A SURVEY ON REAL TIME HUMAN DETECTION AND FACE DETECTION METHODS

[1]Rinu Roy, [2]Ani Sunny
[1]Department of Computer Science and Engineering,
[1]Mar Athanasius College of Engineering, Kothamangalam, Kerala, India.

Abstract: Nowadays the popularity and demand of image processing are increasing due to its immense number of application in various fields. Human detection and face detection in real time has been one of the active research topics in computer vision and machine learning due to potential applications including video surveillance, human computer interface, face recognition, and image database management. There are a different number of works on human detection and face detection. The complete process of human detection and face detection covers in different stages. Various techniques are then needed for all of these stages. Also, these techniques vary from various surrounding factors such as face orientation, expression, lighting, and background. This paper presents the complete study and review of various techniques used in human detection and face detection stages under different conditions.

IndexTerms **– Gaussian Mixture Model(GMM) , Viola- Jones Algorithm , Cascading Classifier.**

I. INTRODUCTION

Identifying the presence of humans in video streams is one of the most important features that need to be extracted. However, this task becomes more complicated in the presence of different variations in brightness, lightings, contrast levels, poses, and backgrounds. In video based surveillance human detection plays a crucial role. The aim of video surveillance is to identify and monitor humans for security purposes in crowded scenes such as airports, train stations, supermarkets etc. The video footage captured by the installed cameras are processed to detect and track the full human body or body parts in the scene. There are different methods used in the past for human detection. This work is to analyze the different methods used for human detection from videos and various algorithms used for face detection and identify the pros and cons of each method.

Biometric-based technologies include identification based on physiological characteristics. Face recognition appears to offer several advantages over other biometric methods. Face recognition can be done passively without any explicit action or participation on the part of the user since face images can be acquired from a distance by a camera. This is particularly beneficial for security and surveillance purposes. There are numerous application areas in which face recognition such as Surveillance, General identity verification, Criminal justice systems, Image database investigations, Smart Card" applications, Multi-media environments with adaptive human computer interfaces and Video indexing.

Face recognition is a challenging and interesting research topic in the field of pattern recognition which has been found a widely used in many applications. Thus many face recognition algorithms have been proposed and different survey in this area can be founded.

II. OVERVIEW OF HUMAN DETECTION METHODS

This section describes various human detection methods. In the history of human object classification various features have been studied. These features can be computed from low-level information such as edge, texture, color or motion.

2.1 Skin Area Detection

Color information of the digital image is utilized to find those areas close to human skin color. This stage helps in reducing the search space and speeds up the simulation by consuming the processing time efficiently. However, skin test is not enough to detect human faces as it will also detect other parts of the body as well as other non face skin colored objects. Thus, other tests to filter out those unwanted areas should be applied. Further stages in the proposed project are designed to gradually eliminate the false detected areas found at this stage. The first test to remove the unwanted skin like areas was chosen to be movement detection[1].

2.2 Shape Based Detection

To encode the shape of an object, edge-base features have been commonly used. The advantage of these features is that shape of an object can be properly captured by the edges. Parallel edge segments and rectangular contours were employed to model human body parts.

2.3 Moving Object Detection

To minimize the errors in face detection we can use the human nature, that human will have at least small amount of movements such as eyes blinking and/or mouth and face boundary movements. We can get this information very easily because we are dealing with the video sequence. It helps to get the whole sequence of the object's movements can be obtained. Taking that point into account we can reduce the error that occurs due to false detection of a human face and minimizes the time of simulation.

However, it is important to take in to account that a change may occur due to several sources such as moving objects, presence or absence of objects camera movement and zooming, brightness changes. This means that some changes are significant and others are not. Although it is difficult to take a decision whether a detected movement is significant or not, it is an important step to remove unwanted changes and focus the processing only on those changes of interest. Hence, movement detection was chosen as a vital stage in the proposed design.

