HUMAN ACTIVITY RECOGNITION USING DIGITAL IMAGE PROCESSING

[1]Dr.P.Sreenivasulu, [2]N.Vinamrutha, [2]Ch.Sai Nithinya, [2]M.Chandhana, [2]M.Vineeth Kumar, [2]G. Jeevan Kiran

[1] Professor, Department of ECE, Narayana Engineering College, Gudur, AP, 524101

[2]UG Student, Department of ECE, Narayana Engineering College, Gudur, AP, 524101

Abstract: The paper presents Haar feature extraction classifier-based image processing technique for the analysis of human facial behavioral expressions. As an improvisation, some of instances misclassified by Haar classifier are further detected by subsequent classifier adaptive algorithm (Adaboost). Careful attention is paid while generating the templates of various facial behavioral expressions including eye open, close, drowsy and smile. The obtained results create the possibilities of adapting an embedded image processor with the developed algorithms for the real time applications. The image processing algorithm is implemented by using MATLAB.

Keywords: Human Facial Expression Identification, HAAR Classifier, Adaboost, Image Processing.

1.INTRODUCTION

The necessity to have an algorithm to analysis as well as automatically recognize human facial behavioral expression from an online image has been understood inevitable. It is also one of the most researched areas and all differs either by various usage of techniques or by rate of accuracy. As face detection is the first step of any face processing system, it has numerous applications in face recognition, face tracking, facial expression recognition, facial feature extraction, gender classification, clustering, attentive user interfaces, digital cosmetics, biometric systems.

While most of the face detection algorithms can be extended to recognize other objects such as cars, humans, pedestrians, and signs, this can be adapted to differentiate and also to recognize the distinct parts and facial features like eyes, nose, mouth, eye brow and then skin. It is well understood that human emotions also have acoustic characteristics. Although the combination of acoustic and visual characteristics promises improved recognition accuracy, the development of effective combination techniques is a challenge, which has not been addressed by many researchers. In this study, real time facial feature tracker has been considered to deal with the problems of face localization and feature extraction in spontaneous expressions.

The main objective of this study is to generate a set of facial features eventually adapt feature extraction classifiers for tracking and recognizing face in an image eventually much facial expression with minimal error rate. The algorithms are preliminarily applied to discriminate the images with face and images without face. The technique equally utilizing webcam with IR light as it has an advantage of being invisible to the naked eye, removing a distraction from the user. The classifiers HAAR and adaboost work together to give more accurate resultant features. The image and feature based methods have been well used by researchers for face recognition. The former one uses trained classifiers with an example set while latter will detect face features like eyes, nose, eye brow and mouth. The combination of two methods simultaneously utilize on the image which contains face. The image in which eye is cropped for analysing whether it is in closed or opened state and then mouth is cropped for analysing it is in yawning, closed or smiling state.

Here we are trying to develop a system to detect human activities by using image processing technique. An efficient face recognition algorithm has to be developed which can recognize face activities efficiently. Also, for image processing we have to have a effective platform to test our algorithm. MATLAB gives the best set of libraries or toolboxes for image processing 3programs. Also, this software gives a user-friendly interface to define functions and create graphical user interface.

2.HEADINGS

Motivation

While most of the face detection algorithms can be extended to recognize other objects such as cars, humans, pedestrians, and signs, this can be adapted to differentiate and also to recognize the distinct parts and facial features like eyes, nose, mouth, eye brow and then skin. It is well understood that human emotions also have acoustic characteristics. Although the combination of acoustic and visual characteristics promises improved recognition accuracy, the development of effective combination techniques is a challenge, which has not been addressed by many researchers. In this study, real time facial feature tracker has been considered to deal with the problems of face localization and feature extraction in spontaneous expressions.

Objectives

The main objective of this study is to generate a set of facial features eventually adapt feature extraction classifiers for tracking and recognizing face in an image eventually much facial expression with minimal error rate.

- The classifiers HAAR and adaboost work together to give more accurate resultant features.
- The image and feature based methods have been well used by researchers for face recognition.
- The former one uses trained classifiers with an example set while latter will detect face features like eyes, nose, eye brow and mouth.
- The combination of two methods simultaneously utilize on the image which contains face.
- The image in which eye is cropped for analyzing whether it is in closed or opened state and then mouth is cropped for analyzing it is in yawning, closed or smiling state.

3.FACE DETECTION

The face detection algorithm proposed by Viola and Jones is used as the basis of our design. The face detection algorithm looks for specific Haar features of a human face. When one of these features is found, the algorithm allows the face candidate to pass to the next stage of detection. A face candidate is a rectangular section of the original image called a sub-window. Generally, these sub-windows have a fixed size (typically 24×24 pixels). This sub-window is often scaled in order to obtain a variety of different size faces. The algorithm scans the entire image with this window and denotes each respective section a face candidate.

