

A Fusion based Approach of Face Detection using Viola – Jones and Skin Color Modeling Technique

Nancy Goyal

U.P.T.U., Lucknow

Pranveer Singh Institute of Technology
Kanpur

Harsh Dev, PhD

Professor, U.P.T.U., Lucknow

Pranveer Singh Institute of Technology
Kanpur

ABSTRACT

Human computer interaction is dealing with different branches of learning that involves the interaction of human with machines directly. For an effective interaction between the machines and humans a user friendly and interactive interface needed. Automatic Face Detection and Recognition proves as a solution for that. Face is a physiological characteristic that indicates what we have in the mind by analyzing the facial expressions such as screaming, happiness and so on. The various factors such as identity, gender, expression, age and pose can be efficiently analyzed by observing the information contained a face. It is very easy for humans to recognize faces because human beings are very good in recognizing process, but it seems to be challenging for machines in the field of computer vision. Therefore, many researchers contribute their interest and attention in this emerging field. In this paper our focus on analyzing the existing faces detection techniques and combining some compatible techniques to form a model with improved efficiency and better detection rate.

Keywords: Face Detection, Face Recognition, Computer-vision, Illumination, Feature and Image.

1. INTRODUCTION

Everyone knew and recognized everyone else in small groups and villages. It is a very basic requirement to identify the individuals precisely. We could easily detect a stranger or identify a potential breach of security. But today, there is more complex society, so it is not simple because more interactions take place electronically, it becomes necessity to have an electronic verification of a person's identity. The magnetic swipe card and password are used for that purpose. The problem related to these forms of electronic identification are that these are not so much secure, as they can be stolen, or lost, or people can failed to remember, and people have found various ways to deceive or forge the authorized ones. It is the major cause of emerging the modern form of electronic verification – Automatic Face detection and Recognition. Automatic Face detection and recognition is an emerging application of computer-vision field that focuses on detecting the human faces in digital images. The selective features of face in the image can be detected, and any other background objects like trees, buildings, bodies, etc. are overlooked. The human face is a strong source of information as we can immediately identify whether person is male or female, what is the approximate age of the person and their emotions also. The detection of facial

region in image, frames or video sequence is usually the first step in various applications such as human computer interaction, video surveillance.

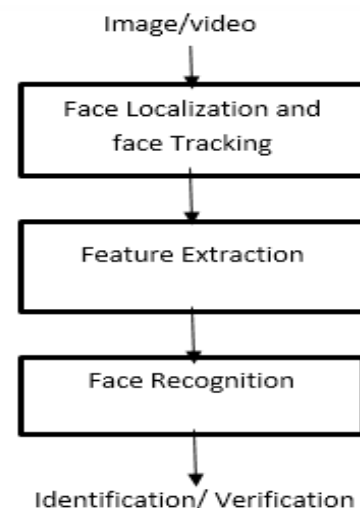


Fig 1: Configuration of a generic face recognition system

This figure shows the steps performed in a generic face recognition system: First the image/video is given as input and then the face is localized and detected in an image. So, face detection is the first step of any face recognition system. Only after face detection, face recognition is performed. There are two type of problems can be raised with it- identification/verification.

In **identification**, the input given to the system is an unknown face, and the system reports the identity of person who he/she is by matching the template to a database of known individuals, whereas in **verification**, the system needs to tell about the claimed identity whether to accept or reject the input face.

2. DIFFERENCES BETWEEN FACE DETECTION AND FACE RECOGNITION

Facial images are the most commonly used physiological characteristics used by human beings to make an individual recognition. Automatic Face Detection and Recognition has extended this idea by machines. Generally, the persons are confused between the terms face

detection and face recognition. This has also led to misinterpretation about the technology. So, we clearly define these terms:

- a) **Face Detection:**
Face detection is a physiological process and it is used for detecting the human faces. The output of this process is used as the input for face recognition.
- b) **Face Recognition:**
Automatic face recognition is a procedure of verifying or identifying an individual automatically from a digital image or from a video sequence. It is performed by matching the template of selective facial features with a repository of face templates. So, it is also called a matching process. Face recognition can identify and remember your face even years after it was first enrolled. It requires the image of face or some selective features that represent the whole data and on the basis of this information face can be identified uniquely and can be compared to other such identity information stored in a repository of facial templates.

