Color Based New Algorithm for Detection and Single/Multiple Person Face Tracking in Different Background Video Sequence

Ranganatha S
Asst. Professor, Dept. of Computer Science and Engineering, Government Engineering College,
Hassan-573201, Karnataka, India
E-mail: ranganath38@yahoo.co.in

Y P Gowramma
Professor, Dept. of Computer Science and Engineering, Kalpataru Institute of Technology, Tiptur-572202,
Karnataka, India
E-mail: gowrikit@gmail.com

Abstract—Due to the lack of particular algorithms for automatic detection and tracking of person face(s), we have developed a new algorithm to achieve detection and single/multiple face tracking in different background video sequence. To detect faces, skin sections are segmented from the frame by means of YCbCr color model; and facial features are used to agree whether these sections contain person face or not. This procedure is challenging, because face color is unique and some objects may have similar color. Further, color and Eigen features are extracted from detected faces. Based on the points detected in facial region, point tracker tracks the user specified number of faces throughout the video sequence. The developed algorithm was tested on challenging dataset videos; and measured for performance using standard metrics. Test results obtained ensure the efficiency of proposed algorithm at the end.

Index Terms—Detection, Face tracking, Different background, Video sequence, YCbCr color, Eigen features, Point tracker.

I. INTRODUCTION

In recent years, processing the images containing person faces has become a growing research concern due to the developments in automatic techniques covering biometrics, video surveillance, security applications, compression etc. Moving visual images constitute a video, whose processing is part of image processing. Face tracking is an attention-grabbing research zone in video processing which tracks face(s) under different technical challenges [1] such as illumination, pose, expression, occlusion, background conditions etc.

Normally, face tracking approaches involve three steps: face detection, feature extraction and face tracking. Faces can be detected using either existing techniques or by self-established techniques. Viola-Jones [2, 3] approach is an existing technique for face detection. It is actually trained to detect frontal posed faces and it needs additional training to detect the faces having other poses. Proposed work has adopted self-established skin color segmentation and region properties based technique for face detection. As some of the objects may also have color similar to faces, it is a challenging task for a machine to classify whether detected portion is face or not.

Feature extraction step is very important, because type of feature and its volume directly influence space and time complexities. Various features are available to extract from the detected faces. Color, intensity and texture are treated as principal features whereas secondary features are formed by fusing principal features; example Haar [3], HOG [4], SIFT [5] and CENTRIST [6] features. Sometimes, color based approaches lack behind while tracking region of interest (ROI) due to the similarity of face color with other regions. Hence, in addition to color, proposed work extracts Eigen features [7] from section of concern. Eigen features are not variant to rotation and translation transformations. This combination helps for the robust tracking of face(s) in video sequence.

During last step, proposed work permit the user to state the number of faces in the video to be tracked. As part of this, point tracker keeps track of specified faces throughout the video sequence. Algorithms are available for tracking person faces in video sequence; but, particular algorithms are unavailable for tracking specified number of faces in different background video sequence. We come across different backgrounds in situations when, 1). Moving face(s) are captured using static camera, 2). Moving face(s) are captured using moving camera, and 3). Static face(s) are captured using moving camera. Proposed work is efficient enough to identify and trace face(s) in videos with different

backgrounds mentioned above.

The upcoming sections of our paper are prepared as stated next. Section 2 houses the associated works that are referred. Section 3 includes face finding, feature mining and face tracing as subsections. Section 4 consist of experimental evaluation of our proposed work. Lastly, the paper halts with conclusion and further work in section 5.

II. RELATED WORKS

Prashanth et al. [8] have proposed an approach for detection and tracing of person face. First, skin regions are segmented using diverse color spaces. Segmented sections are detached out of the frame by applying a threshold; and facial features are used to know whether these regions comprise person face or not. Knowledge based approaches [8, 9] adopt geometrical attributes of human face for detection and tracking. Using RGB color range, it is tough to decide whether the color is human skin color or not. But, YCbCr color space [10] can be adopted to recognize human skin color areas accurately.

