



International Journal of Emerging Technologies in Computational and Applied Sciences (IJETCAS)

www.iasir.net

Facex-Real Time Human Facial Expression Analysis–A Survey

Prof P.D. Moore, Akash Sethi, Ankur Nagar, Ashwin Srivastava, Shradha Bangad.

Dept. of Computer Engineering,
Vishwakarma Institute of Information Technology,
Pune, India

Abstract: With the latest development in the computer field Vision, image processing and pattern recognition, facial expression recognition is becoming more and more feasible for human computer interaction in virtual environments. Facial expression is the facial changes in response to a person's internal emotional states, intentions or social communications. Precise facial feature points are essential to the high level face recognition and expression analysis. This paper talks about extracting the facial Expressions of human beings in real time and classifying them under different categories.

Index terms: Adaptive Model, Viola Jones, Template Matching, Gini Index

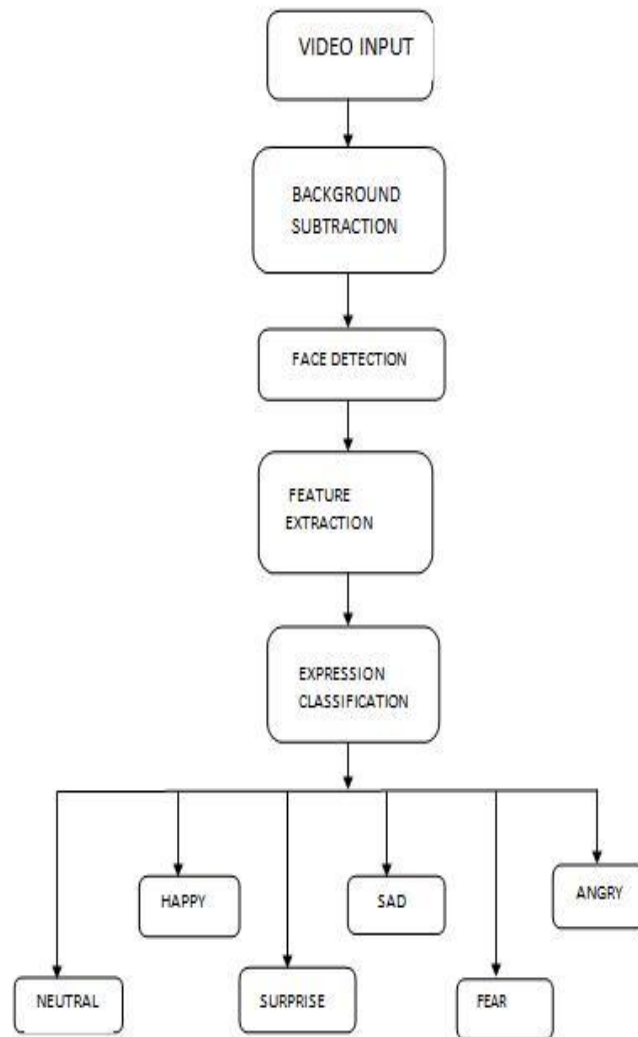
I. Introduction

With the latest development in the computer field Vision, image processing and pattern recognition, facial expression recognition is becoming more and more feasible for human computer interaction in virtual environments. Facial expression is the facial changes in response to a person's internal emotional states, intentions or social communications. Precise facial feature points are essential to the high level face recognition and expression analysis. This paper talks about extracting the facial Expressions of human beings in real time and classifying them under different categories. Facial expression recognition [1] also follows the research framework of the traditional pattern recognition, which is composed of three main aspects: facial expression acquisition, feature extraction, and expression classification. Among them, feature extraction is the most critical. Human beings express their emotions in everyday interactions with others. Emotions are frequently reflected on the face, in hand and body gestures, in the voice, to express our feelings or liking. Recent Psychology research has shown that the most expressive way humans display emotions is through facial expressions.