3. CHALLENGES TO FACE DETECTION AND RECOGNITION

There are following key factors that can remarkably affect the performance of automatic face detection and recognition system:

- a) **Variations in Poses** introduce deformations and self-occlusion which effects the authentication process. The problem is not solved even if 32 head rotation methods are applied.
- b) **Occlusion** factor also affect face recognition performances in a significant way, especially in group photos where faces may be partially occluded by other objects.

- c) **Variations in illumination** are very difficult to handle in automatic facial recognition. Many 2D methods recognizes well when there is moderate illumination variation, but performances decreases when both illumination and pose are changed.
- d) **Facial expression** is another key factor in recognition. The algorithms are relatively robust to facial expression even with exception of extreme expressions such as screaming, crying etc.
- e) **Time delay** is also an important factor, because as time passes, the face changes in a non-linear way. This problem is difficult to solve as compared to others. Age variation is still a growing area for research work.

4. FACE DETECTION TECHNIQUES

The works done by Hjelmås and Low in the field of face detection shows two broad categories, namely Feature-based approach and Image-based approach.

- a) **Feature-Based Approach:**
This approach first extract features of the image and then match it against the knowledge of face features. For this purpose firstly the images are passed for the preprocessing step (like image enhancement) and then the information of the facial features like the eye brows are darker than the surrounded areas such type of information can be enforced to eyes, nose, and mouth.
- b) **Image or View Based Approach:**
These methods were introduced in an attempt to detect the human face without depending on the knowledge of its geometry. View-based methods treat the face detection problem as a pattern recognition problem by having two separate classes, faces and non faces.

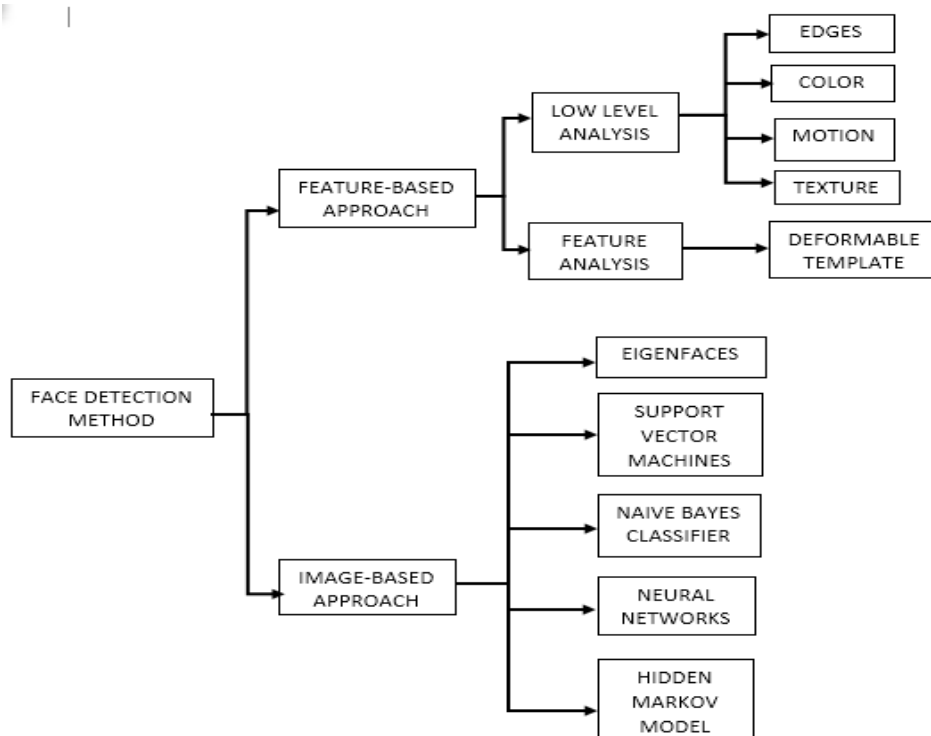


Fig 2: Classification of face detection technique

4.1 Feature- Based Approach

4.1.1 Facial Features

To identify a face from a messy background, a localization method is proposed in [1]. To preserve the face contour the heuristics and edge maps are used by Graf et al. to identify facial features and faces in gray scale images [2]. In [3], Yow and Cipolla presented a method that is based on extracting features from the image and collect their contextual evidence.