Based on the works of Fukunaga et al. [11] and Cheng [12], Bradski [13] proposed an algorithm called CAMSHIFT. This is a familiar motion tracing approach implemented till date. Since, this algorithm is based on chromatic values, color based approaches make use of it for visual tracking. If the target is masked or if there is a similarity between object color and background color, then the accuracy of CAMSHIFT decreases. To overcome these problems, Chunbo et al. [14] have proposed an improved algorithm based on CAMSHIFT. The search window is updated according to the contour features of ROI to weaken the interference of background and stout light. Kanade et al. [7], [15, 16] have contributed an algorithm to the research community, which became popular as KLT. KLT is an efficient point tracking algorithm. But, KLT has the issues of missing fast motion and small search windows.

Ranganatha et al. [17] have proposed a novel approach for tracking person face. Their algorithm was designed by merging centroid of corner points [18] and KLT. Experimental outcomes show that, merged approach is superior compared to KLT alone. Ranganatha et al. [19] have proposed another algorithm for person face tracking using CAMSHIFT and Kalman filter [20-22]. The drawbacks of CAMSHIFT are avoided by integrating it with Kalman filter. Tabulated results indicate the robustness of integrated algorithm compared to CAMSHIFT alone. Celebrities dataset [23] videos were used for testing the algorithms discussed above. The algorithms are proficient of tracing moving lone face caught using a static camera; but, unable to track moving/unmoving single/multiple face captured using a moving camera.

Ranganatha et al. [9], [24, 25] have proposed three algorithms for tracking multiple face in different background videos. The first algorithm was based on BRISK [26] features, the second algorithm was based on image training & LBPH [27], and the third algorithm used image training, corner [18] and FAST [28] features. The algorithms are robust enough to notice and track multiple faces under difficult conditions in different background videos. All the faces present in the video are tracked by default. But, user cannot specify the faces like one, two, three etc. for tracking. Ranganatha et al. [29] have proposed one other technique for tracking selected single face in different background videos using Eigen, FAST and HOG features. But, this algorithm cannot be used for tracking multiple person face.

Till now, we have presented an abstract survey of the works which are relevant to our proposed work. The drawbacks and limitations of the works presented here, have helped in proposing our current novel algorithm. Proposed algorithm uses the videos that are part of the effort in literature [23] for testing. At the end, standard metrics [30-32] are used to measure the attainment of the work proposed.

III. PROPOSED WORK

The proposed work is divided into three sub-sections. Each sub-section includes description and algorithm.

A. Face Detection

Face detection process identifies person faces in digital images. To perform face detection, read a frame from a video. Fig.1 shows a frame of 0241_03_001_anderson_cooper.avi (video) containing two faces.

Fig.1. Frame of a video.

As part of pre-processing, convert the RGB image to YCbCr image. Get the maximum and minimum intensity (luminance) values. Find the index of the pixel with intensity greater than average, using the following formula.

$$Index = Find(YCbCr >= I) . \qquad (1)$$

Where $I = (Min\_lum + 0.95 * (Max\_lum - Min\_lum))$.

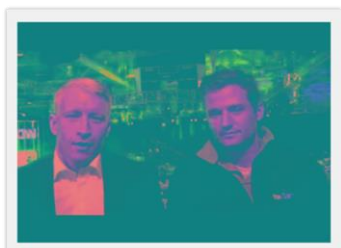The YCbCr image obtained is shown in Fig.2.

Fig.2. YCbCr image.

Assign an average intensity value to all the indexes with value greater than 100; otherwise, don't change.

$$
\left. \begin{array}{c}
if\ (index >= 100) \\
Average = Mean\big(gray(index)\big). \\
else \\
Compensated = Image.
\end{array} \right\}
\qquad (2)
$$

Formula (2) illustrates this step, leading to a light compensated image as made known in Fig.3.

Fig.3. Light compensated image.

Further, the image is transformed to binary image; then morphological procedures like dilation and erosion are enforced. In case of dilation, pixels are added to objects boundary; and erosion does the reverse of dilation. The output of this step, which helps to extract the skin regions is shown in Fig.4.

Fig.4. Morphologically processed binary image.

Next, skin regions are divided into blocks called blobs. Identify the positive and negative regions of the binary image using mathematical calculations. All the blobs in the image are taken out separately based on skin region dimension. Blobs with uneven outlines and are with lowest possibility of being a face are omitted. Fig.5 comprise two blobs that are separated. Further operations are applied individually on the blobs extracted.

Fig.5. Extracted blobs.