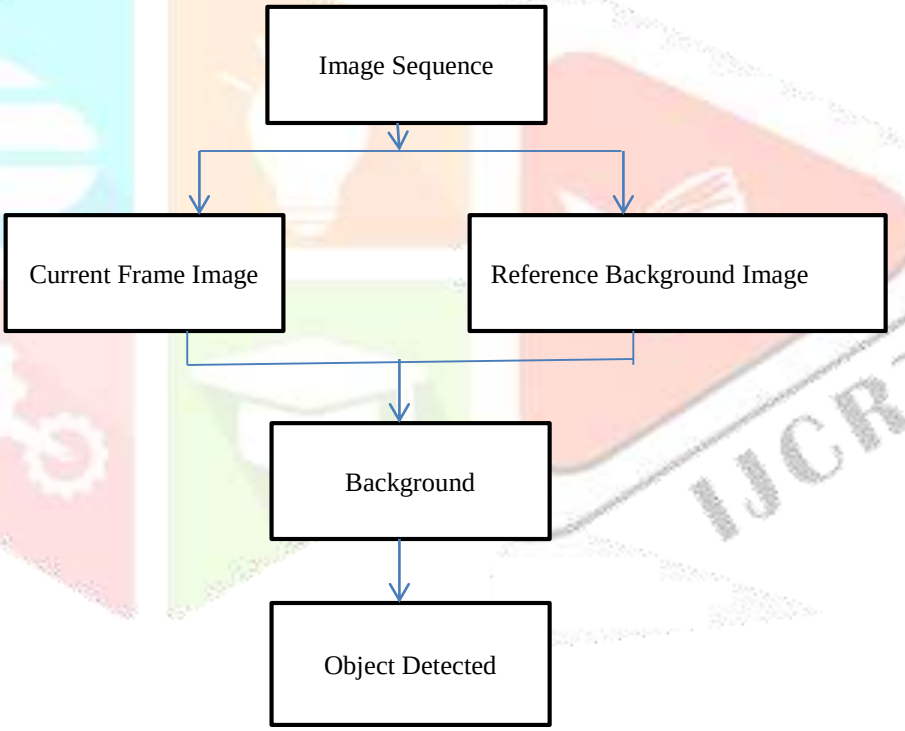

Figure 1: Flowchart of moving object detection

Object detection can be achieved by building a representation of the scene called the background model and then finding deviations from the model for each incoming frame. Any significant change in an image region from the background model signifies a moving object. The pixels constituting the regions undergoing change are marked for further processing. Usually, a connected component algorithm is applied to obtain connected regions corresponding to the objects. This process is referred to as the background subtraction [15].

There are different methods used for finding moving objects. The main 2 methods used are frame difference method and Gaussian Mixture Model.

2.3.1 Frame difference Method

Detection of moving object from a sequence of frames captured from a static camera is widely performed by frame difference method. The objective of the approach is to detect the moving objects from the difference between the existing frame and the reference frame. The frame difference method is the common method of motion detection [10].

$|frame_i - frame_{i-1}| > Threshold$

2.3.2 Optical flow

Optical flow method is to calculate the image optical flow field, and do clustering processing according to the optical flow distribution characteristics of image. This method can get the complete movement information and detect the moving object from the background better, however, a large quantity of calculation, sensitivity to noise, poor anti-noise performance, make it not suitable for real-time demanding occasions.

2.3.3 Gaussian Mixture Model

Gaussian Mixture Model (GMM) is a parametric probability density function. It represented as a weighted sum of Gaussian component densities. GMMs are commonly used as a parametric model of the probability distribution of continuous measurements or features in a biometric system, such as color based tracking of an object in video.

In many computer related vision technology, it is critical to identify moving objects from a sequence of videos frames. In order to achieve this, background subtraction is applied which mainly identifies moving objects from each portion of video frames. By using Gaussian Mixture Model background model, frame pixels are deleted from the required video to achieve the desired results. The application of background subtraction involves various factors which involve developing an algorithm which is able to detect the required object robustly, it should also be able to react to various changes like illumination, starting and stopping of moving objects [10].

