



Face detection and tracking using facial features

¹Ali Ahmed and ²Eman Ahmed

^{1,2}IT Department, Faculty of Computers and Information,
Menoufia University, Shebin El-kom, EGYPT.

Abstract: Face detection and tracking algorithms are of great importance, and an extensively studied field and a very large mass of research have been done. In this paper, a new algorithm combining Viola/Jones algorithm with the unscented Kalman filter (UKF) is introduced. Viola/Jones detector is most effective for face detection. On the other hand, Viola-Jones requires full view frontal upright faces in order to be detected, and the entire face must front to the camera and should not be tilted to either side. Whereas, it seems these above constraints could decrease the algorithm's utility somewhat. UKF is combined with Viola/Jones algorithm to enhance the solution of face tracking problems. The experimental results proved that the proposed method gave high accuracy when compared to other traditional face tracking methods

Keywords: Face Tracking, Viola/Jones framework, Unscented Kalman filter, Camshift algorithm

I. INTRODUCTION

Real-time face tracking plays a significant role in many computer vision applications, such as visual surveillance, perceptual user interfaces, etc. Kalman filter is a well known optimal estimator for linear systems, but, on the other hand, very few systems in the real world are linear. A conventional method to treat this problem is to linearize the system first before using the Kalman filter. The result of this step is a new filter called an extended Kalman filter. This linearization does, however, pose some problems, e.g. it may lead to nonstable estimates [1]. However, an alternative approach has emerged over the last filter, namely the unscented Kalman filter (UKF). This Filter claims both higher accuracy and robustness for nonlinear models. Face detection is an important problem in machine vision and is often the first step in many vision-based applications, which is used to determine the location and sizes of human faces in still images or video frames. It is an essential step used in broad range of applications such as face verification/ identification, lip movement capturing, facial expression, feature extraction, video surveillance, an interaction of human-computer visual systems, gender classification.

Face detection methods have been categorized such as knowledge-based, structural matching, template matching, appearance-based and other feature invariant methods including skin color [2], face detection based Viola/Jones [3]. Among the above-mention approaches, Viola/Jones face detection algorithm is a fast feature computation, an efficient feature selection, a scale, and a location invariant detector. It instead of scaling the image itself (e.g. pyramid-filters), it scales the features, and it may be trained for detection of other types of objects (e.g. cars, hands) [3].

Face Detection is used to extract faces in an image regardless of the illumination, facial expression, occlusion [4], pose, scale, etc. The most important or well-defined approaches include the neural network appearance-based system proposed by Rowley, Baluja and Kanade [5]. A graph matching method to find static portable faces from obtained image was presented by Leung [6]. The obtained graphs or nodes are detected from received features. Another face detection method was introduced by Viola and Jones which computes speedily features using integral images [3]. Another method was used support vector machines to detect faces [7]. The PCA analysis was proposed by Kirby and Sirvoich in which each face is represented based on high variances termed eigen faces [8]. Yang and Huang used knowledge-based method to detect faces [2]. Sakai and Silhouettes used several sub-templates to detect frontal faces in images [2]. This paper presents a new algorithm for robust framework for face tracking. It can track even tilted, rotated faces and further away from the camera, which is not possible in case of Viola/Jones algorithm. This work is organized as follows. Firstly, explains Viola/Jones algorithm and UKF. Finally, the experimental results and concluding remark are given.

II. THE PROPOSED FACE TRACKING SYSTEM

A. Viola-Jones Algorithm for Face Detection

Viola/Jones introduced their object detection framework that can be practiced to human faces. At runtime, the algorithm is capable of processing images extremely rapidly real-time frame rates can be accomplished on consumer hardware. It uses a boosted [10, 13] cascade of some simple classifiers based on rectangular Haar-like features that can be computed quickly. The algorithm has a learning step, and both high detection rate and low false positive rate can be reached with a sufficiently large training set. Viola and Jones had three key contributions. First, it introduced a new efficient method for features calculation, based on an "integral image." second it

presented a method for aggressive features selection, based on AdaBoost learning algorithm; and finally, it proposed an idea of combining classifiers in a "cascade."

