

Multimodal verge for scale and pose variant real time face tracking and recognition

G. Ramkumar, E. Logashanmugam

Department of Electronics and Communication Engineering, Sathyabama Institute of Science and Technology, India

Article Info

Article history:

Received Oct 28, 2018

Revised Nov 21, 2018

Accepted Dec 3, 2018

Keywords:

Biometrics
Face Recognition
Face Tracking
Image Analysis
Simulation

ABSTRACT

In recent times face tracking and face recognition have turned out to be increasingly dynamic research field in image processing. This work proposed the framework DETecting Contiguous Outliers in the LOW-rank Representation for face tracking, in this algorithm the background is assessed by a low-rank network and foreground articles can be distinguished as anomalies. This is suitable for non-rigid foreground motion and moving camera. The face of a foreground person is caught from the frame and then it is contrasted and the speculated pictures stored in the dataset. Here we used Viola-Jones algorithm for face recognition. This approach outperforms the traditional algorithms on multimodal video methodologies and it works adequately on extensive variety of security and surveillance purposes. Results on the continuous demonstrate that the proposed calculation can correctly obtain facial features points. The algorithm is relegate on the continuous camera input and under ongoing ecological conditions.

Copyright © 2019 Institute of Advanced Engineering and Science.
All rights reserved.

Corresponding Author:

G. Ramkumar,
Research Scholar, Department of Electronics and Communication Engineering,
Sathyabama Institute of Science and Technology,
Jeppiaar Nagar, Rajiv Gandhi Salai, Chennai – 600119, Tamilnadu, India
Email: pgrvlsi@gmail.com

1. INTRODUCTION

The major task is to track and detect the face, as correct as conceivable in every one of the frames of a video. A visual person tracking has been a standout amongst the most mainstream look in the computer perception area [1]. Especially, human face tracking over video receive more consideration, which would permit helpful useful applications. In other case, face tracking is as yet a tricky which can't be viewed as fathomed. This is because due to the feature of face appearance change which increase the tracking complexity [2]. Face tracking is not straight forward because more variations in the image appearance, like pose variation (frontal and non frontal), impediment, picture introduction, lighting up condition and facial expression [3]. Face recognition play a vital role in different fields like business, restorative or military frameworks. Face recognition are utilized as a part of real security or religious spots and regions like air terminals and other very delicate. Even though recognition parts facing some issue due to some factors, in that pose variation is one of the major nuisance factor [4]. However, to certainly and effectually utilize the multi-vision video information, we regularly need to appraise the posture of the individual's head. While there are numerous strategies for multi view pose estimation but finding the head position is significant issue, particularly when the quality of the images is poor and the standardization of cameras isn't adequately exact to permit robust multi-vision fusion [5]. Such a situation is particularly valid with regards of surveillance. We propose the face tracking and recognition of person from multi-vision videos.

The effective algorithm for face tracking is done by DECOLOR which approximates the background and the foreground objects simultaneously. Gabor feature extraction is carried out as the next step person identification and recognition is done by Viola Jones Algorithm. For a given multi-vision video arrangements,

we utilize a composite template to track the 3D area of the head utilizing multi-vision information. The proposed method performs better than the existing highlights and algorithms on a multi-vision video database composed utilizing a camera. Face tracking is a crucial preceding step that limits the region of the face in video frames, from which a appropriate feature set can be extracted and consequently supplied as input to the face recognizer. As such, the efficiency of tracking directly controls the ambitious to identify subjects in video.

Face tracking has received special attention in the vision community [6]. Exact tracking is not easy because of changing appearance of targets due to their non rigid structure, 3D motion, interface with other objects and changes in the environment like lighting variations. This method detects human face by the geometric connections between's area of face and hairs in each edge of a video document [7]. Desirable face regions can be figure out using the range of skin color in order to at first confine the face, besides, the plausible squares in an image outline are dictated by methods for spectrums. Consolidated skin and squares conclude applicant face areas in light of the geometric connection. The stage connection movement estimation calculation for the most part used to looks at the successive edges in a video arrangement to group faces that are in movement and track the human appearances from the video record. With 10fps frame rate, the efficiency of single-face tracking is approximately closer to 100%. Video includes more number of datas than images [8]. An immediate method to deal with single view recordings is to exploit the information excess and perform see determination. One hypothetically conceivable arrangement is to apply a brightening standardization strategy to restrict the light varieties impact before tracking. In any case, this isn't a powerful arrangement in light of the fact that the brightening standardization calculations not perform well in low resolution face images. More calculations were created for different applications and utilized unsuccessfully. In any case, these calculations are very troublesome and difficult to face the continuous prerequisites of specific frame-rate. Thus, the proposed can be most likely transplanted to an embedded framework, like the emerging little robot to do dynamic face detection and tracking.