The advantages of GMM are different threshold is selected for each pixel, These pixel-wise thresholds are adapting by time, Objects are allowed to become part of the background without destroying the existing background model and it provides fast recovery. It also contains some disadvantages that are it cannot deal with sudden, drastic lighting changes, initializing the Gaussian is important, and there are relatively many parameters and they should be selected intelligently.

2.4 Improved Template Matching Method

Improved template matching method that employs a generalized distance transform(GDT) and an orientation map in computing the dissimilarity between the template and given object. The GDT allows us to weigh the distance transform more on the strong edge points and the orientation map provides supplementary local orientation information for matching. The orientation map is calculated from local gradients. The improved template matching method by applying to detect humans from given test images [4].

III. FACE DETECTION METHODS

It is a fundamental part of the face recognition system. It has ability to focus computational resources on the part of an image containing face. Face detection involves the separation of image into 2 parts, one containing the face and the other containing the background. It is difficult because although commonalities exist between faces, they can vary considerably in terms of age, skin color and facial expression.

Hjemal and Low [6] divides the face detection techniques into two categories named feature based techniques and image based techniques.

3.1 FEATURE-BASED APPROACH

The feature based approaches use the facial features to their detection process. Hjemal and Low [6] further divide this technique into three categories: low level analysis, feature analysis and active shape model.

3.1.1 Low level analysis: It deals with the segmentation of visual features by using the properties of pixels, gray scale level, and motion information. In [17], implemented an edge representation method for detecting the facial features in line drawings by detecting the changes in pixel properties. In [2], developed this further to detect human head outline. Edge features within the head outline are then subjected to feature analysis using shape and position information of the face. The edge based techniques rely upon the labeled edges which are matched to a face model for verification. In an edge-detection-based approach to face detection, edges need to be labeled and matched to a face model in order to verify correct detections.

Govindaraju [18] accomplishes this by labeling edges as the left side, hairline, or right side of a front view face and matches these edges against a face model by using the golden ratio for an ideal face. Govindaraju's edge-based feature extraction works in the following steps:

a) **Edge detection:** The Marr-Hildreth edge operator
b) **Thinning:** A classical thinning algorithm from Pavlidis
c) **Spur removal:** Each connected component is reduced to its central branch
d) **Filtering:** The components with non-face-like properties are removed
e) **Corner detection:** The components are split based on corner detection
f) **Labeling:** The final components are labeled as belonging to the left side, hairline, or right side of a face

Generally eyebrows, pupils and lips appear darker than surrounding regions and thus extraction algorithms can search for local minima. In contrast, local maxima can be used to indicate the bright facial spots such as nose tips. Detection is then performed using the low-level gray-scale thresholding.

3.1.2 Feature analysis: It uses an additional knowledge about the faces and it removes the ambiguity produces by low level analysis. The first involves sequential feature searching strategies based on the relative positioning of individual facial features [12]. Initially prominent facial features are determined which allows less prominent features to be hypothesized.

3.1.3 Active shape models: It is used to define the actual physical and higher-level appearance of features. These models are developed by Tim Cootes and Chris Taylor in 1995. These models are released near to a feature, such that they interact with the local image, deforming to take the shape of the feature [6]. Active Shape Models are models of the shapes of objects which iteratively deform to fit to an example of the object in a new image. It works in the following two steps: Look in the image around each point for a better position for that point, update the model parameters to best match to these new found positions.

3.2 IMAGE BASED TECHNIQUES

Face detection of facial features by explicit modeling is very trivial approach because it may be troubled by the unpredictability of faces and environmental conditions. So there is a need for more robust techniques, capable of performing in unfriendly environments, such as detecting multiple faces with clutter-intensive backgrounds. Image based face detection has inspired a new research area and by virtue of this face detection is treated as a general pattern recognition problem. The image based approach contains the various approaches like neural networks, example based learning, support vector machine [12] .