This work first requires describing an image regarding features that would be helpful indicators to face existence or absence on a given image. For this framework, Viola/ Jones [3], [9], [10], and [11] uses simple features based on pixel intensities rather than uses pixels directly. Viola/ Jones motivated this choice by two main factors: Features can encode ad-hoc domain knowledge, which otherwise would be hard to learn using limited training data. Features-based system works must faster than a pixel-based system.

B. AdaBoost learning algorithm

The complete set of features is very large 160,000 features per a single 24x24 sub-window. Evaluation of the entire set of features is still extremely expensive, and cannot be performed by a real-time application. The main difficulty was to find these distinctive features. The AdaBoost learning algorithm is used as a feature selection mechanism. It mainly used to enhance classification results of a learning algorithm by combining a collection of weak classifiers to obtain a strong one. The algorithm begins with equal weights for all examples. In each iteration, the weight is updated to give more weight to the misclassified examples.

By drawing a relationship between weak classifiers and the other features, Viola/ Jones decided to use AdaBoost method for aggressive selection of a small number of useful features that have significant variety. Practically, the weak learning algorithm was limited to the set of classification functions, which of each was dependent on a single feature.

C. Detectors Cascade

Detectors Cascade is a method to speed up performance without compromising quality². Each sub-window is treated by a series of detectors, called a cascade. Classifiers are combined sequentially from the simplest to the most complex. The processing of sub-window starts, then, from a simple classifier, which was trained to reject most of the non-face (negative values) frames and to keep almost all face (positive values) frames. The given sub-window proceeds to a more complicated classifier only if it was classified as positive at the preceding stage. If any classifier, in a cascade, rejects a frame, it is ignored, and the system consequently moves to the next sub-window. In case a sub-window is classified as positive by all the classifiers in the cascade, it is reported as containing a face.

As most of the sub-windows do not include faces, only a few number of them are treated by all classifiers in a cascade. As Most of the sub- windows are rejected during first few stages, involving simple classifiers solely, which results can be evaluated very rapidly. Through, the overall detection performance meets the requirements of real-time processing, and yet it yields high detection rates when compared with alternative and much slower techniques.

D. UNSCENTED KALMAN FILTER

The extended Kalman filter (EKF) is probably the most widely used estimation algorithm for nonlinear systems. However, it's hard to implement, hard to tune, and only reliable for almost linear systems on the time scale of the updates. These difficulties mainly arise from its use of linearization. To overcome this constraint, the unscented transformation (UT) has emerged as a method to propagate mean and covariance information through nonlinear transformations [14, 15]. It is more accurate, easier to implement, and uses the same order of calculations as linearization.

E. TRACKING FRAMEWORK COMBINING VIOLA-JONES WITH UKF

The tracking algorithm based on Viola-Jones and UKF can track even tilted, rotated faces and even away from the camera, which is the main challenge in the case of Viola/Jones only. Based on the above analyses, the proposed tracking algorithm (UKF-Viola/Jones algorithm) can be developed as shown in figure 1. The proposed tracking algorithm can be summarized as follows:

Step 1: Read video

Step 2: Acquire a frame at time

Step 3: Extract features by calculating the integral image according .

Step 4: Compute AdaBoost learning algorithm to find distinctive features .

