, II: pp. 9595–9599.
- [10] Z. Sun, G. Bebis, R. Miller, Monocular precrash vehicle detection: features and classifiers [J], *IEEE Trans. Image Process.* 15 (7), 2006, 2019–2034.
- [11] A. Yilmaz, O. Javed, M. Shah, “Object Tracking: A Survey,” *ACM Computing Surveys*, vol. 38, pp. 1-45, 2006.
- [12] P. Tissainayagam, D. Suter, Object tracking in image sequences using point feature, *Pattern Recognit.* 38 (1), 2005, 105–113.
- [13] D. Freedman, T. Zhang, Active contours for tracking distributions, *IEEE Trans. Image Process.* 13 (4), 2004, 518–526.
- [14] W.M. Hu, T.N. Tan, L. Wang, S.A. Maybank, Survey on visual surveillance of object motion and behaviors, *IEEE Trans. Syst. Man Cybern. Part C Appl. Rev.* 34 (3), 2004, 334–352.
- [15] I. Matthews, T. Ishikawa, S. Baker, The template update problem, *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (6), 2004, 810–815.
- [16] H. Nguyen, A Smeulders, Fast occluded object tracking by a robust appearance filter, *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (8), 2004, 1099–1104.
- [17] S. Avidan, Support vector tracking, *IEEE Trans. Pattern Anal. Mach. Intell.* 26 (8), 2004, 1064–1072.

- [18] A. Jepson, D. Fleet, T. El-Maraghi, Robust online appearance models for visual tracking, *IEEE Trans. Pattern Anal. Mach. Intell.* 25 (10), 2003, 1296–1311.
- [19] Kaneko T., Hori O. Feature selection for reliable tracking using template matching, in: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR'03*. Madison, Wisconsin, USA: IEEE Press, 2003. 1:796–802.
- [20] A. Mansouri, Region tracking via level set PDEs without motion computation, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (7), 2002, 947–961.
- [21] T. Drummond, R. Cipolla, Real-time visual tracking of complex structures, *IEEE Trans. Pattern Anal. Mach. Intell.* 24 (7), 2002, 932–946.
- [22] Sun Z., Bebis G., Miller R. Quantized wavelet features and support vector machines for on-road vehicle detection[C], in: *Seventh International Conference on Control, Automation, Robotics and Vision*, 2002, (3): pp. 1641–1646.