I.J. Image, Graphics and Signal Processing, **2022, 4, 70-83**
Published Online August 2022 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijigsp.2022.04.06

Eigen and HOG Features based Algorithm for Human Face Tracking in Different Background Challenging Video Sequences

Ranganatha S
Assistant Professor, Department of Computer Science and Engineering, Government Engineering College, Kushalnagar-571234, Karnataka, India
Email: ranganath38@yahoo.co.in

Y P Gowramma
Professor, Department of Computer Science and Engineering, Kalpataru Institute of Technology, Tiptur-572201, Karnataka, India
Email: gowrikit@gmail.com

Received: 05 February 2021; Accepted: 25 June 2021; Published: 08 August 2022

Abstract: We are proposing a unique novel algorithm for tracking human face(s) in different background video sequences. In the beginning, Eigen features and corner points are extracted from the detected face(s). HOG (Histograms of Oriented Gradients) features are isolated from corner points. Eigen and HOG features are combined together. Using these combined features, point tracker keeps track of the face(s) in the frames of the video sequence. Proposed algorithm is being tested on challenging datasets video sequences with technical challenges such as partial occlusion (e.g. moustache, beard, spectacles, helmet, headscarf etc.), changes in expression, variations in illumination and pose; and measured for performance using standard metrics such as accuracy, precision, recall and specificity. Experimental results clearly indicate the robustness of the proposed algorithm on all different background challenging video sequences.

Index Terms: Tracking human face(s), Different background, Video sequences, Eigen features, Corner points, HOG features, Point tracker, Challenging datasets, and Standard metrics.

1. Introduction

In computer vision, video processing is having wide scope for research. Face tracking comes under video processing; it is considered as one of the fundamental, evergreen and difficult problem. Further, it has several issues which are still open for research.

We say that a video sequence is having different background if it is captured using,

1. Static camera and moving face(s),
2. Moving camera and static face(s),
3. Moving camera and moving face(s).

Due to the challenges that are increasing day by day, old/existing face tracking algorithms are becoming obsolete. As the existing algorithms are limited to certain set of video sequences, particular algorithms are unavailable to track human face(s) in different background challenging videos. The main objective here is to track multiple faces in all the three categories of different background video sequences which pose various technical challenges.

Detection of faces is an initial activity in any tracking setup. Hence, faces are the region of interest (ROI) in this regard. Further, features are extracted from the ROI. Feature extraction reduces dimensionality, minimizes computational space and increases processing speed. At the end, face tracking process keeps track of multiple faces in all the situations irrespective of whether only camera or face is moving or both camera and face are in movement.

The pending portion of this paper is planned as mentioned below. Section 2 includes literature works which have been referred and used during the course of proposed algorithm. Section 3 describes add detections and track face(s) algorithms. Section 4 includes description of datasets and performance evaluation metrics. Section 5 includes detailed experimental results and analysis of different background video sequences. Section 6 includes conclusion of the proposed algorithm and future works which can be done.

2. Related Works

Face detection is an initial feat in any automatic tracking scenario. Various technical challenges one normally come across during this step are part of the work of Ranganatha S et al. [1,28]. The proposed work uses Viola-Jones [2,3] framework for face detection; it detects the face(s) whenever they appear in the frames. Viola-Jones face detection system yields performance analogous to Sung and Poggio [4], Rowley et al. [5] and Roth et al. [6]. The above mentioned algorithms are used for finding unoccluded frontal poses of human faces in different scenes.

Dilation is an operator, used in the field of morphology [7]. Though it is basically applied to binary images, some versions of it also work with grayscale images. Dilation operator enlarges the regions of foreground pixels and reduces the size of holes within those regions. Proposed work converts Detected face(s) into dilated images.

Eigen values are the spectrums or characteristic values of any matrix. J. Shi and C. Tomasi [8] have proposed minimal Eigen points from picture gradient matrices for thresholding at region of interest points as relevant feature values for tracking. The point values produced by this method are not variant to transformations such as rotation and translation; but, variant under projective or affine conditions. Literature in [9] proposes an algorithm for face recognition based on Eigenfaces. The algorithm reduces dimension to allow smaller group of images to specify the training images and produces positive results. Proposed work extracts Eigen features from ROI in the images.

Three major regions in any object of an image are “flat”, “edge”, and “corner”. Harris corner detector [10,11] provides a mathematical approach for deciding which region holds. Harris and Stephens [10] have considered the differential of the corner score instead of shifted patches. Corners are generally firm features over changes of viewpoint; hence, they are better features to match. Corner detection is frequently used in video tracking and object recognition. Proposed work calculates Harris corner points for the ROI in the images.

N. Dalal et al. [12] have generated Histograms of Oriented Gradients (HOG) features by inheriting them from Scale Invariant Feature Transform (SIFT) [13] features. Sidheswar Routray et al. [14] have analysed SIFT, HOG and Speeded-Up Robust Features (SURF) [15] image feature extraction techniques. Their experimental results observation reveals that SIFT performance is good on noisy images. Feature extraction speed of SURF is better among the three methods mentioned above and its performance is nearer to SIFT. HOG perform well during detection of edge, face and texture of image. Proposed work extracts HOG features from the corner points.