4.1.2 Texture

Texture is a very prominent factor to distinguish humans from the other objects. For identify the face-like textures a method is developed by Augusteijn and Skufca [4].

4.1.3 Skin Color

The human skin color has been proven to be an effective feature in many areas of face detection. Despite the actuality that different people have different skin colors, various studies have proven that the difference lies largely between their intensity rather than their brightness [5]. There are many color model have been used to label pixels

as skin including RGB, HSV (or HSI), YCrCb, YIQ and CIE LUV.

4.1.4 Multiple Features

A method is developed by Sobottka and Pitas for localizing the face and extraction of facial features using shape and color [6].

4.1.5 Active Shape Model

In these methods, a typical face pattern (usually frontal) is parameterized manually by a function. Craw et al. presented a localization method based on the contour of a face [7].

4.1.6 Deformable Templates

A model that is based on deformable templates developed by Yuille et al. [8]. In this method firstly a face-shape PDM can be used to locate faces in new images by using active shape model (ASM) search to calculate approximation of the **facial shape parameters**. Then **intensity parameters** are extracted by deforming the face patch to the average shape. These two parameters are collectively used for classification.

Table 1: Advantages and Disadvantages of Feature-Based Approach

Feature-Based Approach	Advantages	Disadvantages
Facial Features	It is easy to find out features like eyes, nose etc.	Some features are difficult to extract.
Skin Color	Avoid exhaustive search for faces in an entire image	Variations in skin color tones
Multiple Features	Face detection is done more efficiently by extracting different facial features.	Time consuming approach.
Active shape model	Computation speed is very high.	Computationally expensive approach
Deformable Templates	Finding out the optimal value without the need for good initialization.	High computational complexity, Unexpected shrinking and rotation of the templates.

4.2 Image or View Based Approach

It is opposite to the template matching methods where the templates are predefined by experts, the “templates” in appearance-based or image-based methods are learned from the examples in the images.

4.2.1 Eigen faces

Facial images can be linearly digitized using a moderate number of basis images by using a model presented by Kirby and Sirovich [9]. It is also called principal component analysis. Turk and Pentland used PCA technique to face recognition and detection [10].

4.2.2 Neural Networks

Neural networks have been used successfully in many pattern recognition problems like object recognition, optical character recognition. A method based on neural networks that is hierarchical in nature was proposed by Agui et al. [11]. By increasing the layers of neural

networks their performance can be improved. Rowley [12] proposed a framework of detecting rotating faces using a router network.

4.2.3 Support Vector Machines

Face detection was first performed with Support Vector Machines (SVMs) in [13] where an efficient method proposed to train an SVM for large scale problems, and used it to face detection.

4.2.4 Naive Bayes Classifier

Schneiderman and Kanade enhanced the Naïve Bayes Classifier method with representation of the wavelet for detecting profile faces and cars [14].

4.2.5 Hidden Markov Model

KLT and HMMs were collectively used to do face localization and recognition in a method presented by Nefian and Hayes [15].

Table 2: Advantages and Disadvantages of Image –Based Approach

Image- Based Approach	Advantages	Disadvantages
Eigenfaces	Without any processing raw intensity data are used directly for recognition, Compression of data is achieved, Simple and fast	Time consuming approach, Sensitive to pixel misalignment, Cannot separate image variances caused by identity and pose variation
Neural Networks	Reduces misclassifications among the neighborhood classes.	Breach of privacy, Expensive to implement
Support Vector Machines	Effective in high dimensional space,	Large memory required, Time consuming approach
Naive Bayes Classifier	Easy to implement, Require a small amount of training data to estimate the parameters, Efficient computation	Strong feature independence assumption therefore loss of accuracy
Hidden Markov Model	Computational speed is fast which makes it practical for real time applications	Expensive approach, Sensitive to geometrical shape

5. EXISTING APPROACHES

5.1 Viola and Jones Algorithm

Viola and Jones presented an algorithm for face detection that has mainly involved 3 techniques:

5.1.1 Haar like features

All human faces have some common properties. By using this knowledge particular features can be constructed that are called Haar Features. Human faces having some general and common properties that are:

- The upper-cheeks are brighter than the eye region.
- The upper bony area of a nose is brighter than the eyes.
- Useful domain knowledge
- Location - Size: eyes & nose bridge region
- Value: darker / brighter

The combination of two or three or four linked black and white rectangles made up a sub window that renowned as Haar like feature. The value of a haar like feature is computed as

$$\text{Rectangle features Value} = \Sigma (\text{pixels in black area}) - \Sigma (\text{pixels in white area})$$

5.1.2 Integral image

Integral images are very useful for calculating the average intensities of an image. Integral image at location(x, y) comprises the sum of the pixel values in the above and left of the x and y [16]. The haar like features of any size in a

rectangle can be calculated quickly using integral image in a constant time.

