A Comprehensive Review of Face Detection and Tracking Techniques

[1]Satbir Kaur, [2]Er. Aayushi Chadha
[1]Research Scholar, [2]Assistant Professor
[1,2]Universal Group of Institutes Lalru, Punjab, 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 using Supervised Descent Method (SDM) based Viola-Jones and Skin color based segmentation, 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 detection and tracking method 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 has been shown that using the base methodologies, some of the points are lost, whereas new method has been proposed to overcome the drawback.**

KEYWORDS: KLT, SDM.

I. INTRODUCTION

Face detection and its tracking in images is among the main issues in the computer vision. Tracking and detection methods may use rigid representation to illustrate the facial region, thus it can be said that they can neither exploited nor captured the non-rigid facial deformations. A face classifier is employed to assign identities to the trajectories.. Generally non-rigid deformations are captured by the locating the position of facial landmarks set such as nose, eyes, mouth etc. This is partially credited to the availability of huge quantity of annotated data, several numbers of which have been given by the first facial landmark localization challenge called as 300-W challenge. These days face detection gives efficient results for those images which are captured randomly (referred as"in-the-wild"). Some of the current face detection systems are prompt enough to be essential parts of very well-liked electronic commodities, like several types of cameras. Generally, for purpose of effectiveness and efficiency, rectangular or ellipse-like shape has been used for rigid representation to illustrate facial region. Face detection and its tracking is become other field of research which has received extensive attention in the past decade. In the most recent generic object tracking benchmarks that carries sufficient amount of short face videos.To realize time period face tracking, this combined motion history image with Mean-Shift to unravel the development of missing and false trailing within the case of existing obstructions. Algorithmic rule method includes quick face detection of initial state and target trailing of continuous state and target lost state. Firstly, we want to extract the predict region of face through interval frame distinction technique, color detection, the proportion of 5 sense organs and alternative data and observe it by Adaboost technique. Then track the sequence image combined with MHI, per the obtained face regional location, size, direction data etc. This technique has been verified by a well-behaved experimental, and therefore the results show that the algorithmic rule can't solely agency one target and multiple targets, however even have a high accuracy, robustness.

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. 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. 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.

Various customers are indispensable information for the management and call public places like large-scale markets, airports, looking centres, museums, stations, etc. once getting knowledge of client variety, managers will apportion resources rationally. On the other hand, it'll bring in smart operation effectiveness; on opposite hand, it can give most effective service for customers. Ancient individuals consider systems embody mechanical system and infrared detector detection system, however they need disadvantages in occupancy of resources and applied math effectiveness. Within the current society, video observation system is progressively applied in several aspects of our life. Getting helpful info by analysis and process of video image has become a very important analysis

direction of video analysis. Video-based individuals tally technology belongs to intelligent video observation scope; its study involves laptop vision, image process, pattern recognition, artificial intelligence, etc. So, this is often a difficult forefront space. Many of us measure researches are created supported video image process. In paper, the author used face detection to search faces, then used Kalman filter to estimate the movement methods of the individuals, finally counted individuals by classifier of movement meth ods. Within the camera ought to be decorated from ceiling of the gate, distinction to extract and victimisation background estimation and track persons, then count once they passes the tally line. In this, grid-based and foreground objects example matching theme is worked to robustly verify every pedestrian. The intensity profile analysis classify an each component of frame as moving, background component or stationary, then classified all non-shadow pixels as stationary or moving kind blobs, ultimately compared the blob size with the standard individuals size to estimate individuals variety. During this method the video observation image entrances of public places like laboratories, school rooms, supermarkets, etc. In these scenes, the lighting conditions are stable comparatively that guarantee individuals are facing observation devices via adjustment of angles of the devices, in order that it'll be comparatively examine for face detection and area-tracking. Once get video frame, first of all a tendency to set the individuals detection and track space, then search if there's anyone passing. Once searches the targets, track them. If middle purpose of track target passes the space, judge it and restart the detection operate in meantime.

II. RELATED WORK

Sagonas, Christos et al. [1] proposed 300 Faces in-the-Wild Challenge (facial landmark localization Challenge). The purpose of this challenge is to evaluate the performance of distinct systems on a new-collected dataset using the identical evaluation protocol and the same mark-up and thus to expand the first standardized benchmark for facial landmark localization.

