

# A Robust Framework for Accurate Face Tracking

Ali Ahmed\*

IT Department, Faculty of Computers and Information, Menoufia University, Shebin El-kom, Egypt.

\* Corresponding author. Tel.: +966541924606; email: ali.ahmed@ci.menofia.edu.eg

Manuscript submitted March 3, 2017; accepted May 17, 2017.

doi: 10.17706/jsw.12.5.326-338

---

**Abstract:** Face detection and tracking are of the most challenging problems of the object tracking field because of the large variability of faces and facial expressions. In this paper, two different algorithms for face tracking based on unscented Kalman filter (UKF) are proposed. The first proposed algorithm is UKF based on Viola-Jones algorithm. Viola-Jones is extremely fast feature computation, efficient feature selection, and scale and location invariant detector. The second proposed algorithm is UKF based on mean shift using the corrected background weighted histogram (CBWH) scheme. This scheme can effectively reduce background's interference in target localization and consequently can guarantee accurate localization of the target. The tracking step is completed using UKF that can estimate the next state with a high level of accuracy. So the two proposed algorithms are used to enhance the solution of face tracking problems. The performance of the two different proposed algorithms is evaluated with other well-known face tracking algorithms.

**Key words:** Mean shift, corrected background weighted histogram, unscented kalman filter, Camshift.

---

## 1. Introduction

The human face is important to our identity. It plays a major role in everyday interaction, communication, and other routine activities. Recently, Face detection step is mainly concentrated on searching for faces on a given image. If a face exists, immediately returns the image location and content of each face. The goal of this approach is to emerge the best face detection approach that satisfies the need for real-time hardware implementation.

The face detection problem is also a very challenging where it needs to account for all possible appearance variation due to a change in illumination, facial features, and partial or full occlusions. Also, the challenging is when needed to detect faces that appear at a different scale, or out of a plane rotation, or with in-plane rotation. Although, all these difficulties, tremendous progress has been made in the last decade, and many systems have shown impressive real-time performance. As this problem is the first step of any face processing system, it is used in many applications of face recognition and tracking, facial feature extraction, attentive user interfaces, gender classification, clustering, digital cosmetics, biometric systems, and Human Computer Interaction system, demographic classification, and surveillance system.

Kalman filter is a well known and widely used optimal estimator for linear systems. This filter predicts the location of a moving object based on its previous state. From a series of noisy measurements, Kalman filter is a recursive adaptive filter that estimates the state of a dynamic system. It has been used for an extensive range of applications in areas such as signal, image processing, and control design. Unfortunately, Kalman filter is widely used to reduce the dimensionality error only when the systems in the real world are linear. A common solution to cope with major problem is to linearize the system first before applying the filter, resulting in a new approach called an extended Kalman filter (EKF). This linearization when does, however, pose some other problems, for

example, the EKF can result in non-stable state estimates [1]. On the other hand, an alternative approach has emerged over the last filter, named unscented Kalman filter (UKF). This UKF filter claims both higher accuracy and robustness for nonlinear systems [2].

This work proposes a robust framework for face tracking based on the unscented Kalman filter algorithm with two different detection methods. The two detection methods are Viola/Jones and mean shift with CBWH algorithm. The two different proposed algorithms can overcome the challenges of face detection such as object appearance variation due to a change in illumination, facial expressions, partial or full occlusions, scale, and rotation.

The paper is organized as follows. Section 2 explains related work about face tracking. Section 3 explains the proposed tracking system. Section 4 the experimental results of faces and concluding remarks are given.

## **2. Related Works**

Face detection in still images is one of the challenge problems in computer vision and is the first step in various vision-based applications that involve interaction between a human and a machine. The main goal of face detection step is to determine whether or not there are any faces in the image and, if exist; return the location and extent of each face. As detection of a face is one of the very challenging problems, many different approaches have been tried to improve it to cope with the new hardware implementation. Viola-Jones method [3-6], based on Haar-like features and weak classifiers are recently the most robust method. Below, a classification of some of other important methods tested during the last decades.

### **2.1. Color-Based Methods**

These methods are based on the color-histogram of a given image. A training step is first needed to learn how the skin color of a face looks like and the ability to identify, whether a face or not, according to the range of colors of a given area. Despite these methods seem fascinating because they are facing pose independent, they have three strong constraints: they only work for color images, they do not work well with all kind of skin color, and they are not very robust under varying lighting conditions in [7, 8].

### **2.2. Edge-Based Methods**

This kind of methods is based on an edge orientation matching followed by a candidate verification using a classifier [23]. The main obstacle of these methods is the difficulty to detect the edges of a face with a complex background. For example using a Canny edge detector to detect a face, which is very clear how important is to have simple backgrounds, being otherwise impossible to distinguish its edges from the face. An alternative solution to this problem is provided by a background subtraction technique.

### **2.3. Feature-Based Methods**

This kind of methods includes a huge group of approaches. Some of them have not achieved any satisfactory at all, but other methods are very popular, different their performance highly depends on the kind of feature that is used. For example, some of the very early approaches use facial features (eyes, mouth, nose, etc.) and they do not achieve success because of the difficulty to define and compare them. On the other hand, other approaches use more sophisticated and person independent features which result in a more accurate system. Viola-Jones method [3]-[6], defined the so-called Haar-like features, which are approved both fast and easy to compute. This method is currently become the most used, although it works only for frontal faces with no occlusions or orientation.