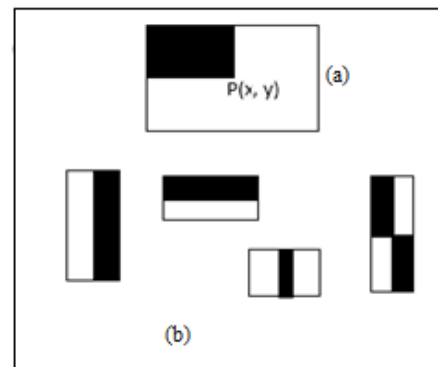


Fig 3: (a) Integral Image (b) Haar-like Features

5.1.3 Adaboost learning rule

Adaboost learning rule is an iterative algorithm, in order to improve the detection accuracy use cascading approach. In this, each stage consists of a weak classifier and together this is called strong classifier based cascading approach. So all the features are grouped into several stages where each stage has certain number of features and as the final output is all the sub widows are grouped as facial region or non-facial region.

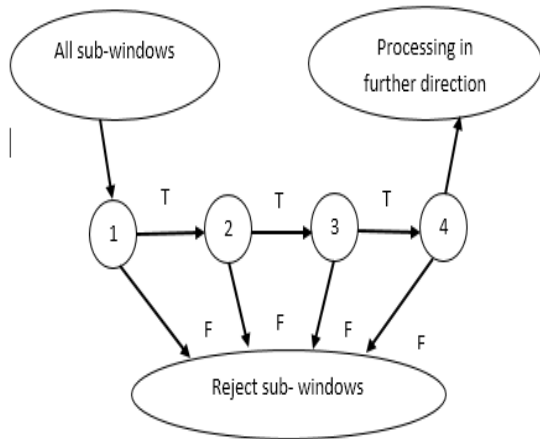


Fig 4: Cascading decision making sub window

5.2 Skin Color Model

For the segmentation of facial region from the image skin color is used as the renowned factor. For this purpose, a stable skin color model required that is adaptable to people having different skin colors and varying illumination conditions. The RGB color space having three components (r, g, and b) that represents colors as well as their intensities also. The lighting of the surroundings affects a person's face resulting in varying brightness condition. It can be removed from the representation of color by using chromatic color space model. Chromatic colors, famous as "pure" colors in the absence of the intensities levels of the colors, are defined by a normalization process as below:

$$r = R/(R+G+B)$$

$$b = B/(R+G+B)$$

There are various applications that use chromatic colors for the segmentation of color images. These colors have been used very effectively to segment skin areas from non-skin areas. In spite of the fact that, skin colors of different people are closely related, but they differ mainly in their intensities levels. As the result, we could develop a skin-color model in the chromatic color space that is renowned as Hue Saturation Value (HSV) model. There are three ingredients included in HSV color space model i.e. hue, saturation and value. Hue indicated as the type of color, where the range lies between 0 and 360, each and every value represents to one color. Saturation refers to the amount of grey lies between black and white. While Value represents the luminance of the color, which also lies between 0 to 100%. Compared with RGB model on the lighting issue, HSV model can achieve better results. Therefore, we use the HSV color space to build up a skin-color model to segment face from non-face region.

