Fast Skin Color Based Face Detection With Improved Morphology And Tracking For Security Applications

[1]Navneet Kaur, [2]Sangeeta Rani
[1]M.Tech Scholar, [2]Assistant Professor
[1,2]Department of Computer Science and Applications
[1,2]Chaudhary Devi Lal University, Sirsa, Haryana, India

Abstract: **Real-time face detection and tracking is a demanding problem in many application circumstances such as faces and gesture recognition in image processing and robotics etc. The ongoing work implements a robust method which can locate various faces at the same time under the varying environment of illumination and complex background in online (real time)by using face detection and tracking in conjunction with depth data. To achieve robustness, the Kanade-Lucas-Tomasi point tracker is used and dedicated to work on facial features by inserting knowledge about the configuration and visual features of the face. The resulting tracker is designed to get better from the loss of points due to tracking drift or temporary occlusion. Performance calculation experiments will be carried out on a set of video sequences of several facial expressions. It will be shown that using the original Kanade-Lucas-Tomasi trackers, some of the points are lost, where as new method has been proposed to overcome the drawback. A performance assessment experiment is carried out on a set of video sequences of several facial expressions.**

Keywords:-**Face tracking, tracker, feature points, Image Processing etc**

I. INTRODUCTION

With human computer interaction technology turning into a newest topic within field of AI, face detection and trailing has become a serious concern analysis direction in computer vision that has broad application prospects within field of human computer interaction, video game so on. The complexness of the face results in a definite degree of issue for quick the detection and trailing. There are various ways for the face detection, for example: mathematics technique, temple t matching technique, the strategy of support vector machine, the strategy of active contour model, the strategy of variability template, etc. however the a lot of thought approach for face detection is Ad boost-based, however, the speed of that doesn't meet the time period in high-resolution video sequence detection. Search candidates face by mistreatment the strategy of color and contour detection, then make sure pattern matching recognition through the algorithmic rule of principal element analysis, the trailing accuracy rate of that is sort of high, however the track result is poor within the event of occlusion. In another thesis, an algorithmic rule, that is combined multiple target hunter with face observe technique, with that is ready to trace the multiple faces at a similar time. This technique has higher time period performance; however the algorithmic rule is a lot of advanced. It takes the Mean-Shift technology to trace moving target that has the quick and economical characteristic. However its trailing window is mounted. Once the target moves at the side of the direction of the camera, it had lost.

Face detection is a crucial requirement in several fields. Detection may be performed by collaboration of people. It can be the case in Human-Machine interactions. It may also be done without the topic being conscious of it. It might be the case in high-security buildings. An event of human central applications explains why face detection has become a broad studied and researched topic. It has inaccurate to mention the matter which has been solved for perspective cameras. However, several ways developed to reach maturity and performance levels that enabled them to makeover from educational labs to daily lives. Nowadays, folk's use of face discovering algorithms while not realizing it took years of a machine equipped with camera will currently detect person's face in time period, that's to mention quick because image is streamed by capturing device. Viola and Jones weren't pioneers within field. Numerous researchers already dedicated themselves to present topic. However, Viola and Jones' scheme has driven by trade-off between detection concerns and hardware concerns. As result, the technique that achieved unexampled classification performance, whereas victimization has extremely economical process step. These approaches continue to inspire various researchers within search an ideal face detector. All the same, always highlight every one way were developed for the perspective cameras. However, many disadvantage of the perspective cameras their field of read. Thus, people got to enter FOV to act with the machine. It can be not ideal for various things. One may cite videoconference things wherever the FOV is thus tiny that only 1 person may be seen within the image.