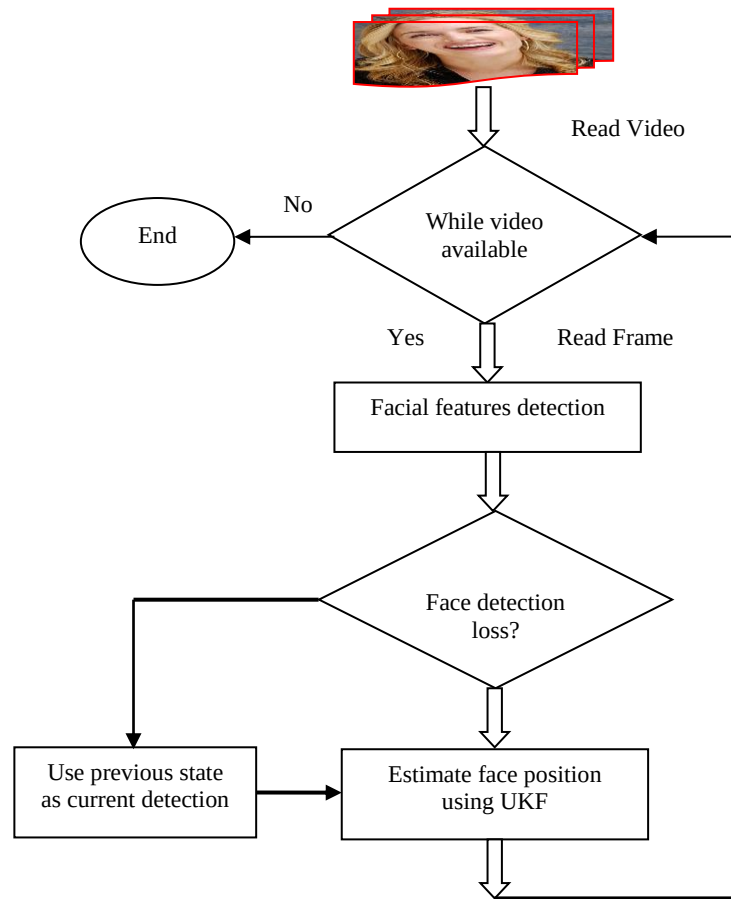
Step 5: Calculate Detectors Cascade (Bbox=detect face(frame)).

Step 6: IF (Bbox<0) THEN Face detected by Viola-Jones algorithm. Estimate face position using UKF.

Step 7: ELSE, use the previous state of UKF as its current state to estimate the next position.

Step 8: Goto step 2.

Figure 1: The proposed face tracking algorithm.



III. EXPERIMENTAL RESULTS

The experiments were conducted using Intel(R) Core(TM) i5 CPU device with 2.03 GHZ and 3.16 GB of RAM and Window 7, and Matlab R2016a are used. Four different video datasets are used the proposed system. First one from done in our laboratory and the second uses Hector's Video. The experiment one use video which consists of 117 frames and every frame has 450 pixels in width and 350 pixels in height. The experiment four, which consists of 159 frames and every frame, has 450 pixels in width and 350 pixels in height. Experiment two and three are also used to proof the effectiveness of the proposed system using low resolution and gray scale datasets. The first, second, and the third experiments are used only to test the proposed system with frontal faces. The fourth experiment is used to detect and track facial features using rotated faces dataset..

a) *Frontal faces expriments*

For testing the proposed system by using dataset 1, the tracking results for the proposed system are compared with the results of Viola/Jones and Camshift algorithm as shown in figure 2. From the figure, the proposed system successfully tracked the face under these challenges. Figures 3 shows the x-axis and y-axis tracking trajectory of the proposed algorithm is compared with other algorithms. Also, the tracking error trajectories using root mean square error (RMSE) of the proposed system and the other methods on x-axis and y-axis are shown in figure 4. Experiment two and three are used to test the proposed system using low resolution and gray scale datasets as shown in figures 5-10. The tracking results show the superior of the proposed system over the traditional algorithms.

Figure 2: Tracking results using dataset1.

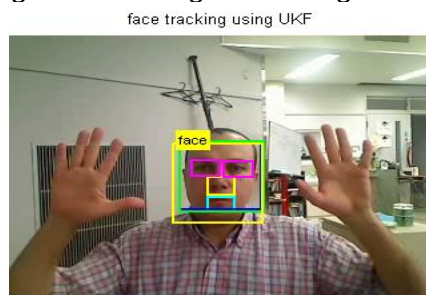


Figure 3: X and Y tracking trajectories using dataset 1.