To consider a region as face, try to find the mouth region. If found, treat it as the face region; otherwise, omit the blob. Based on the size of blob, bounding box points are calculated. Using these, put a box around the face region and label it as a face. In case of detecting multiple faces, bounding box point values are stored in array.
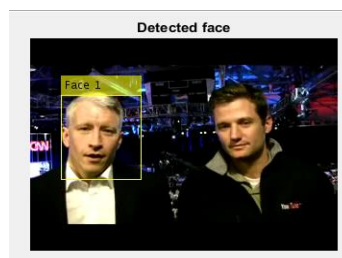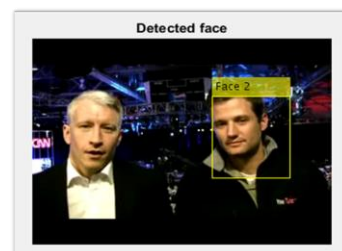

Fig.6. Detection of face #1.

Fig.7. Detection of face #2.

Fig.6 and Fig.7 shows two images with bounding boxes, indicating the detection of face 1 and face 2 respectively. Algorithm 1 includes the steps of face detection process.

Algorithm 1: Face Detection
One. Get the video from user using Graphical User Interface (GUI) and extract the first frame from it.
Two. Convert the obtained RGB color frame into YCbCr color frame.
Three. Equalize the light in the frame.
Four. Convert the frame to binary format, process it morphologically to extract the skin regions.
Five. Extract each separate skin region called blob.
Six. Accept/Reject blobs using the following 3 sub-steps.
1st Step of eliminating blobs: If blob contains eye feature, then select it as possible face and go to next step. Otherwise, reject the blob.
2nd Step of eliminating blobs:

If blob is (narrow or short or wide or (narrow and tall) or (wide and short)), then reject it. Otherwise, select it as possible face and go to next step.

3rd Step of elimination:
Loop over each pixel.

```
For i ← 1 to m
    For j ← 1 to n
        // Look for the mouth region.
        Found = Mouth-Map (blob).
        If (Found)
            │ Accept.
        Otherwise
            │ Reject.
        End
    End
End
```

Seven. Consider all the selected/accepted blobs as faces.

Eight. Calculate the bounding box values for the region(s) detected as face(s).

Nine. Draw bounding box around detected region(s).

B. Feature Extraction and Points Detection

Detected faces are considered as the regions of interest. At first, extract color features from the bounding box, and remove the noise. Next, detect Eigen feature points from the bounding box. Extracted points help to trace the ROI further. Fig.8 and Fig.9 shows the points displayed.
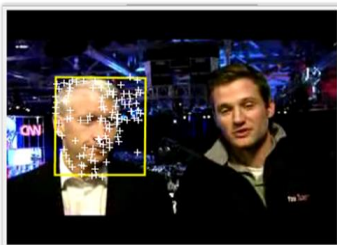

Fig.8. Feature points of face #1.

Fig.9. Feature points of face #2.

Algorithm 2 includes the steps of feature extraction process.

Algorithm 2: Feature Extraction and Points Detection

One. Detect the face as described in Algorithm 1 and extract the color features of ROI from the frame using, frame = Color-feature (frame1, ROI, bounding-box).

Two. Apply the filters to reduce noise in the frame.

Three. Detect the Eigen feature points from ROI and store them in a variable.

Four. Combine points matrix and color features array.

Five. Plot and display the points detected from the facial region of the frame.

C. Single/Multiple Person Face Tracking

Face tracking process permit the user to state the number of faces in the video to be tracked. User can input some number say x (like 1, 2, 3 etc.), to carry out selective tracking.

Fig.10. Tracking of face #1.

Fig.11. Tracking of face #2.

Fig.12. Tracking of both face #1 and face #2.

This way single/multiple person face tracking is accomplished. Fig.10, Fig.11 and Fig.12 illustrates selective tracking of face 1, face 2, and both face 1 and face 2 respectively. Algorithm 3 includes the steps of face tracking process.

Algorithm 3: Face Tracking

One. X = Total number of faces present; ask the user how many faces in the video to be tracked.

Two. For i ← 1 to X
| Num[ ] = Enter the number of faces to track.
End

Three. Initialize the tracker with feature points of ROI.

Four. Initialize the video player to display the face(s) tracked.