KLT is a point tracker, which is due to the works of three popular researchers. Lucas and Kanade [16], Tomasi and Kanade [17] have developed a point tracker by considering the points available in the frames, which is known by the first letters of their names as KLT algorithm. Further, Shi and Tomasi [8] have done broad work on features which are less sensitive to challenges like illumination and aging.

Ranganatha S et al. [18] have proposed an approach for face tracking by fusing centroid, corner points [10] and KLT [8,16,17] algorithm. Their approach is able to track moving face in videos which are captured using static camera, but unable to track multiple faces. It is also unable to track faces in the situations where, i). Only camera is moving and face is static and ii). Both camera and face are in movement. They claim that, attainment of their proposed approach is superior in most of the videos compared to KLT algorithm alone. Ranganatha S et al. [19] have proposed another approach for face tracking by federating CAMSHIFT [20] and Kalman filter [21]. Like their previous work, this approach is able to track only single face captured using static camera. It solves the problem of illumination. For testing and analysis, both the above said approaches [18,19] utilized the videos which are available as part of the work in literature [22]. Ranganatha S et al. have also proposed approaches for tracking multiple [23,29,30], selected single [31] and selected single/multiple [32] faces in video sequences. These approaches are able to track the face(s) in all the situations, irrespective of whether only camera or face is moving or both camera and face are in movement. The work published in literature [23] has also witnessed novel metrics for performance evaluation on the videos of challenging datasets [22,24,25].

3. Methodology

The proposed system architecture is summarized in Fig.1 below. At first, faces(s) are detected as and when they appear in the input video sequence. Eigen features are extracted from ROI in the images. At the same time, Harris corner points are calculated for the ROI in the images, and HOG features are extracted from the corner points being calculated. Eigen and HOG features are joined together. Face tracking process looks for the joined features to keep track of ROI in the upcoming frames of video sequence.

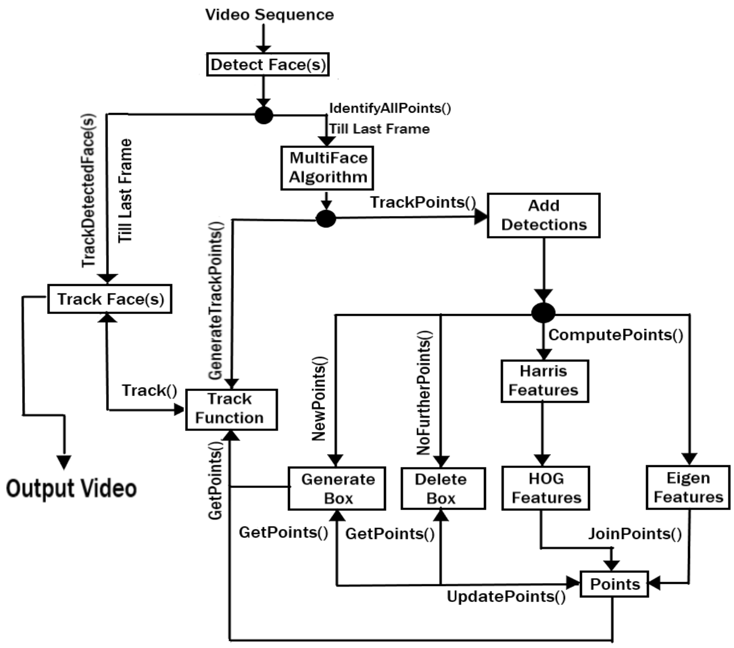

Fig.1. Proposed system architecture.

Proposed algorithm uses novel fused methodology of different points to keep track of facial data. It means that, If Eigen feature points fail to keep track of the face(s) then HOG features does the job and the reverse is also true. It is divided into two major sections:

1. Add detections.
2. Track face(s)

A. Add Detections

Input: Video frames with face(s) detected.
Output: Points generated after computation.

Algorithm 1: Add_Detections()

Repeat
  bboxId ← findMatchingBox among bounding boxes.
  If bboxId = = empty **then**
    Convert image/frame to black and white.
    Dilate the converted image/frame.
    Extract Eigen features and Harris corner points.
    Isolate HOG features from corner points.
    Combine Eigen and HOG feature points.
    Plot points and fill holes with ones.
    Draw bounding box and compute it’s region.
    Set bboxId, PointId and NextId.
  Else // i.e. ***if*** *bboxId != empty*
    Delete old_Box().
    Replace old box with new box.
    Compute the same steps which are computed if bboxId is empty.
    Set oldpoints = newpoints.
  End if
Until Size(bounding boxes).

Algorithm 1 generate points. The points are then used for drawing bounding box. The four points are as shown in Eq. (1).

$$\left.\begin{aligned} X_1 &= min(points(width)), \\ X_2 &= max(points(width)), \\ Y_1 &= min(points(height)), \\ Y_2 &= max(points(height)). \end{aligned}\right\} \tag{1}$$

Eq. (2) computes the bounding box region.

$$Bounding\ Box\ Region = [X_1, Y_1, X_2 - X_1, Y_2 - Y_1]. \tag{2}$$

The four parameters of Eq. (2) are defined as X_1 = X-Coordinate, Y_1 = Y-Coordinate, $X_2 - X_1$ = Width value, and $Y_2 - Y_1$ = Height value.

B. Track Face(s)

Input: Video frames with face(s) detected.
Output: Track faces based on generated points.