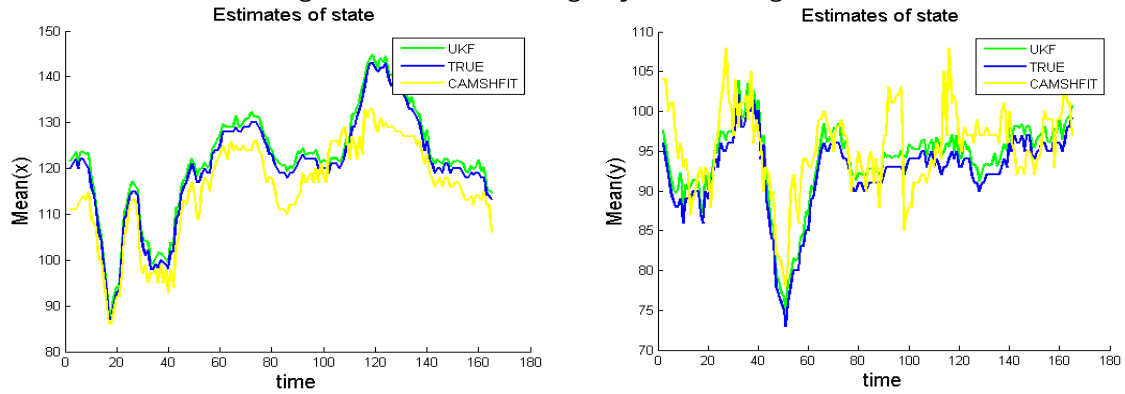


Figure 4: X and Y tracking error trajectories using dataset 1.

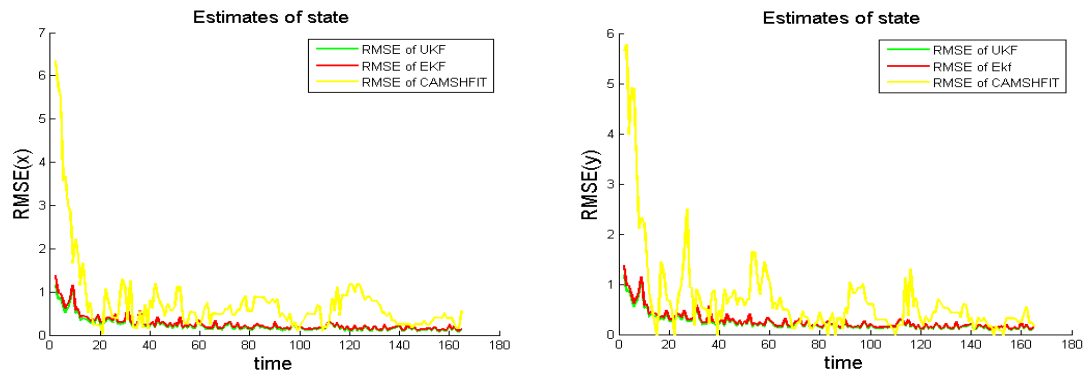


Figure 5: Tracking results using dataset 2.

face tracking using UKF

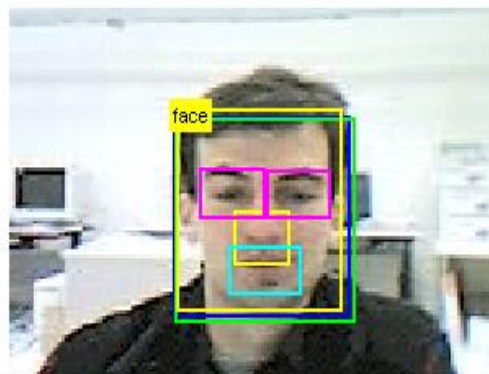


Figure 6: X and Y tracking trajectories using dataset 2.