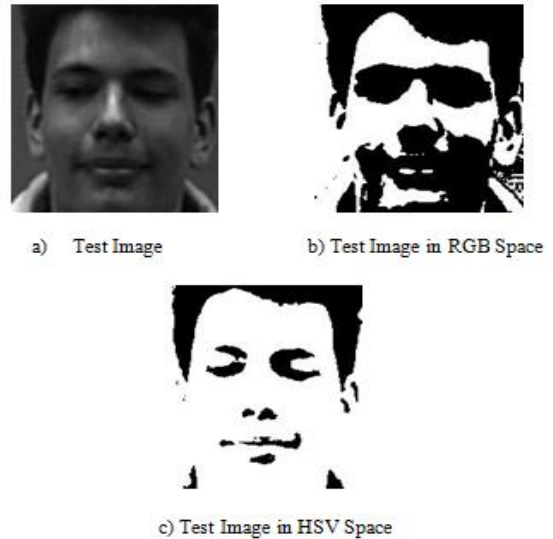


Fig 5: Skin Color Model Technique

5.3 Motion history image

The most commonly used methodology for doing localization on the input scene is MHI. MHI is defined as the cumulative motion over a specified duration and used for the identify object motion [17]. Assuming that the camera is fixed and only the recorded subject's head is moving, the facial part can be easily identified by differentiating the foreground part containing head movement from the fixed background. Generally MHI is classified into two types; point-wise and region-wise. Point-wise MHI is defined as the amount of pixel-wise spatial-temporal variation [18], while region-wise MHI is the amount of region-wise variation [19]. Motion history images, regardless of the type, can be achieved by the absolute difference between adjacent frames to get the frame motion silhouettes, and the profile of these silhouettes is used as a threshold in binary images to get the motion history over the duration. The noise present in the images are removed through morphological operations such as erosion and dilation for improving the robustness of the image.

6. PROPOSED WORK

My proposed work is combining the virtue of each technique to develop a hybrid approach that will achieve a better detection performance than using only a separate technique. As a result, for frontal to frontal face detection, we choose Viola and Jones algorithm, while for profile to frontal face detection, we use Skin Color Model combined by Background Subtraction Technique for overcoming the 'pose- variation' problem and a series of morphological operations are performed to do facial detection [20].

