IJCTA, 9(22), 2016, pp. 173-180

Analysis of Face Feature based Human Detection Techniques

Dushyant Kumar Singh* Dharmender Singh Kushwaha**

Abstract : The problem of human detection has always been an area of interest for the researchers in computer vision domain. The promising applications such as smart surveillance, robotics, virtual reality etc actually drive their interest. A number of different techniques have been developed for human detection to work in both real-time or offline mode. Some of those techniques use face features to detect or distinguish human being from other objects. The paper presents the analysis of such techniques by classifying them two classes based on their characteristics and the constraints of their applicability. The tabular and graph results with corresponding image results justify the analysis done for these techniques.

Keywords : Human detection, face detection, skin colour modelling, Viola-Jones algorithm.

1. INTRODUCTION

Face detection techniques have been studied and utilized to automate human detection process using computer vision. These are used in various real life applications like surveillance, pedestrian detection [1] and guiding blind person [2] etc. Manual techniques have some glitches, since observing the camera is done manually and human being has constraints on his physical capacities of continuous sitting and watching. As well as sufficient man power is required on timely basis in this approach. With the advancement in technology, automation is the need in all real time applications. In human detection, we try to enable computer to look for a human being in a given image. The human face is the most visible and distinguishing part of human body and many algorithms have been proposed to detect human being based on human face features. These face features helps in clearly distinguishing human from other objects in the image/video. These features are face colour and the face texture.

'Face Colour' is nothing but face skin colour and every different living object has different skin colour intensity ranges. By finding an accurate intensity ranges for skin colour of human being, human detection or classification can be done. The second face feature i.e. 'Face Texture' is also a unique characteristic for any living object. Texture information for human face comprises relative intensities of various portions of human face such as eye, eyebrows, lips, cheek, nose etc [3]. Upon analyzing the algorithms proposed in the literature, it is found that face features based human detection algorithms use either of the two features. Based on these two features, the algorithms can be classified into two broad classes, one is '*Skin Colour Modelling*' and the second one is '*Viola & Jones Algorithm*'. Algorithms under skin colour modelling exploits the skin colour feature of the human face while Viola & Jones algorithm uses texture feature of the face.

This paper presents an analysis of the above mentioned classes of face feature based human detection techniques. There are various challenges that algorithms in each of these classes face. These are ambient light illumination affect, varied face poses, occlusion etc. Beside this varying skin colour complexions of human beings on the basis of ethnicity, geographical locations also poses many challenges [4]. A detailed analysis of algorithms is presented considering the various factors stated above. Experimental results in images and comparative analysis using tables and graphs are also presented.

* Department of Computer Science and Engineering, MNNIT Allahabad, Allahabad (INDIA) Email: dushyant@mnnit.ac.in

** Department of Computer Science and Engineering, MNNIT Allahabad, Allahabad (INDIA) Email: ask@gmail.com

2. SKIN COLOR MODELLING

The pixel colour or intensity is the primary information of an image. This pixels intensity varies from low i.e. black to high i.e. white level. There exist 255 different intensity levels for an image with 8-bit pixel size. The number of levels become 255x255x255 for an image with 24-bit pixel size. These all intensity levels correspond to all possible colours existing in nature. Colour of human skin being one among these finds its place in these intensity levels. Skin colour is not a single colour and also not same for all human beings around the world. As skin colour varies based on ethnicity and geographical locations, skin colour occupies a range of some colour intensity levels. So, finding out that range of skin colour intensities is the skin colour modelling [5].