### **2.4. Other Methods**

There have been published many other different methods, but most of them are no longer used. Among these methods, the technique based on template matching, which was one of the first developed and it consists in deforming an average template of a face, which was hand-coded, using the edges of an image. The principal

problem of this method relies on the initialization; to detect a face in a given frame, the face template has to be initialized in a frame close to the current one. Moreover, it mainly works only for frontal faces in [9], [10].

Some other weak face detection methods are the so-called knowledge-base. These methods use some constraints like the center of a face has a uniform intensity value, variation between the average intensity values of the center part and the upper part is significant, or a face often appears with two eyes that are symmetric to each other, a nose and a mouth. These methods that do not work well, basically because of the difficulty to translate human in [11], [12].

Mean shift (MS) tracking algorithm is a well known and a popular method due to its robustness and computational efficiency. MS is classified as nonparametric, iterative method to search a local maximum of a density function. It is considered one of the best color-based algorithms to find the location of a face. The algorithm based on Kalman filter (KF) and Mean Shift (MS) for tracking objects using the object histogram in RGB color space may lose the object when the color distribution is somewhat similar to the target and the background in the paper [13], [14]. In other work [15], [16] when some of the target's features present in the background, the traditional MS do not work well under this obstacle when object size becomes larger. Bhattacharyya coefficient based similarity measurement could not reflect the change of object size accurately. The background-weighted histogram (BWH) mainly aims to decrease background interference in target representation and that it is proposed in [17], [18] is somewhat incorrect and it is proved in [19]. The corrected background-weighted histogram (CBWH) is proved to reduce the interference of background in target localization as proposed in [19], [20].

### 3. Proposed Face Tracking System

Based on our study of the existing systems we have proposed a robust face tracking system based on the UKF with two different detection methods as shown in Fig. 1.

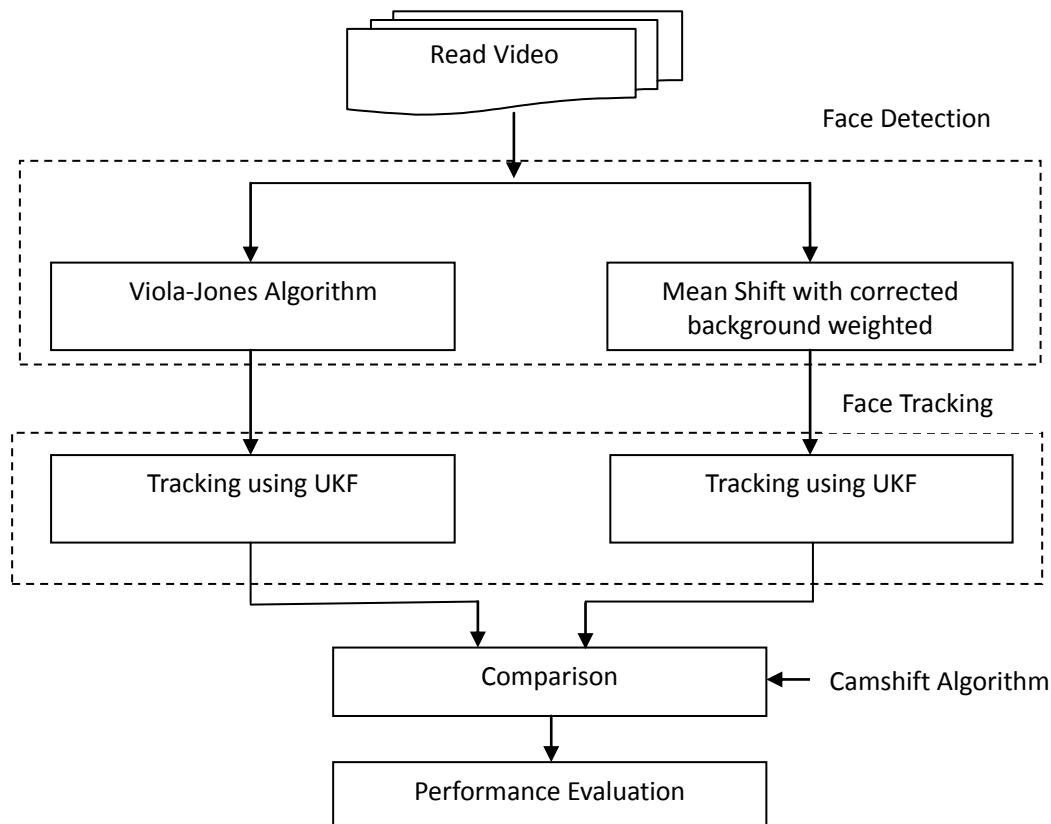


Fig. 1. The block diagram of face tracking system.

The proposed system is implemented in three steps which include face detection, face tracking, and comparison step. The system starts with reading the video and detects the faces using the two different detection methods separately. The first method is Viola-Jones algorithm which extremely fast features computation, efficient feature selection, Scale, and location invariant detector. Viola-Jones method based on Haar-like features and weak classifiers is currently the most robust. The second method is mean shift with CBWH scheme that can reduce background's interference in target localization which consequently can guarantee accurate localization of the target.

For the tracking step, UKF algorithm can estimate the next state with a high level of accuracy. it can predicts the location of a moving object based on its previous values. It's a recursive adaptive filter that accurately estimates the state of a dynamic system from a series of noisy measurements. In the case of using Viola/Jones method as a detection method, if the faces miss detected due to a rotation, tiles of a face or long distance from the camera, the estimated state obtained by UKF is used as its current measurement.