Some image processing techniques extract feature points such as eyes, nose, and mouth and then used as input data toward the application. Various approaches have been proposed to extract these facial points or features from the images. The basic approaches are as follows.

3.2.1 Geometry –based Technique

Here the feature are extracted using the size and the relative position of important components of images. In this technique, firstly the direction and edges of important component is detected. Then building feature vectors from these edges and direction. Canny filter and gradient analysis usually applied in this direction. Second, method is based on the gray scales difference of unimportant components and important components, by using feature blocks, set of Haar-like feature block in Adaboost method [8] to change the gray scales distribution into the feature. In LBP [14] method every face image is divides into blocks and each block has its corresponding central pixel. Then this method examine its neighbor pixels, based on the gray scales value of the central pixel it changes neighbor to 0 or 1. After that a histograms is built for every region. Then these histograms are combined to a feature vector for the face image.

3.2.2 Template Based Technique

This technique extracts the facial feature using appropriate energy function .This method have been proposed by Yuille et al. [21], detecting and describing features of faces using deformable templates. In deformable templates the feature of interest, an eye for example is described by a Parameterized template. These parameterized templates enable a priori knowledge about the expected shape of the features to guide the detection process [21]. An energy function is defined to links the peaks, edges, and valleys in the image intensity with corresponding properties of the template. After that the template matching is done with the image, thereby deforming itself to find the best fit. For the descriptor purpose they used a final parameter value. In the Template based first an eye template is used to detect an eye from the image. Then a correlation is found out between the eye templates with various overlapping regions of the face image. Eye region have the maximum correlation with the template.

3.2.3 Appearance Based Approach:

This method process the image as two dimensional patterns. The concept of "feature" in this approach is different from the simple facial features such as eyes and mouth. Any extracted characteristic from the image is referred to as a feature. This method group found the best performer in facial feature extraction because it keeps the important information of image and rejects the redundant information. Method such as principal component analysis (PCA) and independent component analysis (ICA) are used to extract the feature vector. The main purpose of PCA is to reduce the large dimensionality of observed variable to the smaller intrinsic dimensionality of independent variable without losing the important information [9]. It is observed that many natural signals, including speech, natural images, are better described as linear combinations of sources with super-Gaussian distributions. In that case, ICA method is better than PCA method because:

- ICA provides a better probabilistic model of the data.
- It uniquely identifies the mixing matrix.
- III) It finds an unnecessary orthogonal basic which may reconstruct the data better than PCA in the presence of noise such as variations lighting and expressions of face.
- It is sensitive to high order statistics in the data, not just the covariance matrix

3.2.4 Color Based Method:

Skin region can be detected with the help of different color models like RGB [19]. The image obtained after applying skin color statistics is subjected to binarization. Firstly, it is transformed to gray-scale image and then to binary image by applying suitable threshold. All this is done to eliminate the color and saturation values and consider only the luminance part. After this luminance part is transformed to binary image with some threshold because the features for face are darker than the background colors. After a threshold value is applied noise is removed by applying some opening and closing operation. Then eyes, ears, nose and facial features can be extracted from the binary image by considering the threshold for areas which are darker in the mouth than a given threshold. After getting the triangle, it is easy to get the coordinates of the four corner points that form the potential facial region[16].