The opposite folks round the table can't be seen. Most of time, this disadvantage is overcome by adding some additional cameras, an answer that causes subsequent issues. Within the case of building direction, if one desires to observe a sq. space, one should use several cameras. It is used to increased the FOV of image device will circumvent all this additional bother. In this, such sensors exist already. They're referred to Omni directional vision sensors. As the name indicates, they'll look altogether directions. Omni directional vision device theory and applications area unit well established. Most of applications of this technology may be found within the field of AI. OVS area unit primarily won't to solve navigation and motion estimation issues. The utility of

position vision sensors has conjointly been incontestable within the 3 dimensional reconstructions of scenes. The mathematical tools used area unit supported laptop vision pure mathematics theory. Within the mentioned applications, position pictures area unit regarded and processed as provided by the device. This can be possible once the idea involves matching points, pixels within the case of laptop vision, or tiny regions restricted to the 8-neighborhood of the thought-about element. Several low-level options like Scale Invariant Feature remodel key points or Harris corner detector are with success changed to require into consideration position image geometrical specificity. These recent results have incontestable that position pictures may be directly processed. Object detection will imply point-of-interest matching. Still, the methodology most well-liked by the image process community is predicated on region-based descriptors. The matter analysis to form them becomes advanced on position vision sensors. As a matter of reality, the device resolution may be non-linear on many dimensions. One cannot utilize detectors trained for perspective pictures. In fact, neither the descriptors nor the ensuing classifier will handle this non-linearity. a typical strategy is to synthesize a pseudo-perspective image via a technique referred to as anti-anamorphous. The wrapped position image may be absolutely unwrapped and projected onto a cylinder or a sphere. One may also domestically project the region of interest to be processed on a plane. The ultimate step is to use a detector trained on perspective pictures onto the ensuing footage.

II. RELATED WORK

Dewan, M. Ali Akber et al. [1] proposed an adaptive biometric system. This proposed system is used to solve the problem of single sample face recognition (SSFR) having modular framework. This architecture have one detector per target and for every detector, face model is being produced within the face images. Sequential Karhunen-Loeve technique is used to update the face model using representative face captures that are choosing for operating data by utilizing reliable tracking trajectories. This method is used to produce the intra-class variation of face appearance and improve representativeness of the face models. The efficiency of presented technique is detailed in security surveillance and user authentication using Chokepoint and FIA datasets in SSFR setting.

Karmakar, Dhiman et al. [2] in this paper, facial feature extraction scheme has been proposed. By utilizing the segmentation method, control points are extracted and color images are cropped automatically. In this segmentation method, initially images are cropped automatically and then these cropped images are being applied for detecting the main connected components. The feature points are found by utilizing the geometrical measurement of location and the size of component with no previous information if probabilistic distance between the feature points or using any feature point extraction formula. By containing main feature points, T shaped face image is generated. Finally rate of recognition of unprocessed face images using PCA is reported.

Agrawal, Samiksha et al. [3] in this paper, face detection approach has been proposed. This technique is depending on the Viola and Jones algorithm and principal component analysis. The simulation results demonstrated that the presented approach is efficient and performs better as compared to already existing approach.

Sudha, N., and D. Bharat Chandrahas et al. [4] presented a parallel architecture for face recognition which is suitable for implementing in multicore environment. The proposed system includes the modules for video frame acquisition, PCA, binarization, skin detection on the image frame. Consecutive frame lines are processed in pipeline on multicores. The proposed system attains a frame rate at 8 frames per second for 480x272 image size and the experimental result is area- effective.

Kheirkhah, Esmaeil et al. [5] presented a hybrid technique depending on the adaboost face detection and skin color information. The feature points of hybrid technique are background elimination and down-sizing. The feature point is computed by using segmentation method and utilizing adaptive skin colour classification. In this segmentation method, initially images are cropped automatically and then these cropped images are being applied for detecting the main connected components. Though, Viola-Jones Adaboost-based face detector is being applied in this paper as the final face detector. To verify the precision of presented hybrid technique, then an experiment is performed on some standard datasets and then proposed system is compared with Viola-Jones face detection system. The experiments indicate that the presented technique might effectively enhance the face detection system in terms of precision (98.88%) and detection time (259.59 ms).

Vong, Chi Man et al. [6] proposed sparse Bayesian extreme learning machine (SBELM) for real-time face detection. The proposed method is capable to reduce the size of model without compromise on the fast execution time and precision. Numerous standards for the face datasets are developed for the estimation of SBELM among other traditional approaches. The analysis of results indicates that SBELM attains best ever execution time with high precision over the standard face datasets.