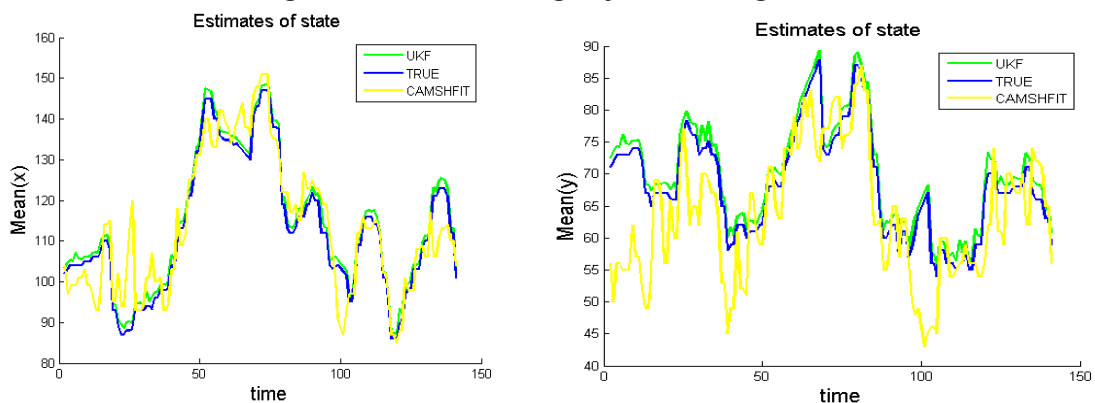


Figure 7: X and Y tracking error trajectories using dataset 2.

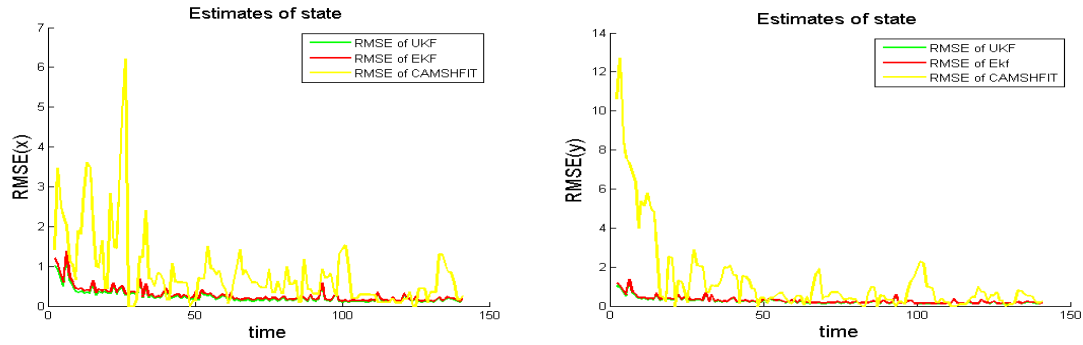


Figure 8: Tracking results using dataset 3.

face tracking using UKF

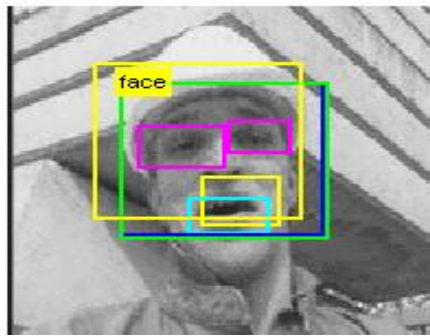


Figure 9: X and Y tracking trajectories using dataset 3.

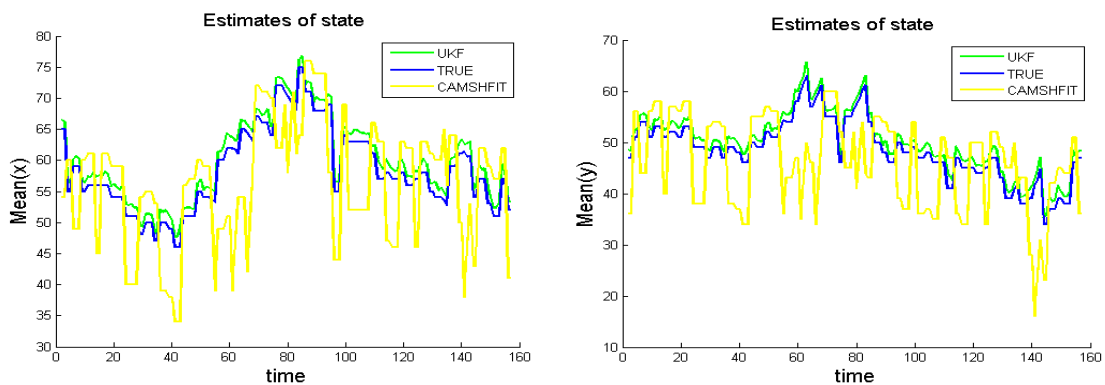
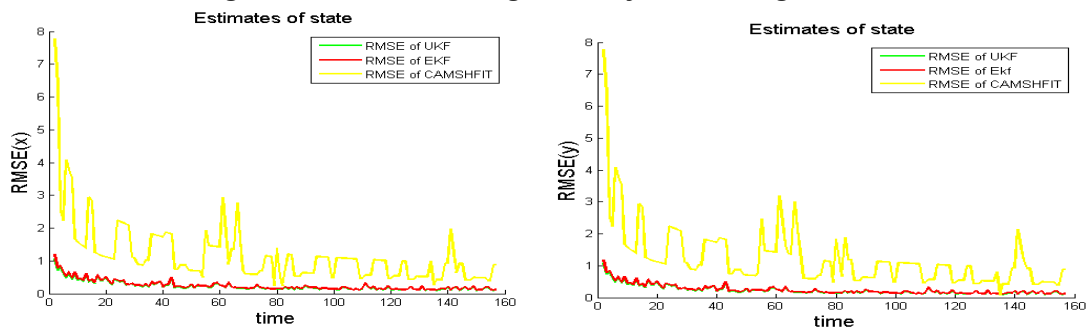