3.3 VIOLA-JONES ALGORITHM

The other method came into exist for face detection is Viola Jones algorithm. This is the first ever real-time face detection system. There are three ingredients working in concert to enable a fast and accurate detection: the integral image for feature computation, Adaboost for feature selection and an attentional cascade for efficient computational resource allocation. Here they propose a complete algorithmic description, a learning code and a learned face detector that can be applied to any color image. Since the Viola-Jones algorithm typically gives multiple detections, a post-processing step is also proposed to reduce detection redundancy using a robustness argument.[20]

This algorithm helps to detect features of a face in a particular frame of a video sequence. This is the first object detection framework which gives a competition to real time detection rates. Paul Viola and Michael Jones are the ones who introduced this algorithm. They made this algorithm mainly by the issue of face detection. There are four steps which have to be followed to detect a face. Firstly, we train the system with the haar features [20]. Haar features are a kind of rectangular boxes which are black and white. Haar features are simple rectangular feature which is the difference of the sum of pixels of areas inside the rectangle. This rectangle can be at any position of the frame and can scale the image. This modified feature set is called 2-rectangle feature. Each feature type can indicate the existence or the absence of certain characteristics in the frame, such as edges or changes in texture.

These haar features are applied to determine the facial features. For example in the Figure 2 (b) is used to detect nose feature of a human face as the black colored part defines the presence of a nose which is located at the center of the face. And the Figure-2 (e) is called 4 rectangle feature. Where the black part is denoted as +1 and the white part is denoted as -1. The result is calculated by subtracting the sum of pixels under the white rectangle from the sum of pixels under black rectangle. Initially some threshold is taken for particular features. Average sum of each black and white is calculated. Then difference is checked with threshold. If the value is above or matches with the threshold then it is detected as relevant feature [11].

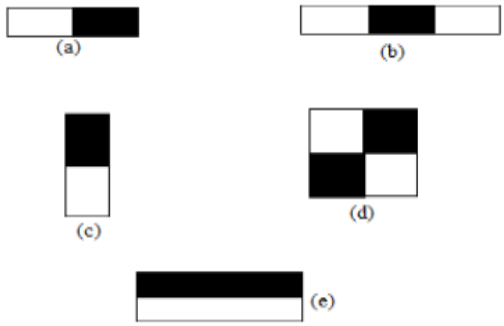

Figure 2 Haar features

There are three methods used to speeded up the detection

a) Integral image

This is a fast way to compute simple features. The integral image part is used to sum all the pixels of a particular box to its left and above ones. The four corner values of the area are to be calculated. This makes avoid summing of each pixel in the region. This integral image conversion process is introduced to speed up the process in calculating pixels [5].

b) Adaboost

In Adaboost the weak learner is nothing but a feature selector. The advantage of this is that if there are *N* weak learners there are merely *N* features to compute. It is a process used to find out the important and unimportant features. It uses the weak classifiers and weights to form a strong classifier. It finds the single rectangular feature and threshold which is the best to separate faces and non-faces in training examples in terms of weighted error. It firstly starts with uniform weights while training. Next it evaluates the weighted error for each feature and takes the best [7]. We reevaluate the examples where incorrect classifiers will have more weight and correct classifiers will have less weight. Finally the classifier will contain the combination of correct classifiers which are having less weight. To reduce the computational time non-faces are discarded.

c) Cascading

Cascaded combination of classifiers is used. Here most of true negatives are rejected very fast at the first few stages. It can keep high detection rate and low false positive rate. This step is introduced to speed up the process and give an accurate result. This step consists of several stages where each stage consists of a strong classifier. All features are grouped into different stages. It detects faces in the frame by sliding a window over the frames [13].

IV. CONCLUSION

This paper discussed various human detection and face detection techniques. Both are the integral and important part of surveillance. Human detection is a crucial task in many computer vision applications including image and video content management, human motion analysis etc. For human detection in real time moving object detection plays an important role. Gaussian Mixture Model helps to find all the moving objects.

Face detection also has different applications nowadays. Feature-based methods help to get the facial features, Template-based methods are easy to implement but it does not represent global face structure. While color segmentation based methods used for skin detection with morphology operation to detect features. So different color model and illumination variation factors can affect performance. Appearance-based methods represent the optimal feature points which can represent global face structure. Viola Jones algorithm the most efficient and commonly used algorithm for face detection.