Algorithm 2: Track()

Get a video and extract frame from frame information.
Create video player instance.
Repeat
    Read a frame and convert it into grayscale.
    Apply Wiener filter on the frame.
    Take next frame.
Until Face detected successfully.
Repeat
    Read a frame and convert it into grayscale.
    Apply Wiener filter on the frame.
    Take next frame.
    If bboxes != empty **then**
        Call Add_Detections().
    Else *// tracking part i.e. Track()*
        Get points and PointId's.
        // generate NewBoxes()
        Get bboxId's and find unique PointId's.
        Fill Bboxes with zeroes.
            Repeat
                Get points with matching PointId's of every face present in the frame.
                Call getBoundingBox(points).
                Draw bounding box surrounding the points.
            Until All the detected faces have bounding boxes in that frame.
        Delete the bounding box details of the face region if it is no longer tracked.
    End if
    Draw rectangle around the face region detected.
    Take next frame.
Until Last frame is executed.

Algorithm 2 compute the points with matching PointId's of faces by calculating area of BB (Bounding Box) as shown in Eq. (3).

$$Area > 0.2 * BB(width) * BB(height). \tag{3}$$

Algorithm 2 compare the $BoxScores$ value of the face region with $minBoxScore$ as depicted in Eq. (4) to decide whether it is getting tracked or not.

$$(BoxScores < 3) - 0.5 < minBoxScore. \tag{4}$$

If the result of Eq. (4) is true, then the face region is not tracked anymore, and the bounding box details of that face is deleted.

4. Datasets and Metrics

A. Datasets

For testing our proposed algorithm, we have considered 34 different challenging video sequences obtained from the following:

1. YouTube Celebrities Face Tracking and Recognition dataset [22]:

This is the major dataset used for conducting the experiments of proposed work. The dataset holds 1910 video sequences of 47 different subjects. The sequences of this dataset are encoded in mpeg4 format at 25 frames per second. Video filenames obey to the convention mmmm_nn_ooo_name.avi, where the meaning of each field is as depicted below.

Field	**Meaning**
Mmmm	Video sequence identifier
nn	Subject clip identifier
ooo	Video sequence segment identifier

For e.g. “0483_02_009_bill_gates.avi” means video sequence identifier 0483, subject clip 02 (out of 3), 9^{th} video sequence segment from the 2^{nd} clip, and subject Bill Gates.

2. VidTIMIT dataset [26]:

This dataset is comprised of video sequences of 43 volunteers (24 male and 19 female). The sequences were recorded in 3 sessions with a gap of 7 and 6 days between session 1 and 2, and session 2 and session 3 respectively. Each video of this dataset is a numbered sequence with a resolution of 512×384 pixels.

3. Manually captured videos:

Few of the video sequences that are captured manually are also considered for computation.

Out of 34, results of 7 video sequences are included in this paper for reference purpose. The seven sequences taken for experimentation comes under different background video sequence categories discussed earlier. Each sequence has various constraints in it, whose description is included in results and analysis section.

B. Metrics

We have chosen the following metrics [27] for performance evaluation.

1. Accuracy (ACC).
2. Precision (also named Positive Predictive Value (PPV)).
3. Recall (also called True Positive Rate (TPR) / Sensitivity).
4. Specificity (also called True Negative Rate (TNR)).

The above metrics are calculated using the following outcomes. Considering an occurrence as well as a classifier, four outcomes are obtainable.

1. True Positive (TP), when occurrence and classification both are positive.
2. False Negative (FN), when occurrence being positive but classification is negative.
3. True Negative (TN), when occurrence and classification both are negative.
4. False Positive (FP), when occurrence being negative but classification is positive.

Along with these outcomes, the following terms help us in metrics formulation.

1. Positive (P), which is a combination of TP+FN.
2. Negative (N), which is a combination of TN+FP.

Finally, we end this subsection with the following equations.

$$ACC = \frac{(TP + TN)}{(P + N)} = \frac{(TP + TN)}{(TP + FN + TN + FP)} . \quad (5)$$

$$PPV = \frac{TP}{(TP + FP)} . \quad (6)$$

$$TPR = \frac{TP}{P} = \frac{TP}{(TP + FN)} . \quad (7)$$

$$TNR = \frac{TN}{N} = \frac{TN}{(TN + FP)} . \quad (8)$$

Equations (5), (6), (7) and (8) compute accuracy, precision, recall and specificity respectively. The results of the equations are probabilistic values between 0 to 1, with 1 being the best and 0 being the worst result value.

5. Experimental Results and Analysis

Proposed algorithm (PRESENT) is compared with KLT [8, 16, 17], CAMSHIFT [20], Ranganatha S et al. [23] (Algorithm 1) and Ranganatha S et al. [18] (Algorithm 2) algorithms. For each video sequence, the outcome values i.e. Time, NF, TP, TN, FP and FN of all the five algorithms are tabulated. ‘Time’ is in terms of seconds and NF stands for Number of Frames. Using the tabulated values, metrics chosen for performance evaluation are calculated. Results of performance evaluation metrics of all the five algorithms are compared and analysed at the end.

A. Static Camera and Moving Face(s) Video Sequences

In this category, camera is stationary and motion in the video sequence is due to human face movement. We have computed and tabulated the results for three different low resolution indoor video sequences of this category.