2. PROPOSED METHODOLOGY

In this work we endeavored to consolidate the face tracking and recognition. First the decoloring is done in that background learning and object detection is done simultaneously. In genuine recordings, the front objects generally are small packages. In this manner, neighboring areas ought to be chosen to be recognized. An image or a frame is captured from a real-time video source [9]. Then the face region is detected and after that the detected face is sent for recognition using Viola-Jones algorithm, as shown in Figure 1.

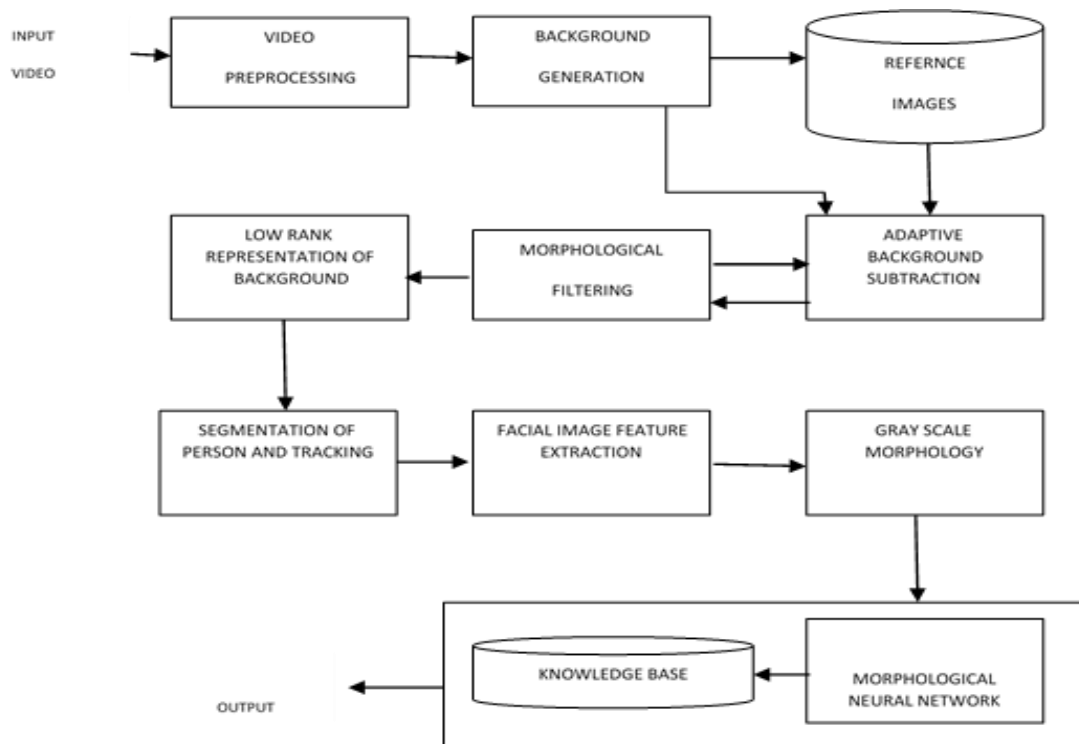


Figure 1. Block Diagram of Proposed System

There are 3 key strides for machine-controlled video examination such as identification, tracking and recognition. In the initial step, object identification targets to trace and section noteworthy questions in a video. At that point, that items can be followed from edge to edge, and the tracks can be prompted to perceive object Behavior. Accordingly, object assumes an indispensable part in practical usances [10]. Object detection in a video is generally accomplished by object identifier or background subtraction procedures. In this, we tend to demonstrate that the above difficulties can be tended to in a unified framework which is call it as sleuthing Contiguous Outliers into the low-rank illustration (DECOLOR). This construction coordinates object identification and background learning into a solitary procedure of incorporation, and it can naturally model complicated background and avoid the complicated computation of foreground motion [11].