Five. Generate the points in the remaining frames of the video; and measure the geometric conversion between previous points and current points.

Six. Reset the point tracker, and go to step Four; repeat the steps from Four to Six up to the last frame of the video sequence.

IV. EXPERIMENTAL EVALUATION

The discussion here is about the outcome of proposed work. This section houses dataset, performance metrics, and results and analysis as sub-sections.

A. Dataset

We have considered ten YouTube Celebrities dataset [23] videos for testing. The videos belong to the different background categories discussed in section 1. Out of ten, four videos have single face and six videos contain multiple (i.e. more than one) faces. The videos pose many technical challenges like low resolution, occlusion (e.g. spectacle, headscarf), face skin colored regions, variations in pose, expression and illumination etc.

B. Performance Metrics

Fig.13 denotes confusion matrix [31], it classifies 4 metrics as follows.

1. True Positive (TP) – Legitimate elements properly grouped as legitimate.
2. False Negative (FN) – Legitimate elements improperly grouped as illegitimate.
3. True Negative (TN) – Illegitimate elements improperly grouped as illegitimate.
4. False positive (FP) – Illegitimate elements improperly grouped as legitimate.

Fig.13. Confusion matrix.

Let $TP = \alpha1$, $TN = \alpha2$, $FP = \delta1$, $FN = \delta2$. Using these, the following metrics can be formulated.

$$Recall\ (R) = \frac{\alpha1}{\alpha1+\delta2} \quad (3)$$

$$Precision\ (P) = \frac{\alpha1}{\alpha1+\delta1} \quad (4)$$

$$True\ Negative\ Rate\ (TNR) = \frac{\alpha2}{\alpha2+\delta1} \quad (5)$$

$$Negative\ Predictive\ Value\ (NPV) = \frac{\alpha2}{\alpha2+\delta2} \quad (6)$$

$$False\ Negative\ Rate\ (FNR) = \frac{\delta2}{\delta2+\alpha1} \quad (7)$$

$$False\ Positive\ Rate\ (FPR) = \frac{\delta1}{\delta1+\alpha2} \quad (8)$$

$$Accuracy\ (A) = \frac{\alpha1+\alpha2}{\alpha1+\alpha2+\delta1+\delta2} \quad (9)$$

$$F1\ Score\ (F1S) = 2\left(\frac{R \times P}{R+P}\right) \quad (10)$$

Equations (3) to (10) compute R, P, TNR, NPV, FNR, FPR, A, and $F1S$ respectively. Along with these, multiple object ('face' in our work) tracking metrics i.e. FN, FP, multiple object tracking accuracy (MOTA), and multiple object tracking precision (MOTP) are also computed. FN and FP have already been defined, MOTA is widely used to assess trackers performance. It syndicates 3 sources of errors as

$$MOTA = 1 - \frac{\sum_z \left(\delta1_z + \delta2_z + MME_z\right)}{\sum_z GT_z} \quad (11)$$

where z is the frame index, MME is the mismatch error, and GT is the count of ground truth objects. MOTP indicates the capability of a tracker to estimate accurate object location. For bounding box similarity, this is calculated as

$$MOTP = \frac{\sum_{z,x} D_{z,x}}{\sum_z C_z} \quad (12)$$

where C_z is the count of matches found at frame z, and $D_{z,x}$ is the bounding box similarity of target x with its given ground truth object. The outcomes of all the above equations are probabilistic values, with 1 being the best and 0 being the worst result value.

C. Results and Analysis

This section talks about experimental results of the proposed algorithm. By means of the performance metrics discussed before, the outcomes of the proposed algorithm are then equated with CAMSHIFT and Algorithm* [19].

At first, results of the single face video sequences are discussed. Fig.14 and Fig.15 shows some of the frames of two low resolution indoor videos. A single face is detected and tracked as shown in the figures. In both the videos, celebrities are giving the interviews; camera is unmoving, and motion in the videos is due to person face movement.

Fig.14. Detection and single person face tracking result of 67 frames video sequence 0272_01_002_angelina_jolie.avi with frames #01, #17, #42, and #65 displayed (from left to right).

Fig.15. Detection and single person face tracking result of 172 frames video sequence 0485_03_002_bill_gates.avi with frames #01, #34, #96, and #149 displayed (from left to right).