Hatem, Hiyam et al. [7] presented a well-organized and fast face detection method in face images and videos. Initially color filtering method is being adopted in order to recognize the face quickly; it skips that area which is not matching with candidate faces. in this paper, haar like features is being used having adaboost algorithm which build a stronger classifier consists of weak classifiers and including facial sections to enhance the recognition precision. In this paper, motion model estimation has been adopted. The result indicates that the computational cost of proposed method is lower and provides better performance in factors which affects appearance of faces like change in expression, goggles, spectacles, long hairs, beard etc.

Shakir, Haidar R [8] presented a method for detection of faces. This proposed method is depending on the learning algorithm to scan the images so as to detect the human face. In such system, issues like pose variation, scale variation, etc. occurs. Initially first

stage in proposed method includes the improvement of image. It will be significant stage specifically if the image has been acquired under unconstrained illumination condition. In the next stage, segmentation of the skin in HSV space is done and hole filling algorithm is utilized to refine the skin segmentation results. The input image edges and skin tone are fused in oder to ensure that all non-face regions are removed from candidate faces,

Zhao, Qian, Shuzhi Sam Ge et al. [9] presented a method to learn a type of saliency features, which merely makes response in face regions. Based on the saliency features, a joint pipeline is designed to detect and recognize faces as a part of human–robot interaction (HRI) system of SRU robot. By restricting false responses, the performance of face verification can be improved, especially when the training and testing are implemented on different dataset. In experiments, the effects of saliency term on face verification and benchmark discriminative ability of saliency features on LFW are analyzed. And the effectiveness of this method in face detection is verified by the experimental results on FDDB.

Aghaei, Maedeh, Mariella Dimiccoli et al. [10] presented multi-face tracking method which produces a group of tracklets and takes benefit of tracklets redundancy to deal with unreliable ones. Same tracklets are the combined into bag of trackeles that is aimed to correspond to a particular person. This method is a widespread dataset of egocentric photo-streams and compared with other methods, indicating its efficiency and robustness.

III. METHODOLOGY

Kanade-Lucas-Tracker Algorithm

It is simple for human eye to follow another person's head while she or he moves around. So, the same challenge for the computer, but there is a different story. In the world of computer vision tracking it's a wonderful task. There are mainly three steps that has to take place while human face tracking:

- Finding the face in an image place
- Opting the features to track
- Tracking the features

The main important step before detection is the selection of tracking features. The keyword during feature selection is texture. Area with a modifying texture pattern, is mostly unique in an image, during linear or uniform intensity areas are common and not unique. It is clear that we must have to look at areas of texture or the first logical step would therefore be an investigation the image intensity gradients.

A texture pattern can only exist if we look at several pixels in an area. The features which we are tracking are accurately described as feature windows including texture information. The area of feature window may be vary, that depending on the total number of features which requires being located.

$$\mathbf{g}\mathbf{g}^{\mathrm{T}} = \begin{bmatrix} \frac{\delta(I)}{\delta x} \\ \frac{\delta(I)}{\delta y} \end{bmatrix} \begin{bmatrix} \frac{\delta(I)}{\delta x} & \frac{\delta(I)}{\delta y} \end{bmatrix} = \begin{bmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{bmatrix}$$

$$Z = \iint_W \begin{bmatrix} g_x^2 & g_x g_y \\ g_x g_y & g_y^2 \end{bmatrix} \omega d\mathbf{x}$$

Now integrates the matrix derived over the area W, that we get :

Tracking facial features

Let assume that human face is already located in an image sequence, and then the tracking is based on Kanade-Lucas tracking equation. In the end of technique can be described as follows: Firstly you have to select all the features which can be tracked from image to image in video image stream. The opting of features is based on texture. A number of fixed sized features are selected on the head of person in first image of sequence. These windows features are tracked from one image to next using the KLT method. The tracking method calculates the total sum of squared intensity differences b/w a feature in the previous image or the features in current image. The replacement of the particular feature is then defined as the replacement that minimizes sum of differences. It is done continuously between sequential pictures so; all the features can be tracked.