An object locator is routinely a classifier that outputs the picture by a sliding window and names each sub picture portrayed by the window as either question or start of a video. Then again, background subtraction system contrasts the images with a foundation model and finds the progressions as articles. It regularly expect that no object shows up in images when assembling the background display. Such arrangements of fundamentals cases for object or background modeling absolutely reduce the application of above-noticed techniques in mechanized video examination. Frequently, an object detector have to need to physically qualified cases for prepare a binary classifier, when background subtraction requires a preparation grouping that has no objects to develop a background model [12]. To preset the exploration, object identification without an individual preparing stage transform into a significant errand. Individuals have endeavored to handle this operation by utilizing movement data. Examinations on both simulated information and real arrangements demonstrates that DECOLOR exceed the best in class techniques and it can work productively on an out scope of complex situations, as shown in Figure 2.

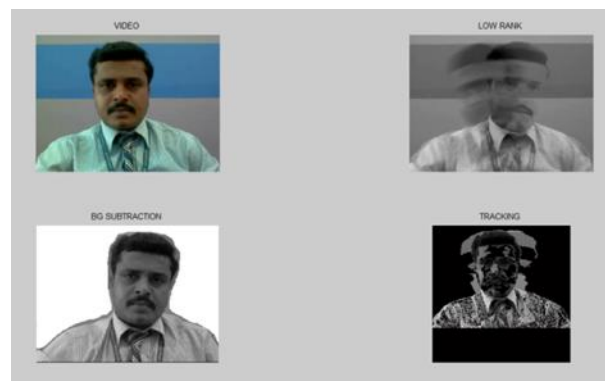


Figure 2. Decoloring process from the input video

The next step - highlight extraction-includes realizing appropriate facial highlights from the information. These highlights could be specific face regions, varieties, edges, which can be human significant or non significant. This area has some different usances like facial feature tracking or feeling recognition [13]. Finally, the framework distinguishes the face. In an acknowledgment undertaking, the framework would report a integrity from a database. Gabor highlights extricate local bits of data which are finally merge to identify an object or locale of intrigue [14].

The primary finding was the dynamic connection engineering which presented Gabor jet idea. A set of Gabor templates with various frequencies and introductions might be useful for extracts the essential features from an picture. Utilization of Gabor filters in feature extraction can be defended by organic discoveries in vision frameworks, common picture statistics and achievement in prevailing applications [15]. Refinement of their determination and convenience advances their use additionally in upcoming applications. Gabor filters leads an important role in the application of computer vision, more practical is due to their success in face detection, recognition, and all the biometric techniques. Feature extraction using gabor template is given by

$$\Psi_{(x,y)} = f_2 \pi \beta \alpha e - (f_2 \beta_2 x_{02} + f_2 \alpha_2 y_{02}) e j^{2\pi f x}$$

$$X_0 = x \cos \theta + y \sin \theta$$

$$Y_0 = -x \sin \theta + y \cos \theta$$

where

f --central frequency of the template,
 θ – Degree of the rotation angle,
 β – Major axis (bandwidth) and
 α – Minor axis (sharpness)

Aspect ratio of the gaussian function is given by α/γ . Frequency domain function for the given form is

$$\Psi(a, b) = e - \pi^2 f_2 (\beta_2 (a'-f)^2 + \alpha_2 b'^2)$$

$$a' = a \cos \theta + b \sin \theta$$

$$b' = -a \sin \theta + b \cos \theta$$

Utilizing a classifier, as basic as Gaussian mixture models in the facial feature models in the facial component used to distinguish and perceive complex genuine or old structures in images. Face recognition is a quickly growing up innovation, generally utilized as a part of criminal recognizable proof, secured access, and jail security [16]. The machine learning and PC designs groups are likewise continuously associated with face recognition. Moreover, there are a more number of business, securities requiring the utilization of face recognition technologies. Face recognition has intrigued much consideration and its exploration has rapidly spread out by engineers as well as neuroscientists

The sample video is given as input. The major step in pre-processing is the input video is converted into frames. Likewise to enhance the picture to guarantee the accomplishment of further procedures. (i.e) enhancing contrast, evacuating noise, identifying the data rich areas [17]. From the input video, at each frame the background varies slightly. These backgrounds are considered as $n, n+1, \dots$ such generated backgrounds in each frame is noted and recorded here. The images of the persons we are tracking will be stored in the database. So that the test image is compared with the reference images stored and the tracking is done. In Adaptive background subtraction, the background of each frames (i.e) $n, n+1, \dots$ in a video is subtracted only the backgrounds alone are separated so that the person/object can tracked down easily. It compares the images with a background demonstrate and identifies the adjustments in object [18]. Morphological filtering is for enhancing the image such as smoothing or simplification, noise suppression. Majorly it contributes in removing the artifacts (noise) that are introduced while processing the image. The actual image was initially changed to RGB-CbCrCg color space. At that point the skin elements were divided in view of the prospective skin detection system portrayed beforehand. Therefore morphological sifting was connected to decrease false positives. At last the face detection recognized utilizing Viola-Jones face detector.