Emotions are feeling or response to particular situation or environment. Emotions are an integral part of our existence, as one smiles to show greeting, frowns when confused, or raises one's voice when enraged. It is because we understand other emotions and react based on that expression only enriches the interactions. Computers are "emotionally challenged". They neither recognize other emotions nor possess its own emotion. To enrich human-computer interface [2] from point-and-click to sense-and-feel, to develop non intrusive sensors, to develop lifelike software agents such as devices, this can express and understand emotion. Since computer systems with this capability have a wide range of applications in different research arrears, including security, law enforcement, clinic, education, psychiatry and Telecommunications. There has been much research on recognizing emotion through facial expressions. We classify the emotion into positive and negative emotions. The six basic emotions are angry, happy, fear, disgust, sad, surprise. One more expression is neutral. Other emotions are Embarrassments, interest, pain, shame, shy, anticipation, smile, laugh, sorrow, hunger, curiosity. In different situation of emotion, Anger may be expressed in different ways like Enraged, annoyed, anxious, irritated, resentful, miffed, upset, mad, furious, and raging. Happy may be expressed in different ways like joy, greedy, ecstatic, fulfilled, contented, glad, complete, satisfied, and pleased. Disgust may be expressed in different ways like contempt, exhausted, peeved, upset, and bored. In emotional expression face, If we are angry, the brows are lowered and drawn together, Vertical lines appear between the brows, lower lid tensed, eyes hard stare or bulging, lips can be pressed firmly together with corners down or square shape as if shouting, nostrils may be dilated, the lower jaw juts out. If we are happy, corners of the lips are drawn back and up, mouth may or may not be parted, teeth exposed, a wrinkle runs from outer nose to outer lip, cheeks are raised, lower lid may show wrinkles or be tense, crows feet near the outside of the eyes. Disgust has been identified as one of the basic emotions. Its basic definition is "bad taste" secondarily to anything which causes a similar feeling, through the sense of smell, touch and even of eyesight.

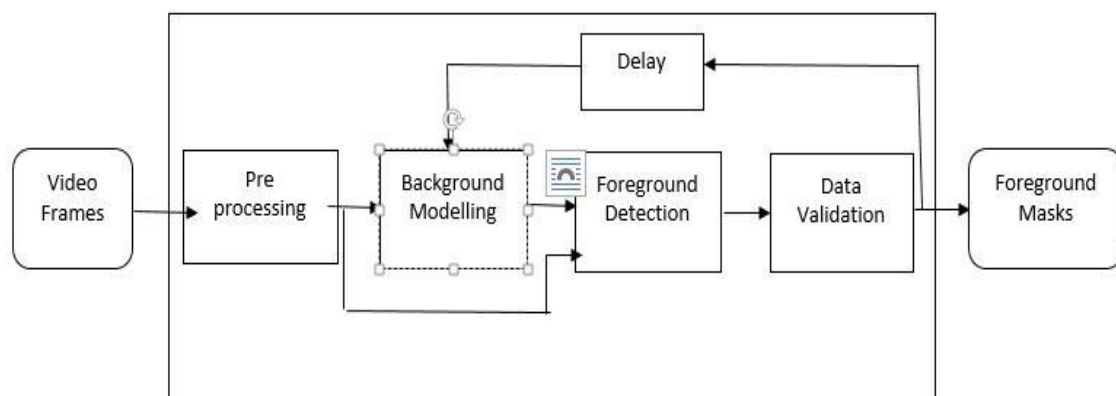
II. Control Flow

In this whole paper we drive through the proposed work described in (Fig. 1). In the initial stage we input the face image or the video whose expression is needed to be recognized. Feature is extracted in the next stage. As the facial expression information is mostly concentrated in the facial regions like, eyes, eyebrows and mouth, hence these features are extracted from the input face images. Once the template is being matched the facial characteristic points is calculated which help in evaluating the facial animation parameters. With the aid of the facial animation parameters and the classification technique Gini index, we are able to recognize the expression of the input face images.



III. Background Subtraction

As the name suggests, background subtraction [3] is the process of separating out foreground objects from the background in a sequence of video frames. Background subtraction is a widely used approach for detecting moving objects from static cameras. Background subtraction should segment objects of interest when they first appear (or reappear) in a scene. An appropriate pixel level stationary criteria should be defined. Pixels that satisfy this criterion are declared background and ignored. The background model must adapt to both sudden and gradual changes in the background.