The main fundamentals of tracking are explained by looking at two pictures in an image sequence. Let's assume that the first picture was captured at one time t and the second picture at time another time t+_. It is important to remember that incremental time _ depends on the frame rate of an image/video camera and should be as small as possible. A higher frame rate lets for better

tracking. A grayscale picture is a pattern of intensities, where intensity values range from 0 to 255. In this stage you can define an error function that has to be minimized in order to minimize the noise effect

$$\epsilon = \iint_W [A(\mathbf{x} - \mathbf{d}) - B(\mathbf{x})]^2 \omega d\mathbf{x}$$

Now, we have covered the mathematics portion, so we look at what happens physically while the execution of the tracking algorithm. The tracking steps are: Firstly we assume that one feature window in the first picture and we want to determine the replacement of this window in the second image. In the first picture, this is used to calculate the Z matrix in 10. The vector is calculated with subtracting in second frame from the first frame and multiplying the conclusion with the gradient calculation and a weighting function applied to. To find the displacement equation has to be minimized. Processing an image with high resolutions is comparatively expensive. Looking that face tracking is a real-time application, we just require to minimize the processing time to keep up with a moving person. The algorithm can be used in a multi-resolution picture pyramid for processing the images.

Algorithm

1. Acquire image from the video
2. Convert from RGB to YCbCr
3. Apply skin color based segmentation
4. Apply morphological operations
5. Obtain face and hand as blobs
6. Apply spatio-temporal segmentation based on velocity of the blobs
7. Obtain the facial region
8. Apply eigen feature extractor
9. Use the feature for the tracking algorithm
10. Apply KLT tracking algorithm and track the face
11. Repeat steps 1-8 for each frame in the video

IV. RESULTS

The results proposed for spatio-temporal segmentation and skin color based segmentation. The result has been applied to produce an application that works on both live videos and recorded. A user can interfere at any time to tag a detected face in frame, when software application is running; then the vector feature of tagged face may be evaluated and stored in gallery. Because of the parallel computing in several levels and different modules of system, critical work has been conducted in implementation to avoid the conflicts among individual threads.

In this, MATLAB's inbuilt functions are utilized in many segment of implementation. The software application with 320x240 frame size has been tested on the videos. This face detection method works once every 20 frames. The computation official feature for face recognition takes 44ms per frame. For the purpose of tracking, least amount of eigen values are taken as feature points. In dynamic collection, every subject has approximately five clusters and each one carry maximum 20 samples.

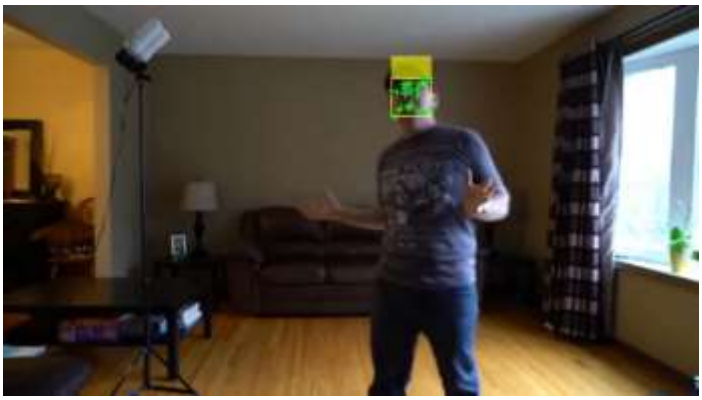

Fig 1.

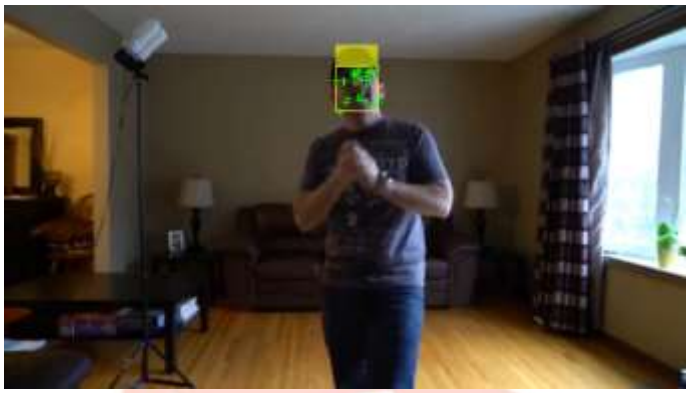

Fig 2.