Asthana, Akshay et al. [2] discussed the issue of updating a discriminative facial deformable model. in this paper, very effective strategies has been proposed in order to update a discriminative model which is trained by cascade of regressors. It is show that is possible to automatically construct robust discriminative person and imaging condition specific models 'in -the-wild' that do better than state-of-the-art generic face alignment strategies.

Baltrusaitis, Tadas et al. [3] proposed Constrained Local Neural Field model for facial landmark detection. In this model probabilistic patch expert (landmark detector) has been introduced which may learn non-linear and spatial relationships among the input pixels and the probability of a landmark being aligned. Moreover, this model is optimised by utilizing a novel Non-uniform Regularised Landmark Mean-Shift optimisation approach.

Dahal, B., Abeer et al. [4] in this system, hybrid system has been presented which uses the skin color for enhancing precision for tracking and detection. In this approach, an image is divided into different part depending on their skin color and also imag e is rescaled. The portioned skin color image is united with the edges before applying the morphological operations. The result indicated that proposed hybrid system outperforms as compared with other traditional system.

Goswami, Gaurav et al. [5] in this paper, 3 dimensional images has been used that may encode many information related to face. But, 3 dimensional images are costly to capture because of the sensor cost. In RGB-D images can be acquired using consumer level devices like Kinect that gives pseudo-depth data. Moreover in this paper, RGB-D face tracking algorithm has been proposed which is depending on extracting biased features using entropy and saliency from RGB-D images. The result demonstrates that proposed algorithm is very efficient and understands the several facets of RGB-D face recognition.

Shi, Xin, Jian Wu et al. [6] presented a face tracking method which is depending on the threshold determination of the positive face sequence. In this method, initially all the faces having distinct angles are find and tracked. And also the position and angel information of face is being noted. In the next stage, different faces are matched with the face in the frame. Then positive faces that are similar to the person are detected. In final stage, results can be attained by using threshold determination method. The result analysis shows that the proposed method has better rate of recognition in case of moderate flow density. Also, the proposed app roach is capable to satisfy the requirement of the real-time system.

Dewan, M. Ali Akber et al. [7] 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 detecto r, 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.

Vong, Chi Man et al. [8] 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.

Negi, Rahul Singh et al. [9] presented a new method in order to implement face recognition using PCA. Although, there will lots of work is done on PCA, the major disadvantages that was happened with the distance issue. In this paper, problem is being solved by utilizing Hausddroff distance algorithm.

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

Human face detection and tracking is an important research topic for many applications in computer vision. However, detecting human face in images or videos is a challenging problem due to high variety of possible configurations of the scenario, such as changes in the point of view, different illumination, and background complexity. The method of face detection and tracking is connection of two stages –

1. Facial feature extraction using Principal Component Analysis and,
2. Motion tracking based on points extracted.

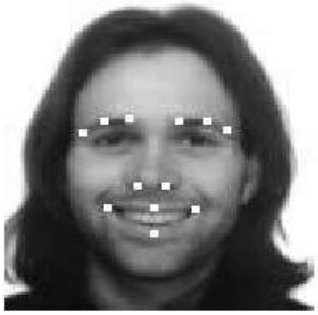

Fig.1 Facial feature extraction

The main basic information has now been build to solve the replacement D of face feature from one image window to next window. For simply, we define the second window as B(x) = I(x,y,t+_) and the first window A(x-d) = I(x-d) = I(x- x, y-y,t) as where x= (x,y). Now, the relationship between the two images is given :

$$B(\mathbf{x}) = A(\mathbf{x} - \mathbf{d}) + n(\mathbf{x}).$$

In the equation, n(x) is a noise function which is caused by the ambient interference like reflections and scattered light. 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, which 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. The tracker works on grey level image sequences of any length. The general tracking process is shown in Fig2.