REFERENCES

[1]. B. Hore, S. Mehrotra, and G. Tsudik, "A Privacy-Preserving Index for Range Queries," Proc. 30th Int'l Conf. Very Large Data Bases (VLDB'04), pp. 720-731, Aug. 2004.

[2]. Craw, I., Ellis, H. and Lishman, "Automatic extraction of face feature".Pattern Recog. Lett. 183- 187 1987.

[3]. D. Hari Hara Santosh, P. Venkatesh, P. Poornesh, L. Narayana Rao, N. Arun Kumar, "Tracking Multiple Moving Objects Using Gaussian Mixture Model", International Journal of Soft Computing and Engineering (IJSCE) ISSN: 2231-2307, Volume-3, Issue-2, May 2013

[4]. Duc Thanh Nguyen " Human detection from Images and Videos", University of Wollongong Research Online 2012.

[5]. Dmitry Mikhaylov, Anton Samoylov, Peter Minin and Alexey Egorov. 2014. Face Detection and Tracking from Image and Statistics Gathering. IEEE Conference on Signal-Image Technology and Internet-Based Systems.

[**6].** Erik Hjelmas, Boon Kee Low, "Face Detection: A Survey," Computer Vision and Image Understanding, 83, 236-274 April 2001

[7]. L. Stan and Z. Zhang" FloatBoost learning and statistical face detection", IEEE Trans. On Pattern Analysis and Machine Intelligence. 26(9). 2004

[8]. M. Jones and P. Viola, "Face Recognition Using Boosted Local Features", IEEE International Conference on Computer Vision,2003

[9]. M. Turk and A. Pentland. Eigenfaces for recognition Journal of Cognitive Neuroscience, 3(1):71–86, 1991.

[10]. Nishu Singla "Motion Detection Based on Frame Difference Method". ISSN 0974-2239 Volume 4, Number 15 (2014), pp. 1559-1565.

[11]. Peter Gejguš and Martin Šperka," Face tracking for expressions simulations", International Conference on Computer Systems and Technologies, 2003.

[12]. Phil Brimblecombe, "Face detection using neural networks".

[13]. Ragini Choudhury Verma, Cordelia Schmid and Krystian Mikolajczyk"Face Detection and Tracking in a Video by Propagating Detection Probabilities" IEEE Transactions on Pattern Analysis And Machine Intelligence. 25(10),2008

[14]. Sanjeev Dhawan, himanshu Dogra, "Feature Extraction Techniques for Face Recognition", International Journal of. Engineering, Business and Enterprise Applications (IJEBEA), 2012

[15]. Sarika Patel, Prof. Udesang K. Jaliya , Prof. Mahashweta J. Joshi, "A Review : Object Detection and Object Tracking Methods", International Journal of Innovative and Emerging Research in Engineering Volume 2, Issue 1, 2015

[16]. Shu Liao, Wei Fan, Albert C. S. Chung and Dit-Yan Yeung, "Facial Expression Recognition Using Advanced Local Binary Patterns, Tsallis Entropies And Global Appearance Features", IEEE International Conference on Image Processing, pp.665-668, 2006

[17]. Toshiyuki Sakai, M. Nagao, Takeo Kanade, "Computer analysis and classification of photographs of human face," First USA Japan Computer Conference, 1972.

[18]. V. Govindaraju, Locating human faces in photographs, *Int. J. Computer. Vision* **19**, 1996.

[19]. V.Vezhnevets, V.Sazonov ,A. Andreeva ," A Survey on Pixel- Based Skin Color Detection Techniques", Graphics and Media Laboratory, Moscow State University, Moscow, Russia

[20]. Yi-Qing Wang "An Analysis of the Viola-Jones Face Detection Algorithm", Image Processing On Line,2014

[21]. Yuille, A. L., Cohen, D. S., and Hallinan, P. W., "Feature extraction from faces using deformable templates", Proc. of CVPR,1989

.

.