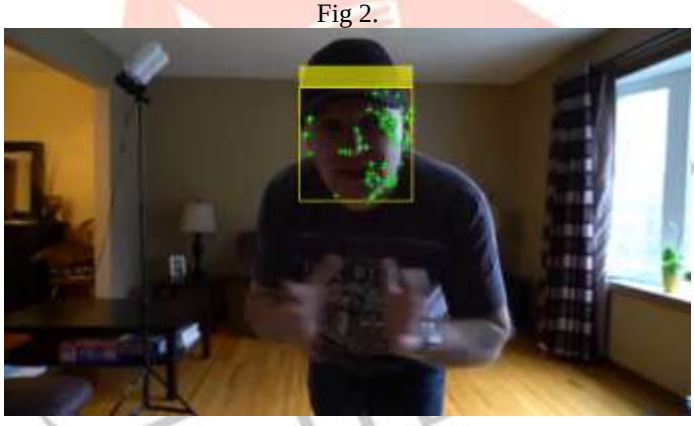

Fig 3

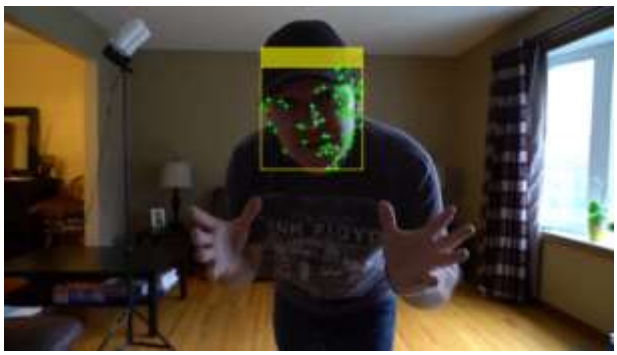

Fig. 4

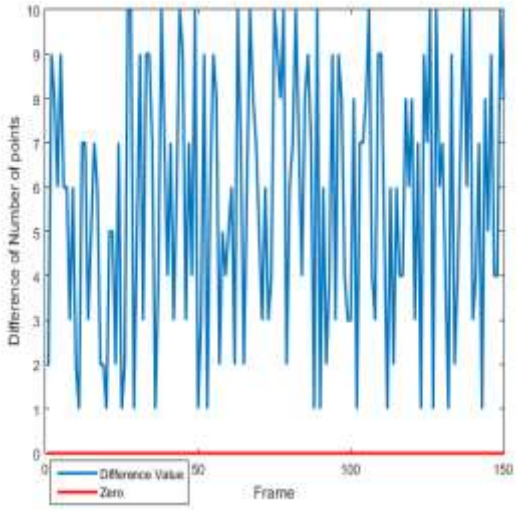

Fig.5 Comparison graph for both method

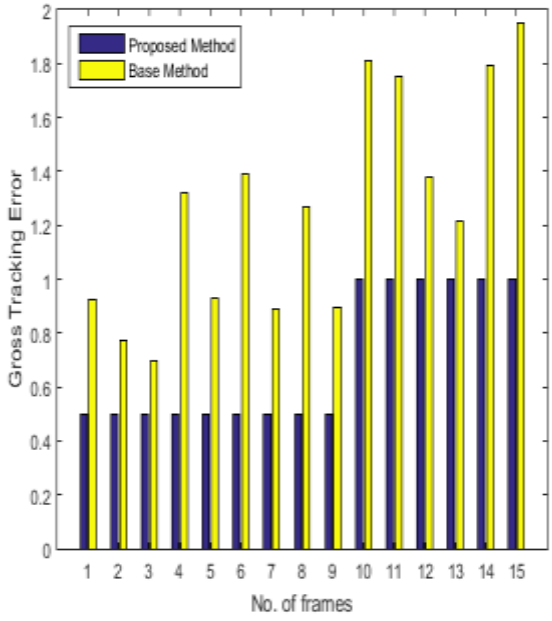

Fig. 6 Comparison graph for both method

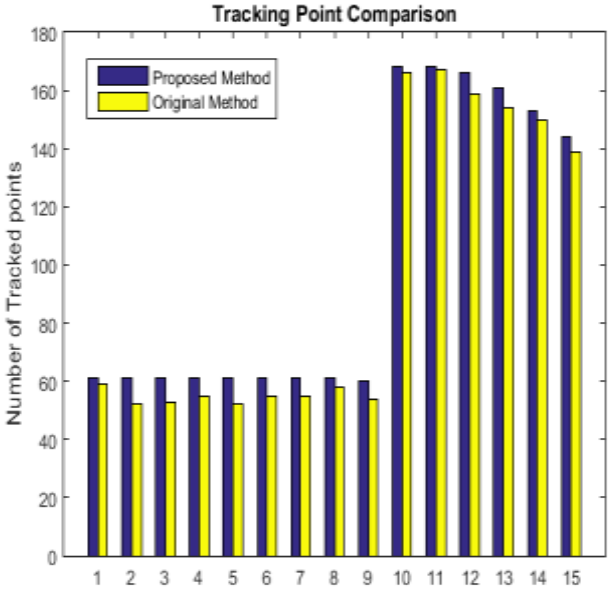

Fig. 7 Comparison graph for both method

V. CONCLUSION

In this work, a facial identification and tracking system for security purpose use is presented. This new method is intended for use in robust recognition of facial information such as the identity or facial expression of a person. Therefore methods are required to overcome the problems and develop a new system for detection as well as tracking. this paper formulate the issues in the existing method that is degrading the quality of either detection or tracking of the facial region and to further implement a new and improved methodology to overcome the problem occurred earlier. MATLAB are used for the generation of the results which is published in this paper and it allows immediate identification and strong detection of face in high-resolution videos. Future research is expected to be carried out in a different of aspects to improve this work. Also, the system can be made robust against occlusion.

REFERENCES

[1] Dewan, M. Ali Akber, Dan Qiao, Fuhua Lin, and Dunwei Wen. "An Approach to Improving Single Sample Face Recognition Using High Confident Tracking Trajectories."In *Canadian Conference on Artificial Intelligence*, pp. 115-121.Springer International Publishing, 2016.

[2] Karmakar, Dhiman, and C. A. Murthy. "Face Recognition using Face-Autocropping and Facial Feature Points Extraction."In *Proceedings of the 2nd International Conference on Perception and Machine Intelligence*, pp. 116-122.ACM, 2015.