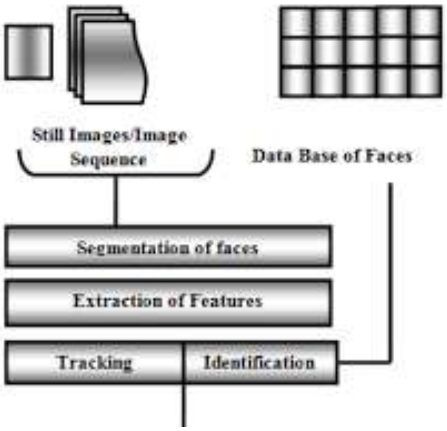

Fig 2. Flow chart

Algorithm

1. Acquire image from the video
2. Apply skin color based segmentation
3. Obtain face and hand as blobs
4. Apply SDM Viola Jones with skin color segmentation based on velocity of the blobs
5. Obtain the facial region
6. Apply eigen feature extractor
7. Use the feature for the tracking algorithm
8. Apply KLT tracking algorithm and track the face
9. Repeat steps 1-8 for each frame in the video

IV. RESULTS

The results presented are there for the proposed skin color based segmentation along with SDM Viola Jones with skin color segmentation. The solution has implemented to create an application which works on both recorded and live videos. When software application is running, a user can interfere at any time to tag a detected face in frame. A vector feature of tagged face can be computed and saved in gallery. Due to the parallel computing in various levels and distinct modules of system, critical work has been conducted in implementation to avoid the conflicts among individual threads. In this, MATLAB's inbuilt functions are used in multiple portions of implementation. The software application with 320x240 frame size has been tested on the videos. These experiments made on various mainstream laptops, desktop and PCs, each contains up to two multi-core processors. Consistent performance has been attained from PCs. Face detection process works once every 20 frames. The calculation official feature for face recognition takes 44ms per frame. For tracking purpose, minimum eigen values are considered as feature points. Dynamic face modeling can be called once every seven frames. In dynamic collection, each subject has almost five clusters and each one contain maximum 20 samples. In the dynamic face modeling, subject has been identified much faster when someone reappears in video of standard benchmark videos, it is not practical to compare the detection accuracy with prior arts.