The final step is to measure the performance evaluation by comparing the two methods, the UKF based on Viola/Jones algorithm and UKF based on mean shift with CBWH with one of the well-known algorithms named Camshift algorithms. The accuracy of the two different proposed systems is also evaluated using true and false positive.

The different steps of the proposed tracking system are discussed in details in the following sections.

### **3.1. Mean Shift with CBWH for Face Detection**

Mean shift algorithm is considered a semi-automatic tracking method. The color histogram is used to represent the target, and the target object is typically defined as a rectangle region in the frame. The mean shift algorithm assumes that the initialization point falls when there is a significant movement between two consecutively processed frames. The mean shift algorithm, using the histogram in RGB color space, may miss object when being used in the environment where the color distribution is much similar to the target and the background.

The limitation of traditional mean shift method is prone to local minima when some of the target's features present in the background. The kernel bandwidth in conventional MS algorithm is solved, which cannot be adjusted to changing of object size. Although worthy research results in adaptive kernel bandwidth were conducted, they mostly use Bhattacharyya coefficient as a similarity measurement to adjust the original kernel bandwidth [15, 17]. However, these methods perform satisfactorily only when the object becomes smaller. Because of Bhattacharyya coefficient may not be larger with a large window than with small window, these methods do not do well under this circumstance. Since Bhattacharyya coefficient based on similarity measurement, it could not accurately reflect the change of object size.

The BWH transformation is mainly used to reduce the effects of prominent background features in the target candidate region on the target localization. The BWH cannot achieve this goal because it is equivalent to the common target representation under the mean shift tracking framework was proven in [15]. Although the idea of BWH is useful, the BWH algorithm does not improve the target localization. An improvement to BWH to achieve its goal, a proposed transformation method, called the corrected BWH (CBWH) algorithm. CBWH is proposed to transform only the target model but not the target candidate model. CBWH mainly reduce the prominent background features only in the target model but not in the target candidate model. CBWH was explained with details in [19], [20].

### **3.2. Viola-Jones Algorithm for Face Detection**

Viola/Jones introduced their object detection framework that can be applied to detect human faces. The algorithm is capable of processing images rapidly real-time frame rates, and it can be achieved on consumer hardware. It mainly uses a boosted cascade of simple classifiers that are based on rectangular Haar-like features, and it can be computed quickly as in [3], [4]. Also, the algorithm has a learning stage, and both high detection rate and low false positive rate can be achieved with a sufficiently large training set. Viola/Jones algorithm had

three key contributions. The first contribution is to propose a new efficient method for features calculation based on a so-called integral image. Secondly, it suggests a method for aggressive features selection based on AdaBoost learning algorithm. Finally, the algorithm proposed an idea of combining classifiers in a "cascade."

The value of the so-called integral image at location  $(x, y)$  is evaluated by the sum of the total pixels from the original image above and to the left of location  $(x, y)$  inclusively. Where  $ii(x, y)$  an integral is image and  $i(x, y)$  is an original image

$$ii(x, y) = \sum_{\tilde{x} \leq x, \tilde{y} \leq y} i(\tilde{x}, \tilde{y}) \quad (1)$$

The AdaBoost learning algorithm is used as a feature selection method. It is used to combine a collection of weak classifiers to form a strong classifier to improve classification results of a learning algorithm as in [3]-[6].

Detectors Cascade is a method to speed up performance without compromising quality (see Fig. 2). A series of detectors called a cascade, is to treat each sub-window. Then, classifiers are combined sequentially from simplest to the most complex. The processing of a sub-window start, from a simple classifier, which was trained to reject most of the negative (non-face) frames, while keeping almost all positive (face) frames. A sub-window only proceeds to the following more complex classifier if it was classified as positive at the preceding stage. If one of the classifiers in a cascade rejects a frame, it is ignored, and the system continues to the next sub-window. As a result, if a sub-window is classified as positive by all the classifiers in the cascade, it is represented as containing a face.

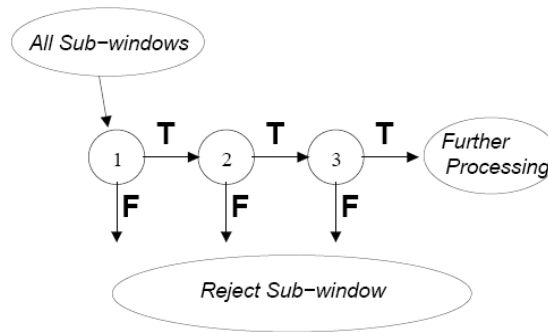


Fig. 2. Detectors cascade of classifiers.

### 3.3. Unscented Kalman Filter for Face Tracking

UKF iteratively estimates the state of a dynamic system based on a sequence of observations and control information. The state distribution can be represented using a Gaussian random variable. It is specified using a minimal set of carefully chosen sample points called unscented transformation (UT) as in [1], [2]. The unscented transformation is a well-known method for calculating the statistics of a random variable that undertake a nonlinear transformation. The main principle of the unscented transformation is that it is easier to approximate a probability distribution than an arbitrary nonlinear function. Assume  $x_k$  denote the state of the system at a given time  $k$ . The two terms,  $u_k$  and  $z_k$ , are the control input and observation at time  $k$ , respectively. The dynamic system develops according to a state transition function

$$x_k = g(u_{k-1}, x_{k-1}) + \varepsilon_k$$

where  $\varepsilon_k$  is additive and zero-mean Gaussian noise with covariance  $Q_k$  that is  $\varepsilon_k \sim (0, Q_k)$ . Also, the