Machine learning algorithm is used to learn the possible skin colour intensities from the training dataset of human images. This learning process is also termed as modelling. The trained model is used for detecting or distinguishing the skin colour from non-skin colour. Model alternatively is also called as the classifier model. A number of classifier models have been proposed in literature for skin colour classification [4, 5]. Linear /non-linear rules in three dependent random variables are used for classification. These dependent random variables are 3 planes of a colour space example RGB [6, 7], HSV [8], YCbCr [7] etc. The accuracy of classifier is measured by the parameters like True Positive Rate (TPR), False Positive Rate (FPR) and Accuracy [5]. Achieving high accuracy is always a challenge. Some parametric [9, 10] and non-parametric [11, 12] classifier models are also proposed in literature which delivers higher accuracy than simple linear rules. The accuracy of classifier is improved in later approaches but the classification complexity increases side by side for these approaches.

The classifier model proposed in skin colour modelling are so designed that the skin colour range they classify clearly distinguishes the nearby colours in the colour intensity spectrum. This results in high classifier performance. But the exceptional cases can’t be ignored where background objects colour exactly lies in the same skin colour range as used by classifier model. Such background objects are tree barks, ripen leaves, ground areas etc. Here, for the analysis sake linear classifier model is deployed, which gives good result.

Pixel colour based image processing techniques face the problem of ambient light affects. Due to varying ambient lighting, threshold values have to be changed in a problem of colour based image segmentation. Since skin colour modelling uses pixel colour as feature, this too have high impact of changing ambient illumination. This problem first incurred while using the RGB colour space for processing. Later, from other researches [10, 13] done in past, it is found that colour image processing done in HSV and YCbCr spaces have comparatively lesser effect of changing illumination. Experimental results of analysis are shown in section 4 of this paper.

Original videos available at various sources are in RGB format and the same is tone for videos captured from camera. Processing of skin classifier algorithms can be done directly on this format. For processing in other parallel spaces like HSV and YCbCr, the video from RGB format needs to be first converted to HSV or YCbCr. Classifier models are different for doing skin classification in these three colour spaces. The linear rules for classification in these spaces are detailed next.

(*a*) **RGB colour space :** The rules for skin colour in RGB colour space as given in [6, 7] are

$$(R > 95)\ \&\ (G > 40)\ \&\ (B > 20)\ \text{and}$$

$$(\max\{R, G, B\} - \min\{R,G, B\} > 15)\ \text{and}$$

$$(|R\text{-}G| > 15)\ \&\ (R > G)\ \&\ (R > B)$$

These rules were derived after a rigorous test on different skin colours. This colour space is very sensitive to illumination changes. So, classification accuracy with these rules reduces for images/videos takes in shadow or under bad weather conditions i.e. improper ambient lighting.

(*b*) **HSV colour space :** In HSV colour space, it is the Hue which determines the colour while Saturation determines the colourfulness of a specific colour and the Value determines the darkness or lightness of the colour. Hence Hue is the parameter which determines the colour component along with saturation. The classification rules for skin colour in HSV are proposed in [8] and these are:

$$0 \le H \le 50 \quad \text{and} \quad 0.23 \le S \le 0.68$$

(*c*) **YCbCr colour space :** The Y component of YCbCr is the luminance component which is poses the visual information of the image. Cb and Cr components are chrominance components which determine the colours in image and the Y component merely determines the greyscale intensity. Cb and Cr are resistant to changing illumination effects therefore this colour space is recognized to be the best compared to others for colour image processing under these circumstances. Therefore the classification rules only involve conditions on Cb and Cr values and these are [7]:

$90 \leq Cb \leq 130$ and $137 \leq Cr \leq 177$

3. VIOLA & JONES ALGORITHM

The Viola Jones algorithm was initially designed for face detection. The classifier in this algorithm was trained for human face only. Later the classifier was trained for other objects and has been used in many different problems. The same is extended to the human detection as well, which uses face features to detect human being. Some key characteristics of a human face that were exploited to design algorithm are [14, 15 & 16]:

(*a*) The eye region is darker than upper cheek region,

(*b*) The nose ridge is brighter than surrounding vertical area, and

(*c*) Intensity variations at other face components as mouth, forehead, ears etc.