The proposed work for the face detection are implemented by Matlab software. Note that the morphological operators were actualized utilizing the capacities (imerode and imfill) worked in Image Processing Toolbox, while Viola-Jones algorithm was given by Computer Vision System Toolbox. From the algorithm we use, DECOLOR where the background is approximate by the low rank matrix [19]. The person/object from the image is segmented to track. Facial image extraction provides the features of the tracked person. If the input data is too large, then it can be it can be changed into a diminished arrangement of features.

Generally the extracted features contain applicable information from the input; with the goal that the favored undertaking should be possible by utilizing this reduced representation rather than the entire initial data. The test image is converted into gray scale in order to process it in this step. In this, the data base is trained to identify the person we are tracking by providing the images that are stored [20]. In knowledge base the where images are stored and where we compare the test image and then they provide the result whether it matches with the database image or not. Finally the face of the tracked person is recognised and provides the authentication if the image matches with the image put away in dataset. If both images are coordinated, the access is conceded. Else the access will be denied to the specific person.

3. SIMULATION RESULTS AND PERFORMANCE ANALYSIS

The improved form implements an arrangement of channels self-comparative, i.e. measured and turned variants of each other, inspite of the recurrence f and orientation θ . Gabor features, alluded to as Gabor jet, multi-resolution Gabor feature, are developed from reactions of Gabor filters in the above conditions by utilizing numerous channels on few frequencies f_m and orientations θ_n . Crude features are the complex esteemed reactions of an arrangement of multi determination Gabor channels as lit up in Figure 3.

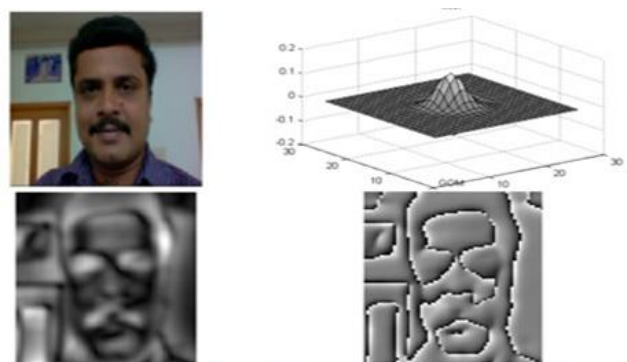


Figure 3. Gabor Feature Extraction

4. CONCLUSION

In our work, we proposed an approach for face recognition by integrating gabor feature extraction technique with viola jones algorithm. We additionally proposed an algorithm for tracking to control the highlight capturing in a camera arrangement setting. There is most likely a single face, or if there are different faces, the biggest will be the principle user of the computer and the one of intrigue. Therefore, we can limit our detection process to a single face and quit preparing once a single face is found. We showed the execution of our work on a moderately uncontrolled multi-vision video database. In Table 1, this execution outperforms the traditional algorithms on multimodal video methodologies interms of accuracy, speed, efficiency and it works adequately on extensive variety of security and surveillance purposes.

Table 1. Overall Performance



Method	Accuracy	Speed	Efficiency
LPP	56.1%	58.8%	65.9%
LDA	37.3%	40.6%	47.4%
SH-PCA	40.7%	39.3%	52.2%
Proposed	65.3%	79.2%	87.3%