Figure 10: X and Y tracking error trajectories using dataset 3.



b) *Rotated faces experiment*

By testing the proposed system by Hector's Video, the system faces some challenges. Some of these challenges are, changing the facial expressions such as tilted, rotated face or faces occluded as seen in figure 11. The tracking results for the proposed system are compared with the results of Viola/Jones and Camshift algorithm as shown in

figure 11. From the figure, the proposed system successfully tracked the face under these challenges. Figure 12 and 13 show the x-axis and y-axis tracking trajectory and tracking error of the proposed algorithm is compared with other algorithms respectively.

Figure 11: Facial features tracking results using dataset 4.

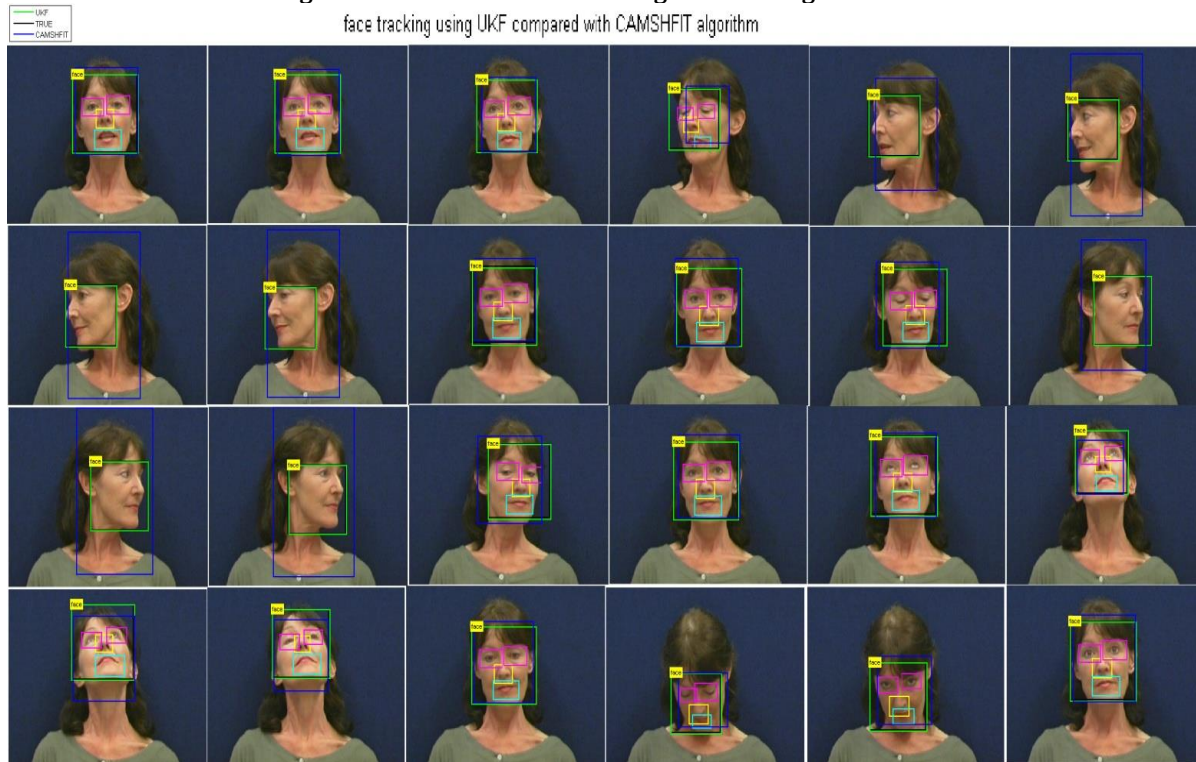


Figure 12: Tracking trajectory using dataset 4.