Fig.16 and Fig.17 shows few of the frames of two single face low resolution outdoor videos. In both the videos, camera is moving along with the celebrity who is driving a car.

Fig.16. Detection and single person face tracking result of 23 frames video sequence 0193_01_004_alanis_morissette.avi with frames #01, #10, #19, and #22 displayed (from left to right).

Fig.17. Detection and single person face tracking result of 180 frames video sequence 0199_01_010_alanis_morissette.avi with frames #01, #85, #143, and #175 displayed (from left to right).

Close observation of Fig.16 and Fig.17 reveals the following facts. Due to the dynamic movement of face, first video (0193_01_004_alanis_morissette.avi) belongs to moving face and moving camera video sequence category. In the second video (0199_01_010_alanis_morissette.avi), though the face is moving along with the camera, it is almost in the position of rest, making it more suitable for static face and moving camera category video sequence.

Now, the results of multiple person face video sequences are elaborated. Out of six multiple face videos considered for testing, detection and tracking result of 0241_03_001_anderson_cooper.avi has already discussed as part of proposed work. Detection and tracking results of the remaining five video sequences in which the moving faces are captured using a static camera, are discussed next.

Fig.18 shows the frames of a low resolution video sequence. It contains two persons talking to each other, who are part of a television show. The frames in figure clearly shows the detection of faces and selective tracking being carried out.

Fig.18. Detection and single/multiple person face tracking result of 106 frames video sequence 0092_02_013_al_gore.avi with frames #01, #18, #103, and #95 displayed (from left to right).

Fig.19 shows five frames; each frame containing two persons, who are talking to each other in a television news channel program. Detection of face 1 and face 2, and selective tracking of face 1, face 2, and both face 1 and face 2 are shown in the figure clearly.

Fig.19. Detection and single/multiple person face tracking result of 303 frames video sequence 0234_01_008_anderson_cooper.avi with frames #01, #01, #190, #256, and #87 displayed (from left to right).

Fig.20 shows low resolution video frames where two celebrities are giving an interview. The video contains background poster and there are chances of tracking non-faces; but, proposed algorithm continues to track only the selected face(s). In Fig.20, we can observe that proposed algorithm is clever to track a partially visible face, facing the challenge of abrupt motion from frame to frame.

Fig.20. Detection and single/multiple person face tracking result of 196 frames video sequence 0818_02_013_hugh_grant.avi.avi with frames #01, #34, #123, and #37 displayed (from left to right).

Fig.21 shows detection and selective tracking of person face(s). We can observe multiple colors at several places in every frame, which poses challenges to a color based face tracker. Fig.21 shows the detection of three faces; and selective tracking of face 1, face 2, face 3, face 1 and face 2, face 1 and face 3, face 2 and face 3, and all the three faces, correspondingly from left-to-right and top-to-bottom.
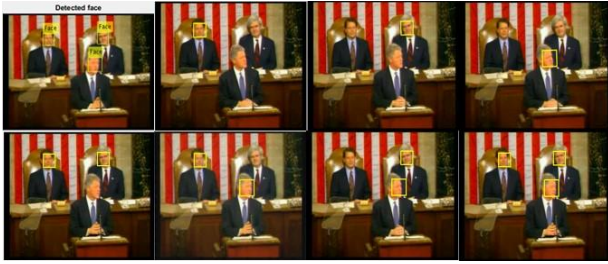
Fig.21. Detection and single/multiple person face tracking result of 126 frames video sequence 0467_03_018_bill_clinton.avi with frames #01, #18, #21, #37, #49, #63, #86, and #102 displayed (from left-to-right and top-to-bottom row wise).

Fig.22 houses frames of a video in which some of the faces are merged. It contains multiple persons who are gathered together in an event. Similar to Fig.21, Fig.22 shows detection and selective tracking of three faces.
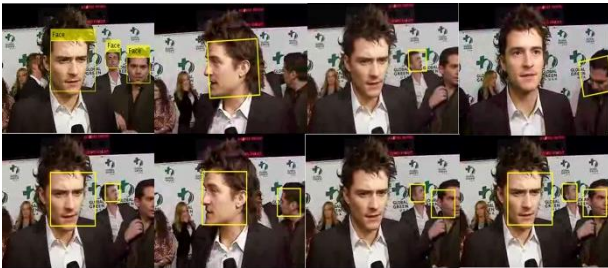

Fig.22. Detection and single/multiple person face tracking result of 311 frames video sequence 1534_02_001_orlando_bloom.avi with frames #01, #100, #212, #308, #213, #98, #214, and #215 displayed (from left-to-right and top-to-bottom row wise).