This information accounts to the texture on a 2D representation of the face. The texture information is mathematically modelled by Haar-like features. Haar-like features are combination of black and white rectangular boxes where each box could be a single pixel or group of pixels. Such features could be 2-rectangle feature; 3-rectangle feature and 4-rectangle feature. These features differ on the way how rectangle boxes are combined and total of such features is a combinatorial set. Black and white are not exact black & white intensity levels rather these are relative levels, *i.e*. a lesser intensity value is black and a higher intensity level is white for any haar-like feature. More details of Haar-like features can be seen in [14, 17].

The classifier in the Viola-Jones algorithm is a cascaded classifier trained with the face image training set. The input to the classifier is not the pixel intensity information of face images rather it's the Haar-like features generated for each face image in the training set. The trained classifier is used for classifying/detecting a face in unknown sample. The peculiarity of cascade classifier is the fast and utmost accurate classification. The details of cascade classifier can be seen in [15, 17].

For processing an unknown sample, the image is taken in grayscale format, and the Haar features are convoluted on the image. The pixels under the black and white box regions are summed separately and subtracted. The resultant is then compared to a threshold and if the resultant is more, then the region is marked passed. Different Haar features are used to detect different parts of human face. Cascade of Haar features is used to detect facial features accurately over a given area. The region where all the facial features are successful found, that region is marked as a face.

4. EXPERIMENTAL RESULTS

As discussed in the previous sections, the experiments are performed for both the classes of algorithms *i.e.* Skin Colour Modelling and Viola & Jones algorithm. The experiment is done on a video of 50 sec randomly taken from “Youtube” [18]. Frame rate of video is 10 FPS. Sampling of 1 out of 5 frame is done and finally we have 100 frames to process. Each of these frames has one or more than one face. The results are shown for detecting human face using both the techniques.

Frames in first row of figure 1 are the original video frames transformed in HSV space and frames in next row shows the detected faces in corresponding frames in HSV space. The percentage of face detected either frontal or tilted is presented in table 1. The result in table is shown for randomly selected 10 frames.

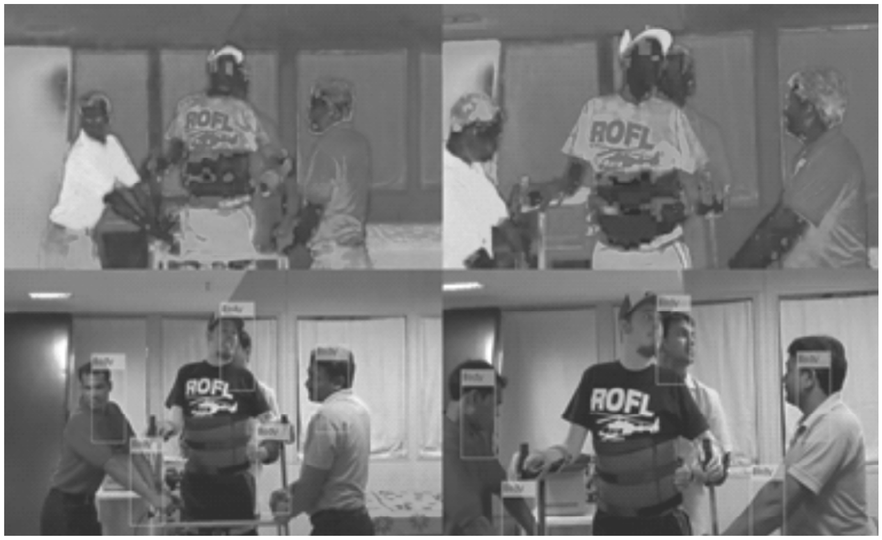

Fig. 1. Skin colour modelling in HSV space for face detection.