Video 1: 0286_01_016_angelina_jolie.avi

This video has single face where an actress is giving interview and it is chosen from YouTube Celebrities Face Tracking and Recognition dataset. We can observe the changes in illumination, expression as well as pose variations from frame to frame throughout the video sequence. These issues pose challenges to the best existing algorithms to keep track of ROI. But, proposed algorithm is able to track face region facing the technical challenges. Randomly chosen four frames of this video sequence containing the face(s) detected and tracked are shown in Fig.2 below.

Fig.2. Frames of 0286_01_016_angelina_jolie.avi.

Values of different outcomes obtained by executing each of the five algorithms are tabulated in Table 1 below.

Table 1. Outcomes of 0286_01_016_angelina_jolie.avi.

Algorithm	Time	NF	TP	TN	FP	FN
KLT	14.3	185	184	0	0	1
CAMSHIFT	Failed					
Algorithm 1	9.2	185	183	0	0	2
Algorithm 2	4.6	185	184	0	0	1
PRESENT	7.3	185	184	0	0	1

Using the values tabulated in Table 1, performance evaluation metrics i.e. TPR, TNR, PPV and ACC are calculated for all the five algorithms. Fig.3 contains comparison of metrics results obtained. We can observe that the values accuracy, precision, recall and specificity of our proposed algorithm are either equal to or greater than other algorithms used for comparison.

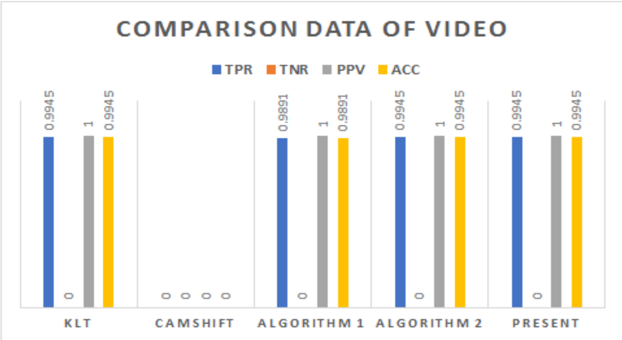

Fig.3. Column chart of 0286_01_016_angelina_jolie.avi.

Video 2: mcem0_head.mpg

This video has single face with extreme pose variations from frame to frame and it is chosen from VidTIMIT dataset. Fig.4 contains five different extreme poses of the face being detected and tracked.

Fig.4. Frames containing extreme poses of mcem0_head.mpg.

Values of different outcomes obtained by executing each of the five algorithms are tabulated in Table 2 below.

Table 2. Outcomes of mcem0_head.mpg.

Algorithm	Time	NF	TP	TN	FP	FN
KLT	26	364	364	0	0	0
CAMSHIFT	8.3	364	360	0	0	4
Algorithm 1	18.9	364	364	0	0	0
Algorithm 2	7.6	364	363	0	0	1
PRESENT	10.1	364	364	0	0	0

Using the values tabulated in Table 2, performance evaluation metrics are calculated for all the five algorithms. Fig.5 contains comparison of metrics results obtained. We can observe that the metrics values of our proposed algorithm are ideal i.e. TNR remains 0 and the rest are equal to 1.

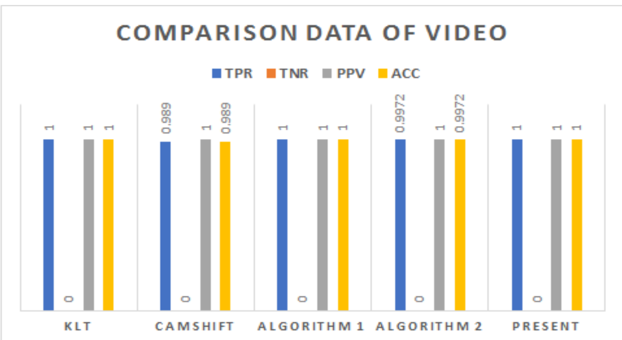

Fig.5. Column chart of mcem0_head.mpg.

Video 3: 1632_02_028_ronald_reagan.avi

This video houses three celebrities who are addressing the gathering and it is chosen from YouTube Celebrities Face Tracking and Recognition dataset. The video has many challenges pertaining to each celebrity. The face of the celebrity who is standing in the middle is moving abruptly from frame to frame. The two face-skin colored arms of the celebrity (lady) who is standing on the right are visible outside; here, there are chances of tracking the skin colored regions. The face of the celebrity who is standing on the left is partially occluded, because it contain moustache; here, there are chances for tracking the face improperly. In some of the frames, one side (or portion) of the face(s) is not visible i.e. visible only partially. Apart from human faces, the video contains background poster and a trophy placed on the podium; here, there are chances of tracking non-faces. In spite of all these challenges, proposed algorithm continues to track only the detected faces. Fig.6 comprises four frames, each frame shows the detection and tracking of three person faces.

Fig.6. Frames of 1632_02_028_ronald_reagan.avi.

Table 3 shows values of different outcomes obtained.

Table 3. Outcomes of 1632_02_028_ronald_reagan.avi.