For simplicity and clear representation of tabular results, let us denote 0092_02_013_al_gore.avi, 0193_01_004_alanis_morissette.avi, 0199_01_010_alanis_morissette.avi, 0234_01_008_anderson_cooper.avi, 0241_03_001_anderson_cooper.avi, 0272_01_002_angelina_jolie.avi, 0467_03_018_bill_clinton.avi, 0485_03_002_bill_gates.avi, 0818_02_013_hugh_grant.avi, and 1534_02_001_orlando_bloom.avi with numbers from 1 to 10 respectively. Using the metrics discussed, outcomes of three algorithms are presented in Table 1, Table 2, and Table 3. The "-" in the tables denote failure in tracking of that particular video by the given algorithm.

Table 1. Quantifiable probabilistic outcomes of the proposed algorithm for 10 video sequences.

VIDEO	1	2	3	4	5	6	7	8	9	10
R	1	1	1	1	1	1	1	1	1	1
P	1	1	1	1	1	1	1	1	1	1
TNR	0	0	0	0	0	0	0	0	0	0
NPV	0	0	0	0	0	0	0	0	0	0
FNR	0	0	0	0	0	0	0	0	0	0
FPR	0	0	0	0	0	0	0	0	0	0
A	1	1	1	1	1	1	1	1	1	1
F1S	1	1	1	1	1	1	1	1	1	1
MOTA	1	1	1	1	1	1	1	1	1	1
MOTP	1	1	1	1	1	1	1	1	1	1

Table 2. Quantifiable probabilistic outcomes of the CAMSHIFT algorithm for 10 video sequences.

VIDEO	1	2	3	4	5	6	7	8	9	10
R	-	1	-	-	-	1	-	-	1	-
P	-	1	-	-	-	1	-	-	1	-
TNR	-	0	-	-	-	0	-	-	0	-
NPV	-	0	-	-	-	0	-	-	0	-
FNR	-	0	-	-	-	0	-	-	0	-
FPR	-	0	-	-	-	0	-	-	0	-
A	-	1	-	-	-	1	-	-	1	-
F1S	-	1	-	-	-	1	-	-	1	-
MOTA	-	1	-	-	-	1	-	-	1	-
MOTP	-	1	-	-	-	1	-	-	1	-

Table 3. Quantifiable probabilistic outcomes of the Algorithm* [19] algorithm for 10 video sequences.

VIDEO	1	2	3	4	5	6	7	8	9	10
R	-	-	-	-	-	-	-	1	-	-
P	-	-	-	-	-	-	-	1	-	-
TNR	-	-	-	-	-	-	-	0	-	-
NPV	-	-	-	-	-	-	-	0	-	-
FNR	-	-	-	-	-	-	-	0	-	-
FPR	-	-	-	-	-	-	-	0	-	-
A	-	-	-	-	-	-	-	1	-	-
F1S	-	-	-	-	-	-	-	1	-	-
MOTA	-	-	-	-	-	-	-	1	-	-
MOTP	-	-	-	-	-	-	-	1	-	-

In the results tabulated, we can observe that the proposed algorithm has produced best results for all the undertaken videos. Considering all the 10 video sequences, the averages of R, P, A, F1S, MOTA, and MOTP are tabulated as shown in Table 4.

Table 4. Assessable results of three algorithms.

Algorithm	R (%)	P (%)	A (%)	F1S (%)	MOTA (%)	MOTP (%)
Proposed	100	100	100	100	100	100
CAMSHIFT	30	30	30	30	30	30
Algorithm*	10	10	10	10	10	10

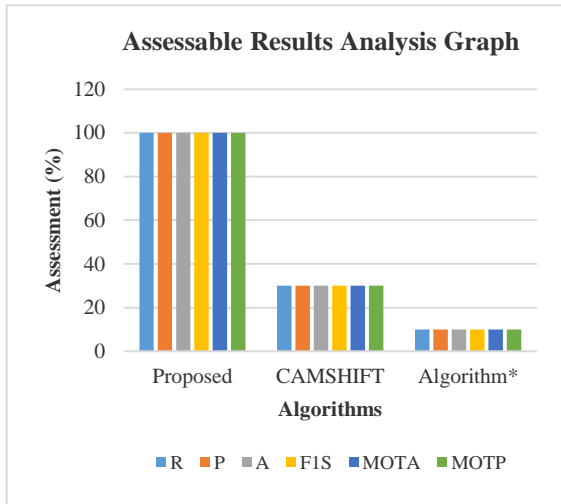Fig.23 illustrates assessable results analysis of three algorithms graphically.