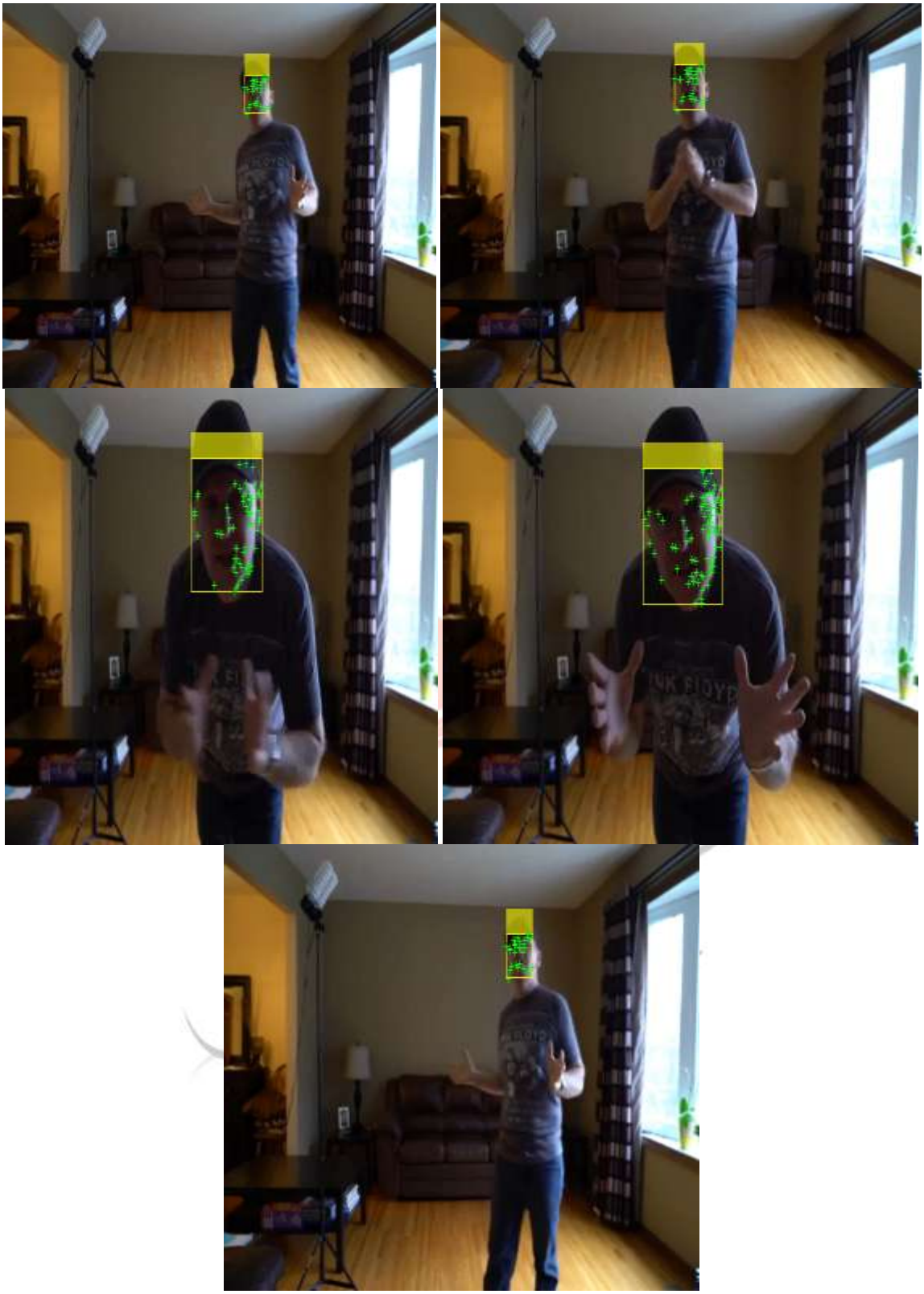

Fig 3 Image frames with tracking features

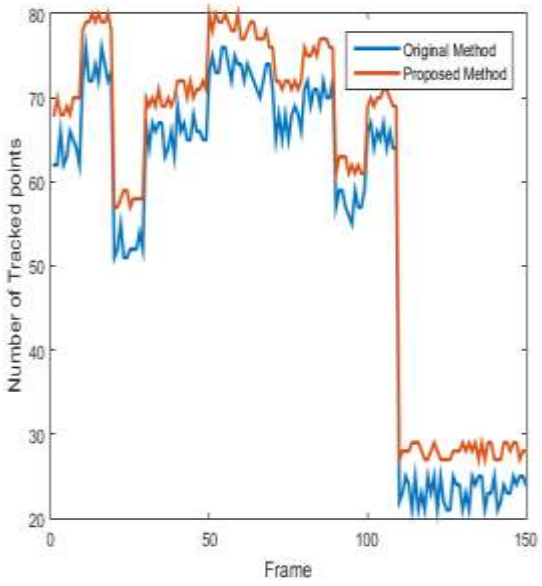

Fig.4 Comparison graph for both method

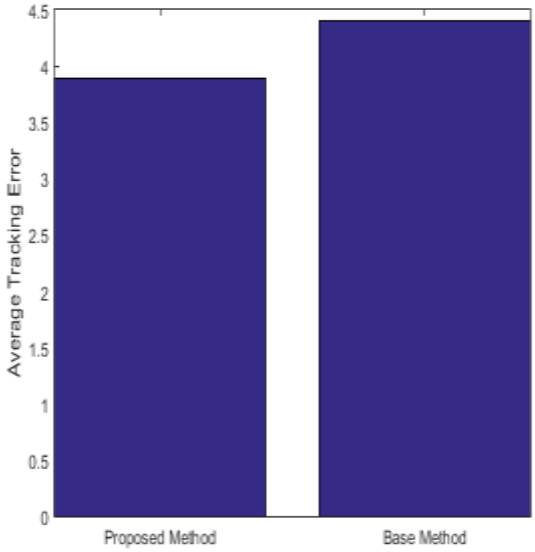

Fig. 5.Error Comparison graph for both method

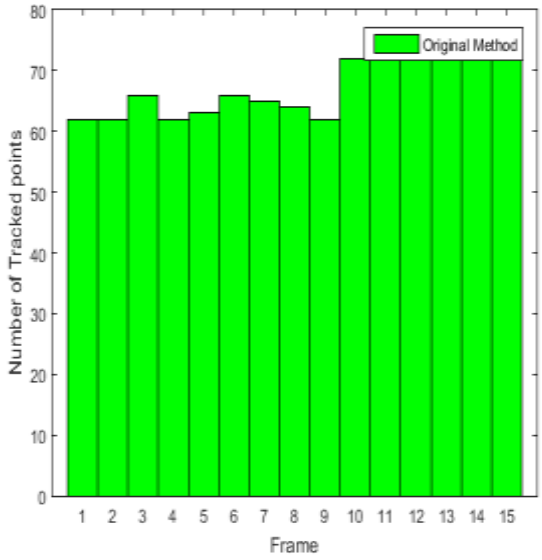

Fig.6 Tracked points graph for base method

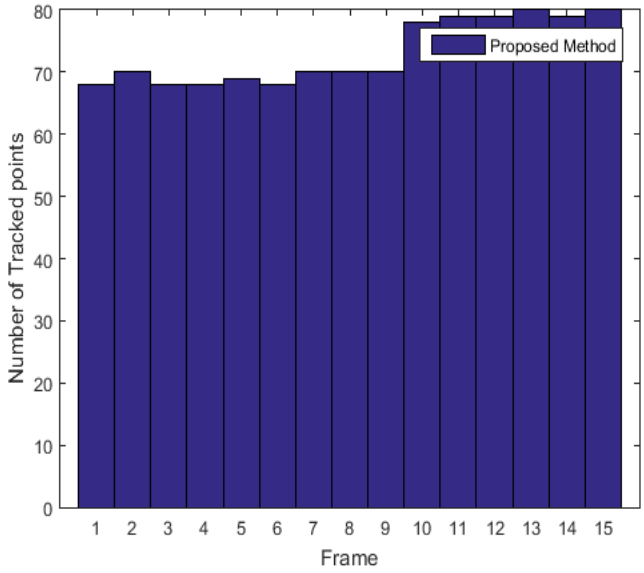

Fig. 7. Tracked points graph for proposed method

V. CONCLUSION AND FUTURE SCOPE

In this work, a facial identification and tracking system for security purpose use is presented. A novel method, of face detection using SDM Viola Jones with skin color segmentation is proposed and innovations were made to image examination modules for effectiveness and robustness. The results are published in this paper and are generated using MATLAB which 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. More efforts can be made to enhance up the tracking method so that the system run faster, and specifically, the system's cost may reduce. Also, the system can be made robust against occlusion.

REFERENCES

[1] Sagonas, Christos, Georgios Tzimiropoulos, Stefanos Zafeiriou, and Maja Pantic. "300 faces in-the-wild challenge: The first facial landmark localization challenge." In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 397-403. 2013.
[2] Asthana, Akshay, Stefanos Zafeiriou, Shiyang Cheng, and Maja Pantic. "Incremental face alignment in the wild." In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1859-1866. 2014.