<table>
<tr><th>Algorithm</th><th>Time</th><th>NF</th><th>TP</th><th>TN</th><th>FP</th><th>FN</th></tr>
<tr><td>KLT</td><td colspan="6" rowspan="4">Failed</td></tr>
<tr><td>CAMSHIFT</td></tr>
<tr><td>Algorithm 1</td></tr>
<tr><td>Algorithm 2</td></tr>
<tr><td>PRESENT</td><td>49</td><td>318</td><td>311</td><td>0</td><td>0</td><td>7</td></tr>
</table>

Table 3 discloses the failure of all the four algorithms considered for comparison. KLT and CAMSHIFT detect all the three faces; but, they are unable to keep track of them further. Surprisingly Algorithm 1 has failed, though it is capable of tracking multiple faces. Failure of Algorithm 2 is obvious, due to the fact that it is capable of tracking only single face. Fig.7 shows the domination of proposed algorithm againt other tracking techniques.

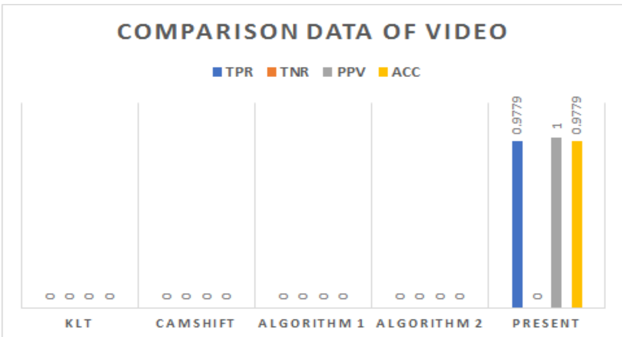

Fig.7. Column chart of 1632_02_028_ronald_reagan.avi.

B. Moving Camera and Static Face(s) Video Sequence

In this category, camera moves and face(s) remain static. Here, we have computed and tabulated the results for an indoor video and a corridor video.

Video 1: 0676_01_007_gloria_estefan.avi

The video being chosen in this section is from YouTube Celebrities Face Tracking and Recognition dataset. The video has frames of an actress who is in the position of rest and singing a song; but, camera is moving around her. Randomly chosen four frames containing the face detected and tracked are shown in Fig.8.

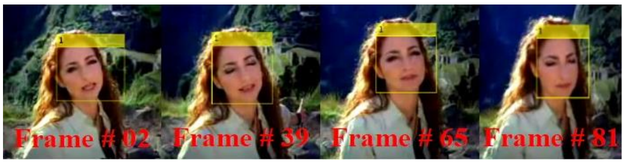

Fig.8. Frames of 0676_01_007_gloria_estefan.avi.

Table 4 shows values of different outcomes obtained.

Table 4. Outcomes of 0676_01_007_gloria_estefan.avi.

Algorithm	Time	NF	TP	TN	FP	FN
KLT	Failed					
CAMSHIFT						
Algorithm 1	6.3	82	76	0	0	6
Algorithm 2	Failed					
PRESENT	8.1	82	82	0	0	0

Table 4 reveals that KLT, CAMSHIFT and Algorithm 2 are unable to detect face in the video sequence; because they are trained to detect and track face(s) appearing from first frame only. But, our proposed algorithm detect and track face(s) even if the face(s) do not appear in the first frame itself. Using Table 4 values, metrics are calculated. Fig.9 contains comparison of metrics results obtained. We can observe a clear win of our proposed algorithm with others.

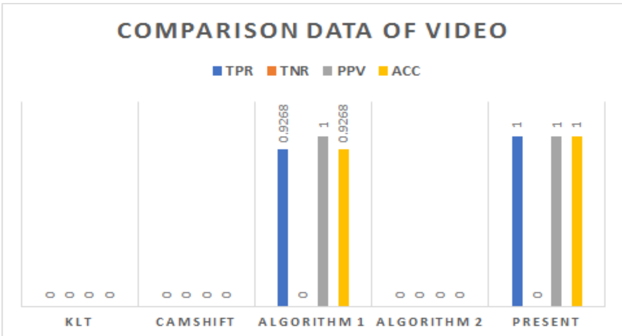

Fig.9. Column chart of 0676_01_007_gloria_estefan.avi.

Video 2: manually captured video sequence

Here, due to the lack of “moving camera and static face(s)” video sequences, we have considered a manually captured corridor video sequence for computation. The video contains two partially occluded faces. The first face is occluded because of glasses, and the second face is occluded due to helmet. The video background changes from frame to frame, but the faces remain static. The four frames of Fig.10 depicts tracking result of this video sequence.

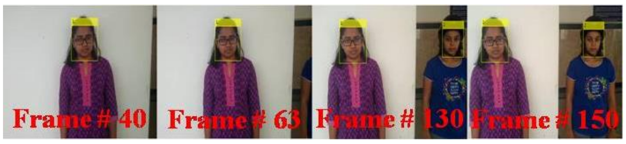

Fig.10. Frames of manually-captured-video.mp4.

Values of different outcomes obtained by executing each of the five algorithms are tabulated in Table 5 below.

Table 5. Outcomes of manually-captured-video.mp4.

Algorithm	Time	NF	TP	TN	FP	FN
KLT	Failed					
CAMSHIFT						
Algorithm 1	57	238	234	0	1	3
Algorithm 2	Failed					
PRESENT	45	238	235	0	1	2

Table 5 ensures the failure of KLT, CAMSHIFT and Algorithm 2. KLT algorithm tracks only the first face. CAMSHIFT and Algorithm 2 detects the first face, but fails in tracking both the faces. Out of the results tabulated, Algorithm 1 and proposed algorithm have produced best results for the undertaken video. When we plot the data of Table 5, we witness the graph shown in Fig.11 below.