observation  $z_k$  is a function  $h$ , of the current state corrupted by additive Gaussian noise  $\delta_k$  with covariance  $R_k$ ,

$$z_k = h(x_k) + \delta_k \quad (3)$$

where  $g$  and  $h$  are non-linear functions. As a conclusion, even when the estimate of the state  $x_{k-1}$  is a Gaussian, the estimate after transferring the state through the transition function  $g$  is no longer Gaussian. To estimate posteriors over the state space using an efficient Kalman filtering, one, therefore, has to linearize the functions  $g$  and  $h$ . While extended Kalman filters (EKF) do the above linearization using Taylor series expansion around the very current estimate, UKFs apply a more accurate, stochastic approximation, also called the unscented transform [1]. An  $n$ -dimensional random variable,  $x$ , distributed according to Gaussian with mean  $\bar{x}$  and covariance  $p_x$ . The main goal is to estimate a Gaussian approximation of the distribution over  $f(x)$ , where the function  $f$  is a probably non-linear function. This approximation performed using unscented transform (UT) by extracting sigma points  $X$  from the Gaussian estimate and then passing them through. Setting sigma points at the mean and at the mean plus or minus one standard deviation in each dimension (using  $2n + 1$  sigma points) [1], [2].

#### Formulation

Assume  $x$  has mean  $\bar{x}$  and covariance  $p_x$ , a set of  $(2n + 1)$  weighted samples or sigma points are chosen as:

$$x^{[0]} = \bar{x}, \quad (4)$$

$$x^{[i]} = \bar{x} + (\sqrt{n + \lambda} p_x)^i \quad \text{for } i = 1, \dots, n, \quad (5)$$

$$x^{[i]} = \bar{x} - (\sqrt{n + \lambda} p_x)^i - n \quad \text{for } i = n + 1, \dots, 2n. \quad (6)$$

$$w_m^{[0]} = \left( \frac{\lambda}{n + \lambda} \right), \quad (7)$$

$$w_c^{[0]} = w_m^{[0]} + (1 - \alpha^2 + \beta) \quad (8)$$

$$w_m^{[i]} = w_c^{[i]} = \left( \frac{1}{2(n + \lambda)} \right), i = 1, \dots, 2n.. \quad (9)$$

Here  $(\sqrt{n + \lambda} p_x)^i$  is  $i^{\text{th}}$  row (or column) of the matrix square root,  $\sqrt{n + \lambda} p_x$  can be measured from the lower-triangular matrix of the Cholesky factorization. The term  $\lambda = \alpha^2(n + k) - n$  is a scaling parameter;  $\alpha$  actually determines the distribution of the sigma points around  $\bar{x}$  and is usually set to a small positive (e.g.,  $1e - 4 \leq \alpha \leq 1$ ).  $k$  is a second scaling parameter (usually set as 0).  $\beta$  is used to incorporate prior knowledge of the distribution of  $x$  (When  $x$  is normally distributed,  $\beta = 2$  is an optimal value).  $w_m^{[i]}$  is the weight for the mean associated with the  $i^{\text{th}}$  point, and  $w_c^{[i]}$  is the weight for the covariance associated with the  $i^{\text{th}}$  point.

The sigma points are then moved through the function  $f$ , thereby probing how changes the shape of the Gaussian:

$$y^{[i]} = f(x^{[i]}) \quad (10)$$

The estimated mean and covariance of  $y^{[i]}$  are computed according to

$$y' = \sum_{i=0}^{2n} w_m^{[i]} y^{[i]}, p_{y'} = \sum_{i=0}^{2n} w_c^{[i]} (y^{[i]} - y')(y^{[i]} - y')^t \quad (11)$$



where the weights  $w_m^{[i]}$  and  $w_c^{[i]}$  are chosen appropriately (see for more details [1], [2]). The UKF inherits the benefits of the unscented transform for linearization and hence is highly efficient filter.

## 4. Experimental Results

The experiments were conducted using Intel(R) Core(TM) i5 CPU device with 2.03 GHZ and 4 GB of RAM and Window 7 operating system, and Matlab R2014a are used. To test the performance of the system, a comparison between face tracking algorithms is done using three different video datasets. Experiment one, two from vision.ucsd.edu, and the third from University of Stirling Face.

The experiment one for Hector's Video, which composed of 117 frames and every frame has 450 pixels width and 350 pixels height. The experiment two using Behzad's Video, which composed of 159 frames and every frame has a dimension of 450 pixels width and 350 pixels in height. Experiment three using M1044\_V2 Video, which consists of 140 frames and every frame, has a dimension of 400 pixels width and 300 pixels height. The green box refers to the UKF based on Viola-Jones algorithm, and the red box shows UKF based on mean shift with CBWH algorithm. The blue one refers to Camshift algorithm.

### 4.1. Tracking Results Using Hector's Video

By comparing the face tracking algorithms using Hector's Video, the system faces some challenges due to changing the facial expressions, face orientation, and partial or full occlusion as shown in Fig. 3. The tracking result of the proposed UKF based on Viola/Jones method successfully tracked the face under these challenges as declared in Fig. 3.

Also from Fig. 3, the tracking results show that the proposed UKF based mean shift with CBWH successfully tracks the face in the video sequence under the above-mentioned challenges. As clear from the figures, the Camshift algorithm is partially or completely loses face in some frames (frame 87 and 98) due to face rotation and face tiles.