[3] Baltrusaitis, Tadas, Peter Robinson, and Louis-Philippe Morency. "Constrained local neural fields for robust facial landmark detection in the wild." In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pp. 354-361. 2013.
[4] Dahal, B., Abeer Alsadoon, P. W. C. Prasad, and Amr Elchouemi. "Incorporating skin color for improved face detection and tracking system." In *2016 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, pp. 173-176. IEEE, 2016.
[5] Goswami, Gaurav, Mayank Vatsa, and Richa Singh. "Face Recognition with RGB-D Images Using Kinect." In *Face Recognition Across the Imaging Spectrum*, pp. 281-303. Springer International Publishing, 2016.
[6] Shi, Xin, Jian Wu, Xin Ling, Qiao-li Zheng, Xuan-qi Pan, and Zhi-qiang Zhao. "Real-Time Face Recognition Method Based on the Threshold Determination of the Positive Face Sequence." In *Proceedings of the 22nd International Conference on Industrial Engineering and Engineering Management 2015*, pp. 125-136. Atlantis Press, 2016.
[7] 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.
[8] Vong, Chi Man, Keng Iam 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.
[9] Negi, Rahul Singh, and Rekha Garg. "Face Recognition Using Hausdroff Distance as a Matching Algorithm." *International Journal of Scientific Research* 4, no. 9 (2016).
[10] Aghaei, Maedeh, Mariella Dimiccoli, and Petia Radeva. "Multi-face tracking by extended bag-of-tracklets in egocentric photo-streams." *Computer Vision and Image Understanding* 149 (2016): 146-156.