Integral Image

The integral image is defined as the summation of the pixel values of the original image. The value at any location (x, y) of the integral image is the sum of the image's pixels above and to the left of location (x, y). "Fig. 1" illustrates the integral image generation.

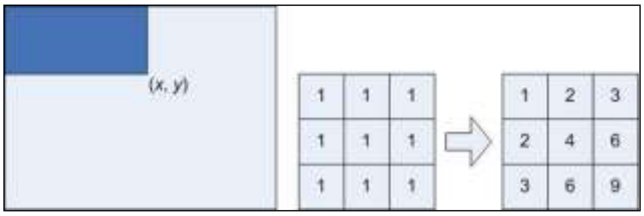

Figure 1. Integral image generation.

Haar Features

Haar features are composed of either two or three rectangles. Face candidates are scanned and searched for Haar features of the current stage. The weights are constants generated by the learning algorithm. There are a variety of forms of features as seen below in "Fig. 2".

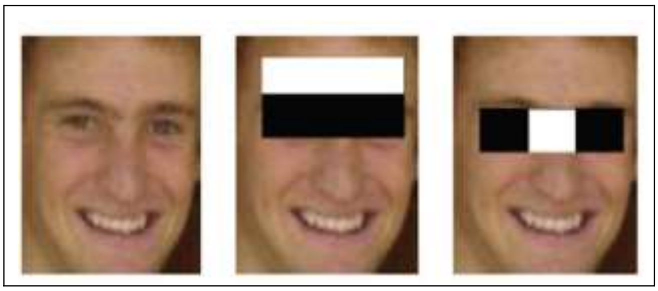

Figure 2. Examples of Haar features.

Areas of white and black regions are multiplied by their respective weights and then summed in order to get the Haar feature value. Each Haar feature has a value that is calculated by taking the area of each rectangle, multiplying each by their respective weights, and then summing the results. The area of each rectangle is easily found using the integral image. The coordinate of the any corner of a rectangle can be used to get the sum of all the pixels above and to the left of that location using the integral image Since L1 is subtracted off twice it must be added back on to get the correct area of the rectangle. The area of the rectangle R, denoted as the rectangle integral, can be computed as follows using the locations of the integral image: L4-L3-L2+L1.

Haar Feature Classifier

A Haar feature classifier uses the rectangle integral to calculate the value of a feature. The Haar feature classifier multiplies the weight of each rectangle by its area and the results are added together. Several Haar feature classifiers compose a stage. A stage comparator sums all the Haar feature classifier results in a stage and compares this summation with a stage threshold.

The threshold is also a constant obtained from the AdaBoost algorithm. Each stage does not have a set number of Haar features. For example, Viola and Jones" data set used 2 features in the first stage and 10 in the second.

All together they used a total of 38 stages and 6060 features [6]. Our data set is based on the OpenCV data set which used 22 stages and 2135 features in total.

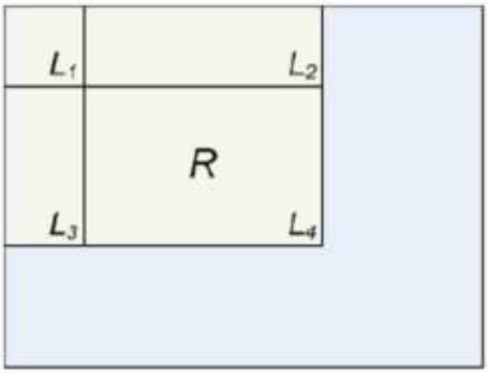

Figure 3. Calculating the area of a rectangle R is done using the corner of the rectangle: L4-L3-L2+L1.

Cascade

The Viola and Jones face detection algorithm eliminates face candidates quickly using a cascade of stages. The cascade eliminates candidates by making stricter requirements in each stage with later stages being much more difficult for a candidate to pass. Candidates exit the cascade if they pass all stages or fail any stage. A face is detected if a candidate passes all stages. This process is shown in "Fig. 4".

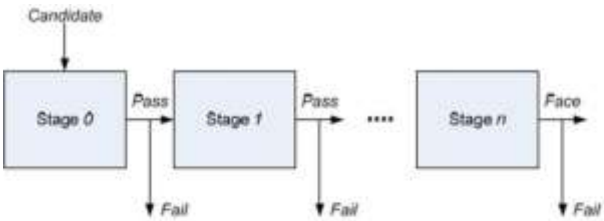

Figure 4. Cascade of stages. Candidate must pass all stages in the cascade to be concluded as a face.