Fig. 3. Tracking results of the two proposed methods (UKF based on Viola-Jones, and UKF based on MS with CBWH) are compared with Camshift algorithm for experiment one.

The tracking trajectories, as seen in Fig. 4, of the two proposed methods comparing with Camshift algorithm, show that the Camshift algorithm trajectory show diversion in some points away from the other two proposed trajectories due lose faces in some frames.

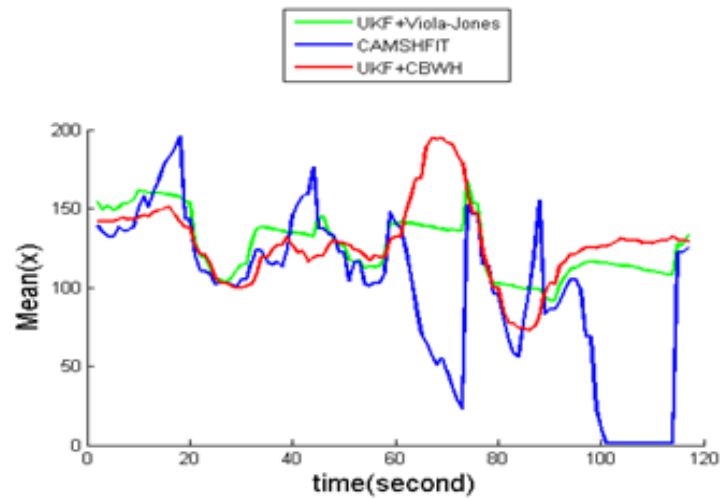


Fig. 4. Tracking trajectories of the two proposed methods (UKF based on Viola-Jones, and UKF based on MS with CBWH) are compared with Camshift algorithm for experiment one.

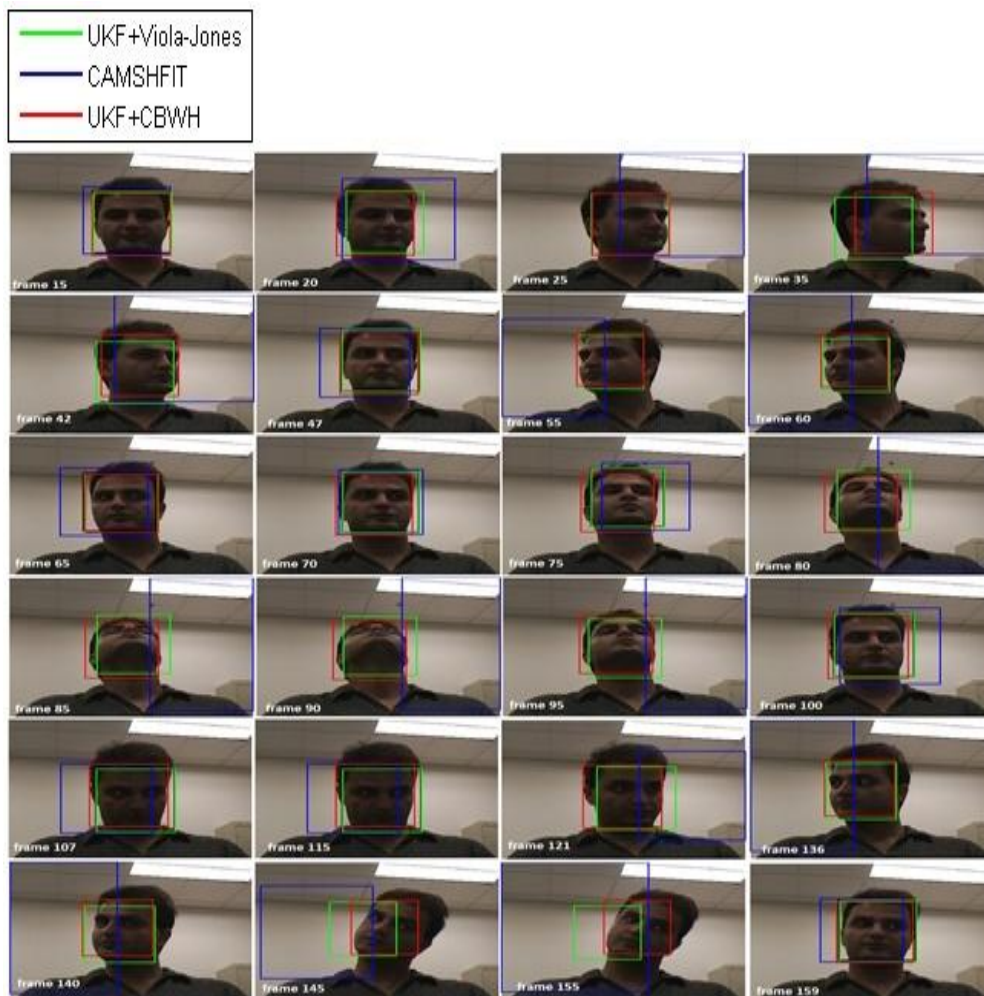


Fig. 5. Tracking results of the two proposed methods (UKF with Viola-Jones, and UKF with MS with CBWH) are compared with Camshift algorithm for experiment two.



## 4.2. Tracking Results Using Behzad's Video

Experiment two is carried to prove the performance of the proposed systems in a complex environment where the orientation of the face (tilts, rotates) and lighting variation. The tracking results for the two proposed algorithms (UKF based on Viola-Jones, and UKF based on MS with CBWH) comparing with Camshift algorithm are shown in Fig. 5. As declared in the figure, the proposed system accurately tracks faces in the video sequence. On the other hand, Camshift algorithm partially loses the faces in some frames (e.g. frames 80, 121, 123 and 155) and completely losses the faces in some frames (e.g. frames 85, 90, 95 and 145) in such environment when the faces are tilts, rotates and lighting variation.