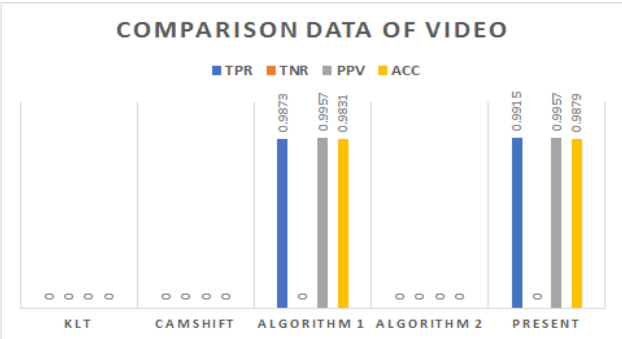

Fig.11. Column chart of manually-captured-video.mp4.

C. Moving Camera and Moving Face(s) Video Sequence

In this category, both camera and face(s) are in movement. We have computed and tabulated the results for one indoor and one outdoor video sequences of this category.

Video 1: 0193_01_004_alanis_morissette.avi

Fig.12 shows few of the frames of a single face low resolution outdoor video sequence. In this video, camera is moving along with the celebrity who is driving a car and it is chosen from YouTube Celebrities Face Tracking and Recognition dataset. The face of the celebrity is occluded because of headscarf. Since the camera and face are in motion throughout the video sequence, both face as well as video background change dynamically from frame to frame.

Fig.12. Frames of 0193_01_004_alanis_morissette.avi.

Table 6 shows values of different outcomes obtained.

Table 6. Outcomes of 0193_01_004_alanis_morissette.avi.

Algorithm	Time	NF	TP	TN	FP	FN
KLT	Failed					
CAMSHIFT						
Algorithm 1	7.2	21	20	0	0	1
Algorithm 2	Failed					
PRESENT	12.3	21	21	0	0	0

Table 6 reveals the success of only two algorithms as shown in Fig.13. The remaining algorithms failed to detect the face itself.

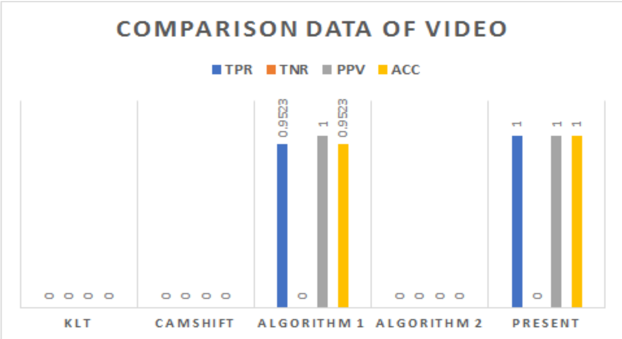

Fig.13. Column chart of 0193_01_004_alanis_morissette.avi.

Video 2: 1167_01_011_john_kerry.avi

This is a technically challenging, low resolution indoor video sequence. It contains multiple faces with camera and face(s) in motion throughout the video sequence and it is chosen from YouTube Celebrities Face Tracking and Recognition dataset. Since the camera and face(s) are in motion throughout the video sequence, both face(s) as well as video backround change dynamically from frame to frame. In this video sequence, some of the faces are merged, and some are partially occluded because of spectacles and beard. Randomly chosen four frames containing five human face(s) detected and tracked are shown in Fig.14.

Fig.14. Frames of 1167_01_011_john_kerry.avi

Values of different outcomes obtained i.e. 'Time', NF, TP, TN, FP and FN by executing each of the five algorithms are tabulated in Table 7 below.

Table 7. Outcomes of 1167_01_011_john_kerry.avi.

Algorithm	Time	NF	TP	TN	FP	FN
KLT	Failed					
CAMSHIFT						
Algorithm 1						
Algorithm 2						
PRESENT	6.6	119	113	0	1	5

Table 7 shows failure of four algorithms. KLT and CAMSHIFT algorithms detect four faces; but, both of them fail in tracking the faces detected. Algorithm 1 is a multiface tracker, it is developed to detect and track faces in different background video sequences. But, it fails again due to the complexity and inherent technical challenges in the video sequence. Failure of Algorithm 2 is apparent, because it is trained to detect and track single face only in "static camera and moving face(s)" category of video sequences. But, our proposed algorithm detect and track multiple faces in all the three categories of different background video sequences. Fig.15 contains comparison of metrics results obtained. Graphical data shows an almost ideal type of success for the proposed algorithm against other tracking techniques.

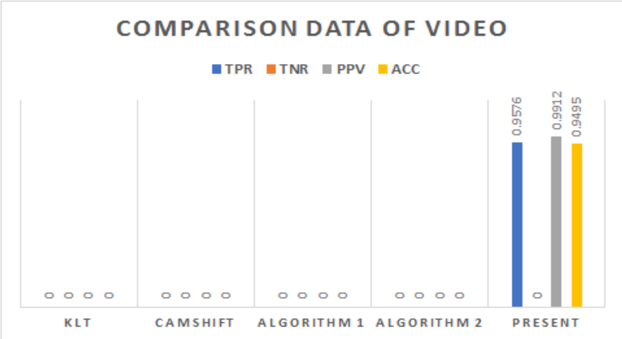

Fig.15. Column chart of 1167_01_011_john_kerry.avi.