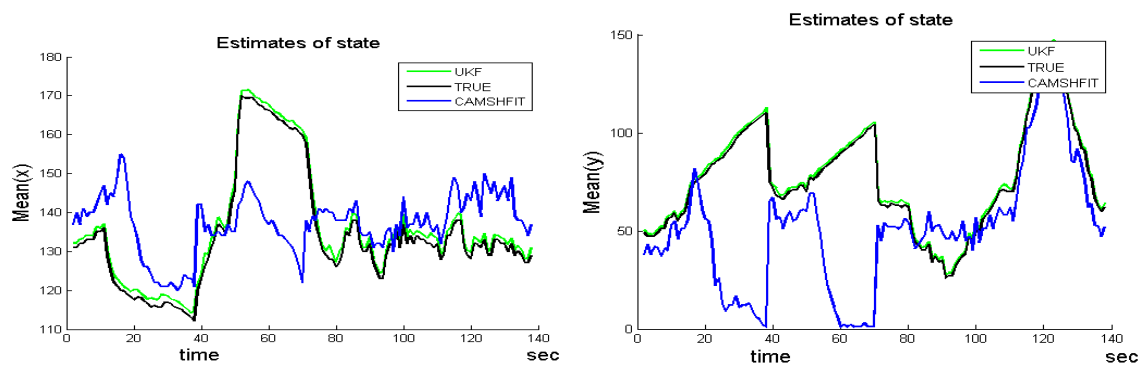
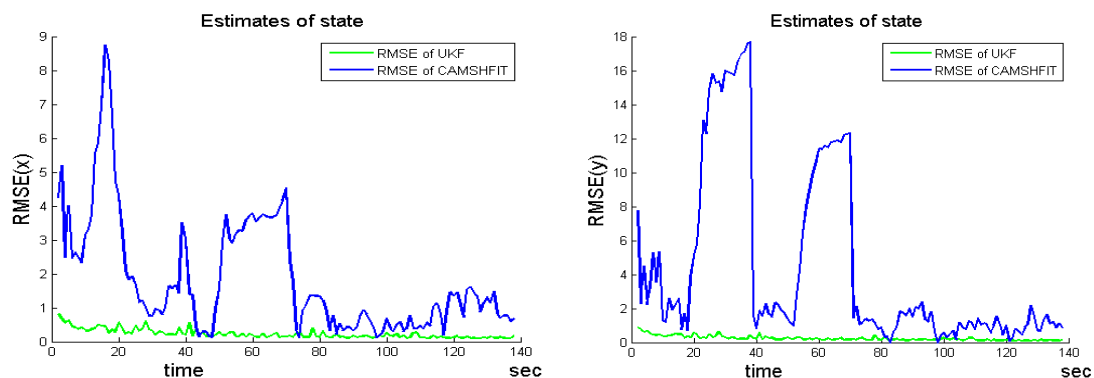


Figure 13: Tracking error trajectories using dataset 4.



As clear from the figures, the Viola-Jones can't detect the face when extremely tilts and rotates. The Camshift loses face in some frames when face rotates or tilts, since face is colored too similarly to the background, and the Camshift depends on hue value which is not sufficient to discriminate the face from the background. On the other hand, the proposed algorithm successfully tracks the face in the video sequence.

c) **System performance**

The performance of the proposed system is measured using the mean and variance of the root mean square error (RMSE). To proof the effectiveness of the UKF- Viola/Jones algorithm, a comparison with EKF and CAMShift algorithms is done in TABLE 1. Using the four experiments, the proposed UKF- Viola/Jones algorithm gives high performance over the other algorithms as seen on the following table.