REFERENCES

- [1] Feifei Zhang et al. Pose-robust feature learning for facial expression recognition, *Frontiers of Computer Science*, 2016, Vol 10, Issue 5, pp 832-844.
- [2] G.Ramkumar, E.Logashanmugam " Hybrid Framework for detection of human face based on haar-like feature" *International Journal of Engineering & Technology*, 7 (3) 2018, 1786-1790
- [3] X. Li, D. Li, Z. Yang and W. Chen, "A Patch-Based Saliency Detection Method for Assessing the Visual Privacy Levels of Objects in Photos," in *IEEE Access*, 2017, vol. 5, pp. 24332-24343., doi: 10.1109/ACCESS.2017.2767622.
- [4] Chanintorn Jittawiriyankoon "Proposed algorithm for image classification using regression-based pre-processing and recognition models" *International journal of Electrical and Computer Engineering (IJECE)* Vol 9 No. 2 2018.
- [5] C. A. Corneanu, M. O. Simón, J. F. Cohn and S. E. Guerrero, "Survey on RGB, 3D, Thermal, and Multimodal Approaches for Facial Expression Recognition: History, Trends, and Affect-Related Applications," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, vol. 38, no. 8, pp. 1548-1568.
- [6] Agung Nugroho et al " Comparison Analysis of Gait Classification For Human Motion Identification using Embedded Computer" *International journal of Electrical and Computer Engineering (IJECE)* Vol 8 No. 6 2018.
- [7] M. Du, A. C. Sankaranarayanan and R. Chellappa, "Robust Face Recognition From Multi-view Videos," in *IEEE Transactions on Image Processing*, 2014, vol. 23, no. 3, pp. 1105-1117.
- [8] G. Ramkumar, E. Logashanmugam, "An effectual face tracking based on transformed algorithm using composite mask," *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICIC)*, Chennai, 2016, pp. 1-5..
- [9] Eng-jon et al, Robust Facial Feature Tracking Using Shape-Constrained Multiresolution-Selected Linear Predictors Pattern Analysis and Machine Intelligence, *IEEE Transactions on* 2011, (Volume:33 , Issue: 9)
- [10] X. Chai, S. Shan, X. Chen, W. Gao, "Locally linear regression for pose-invariant face recognition," *IEEE Trans. Image Process.*, 2007, vol. 16, no. 7, pp. 1716–1725.
- [11] Fabian Parsia George et al "Recognition of emotional states using EEG signals based on time-frequency analysis and SVM classifier" *International journal of Electrical and Computer Engineering (IJECE)* Vol 9 No. 2 2018.
- [12] S. Li, X. Liu, X. Chai, H. Zhang, S. Lao, S. Shan, "Morphable displacement field based image matching for face recognition across pose," in *Proc. Eur. Conf. Comput. Vis.*, 2012, pp. 102–115.

- [13] R. Wang, S. Shan, X. Chen, G. Wen, "Manifold-manifold distance with application to face recognition based on image set," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2008, pp. 1–8.
- [14] G. Aggarwal, A. K. Roy-Chowdhury, R. Chellappa, "A system identification approach for video-based face recognition," in *Proc. Int. Conf. Pattern Recognit.*, 2004, pp. 175–178.
- [15] Fan Ou et al., Evaluation and Selection of Discriminating Gabor Features for Face Recognition Automatic Identification Advanced Technologies, *IEEE Workshop 2007*
- [16] Kamarainen, Machine Vision & Pattern Recognition Lab., (LUT Kouvola), Lappeemranta, Finland Gbor feature in image analysis 2012, 15-18 *IEEE 2154-5111*
- [17] G. Ramkumar, E. Logashanmugam, "Study on impulsive assessment of chronic pain correlated expressions in facial images" *Biomedical Research* - 2018, Volume 29, Issue 16
- [18] Yanmei Wang et al., "Facial recognition based on kernal PAC" *Third international Conference on Intelligent Networks and Intelligent Systems 2010*, 978-0-7695-4249- 2/10 IEEE
- [19] Guan Chun Luh., Face detection using combination of skin color pixel detection and viola-jones face detector, 2014 978-1-4799-4215-2/14/\$31.00, *IEEE*
- [20] Ming Du, Aswin C. Sankaranarayanan Rama chellappa. Robust face recognition from Multiview Videos © 2014, P No. 1105 – 1117 Doi 10.1109/TIP.2014.2300812 *IEEE*.

BIOGRAPHIES OF AUTHORS

	<p>G. Ramkumar is currently working as an Assistant Professor in Jeppiaar Maamalan Engineering College, Chennai. He has completed his Bachelor Degree in Electronics and Communication Engineering from Sathyabama Institute of Science and Technology in the year 2009 and Master's in VLSI Design from Sathyabama Institute of Science and Technology in the year 2012. He is pursuing his Ph.D in Sathyabama Institute of Science and Technology. His research area is image processing. He has Several publications in National / International Journals /Conferences</p>
	<p>Dr. E. Logashanmugam, Director Administration, Department of Electronics and Communication, has completed his Bachelor Degree in Electronics and Communication Engineering from Madurai Kamarajar University in the year 1991 and Master's in Electronics Engineering from MIT Anna University in the year 2002. He received his Ph.D from Sathyabama Institute of Science and Technology in the year 2009. He joined as a Lecturer in Sathyabama Institute of Science and Technology in the year 1995. His research interests are video signal processing, Embedded Systems, Image Processing etc., He has Several research publications in National / International Journals /Conferences</p>