The estimated trajectories of the two proposed systems, as well as the Camshift algorithm trajectory of the face in the video sequence, are shown in Fig. 6. As seen from the figure, the two proposed system's trajectories are nearly closed to each other with little variations. The other hand Camshift trajectory shows high variations far from the two proposed trajectories due to the challenges mentioned above.

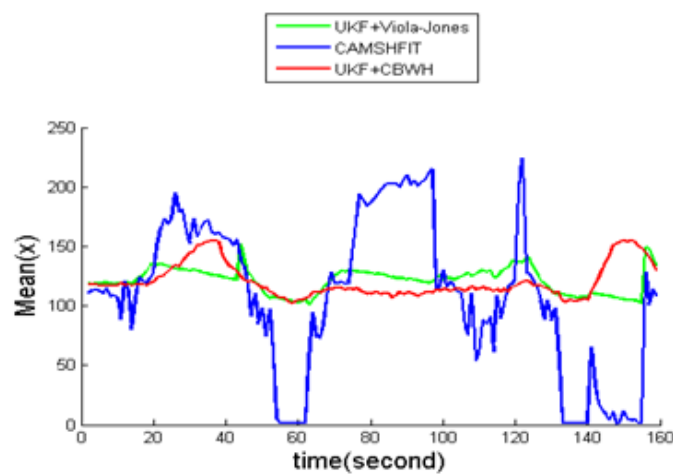


Fig. 6. Tracking trajectory of the two proposed methods (UKF based on Viola-Jones, and UKF based on MS with CBWH) are compared with Camshift algorithm for experiment two.

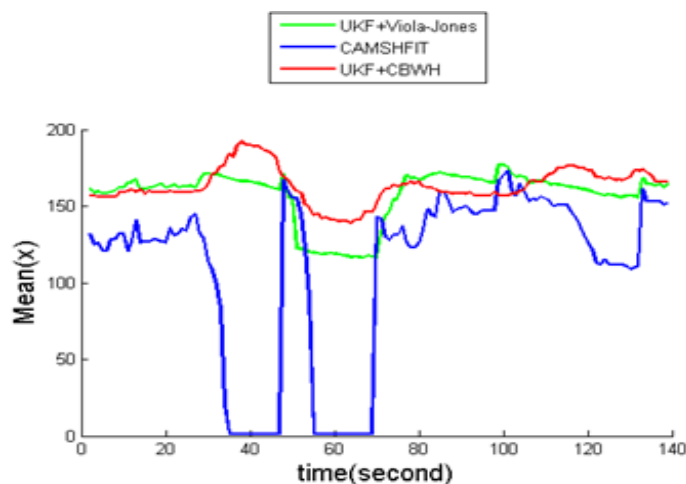


Fig. 8. Tracking trajectory of proposed methods (UKF based on Viola-Jones, and UKF based on MS with CBWH) are compared with Camshift algorithm for experiment three.

## 4.3. Tracking Results Using M1044\_V2's Video

Experiment three is designed to show the robustness of our proposed system in a complex environment

where the face rotates (left, right, up, down), and lighting variation.

Fig. 7 shows face tracking results of the three algorithms. The two proposed algorithms successfully track the faces in the video sequence in a complex environment while the Camshift algorithm loses face partially or completely in some frames (e.g. frames 35, 40 and 46). The estimated trajectories of the two proposed systems appear to be close to each other while the Camshift algorithm trajectory shows high variation due to losing of faces at some tracking time as shown in Fig. 8.