Table 1. Face detected in HSV domain skin colour modelling

Frame No.	*Number Frontal face*	*Number Tilted face*	*Total face*	*% of frontal face detected*	*% of tilted face detected*	*% total face detected*
1	2	1	3	50	100	100
9	3	1	4	66.67	0	50
13	2	2	4	50	50	50
25	3	1	4	33.33	0	25
34	3	1	4	33.33	100	50
46	3	1	4	66.67	0	50
53	2	2	4	50	50	50
61	0	4	4	100	75	75
75	1	3	4	0	100	75
89	2	2	4	50	50	50

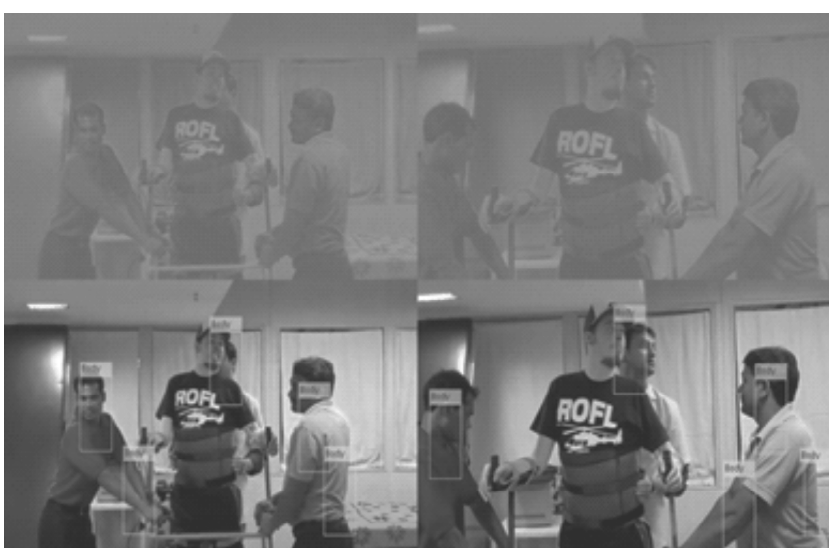

Fig. 2. Skin colour modelling in YCbCr space for face detection.

Frames in first row of figure 2 are the original video frames transformed in YCbCr space and frames in next row shows the detected faces in corresponding frames in YCbCr space. The percentage of face detected either frontal or tilted is presented in table 2.

Table 2. Face detected in YCbCr domain skin colour modelling.

Frame No.	*Number Frontal face*	*Number Tilted face*	*Total face*	*% of frontal face detected*	*% of tilted face detected*	*% total face detected*
1	2	1	3	100	0	66.67
9	3	1	4	66.67	0	50
13	2	2	4	50	0	25
25	3	1	4	33.33	0	25
34	3	1	4	66.67	0	50
46	3	1	4	66.67	0	50
53	2	2	4	50	100	75
61	0	4	4	100	75	75
75	1	3	4	0	100	75
89	2	2	4	50	100	75

Fig. 3. Face detection using Viola-Jones Algorithm

Frames in first row of figure 3 are the original video frames and frames in next row shows the detected faces in corresponding frames after applying Viola-Jones algorithm. The percentage of face detected either frontal or tilted is presented in table 3.

Table 3. Face detected in video using Viola-Jones algorithm.

Frame No.	*Number Frontal face*	*Number Tilted face*	*Total face*	*% of frontal face detected*	*% of tilted face detected*	*% total face detected*
1	2	1	3	50	0	66.67
9	3	1	4	100	0	75
13	2	2	4	100	0	50
25	3	1	4	66.67	0	50
34	3	1	4	66.67	0	50
46	3	1	4	66.67	0	50
53	2	2	4	100	0	50
61	0	4	4	100	25	25
75	1	3	4	100	33.33	50
89	2	2	4	100	0	50

The above shown results illustrated in table 1, 2 and 3, for face detection are compiled and analyzed for finding the performance of these approaches. The performance is measured in terms of TPR, FPR and Accuracy, which are calculated by following equations.