Fig.23. Assessable results analysis of three algorithms.

Time plays a major role during evaluation of performance; hence, time taken by the three algorithms has been tabulated in Table 5 below.

Table 5. Time taken to process videos (in seconds).

Video	Proposed	CAMSHIFT	Algorithm*
1	1.9	-	-
2	0.4	0.4	-
3	2.6	-	-
4	6.1	-	-
5	3.4	-	-
6	0.8	2.3	-
7	1.4	-	-
8	4.14	-	4.04
9	4.7	6.3	-
10	3.2	-	-

When we plot the data of Table 5, we witness the graph shown in Fig.24 below. The straight lines (of CAMSHIFT and Algorithm*) on x-axis (horizontal i.e. video title axis) are the result of "-" (False/Failure).
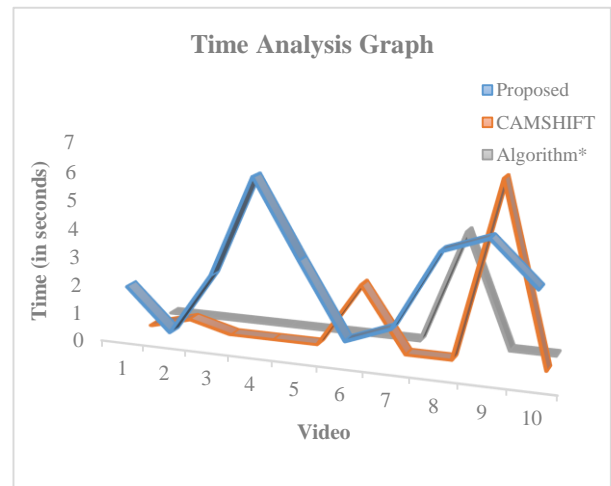

Fig.24. Time analysis of different algorithms.

V. CONCLUSION AND FURTHER WORK

Video and video types are changing day by day; due to which, video processing is becoming complex time to time. An algorithm which was used few years ago, is failing to overcome the challenges that are being faced nowadays. Here, a color centred customized approach has been proposed for finding and selective tracking of human faces in video sequences. Our algorithm is robust and invariant to noise; but, challenging, as face color may match with the objects color.

Our proposed algorithm can be improved further, if the following enhancements are included; they are:

1. The algorithm can be extended to perform tracking in real time applications by carrying out feature extraction in real time.
2. The proposed algorithm can be extended to images with very complex background, wherein the size of the face is very small.
3. While tracking, for the faces which come in between the video, a detection algorithm can be implemented to ensure that no faces are missed.

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), Kauai, USA, Vol.1, pp.511-518, December 2001. DOI: 10.1109/CVPR.2001.990517

[4] 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. DOI: 10.1109/CVPR.2005.177

[5] 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.

[6] J. Wu and J. M. Rehg, "CENTRIST: A Visual Descriptor for Scene Categorization", in IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), Vol.33, No.8, pp.1489-1501, August 2011. DOI: 10.1109/TPAMI.2010.224.

[7] 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. DOI: 10.1109/CVPR.1994.323794

[8] Prashanth Kumar G and Shashidhara M, "Real Time Detection and Tracking of Human Face using Skin Color Segmentation and Region Properties", International Journal of Image, Graphics and Signal Processing (IJIGSP), Vol.6, No.8, pp.40-46, July 2014. DOI: 10.5815/ijigsp.2014.08.06

[9] 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

[10] Mohammad Saber Iraji and Azam Tosinia, "Skin Color Segmentation in YCBCR Color Space with Adaptive Fuzzy Neural Network (Anfis)", International Journal of Image, Graphics and Signal Processing (IJIGSP), Vol.4, No.4, pp.35-41, May 2012. DOI: 10.5815/ijigsp.2012.04.05