[3] Agrawal, Samiksha, and Pallavi Khatri. "Facial expression detection techniques: based on Viola and Jones algorithm and principal component analysis." In *2015 Fifth International Conference on Advanced Computing & Communication Technologies*, pp. 108-112.IEEE, 2015.

[4] Sudha, N., and D. Bharat Chandrahas. "A pipelined memory-efficient architecture for face detection and tracking on a multicore environment."In*VLSI Design and Test (VDAT), 2015 19th International Symposium on*, pp. 1-2. IEEE, 2015.

[5] Kheirkhah, Esmaeil, and Zahra Sadri Tabatabaie. "A hybrid face detection approach in color images with complex background." *Indian Journal of Science and Technology* 8, no. 1 (2015): 49.

[6] Vong, Chi Man, KengIam Tai, Chi Man Pun, and Pak Kin Wong. "Fast and accurate face detection by sparse Bayesian extreme learning machine."*Neural Computing and Applications* 26, no. 5 (2015): 1149-1156.

[7] Hatem, Hiyam, ZouBeiji, and RaedMajeed. "Human Facial Features Detection and Tracking in Images and Video." *Journal of Computational and Theoretical Nanoscience* 12, no. 11 (2015): 4242-4249

[8] Shakir, Haidar R., Loay E. George, and Ghadah K. Tuama. "Single Face Detection Based on Skin Color and Edge Detection." (2015).

[9] Zhao, Qian, Shuzhi Sam Ge, Mao Ye, Sibang Liu, and Wei He. "Learning Saliency Features for Face Detection and Recognition Using Multi-task Network." *International Journal of Social Robotics* (2016): 1-12.

[10] Aghaei, Maedeh, MariellaDimiccoli, and PetiaRadeva. "Multi-face tracking by extended bag-of-tracklets in egocentric photo-streams." *Computer Vision and Image Understanding* 149 (2016): 146-156.