$$TPR = TP / (TP + FN) \tag{6}$$

$$FPR = FP / (FP + TN) \tag{7}$$

$$Accuracy = (TP + TN) / (TP + FN + TN + FP) \tag{9}$$

Where TP, FP, TN & FN are defined in compliance with problem as,

TP $\rightarrow$ is number of frames in which face detected when actually it is present

FP $\rightarrow$ is number of frames in which face detected when actually it is not present

TN $\rightarrow$ is number of frames in which face not detected when actually it is not present

FN $\rightarrow$ is number of frames in which face not detected when actually it is present

Table 4. Comparative analysis of face feature based detection techniques

Performance Measures	*Skin color modeling in HSV*	*Skin color modeling in YCbCr*	*Viola-Jones algorithm*
TPR	0.50	0.53	0.434
FPR	0.53	0.65	0.064
Accuracy	0.49	0.47	0.59

From the results of table 4, it can be concluded that Viola-Jones algorithm has accuracy than skin colour modelling but still the results of table 3 derive an important conclusion that the percentage of tilted (or rotated) face detected is lesser in Viola-Jones compared to skin colour modelling. The graph for above performance results is shown in fig 4. Few facts of analysis finally derived are discussed in next sub-section.

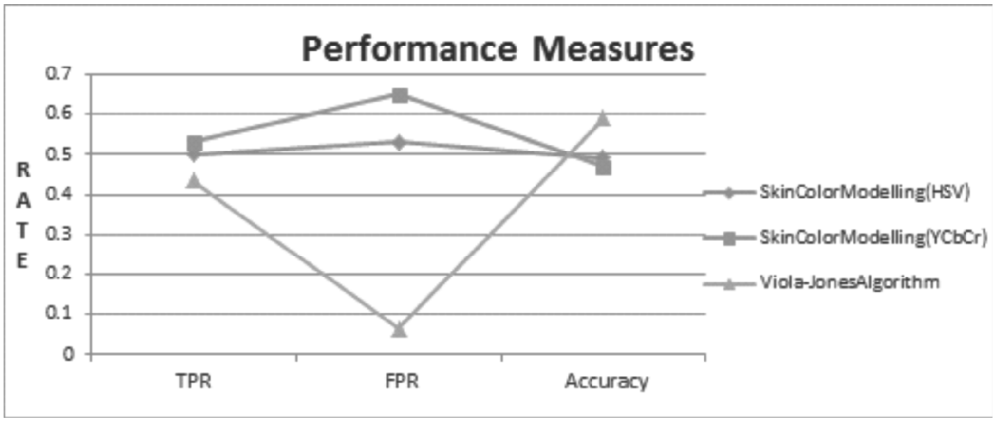

Fig. 4. Performance measurement of the analyzed techniques.

4.1. Result Analysis

The Viola-Jones algorithm outperforms Skin colour modelling in following cases:

1. When the background has skin colour objects, so the skin colour modelling will have very high FPR, while Viola-Jones will be unaffected.
2. Viola-Jones algorithm works faster than skin colour modelling.
3. Viola-Jones algorithm has higher accuracy and TPR than skin colour modelling

The Skin colour modelling outperforms Viola-Jones algorithms in following cases :

1. If the human face is tilted at an angle or the side of the face is visible, then Viola-Jones fails to detect it while the Skin colour modelling succeeds.
2. Skin colour modelling can detect a person even with the back view of the person while Viola-Jones needs a fully visible face.
3. Skin colour modelling works best with background having non-skin colour objects with at least some part of skin being visible while Viola-Jones would not be able to detect a human being just by some part of human body being visible.

5. CONCLUSION

The conclusion drawn from the analysis done in this paper reveals certain interesting facts. One class of algorithm stands good at some design constraints while second is good for other constraints. The application where human being has frontal and clearly visible view then Viola-Jones algorithm is suitable. Example of such an application could be auto video recording of some interviews. Detection in this case will even not be effected by background object colours. But if the application is of type surveillance at public places where persons moving are not intentionally looking at camera i.e. face is not in frontal view, then skin colour modelling is better. Above these, if demerits of these techniques can be used as merit of other and some hybrid solution is derived then it will be a more accurate and constraint free solution to human detection problem based on face features.