6. Conclusion and Future Work

The Proposed algorithm is capable of tracking face(s) appearing in any frame of an indoor/outdoor and difficult/cluttered different background video sequence. It tackles the problems caused by partial occlusion, changes in illumination, expression and pose variations efficiently. The Results obtained shows a clear win of our proposed algorithm with respect to tracking in all category of video sequences compared to other algorithms.

In video processing, face tracking has major scope for research. Although existing algorithms fulfil this purpose, technological developments are posing wide range of challenges while processing the videos. Face tracking in different background video sequence has found many uses in real life circumstances and is trivial from the point of marketable and law enforced applications. Proposed algorithm is capable of tracking both static/dynamic face(s) which are recorded using static/dynamic sensors. Keeping these points in mind, some of the applications of proposed algorithm and face technology in general include Surveillance, Access and Security, Criminal Identification, Marketing and Advertising, and Healthcare to name a few. Further, face tracking has numerous issues that are ever open for research.

In order to detect and track face(s), future work involves training and testing of video sequences using classifiers. Future work also tries to tackle the problems caused by complete occlusion and aging.

References

[1] Ranganatha S and Dr. Y P Gowramma, “Face Recognition Techniques: A Survey”, International Journal for Research in Applied Science and Engineering Technology (IJRASET), ISSN: 2321-9653, Vol.3, No.4, pp.630-635, April 2015.

[2] P. Viola and M. Jones, “Robust Real-Time Face Detection”, International Journal of Computer Vision (IJCV), Vol.57, pp.137-154, 2004.

[3] P. Viola and M. Jones, “Rapid Object Detection Using a Boosted Cascade of Simple Features”, in Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Hawaii, USA, Vol.1, pp.511-518, 2001.

[4] Kah-Kay Sung and Tomasi Poggio, “Example-Based Learning for View-Based Human Face Detection”, in IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), Vol.20, No.1, pp.39-51, January 1998.

[5] H. Rowley, S. Baluja and T. Kanade, “Neural Network-Based Face Detection”, in IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), Vol.20, No.1, pp.22-38, January 1998.

[6] D. Roth, M. Yang and N. Ahuja, “A Snow Based Face Detector”, Advances of Neural Information Processing Systems, MIT Press, pp.855-861, 2000.

[7] Yasunori Kudo and Yoshimitsu Aoki, “Dilated Convolutions for Image Classification and Object Localization”, in Proc. of International Conference on Machine Vision Applications (MVA), Nagoya University, Nagoya, Japan, pp.452-455, May 2017.

[8] Jianbo Shi and Carlo Tomasi, “Good Features to Track”, in Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.593-600, June 1994.

[9] Maria Rosario D. Rodavia, Orlando Bernaldez and Maylane Ballita, “Web and Mobile Based Facial Recognition Security System Using Eigenfaces Algorithm”, in Proc. of IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE), pp.86-92, December 2016.

[10] C. Harris and M. Stephens, “A Combined Corner and Edge Detector”, in Proc. of 4th Alvey Vision Conference, Manchester, UK, pp.147-151, 1988.

[11] R. Manoranjitham and P. Deepa, “Novel Interest Point Detector Using Bilateral-Harris Corner Method”, in Proc. of IEEE International Conference on Advanced Computing and Communication Systems (ICACCS), pp.1-4, January 2017.

[12] N. Dalal and B. Triggs, “Histograms of Oriented Gradients for Human Detection”, in Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Vol.1, pp.886-893, June 2005.

[13] David G. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints”, International Journal of Computer Vision (IJCV), Vol.60, No.2, pp.91-110, November 2004.

[14] Sidheswar Routray, Arun Kumar Ray and Chandrabhanu Mishra, “Analysis of Various Image Feature Extraction Methods against Noisy Image: SIFT, SURF and HOG”, in Proc. of IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), pp.1-5, February 2017.

[15] Herbert Bay, Andreas Ess, Tinne Tuytelaars and Luc Van Gool, “Speeded-Up Robust Features”, Computer Vision and Image Understanding (CVIU), Vol.110, No.3, pp.346-359, June 2008.