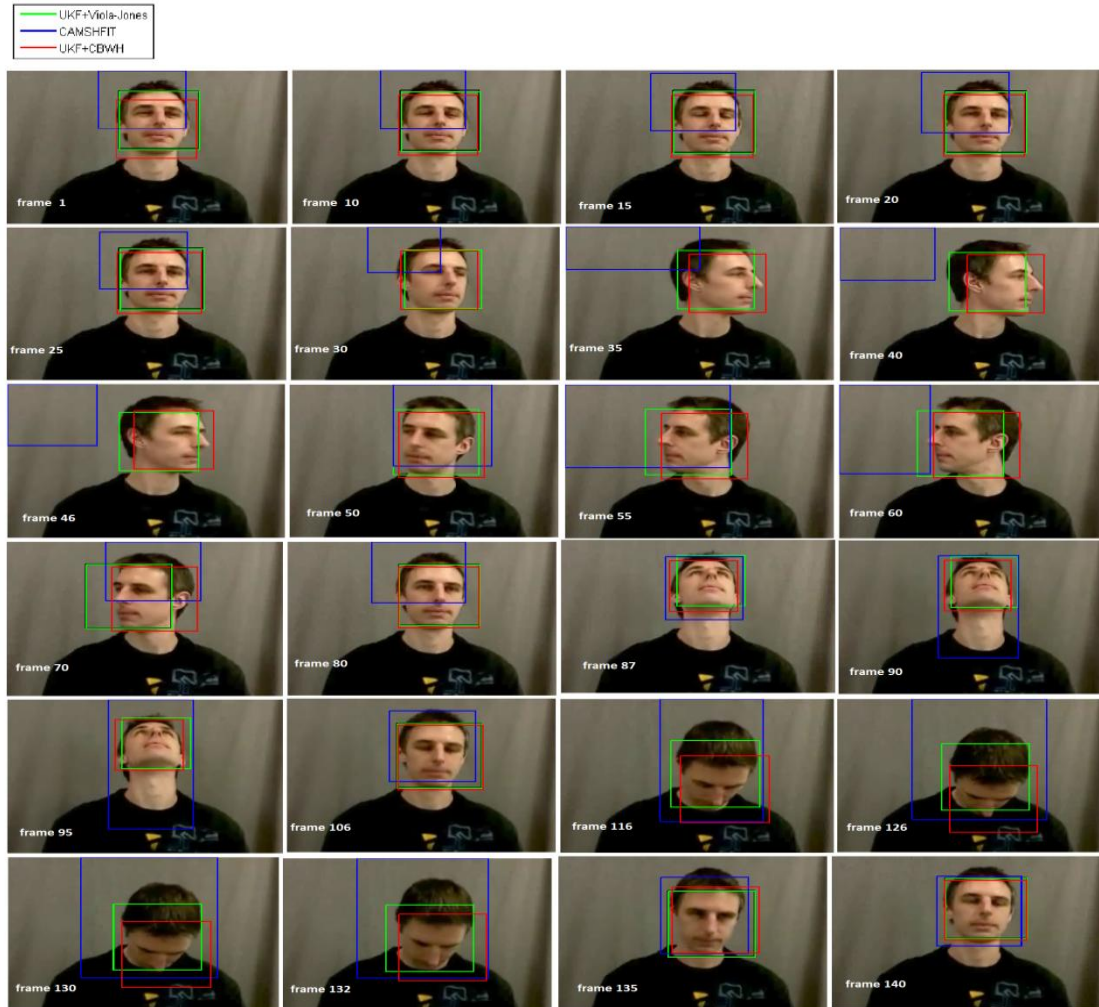


Fig. 7. Tracking results of the two proposed methods (UKF with Viola-Jones, and UKF with MS with CBWH) are compared with Camshift algorithm for experiment three.

#### 4.4. System performance

The system performance of the two proposed methods (UKF based on Viola-Jones, and UKF based on MS with CBWH) and Camshift algorithm are measured using the true positive (TP), true negative (TN), false positive (FP), false negative (FN). The accuracy can be calculated using eq. 12. Measuring the performance using texture features with Co-occurrence matrix is measured similarly to those used in ref. [20, 21].

$$accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (12)$$

As shown from table 1, the two proposed methods give high performance over Camshift algorithm for all the

experiments. Also from the table, the UKF based on Mean Shift gives high accuracy over UKF based on Viola-Jones in experiment one and two when the faces tilt.

Table 1. Performance Evaluation of the Proposed Systems Compared with Camshift Algorithm

Exp.	Num. of frames	Method	TP	FP	Accuracy
Hector	117	Viola-Jones	60	57	51.2821 %
		UKF+Viola-Jones	112	5	95.7265 %
		UKF+CBWH	115	2	98.29059 %
		CAMSHFT	85	32	72.6496 %
Behzad	159	Viola-Jones	77	82	48.4277 %
		UKF+Viola-Jones	156	3	98.1132 %
		UKF+CBWH	158	1	99.37106 %
		CAMSHFT	91	68	57.2327 %
M1044_V2	140	Viola-Jones	68	72	48.5714 %
		UKF+Viola-Jones	140	0	100 %
		UKF+CBWH	140	0	100%
		CAMSHFT	80	60	57.1428 %

Also, as shown from Fig. 9, the first algorithm (UKF based on MS with CBWH) is superior of all for accurate face tracking, then the second algorithm (UKF based on Viola-Jones). Camshift algorithm result is the weakest among all due to losing faces where the face rotates (left, right, up, down), and lighting variation.

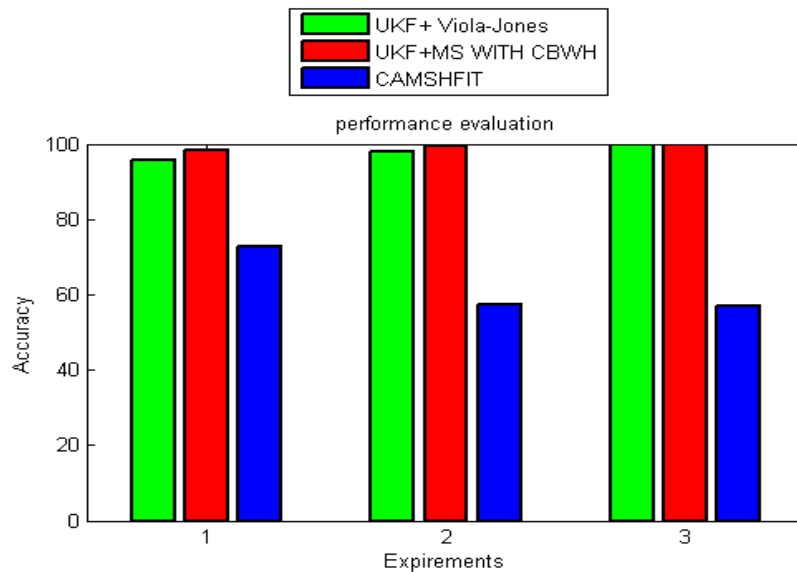


Fig. 9. Performance evaluation of the two proposed algorithm (UKF based on Viola-Jones, and UKF based on MS with CBWH) compared with Camshift algorithm for different datasets.