6. REFERENCES

1. Tan, Wei Ren, et al. "A fusion approach for efficient human skin detection."*Industrial Informatics, IEEE Transactions on* 8.1 (2012): 138-147.
2. Vadakkepat, Prahlad, et al. "Multimodal approach to human-face detection and tracking." *Industrial Electronics, IEEE Transactions on* 55.3 (2008): 1385-1393.
3. Peer, Peter, and Franc Solina. "An automatic human face detection method." (1999): 122-130.
4. Zhao, Xiaohui. "Research of self-adaptive chroma space model skin-color algorithm based on illumination and its application to erotic webpage filtering."*Networking and Digital Society (ICNDS), 2010 2nd International Conference on*. Vol. 2. IEEE, 2010.

5. Jones, Michael J., and James M. Rehg. “Statistical color models with application to skin detection.” *International Journal of Computer Vision* 46.1 (2002): 81-96.

6. Vezhnevets, Vladimir, Vassili Sazonov, and Alla Andreeva. “A survey on pixel-based skin color detection techniques.” *Proc. Graphicon*. Vol. 3. 2003.

7. Amjad, Ammar, Andrew Griffiths, and Mohammad Patwary. “Multiple face detection algorithm using colour skin modelling.” *Image Processing, IET* 6.8 (2012): 1093-1101.

8. Hashem, Hassan Fahmy. “Adaptive technique for human face detection using HSV color space and neural networks.” *Radio Science Conference, 2009. NRSC 2009. National*. IEEE, 2009.

9. Chai, Douglas, Son Lam Phung, and Abdesselam Bouzerdoum. “A Bayesian skin/non-skin color classifier using non-parametric density estimation.” *Circuits and Systems, 2003. ISCAS’03. Proceedings of the 2003 International Symposium on*. Vol. 2. IEEE, 2003.

10. Ma, Zhanyu, and Arne Leijon. “Human skin color detection in RGB space with Bayesian estimation of beta mixture models.” *18th European Signal Processing Conference (EUSIPCO-2010)*. 2010.

11. Phung, Son Lam, Douglas Chai, and Abdesselam Bouzerdoum. “A universal and robust human skin color model using neural networks.” *Neural Networks, 2001. Proceedings. IJCNN’01. International Joint Conference on*. Vol. 4. IEEE, 2001.

12. Lee, Jae Y., and Suk I. Yoo. “An elliptical boundary model for skin color detection.” *Proc. of the 2002 International Conference on Imaging Science, Systems, and Technology*. 2002.

13. Shih, Frank Y., et al. “Extracting faces and facial features from color images.”*International Journal of Pattern Recognition and Artificial Intelligence* 22.03 (2008): 515-534.

14. Viola, Paul, and Michael Jones. “Rapid object detection using a boosted cascade of simple features.” *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on*. Vol. 1. IEEE, 2001.

15. Jones, Michael, and Paul Viola. “Fast multi-view face detection.” *Mitsubishi Electric Research Lab TR-20003-96* 3 (2003): 14.

16. Jain, Astha, Jyoti Bharti, and M. K. Gupta. “Improvements in OpenCV’s Viola Jones Algorithm in Face Detection-Tilted Face Detection.” *International Journal on Signal and Image Processing* 5.2 (2014): 21.

17. Singh, Dushyant Kumar, and Dharmender Singh Kushwaha. “Tracking Movements of Humans in a Real-Time Surveillance Scene.” *Proceedings of Fifth International Conference on Soft Computing for Problem Solving*. Springer Singapore, 2016.

18. https://www.youtube.com/watch?v=IvQmBwoxp2w