4.IMPLEMENTATION

System Overview

We proposed architecture for a real-time face detection system. "Fig. 5" shows the overview of the proposed architecture for face detection. It consists of five modules: variant pose, illumination condition, Facial Expression, Occulsion, Uncontrolled Background, display. Face Detection systems are not only detected faces on uniform environment. In reality, Peoples are always located on complex background with different texture and object. These „thing" are the major factors to affect the performance of face detection system.

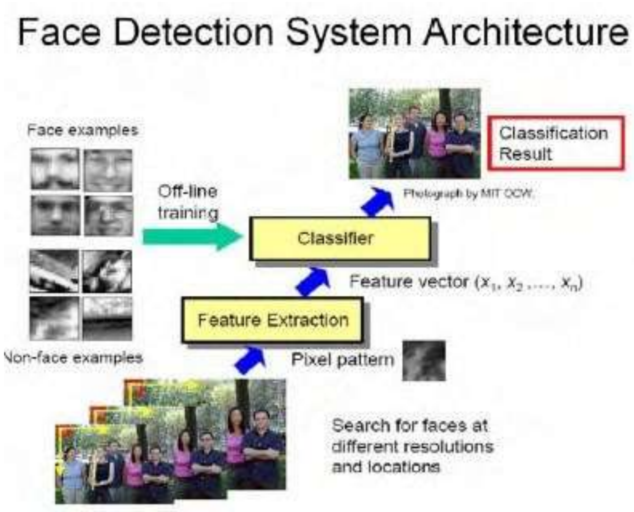

Figure 5. Block diagram of proposed face detection system.

Architecture for Face Detection

Variant Pose

Variant pose is occurred because of peoples not always orient to camera. The image sync signal and the color image data are transferred from the image interface module. The image cropper crops the images based on the sync signals.

Illumination Condition

Different lighting and the quality of camera directly affect the quality of face. Sometimes it can be varied greater than facial expression and occlusion.

Facial Expression

The integral image generation requires substantial computation. A general purpose computer of Von Neumann architecture has to access image memory at least width×height times to get the value of each pixel when it processes an image with width×height pixels. For the incoming pixel where the coordinate is (x, y), the image line buffer controller.

Occulsion

Face detection not only deals with different faces, however, it need deal with any optional object. E.g. Hairstyle, sunglasses are the example of occlusion in face detection.

Integral Image

For the incoming pixel where the coordinate is (x, y), the image line buffer controller performs operations such as in “(1)”, where n is the image window row size, p(x, y) is the incoming pixel value, and L(x, y) represents each pixel in the image line buffer.

$$L(x, y-k) = L(x, y-(k-1)), \text{ where } 1 \leq k \leq n-2 \qquad (1)$$

$$L(x, y-k) = p(x, y), \text{ where } k = 0$$

The image window buffer stores pixel values moving from the image line buffer and its controller generates control signals for moving and storing the pixel values.

$$
\begin{aligned}
& I(i-k,j) = I(i-(k-1),j), \text{ where } 1 \le k \le m-1 \qquad (2) \\
& I(i,j-l) = L(x,y-(l-1)), \text{ where } 1 \le l \le n-1 \\
& I(i-k,j-l) = p(i,j) = p(x,y), \text{ where } k = l = 0, \\
& \text{when } k+l = n-1\ ,\ 1 \le k \le n-1\ ,\ 0 \le l \le n-2\ ,\ m = 2n\ , \\
& I(i-k,j-l) = I(i-(k-1),j-l) + I(i-(k-1),j-(l+1))
\end{aligned}
$$

For the incoming pixel with coordinate (x, y), the image window buffer controller performs operation as in "(2)" where n and m are the row and column size of the image window buffer, respectively. p(i, j) is the incoming pixel value in the image window buffer; p(x, y) is the incoming pixel value; I(i, j) represents each of the pixels in the image window buffer; and L(x, y) represents each of the pixels in the image line buffer.

Since pixels of an integral image window buffer are stored in registers, it is possible to access all integral pixels in the integral image window buffer simultaneously to perform the Haar feature classification.

"Fig. 6" shows all of the actions in the proposed architecture to generate the integral image. For every image from the frame grabber module, the integral image window buffer is calculated to perform the feature classification using the integral image.

5.RESULTS AND TABLES

The Fig. 6 illustrates the sequence of operations performed in recognition of face, eye and mouth. The captured image is converted to grayscale.

Fig 6. Face detection process

The choice of a good database plays vital role for the learning step. Especially to detect significant eye detection those are robust under changing conditions. Non-frontal images can introduce noise in the data, because some eye detection has different appearances in different poses .