## 5. Conclusion

In this paper, the two different proposed methods (UKF with Viola-Jones and UKF with MS with BWH) are used. The first method is Viola/Jones algorithm which extremely fast features computation, efficient feature selection, Scale and location invariant detector. Viola-Jones method based on Haar-like features and weak classifiers are currently the most efficient. The second method is mean shift algorithm with CBWH scheme that can effectively reduce background's interference in target localization which consequently can guarantee accurate localization of the target; the UKF algorithm can estimate the coming state with high level of accuracy,

which predicts the position of a moving object based on its past value. The experimental results that the two proposed methods (UKF based on Viola-Jones, and UKF based on MS with CBWH), can clearly overcome the problems of occlusion, lighting variation, orientation and tilts compared with one of the well-known algorithms named Camshift algorithm. Also, UKF based on MS with CBWH show high accuracy than UKF based on Viola-Jones.

As almost all faces successfully tracked by UKF with Viola-Jones, few faces are often not tracked successfully due to the above-mentioned challenges. To improve this system, we can train the classifier with Haar features on these faces, which would develop for a more general face detector. Also, the feature extraction methods will be considered more closely, and more stable features should be found. However, the implemented system provides a reliable basis for further development.

## References

- [1] Julier, S., & Uhlmann, J. (1997). A new extension of the Kalman filters to nonlinear systems. *Proceeding of the 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls*.
- [2] Julier, S. J., & Uhlmann, J. K. (2004). Unscented filtering and nonlinear estimation. *Proceedings of the IEEE*.
- [3] Viola, P., & Jones, M. (2001). Robust real-time object detection. *Proceeding of the Second International Workshop On Statistical and Computational Theories of Vision-Modelling Learning, Computing and Sampling*.
- [4] Ephraim, T., Himmelman, T., & Siddiqi, K. (2009). Real-time viola-jones face detection. *Proceeding of Canadian Conference on Computer and Robot Vision*.
- [5] Viola, P., & Jones, M. (2004). Robust real-time face detection. *Proceeding of International Journal of Computer Vision*.
- [6] Wang, S., & Abdel-Dayem, A. (2012). Improved viola-jones face detector. *Proceedings of the 1st Taibah University International Conference on Computing and Information Technology*.
- [7] Yang, M., Kriegman, D., & Narendra, A. (2002). Detecting faces in images: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- [8] Patilkulakarni, S., & H. Lakshmi. (2010). Segmentation algorithm for multiple face detection in color images with skin tone regions using color spaces and edge detection techniques. *International Journal of Computer Theory and Engineering*, 2(4).
- [9] Rowley, H., Baluja, S., & Giroso, F. (1998). Neural network based face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(1), 23-28.
- [10] Leung, T., Burl, M., & Perona, P. (1995). Finding faces in cluttered scenes using random labeled graph matching. *Proceeding of ICCV*.
- [11] Osuna, E., Freund, R., & Giroso, F. (2010). Training support vector machines: Application to face detection. *Proceeding of IEEE CVPR*.
- [12] Jindal, N., & Kumar, V. (2013). Enhanced face recognition algorithm using PCA with artificial neural networks. *International Journal of Advanced Research in Computer Science and Software Engineering*, 3(6).
- [13] Comaniciu, D., & Ramesh, V. (2000). Mean shift and optimal prediction for efficient object tracking. *Proceeding of International Conference on Image Processing*.
- [14] Jie, L., Jing, W., & Hui-Jia, S. (2008). A motion tracking method based on Kalman filter combined with mean-shift. *Proceeding of International Conference on Wavelet Analysis and Pattern Recognition*.
- [15] Comaniciu, D., Ramesh, V., & Meer, P. (2003). Kernel-based object tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.
- [16] Collins, R. (2003). Mean-shift blob tracking through scale space. *Proceeding of IEEE Conf. Computer Vision and Pattern Recognition*.
- [17] Li, Y., & Zhao, W. (2012). Adaptive updating of kernel band-width for mean-shift tracking. *Journal of Computational Information Systems*.
- [18] Jeyakar, J., Babu, R., & Ramakrishnan, K. (2009). Robust object tracking with background-weighted local

- kernels. *Journal of Computer Vision and Image Understanding*, 112(3), 296–309.
- [19] Yang, Y., Jia, Y., Ong, C., Zhu, Y., Wang, Y., Yue, Z., & Gao, Z. (2013). Object tracking based on corrected background-weighted histogram mean shift and kalman filter. *Journal of Advanced Materials Research*.
- [20] Ning, J., Zhang, L., & C. Wu. (2010). Robust mean shift tracking with corrected background-weighted histogram. *IET Computer Vision*
- [21] Gotlieb, C., & Kreyszig, E. (1990). Texture descriptors based on co-occurrence matrices. *Computer Vision, Graphics and Image Processing*, 51, 70–86.
- [22] Haralick, R. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE* 67.
- [23] Bernhard, F., & Christian, K. (2002). Robust face detection at video frame rate based on edge orientation features. *Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition*.



**Ali Ahmed** is an assistant professor of information technology with the Menoufia University, Faculty of Computers and Information. He holds a B.Sc. in pure mathematics and computer science from Faculty of Science at Menoufia University, an M.Sc. in computer science from Faculty of Science at Suez Canal University and a Ph.D. in computer science from Tokushima University, Japan in 2011. He now works as assistant professor at college of computer, Qassim private colleges, Kingdom Saudi Arabia. His research interests include computer vision, pattern recognition, target motion analysis, and image processing.