[11] K. Fukunaga and L. D. Hostetler, "The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition", in IEEE Trans. on Information Theory, Vol.21, No.1, pp.32-40, January 1975. DOI: 10.1109/TIT.1975.1055330

[12] Yizong Cheng, "Mean Shift, Mode Seeking, and Clustering", in IEEE Trans. on Pattern Analysis and Machine Intelligence (PAMI), Vol.17, No.8, pp.790-799, August 1995. DOI: 10.1109/34.400568

[13] G. Bradski, "Computer Vision Face Tracking for Use in a Perceptual User Interface", Intel Technology Journal, pp.12-21, 1998.

[14] Chunbo Xiu, Xuemiao Su, and Xiaonan Pan, "Improved Target Tracking Algorithm based on Camshift", in Proc. of IEEE Chinese Control and Decision Conference (CCDC), pp.4449-4454, June 2018. DOI: 10.1109/CCDC.2018.8407900

[15] 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.

[16] Carlo Tomasi and Takeo Kanade, "Detection and Tracking of Point Features", Carnegie Mellon University Technical Report CMU-CS-91-132, April 1991.

[17] 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

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

[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] J. Strom, T. Jebara, S. Basu, and A. Pentland, "Real Time Tracking and Modeling of Faces: An EKF-based Analysis by Synthesis Approach", in Proc. of IEEE International Workshop on Modelling People (MPeople), pp.55-61, September 1999. DOI: 10.1109/PEOPLE.1999.798346

[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] Natalia Chaudhry and Kh. M. Umar Suleman, "IP Camera Based Video Surveillance Using Object's Boundary Specification", International Journal of Information Technology and Computer Science (IJITCS), Vol.8, No.8, pp.13-22, August 2016. DOI: 10.5815/ijitcs.2016.08.02

[23] 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), pp.1-8, June 2008. DOI: 10.1109/CVPR.2008.4587572

[24] 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, September 2018.

[25] 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

[26] Stefan Leutenegger, Margarita Chli, and Roland Y. Siegwart, "BRISK: Binary Robust Invariant Scalable Keypoints", in Proc. of IEEE International Conference on Computer Vision (ICCV), pp.2548-2555, November 2011. DOI: 10.1109/ICCV.2011.6126542

[27] Md. Abdur Rahim, Md. Najmul Hossain, Tanzillah Wahid, and Md. Shafiul Azam, "Face Recognition using Local Binary Patterns (LBP)", Global Journal of Computer Science and Technology, Vol.13, No.4, pp.1-8, 2013.

[28] E. Rosten and T. Drummond, "Fusing Points and Lines for High Performance Tracking", in Proc. of IEEE International Conference on Computer Vision (ICCV), Vol.2, pp.1508-1515, October 2005. DOI: 10.1109/ICCV.2005.104

[29] 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), in press.

[30] D. M. W. Powers, "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation", Journal of Machine Learning Technologies (JMLT), Vol.2, No.1, pp.37-63, 2011.

[31] T. Fawcett, "An Introduction to ROC Analysis", Pattern Recognition Letters, Vol.27, No.8, pp.861-874, June 2006. DOI: 10.1016/j.patrec.2005.10.010

[32] Keni Bernardin and Rainer Stiefelhagen, "Evaluating Multiple Object Tracking Performance: The CLEAR MOT Metrics", EURASIP Journal on Image and Video Processing, pp.1-10, May 2008. DOI: 10.1155/2008/246309

Authors' Profiles

Ranganatha S received B.E degree (Branch: CS & E, College: KIT, Tiptur, Karnataka, India) and M.Tech degree (Specialization: CS & E, College: RVCE, Bangalore, Karnataka, India) in the years 1998-2002 and 2004-2006 respectively from Visvesvaraya Technological University (VTU), Belagavi, Karnataka, India and currently pursuing his PhD in image processing at VTU Research Resource Centre. He has more than 16 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, Hassan, Karnataka, India. He has published papers in various national and international journals and conferences. His current research interests include image processing and algorithms.

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 PhD degree (University: VTU, Belagavi, Karnataka, India) in the years 1994, 1998 and 2014 respectively. She has more than 24 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.