Table 01: Performance evaluation of the proposed algorithm compared with the other algorithms.

Experiment	Method	RMSE	
		Mean	Variance
Exp. One	UKF	0.22171	0.030293
	EKF	0.25674	0.038223
	CAMShift	0.75289	0.79268
Exp. two	UKF	0.24546	0.031074
	EKF	0.28077	0.03778
	CAMShift	0.83021	0.86925
Exp. Three	UKF	0.22967	0.021723
	EKF	0.24738	0.025288
	CAMShift	1.1644	0.9326
Exp. Four	UKF	0.23137	0.028189
	EKF	0.27413	0.037244
	CAMShift	1.9409	2.9802

IV. CONCLUSION

A robust face tracking algorithm using facial features is proposed, which combines the UKF and Viola/Jones algorithm. The proposed tracking algorithm based on Viola/Jones and UKF can track even tilted, rotated faces and further away from the camera, which is the main challenge in the case of Viola/Jones. Only the experimental results demonstrate that the proposed UKF-Viola/Jones algorithm can efficiently overcome these problems due to the UKF algorithm can estimate the future state with high level of accuracy. Camshift tracking tends to fail when the target face is colored too similarly to the background, or when the target face's color histogram rapidly changes, or hue value is not sufficient to discriminate the object from the background.

REFERENCES

- [1] S.J. Julier and J.K. Uhlmann, A new extension of the Kalman filter to nonlinear systems. In Proc. of AeroSense: The 11th International Symposium on Aerospace/Defense Sensing, Simulation and Controls, 1997.
- [2] Ming-Hausan Yang, David.J.Kriegman, Narendra Ahuja, Detecting Faces in Images: A Survey, IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(1),2002.
- [3] P.Viola, M.Jones, —Robust Real-time Object Detection: Second International Workshop On Statistical and Computational Theories of Vision-Modelling Learning, Computing and Sampling, 2001
- [4] S.Patilkulakarni, H.C. Vijay Lakshmi, 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), 2010.
- [5] H.Rowley, S.Baluja , F.Girosi, Neural network based face detection, IEEE Transactions on Pattern Analysis and Machine Intelligence, 20(1),1998, 23-28.
- [6] E. Osuna, R. Freund, F.Girosi, Training Support Vector Machines: application to face detection, IEEE CVPR, 1997, 130-136.
- [7] T.Leung, M.C.Burl, P.Perona, Finding faces in Cluttered Scenes using Random Labeled Graph Matching, Proceedings OF Fifth International Conference on Computer Vision, 1995.
- [8] Navneet Jindal, Vikas Kumar, Enhanced Face Recognition Algorithm using PCA with Artificial Neural Networks, International Journal of Advanced Research in Computer Science and Software Engineering, 3(6), June 2013.
- [9] Theo Ephraim, Tristan Himmelman, and Kaleem Siddiqi. Real-time viola-jones face detection in a web browser: In Proceedings of the 2009 Canadian Conference on Computer and Robot Vision, CRV 2009, 321-328.
- [10] Paul Viola and Michael Jones. Robust real-time face detection. International Journal of Computer Vision, VOL. 154, 2004, 57:137.
- [11] ShanshanWang and Amr Abdel-Dayem. Improved Viola-Jones face detector: In Proceedings of the 1st Taibah University International Conference on Computing and Information Technology, ICCIT '12, 2012, 321-328.
- [12] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation: Proceedings of the IEEE, 92(3), 2004, 401-422.
- [13] S. Thrun , W. Burgard , and D. Fox. Probabilistic Robotics (MIT Press ,Cambridge, MA, September 2005).
- [14] Simon J Julier and Jeffrey K Uhlmann, New extension of the Kalman filter to nonlinear systems: Int. Symp. Aerospace/Defense Sensing, Simul. and Controls, vol. 3, 1997, 182-193.
- [15] Eric Wan, Ronell Van Der Merwe, and others, The unscented Kalman filter for nonlinear estimation, in Adaptive Systems for Signal Processing, Communications, and Control Symposium, AS-SPCC. The IEEE , 2000, 153-158.