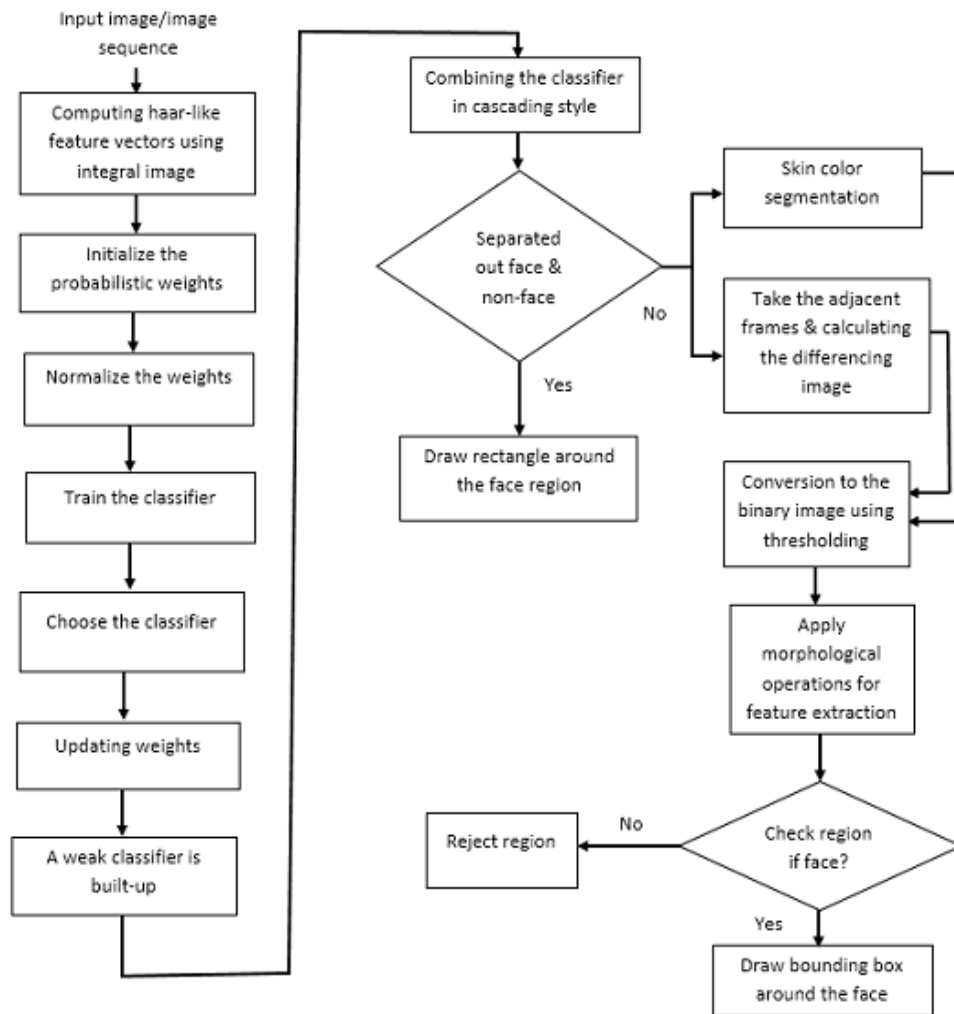


Fig 6: Proposed Framework for Face Detection

7. EXPERIMENTAL RESULT

We have implemented the above introduced hybrid approach combined by Viola and Jones algorithm, Skin Color Model and Motion History Image on MATLAB for performing facial detection. For MHI based face detection, first frame is used as a still image and then the current frame is subtracted with the still frame to find the moving object(face), the connected objects are separated and noises are removed by the morphological operations, and finally, we mask the subtracted frame containing object(face). For skin based detection, we first build a skin pixel model based on a sample skin image and then converted this Image into HSV format, then analyzed each pixel among this image and finally computed the mean/ variance for all three channels for a skin pixel.

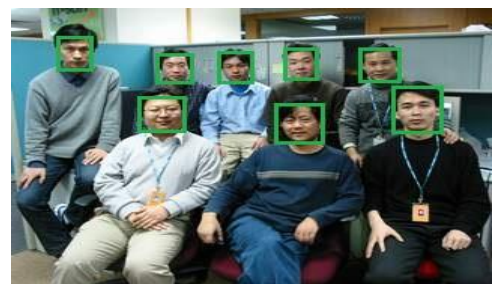


Fig 7: The typical input image through the detector



Fig 8: Gray scale image detected through proposed detector

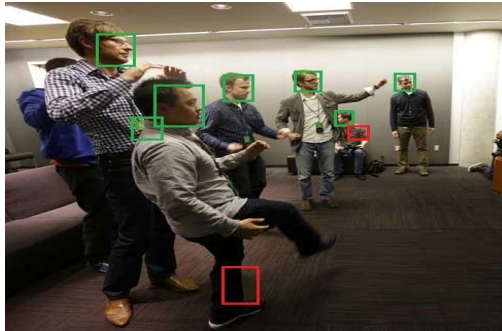


Fig 9: Test image with merged detection

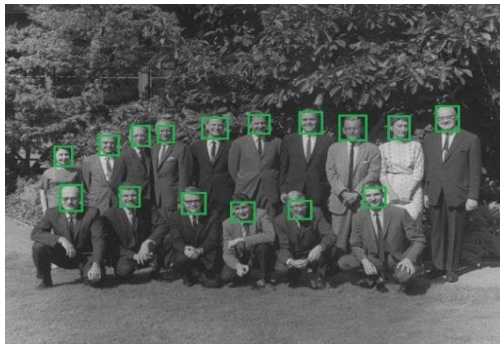


Fig 10: Face detection result set

Face Detection Approaches	Detection Rate
Viola and Jones Algorithm	20%
Skin Color Model Technique	30%
Skin Color Model Technique with Background Subtraction Technique	50%
Proposed Technique	74.33%

Table 3: Model Summary of Proposed Work

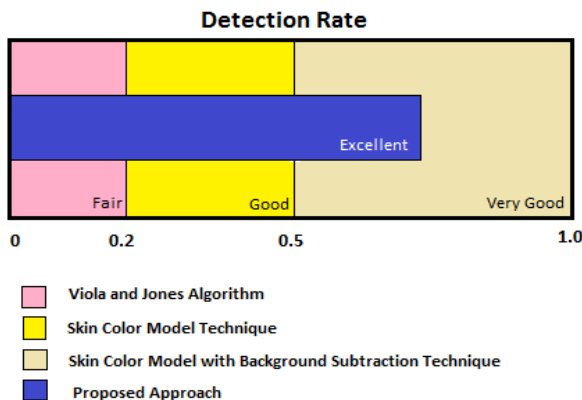


Fig 11: Silhouette Measure of Proposed Work

8. CONCLUSION AND FUTURE SCOPE

Here we have analyzed the different techniques which have been used earlier in automatic face detection. Like in feature based approach we try to extract features of the face features, while in image based approach the human face is detected without depending on the knowledge of its geometry. We have combined these 3 techniques Viola and Jones Algorithm, Skin Color Model and Background Subtraction Technique to get the combined advantages of them and to generate more efficient multimodal face detection system, which will have the advantages of these combined methodologies and thereby improving the efficiency by better detection performance. In future, this system can be used online and provide an efficient detection rate.