In this study, the algorithm is tested for both face and non- face images. The four different face patterns are tested, such as,

Pattern 1: Normal human face i.e. eyes are in opened position

Pattern 2: Face differs from common faces i.e. eyes are in closed position

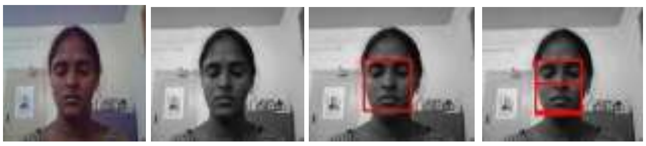

Fig 7. Detection of eye in a closed position

For each pattern, the technique is applied for 37 numbers of images. The detection rate is shown in Table 1. In this study, the difficulty is encountered when adapting the algorithm to the eyes in half opened and/or closed state as shown in Fig. 4. The reason for the difficulty come from that the eye is actually in opened condition but the iris is not visible. The nearly closed eye is regarded as closed and it seems to be adoptable in many practical applications. It is well understood that identification of facial features using positions relative to the face image is difficult because of multiple variations of features possible at different occasions. These variations are due to changing orientation, emotion and especially identity. A Haar classifier should be used to identify these regions.

Table.1. Eye detection rate

Eye State	Correct Rate (%)	Misses Rate (%)
Opened	99.19	**0.81**
Closed	98.89	**1.11**
Lightly Opened	97.14	**2.86**
With Glasses	**89.12**	**10.88**

The misses rate is significantly higher for pattern 4 i.e. eyes are with glasses as it encounters with glaring phenomenon. Further study is under way to improve the detection accuracy for the faces with spectacles by reducing the effect of reflection.

The technique is also tested to classify the face images with facial expressions. In this operation, the mouth detection is under taken. The face images are collected in three patterns such as,

Pattern 1: Normal human face i.e. mouth in closed position

Fig 8. Detection of mouth in closed position

The developed algorithm is applied for 30 numbers of images and rate of detection is given in Table 2. The results create the possibility of adapting this technique for real time both face and facial expression analysis. It is understood, each pixel in the face image is associated to an appearance of cluster. The cluster stands for robust face features which are: eyes, mouth.

Table.2. Mouth detection rate

Mouth State	Correct Rate %	Misses Rate %
Closed(24)	98.70	**1.30**
Opened(30)	93.72	**6.28**
Smile(25)	**91.24**	**8.76**

It is tried to extract and utilize the maximum of information contained on a single image of a face. These regions can be enclosed with rectangles in the image. The embedded image processor with this algorithm is being studied.

6.CONCLUSION

We present face detection based on the AdaBoost algorithm using Haar features. In our architecture, the scaling image technique is used instead of the scaling sub-window, and the integral image window is generated instead of the integral image contains whole image during one clock cycle.

The Haar classifier is designed using a pipelined scheme, and the triple classifier which three single classifiers processed in parallel is adopted to accelerate the processing speed of the face detection system.

Finally, the proposed architecture is implemented on a Modelsim Altera 6.3 and its performance is measured and compared with an equivalent hardware implementation. We show about 35 times increase of system performance over the equivalent software implementation. We plan to implement more classifiers to improve our design. We have demonstrated that this face detection, combined with other technologies, can produce effective and powerful applications.

7.REFERENCES

1. Z. Guo, H. Liu, Q. Wang and J. Yang, " A Fast Algorithm of Face Detection for Driver Monitoring, " In Proceedings of the Sixth International Conference on Intelligent Systems Design and Applications, vol. 2, pp.267 - 271, 2006.
2. M. Yang, N. Ahuja, "Face Detection and Gesture Recognition for Human-Computer Interaction," The International Series in Video Computing, vol.1, Springer, 2001.
3. Z. Zhang, G. Potamianos, M. Liu, T. Huang, "Robust Multi- View Multi-Camera Face Detection inside Smart Rooms Using Spatio-Temporal Dynamic Programming," International Conference on Automatic Face and Gesture Recognition, pp.407-412, 2006.
4. W. Yun; D. Kim; H. Yoon, "Fast Group Verification System for Intelligent Robot Service," IEEE Transactions on Consumer Electronics, vol.53, no.4, pp.1731-1735, Nov. 2007.
5. V. Ayala-Ramirez, R. E. Sanchez-Yanez and F. J. Montecillo-Puente "On the Application of Robotic Vision Methods to Biomedical Image Analysis," IFMBE Proceedings of Latin American Congress on Biomedical Engineering, pp.1160-1162, 2007.
6. P. Viola and M. Jones, "Robust real-time object detection," International Journal of Computer Vision, 57(2), 137-154, 2004.