[16] Bruce D. Lucas and Takeo Kanade, “An Iterative Image Registration Technique with an Application to Stereo Vision”, in Proc. of International Joint Conference on Artificial Intelligence, Vol.2, pp.674-679, August 1981.
[17] Carlo Tomasi and Takeo Kanade, “Detection and Tracking of Point Features”, Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991.
[18] Ranganatha S and Y P Gowramma, “A Novel Fused Algorithm for Human Face Tracking in Video Sequences”, in Proc. of IEEE International Conference on Computation System and Information Technology for Sustainable Solutions (CSITSS), pp.1-6, October 2016. DOI: 10.1109/CSITSS.2016.7779430
[19] Ranganatha S and Y P Gowramma, “An Integrated Robust Approach for Fast Face Tracking in Noisy Real-World Videos with Visual Constraints”, in Proc. of IEEE International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp.772-776, September 2017. DOI: 10.1109/ICACCI.2017.8125935
[20] G. Bradski, “Computer Vision Face Tracking for Use in a Perceptual User Interface”, Intel Technology Journal, pp.12-21, 1998.
[21] Douglas Decarlo and Dimitris N.Metaxas, “Optical Flow Constraints on Deformable Models with Applications to Face Tracking”, International Journal of Computer Vision (IJCV), Vol.38, No.2, pp.99-127, July 2000.
[22] M. Kim, S. Kumar, V. Pavlovic and H. Rowley, “Face Tracking and Recognition with Visual Constraints in Real-World Videos”, in Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2008.
[23] Ranganatha S and Y P Gowramma, “Development of Robust Multiple Face Tracking Algorithm and Novel Performance Evaluation Metrics for Different Background Video Sequences”, International Journal of Intelligent Systems and Applications (IJISA), Vol.10, No.8, pp.19-35, August 2018. DOI: 10.5815/ijisa.2018.08.03
[24] Yongkang Wong, Shaokang Chen, Sandra Mau, Conrad Sanderson and Brian C. Lovell, “Patch-Based Probabilistic Image Quality Assessment for Face Selection and Improved Video-Based Face Recognition”, in proc. of IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp.74-81, June 2011.
[25] Ivan Laptev, Marcin Marszalek, Cordelia Schmid and Benjamin Rozenfeld, “Learning Realistic Human Actions from Movies”, in Proc. of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp.1-8, June 2008.
[26] C. Sanderson and B.C. Lovell, “Multi-Region Probabilistic Histograms for Robust and Scalable Identity Inference”, Lecture Notes in Computer Science (LNCS), Vol.5558, pp.199-208, 2009.
[27] T. Fawcett, “An Introduction to ROC Analysis”, Pattern Recognition Letters, Vol.27, No.8, pp.861-874, 2006.
[28] Ranganatha S and Y P Gowramma, “A Comprehensive Survey of Algorithms for Face Tracking in Different Background Video Sequence”, International Journal of Computer Applications, Vol.181, Issue 27, pp.43-49, 2018. DOI: 10.5120/ijca2018918134
[29] Ranganatha S and Y P Gowramma, “Image Training, Corner and FAST Features based Algorithm for Face Tracking in Low Resolution Different Background Challenging Video Sequences”, International Journal of Image, Graphics and Signal Processing (IJIGSP), Vol.10, No.8, pp.39-53, August 2018. DOI: 10.5815/ijigsp.2018.08.05.
[30] Ranganatha S and Y P Gowramma, “Image Training and LBPH Based Algorithm for Face Tracking in Different Background Video Sequence”, International Journal of Computer Sciences and Engineering (IJCSE), Vol.6, No.9, pp.349-354, September 2018. CrossRef-DOI: https://doi.org/10.26438/ijcse/v6i9.349354.
[31] Ranganatha S and Y P Gowramma, “Selected Single Face Tracking in Technically Challenging Different Background Video Sequences Using Combined Features”, ICTACT Journal on Image and Video Processing (JIVP), Vol.9, No.2, pp.1911-1918, November 2018. DOI: 10.21917/ijivp.2018.0271.
[32] Ranganatha S and Y P Gowramma, “Color Based New Algorithm for Detection and Single/Multiple Person Face Tracking in Different Background Video Sequence”, International Journal of Information Technology and Computer Science (IJITCS), Vol.10, No.11, pp.39-48, November 2018. DOI: 10.5815/ijitcs.2018.11.04.

Authors’ Profiles

Dr. Ranganatha S received B.E degree (Branch: CS & E, College: KIT, Tiptur, Karnataka, India), M.Tech degree (Specialization: CS & E, College: RVCE, Bangalore, Karnataka, India) and Ph.D. degree (Area: Image/Video Processing) in the years 2002, 2006 and 2020 respectively from Visvesvaraya Technological University (VTU), Belagavi, Karnataka, India. He has 20 years of teaching experience in reputed organizations at graduate and post graduate level. He is currently working as Assistant Professor, Department of Computer Science & Engineering, Government Engineering College, Kushalnagar, Kodagu district, Karnataka, India. He has published papers in various national and international journals and conferences. His areas of interest include Image and Video Processing, Algorithms, Data Mining, Natural Language Processing, Operating Systems and Educational Technology.

Dr. Y P Gowramma received B.E degree (Branch: CS &E, College: AIT, Chikmagalore, University: Mysore, Karnataka, India), M.Tech degree (Specialization: SACA, College: NITK, Karnataka, India) and Ph.D. degree (University: VTU, Belagavi, Karnataka, India) in the years 1994, 1998 and 2014 respectively. She has more than 25 years of teaching experience in reputed organizations at graduate and post graduate level. She is currently working as Professor, Department of Computer Science & Engineering, Kalpataru Institute of Technology, Tiptur, Karnataka, India. She has published papers in various national and international journals and conferences. Her current research interests include image processing and algorithms.

How to cite this paper: Ranganatha S, Y P Gowramma, "Eigen and HOG Features based Algorithm for Human Face Tracking in Different Background Challenging Video Sequences", International Journal of Image, Graphics and Signal Processing(IJIGSP), Vol.14, No.4, pp. 70-83, 2022. DOI:10.5815/ijigsp.2022.04.06