9. REFERENCES

- [1] S.A. Sirohey, "Human Face Segmentation and Identification," Technical Report CS-TR-3176, Univ. of Maryland, 1993.
- [2] H.P. Graf, T. Chen, E. Petajan, and E. Cosatto, "Locating Faces and Facial Parts," Proc. First Int'l Workshop Automatic Face and Gesture Recognition, pp. 41-46, 1995.
- [3] K.C. Yow and R. Cipolla, "Feature-Based Human Face Detection, Image and Vision Computing," vol. 15, no. 9, pp. 713-735, 1997.
- [4] M.F. Augusteijn and T.L. Skujca, "Identification of Human Faces through texture-based feature recognition and Neural Network Technology," Proc. IEEE Conf. Neural Networks, pp. 392-398, 1993.
- [5] D. Chai and K.N. Ngan, "Locating Facial Region of a Head-and- Shoulders Color Image," Proc. Third Int'l Conf. Automatic Face and Gesture Recognition, pp. 124-129, 1998.
- [6] K. Sobottka and I. Pitas, "Face Localization and Feature Extraction Based on Shape and Color Information," Proc. IEEE Int'l Conf Image Processing, pp. 483-486, 1996.
- [7] I. Craw, H. Ellis, and J. Lishman, "Automatic Extraction of Face Features," Pattern Recognition Letters, vol. 5, pp. 183-187, 1987.
- [8] A. Yuille, P. Hallinan, and D. Cohen, "Feature Extraction from Faces Using Deformable Templates," Int'l J. Computer Vision, vol. 8, no. 2, pp. 99-111, 1992.
- [9] M. Kirby and L. Sirovich, "Application of the Karhunen-Loeve Procedure for the Characterization of Human Faces," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 12, no. 1, pp. 103-108, Jan 1990.
- [10] M. Turk and A. Pentland, "Eigenfaces for Recognition," J. Cognitive Neuroscience, vol. 3, no. 1, pp. 71-86, 1991.

- [11] T. Agui, Y. Kokubo, H. Nagashashi, and T. Nagao, "Extraction of Face Recognition from Monochromatic Photographs Using Neural Networks," Proc. Second Int'l Conf. Automation, Robotics, and Computer Vision, vol. 1, pp. 18.8.1-18.8.5, 1992.
- [12] H. Rowley, S. Baluja, and T. Kanade, "Rotation Invariant Neural Network-Based Face Detection," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 38-44, 1998.
- [13] E. Osuna, R. Freund, and F. Girosi, "Training Support Vector Machines: An Application to Face Detection," Proc. IEEE Conf. Computer Vision and Pattern Recognition, pp. 130-136, 1997.
- [14] H. Schneiderman and T. Kanade, "A Statistical Method for 3D Object Detection Applied to Faces and Cars," Proc. IEEE Conf. Computer Vision and Pattern Recognition, vol. 1, pp. 746-751, 2000.
- [15] A.V. Nefian and M. H. H, "Face Detection and Recognition Using Hidden Markov Models," Proc. IEEE Int'l Conf. Image Processing vol. 1, pp. 141-146, 1998.
- [16] Wilson, P. I. and Fernandez, "Facial Feature Detection using Haar Classifiers," JCSC 21, April 2006.
- [17] Boblck, A. F. and Davis, J. W., "The Recognition of Human Movement Using Temporal Templates," IEEE Trans. on PAMI vol. 23, pp. 257- 267, 2001.
- [18] Essa, I. and Pentland, "Coding Analysis, Interpretation and Recognition of Facial Expressions," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 7, pp. 757- 763, 1997.
- [19] Park H. and Park J. I., "Analysis and Recognition of Facial Expression Based on Point-wise Motion Energy," Springer Berlin/Heidelberg, vol. 3212, 2004.
- [20] Nidhi Srivastava, Dr. Harsh Dev and S. Qamar Abbas, "Framework for Face Detection," International Journal of Computer Applications (0975 – 8887) Vol. 58– No.17, November 2012.