In this we have the video frame, then the preprocessing happens. After the preprocessing, background modeling is done, or depending upon the video input we can directly do the foreground detection. Once these processes are done then comes the validation of the input frame. If we have any delay then it goes back again.

to the background modeling and the process continues till we get the foreground objects. In this paper, we are using Adaptive Gaussians mixture model algorithm for the background subtraction. Background subtraction is common method for real time segmentation of moving regions in image sequences. This method differs from other models for the background subtraction and enhances the efficiency for the real time segmentation of the moving regions. In this there is modeling of each pixel as a mixture of Gaussians and also using an online segmentation in order to update the model. Evaluation is done in order to determine which are most likely to result from background process of the adaptive mixture model.

IV. Viola Jones Algorithm

Viola Jones algorithm [4] is a face detection algorithm developed by Paul Viola and Michael Jones. This algorithm is one of the fastest algorithms for detecting faces. In this algorithm, a detector or window is used like Haar wavelet for feature extraction that detector moves all over the face and finds out the features of the face, the window size minimally should be of 24 x 24 pixels. This algorithm is one of the fastest algorithms to detect face using skin color as a feature. This algorithm also uses AdaBoost for feature selection mechanism. At first, the input image is converted into integral images. This conversion is done by making each of the pixels equal to the sum of all the pixels and this allows the calculation of all the pixels using only four values of the rectangle. Viola Jones uses AdaBoost as a machine learning boosting algorithm which is used as an enhancement algorithm, all the weak classifiers are combined together to form a strong classifier. A weak classifier is mathematically described as:

$$h(f, x, p, \theta) = \begin{cases} 1 & \text{if } pf(x) > p\theta \\ 0 & \text{otherwise} \end{cases}$$

where x is a 24*24 pixel sub-window, f is the applied feature, p the polarity and θ the threshold that decides whether x should be classified as a positive (a face) or a negative (a non-face).

V. Feature Extraction

Dimensionality reduction is the basic use of Feature Extraction. In order to perform the desired task of extracting [5] the facial characteristics points from the features being matched, set of facial features, e.g. eyebrows, eyes, mouth etc are deduced from the image taken as input. For this, shape based feature extraction technique known as Template Matching is used. Template Matching is a novel technique to track patterns from the input image based on correlation.

The steps for accomplishing template matching is:

- Step 1: Take template and the target image as input.
- Step 2: Convert both of them to gray scale images.
- Step 3: Now obtain the 2-d convolution of the target and the template image.
- Step 4: Now obtain the mean and the variance of the template image.
- Step 5: Obtain the correlation score 'C'.
- Step 6: Obtain the pixel value having the highest value of the correlation score 'C' from the search image.
- Step 7: Using the index of the correlation score matched make the bounding rectangles.

VI. Classification

Classification is the task of assigning patterns to one of the predefined class. According to definition, Classification is the task of learning a target function. In this paper we have talked about Gini index [6] for the classification of different expressions. Classification will be done on the basis of expressions that an individual is going to make. The different expressions that can be are - sad, happy, grief, angry, surprised. Gini index approach is a classification technique fall under the decision tree category. In a Gini index decision tree classification, each leaf node is assigned as a class label. The non-terminal nodes includes, the root and other internal nodes, that contain attributes test conditions to split the data set that have different characteristics. The splitting is performed based on the basis of measure of degree of impurity of the child nodes. The steps are as follows:-

- Step 1: If all records are having either same class labels or having same attribute values, then go to step 2 and STOP otherwise go to step 3.
- Step 2: Specify node as a leaf node; classify datasets D into the attribute set 'A' and return leaf node
- Step 3: Specify node as a root node and got to step 4.
- Step 4: Find the best split using equation (2) by selecting the best attribute having least value of GINI index calculated using equation (1) and calculate the GINI gain using equation (3).
- Step 5: Specify the possible outcomes of step 4 as 'O'.
- Step 6: For each O do step 7 to step 9
- Step 7: Traverse through the edge E according to the condition specified in the step 4.
- Step 8: Again call Gini index ($E(o)$, 'A') as a child node
- Step 9: Add the child node as descendent node of root and label the edge (root->child) as o

VII. Conclusion

In this paper, we have presented a survey of the various steps required for facial expression analysis of a human in real-time video. The sequence of steps that we have followed are Background Subtraction using Adaptive Mixture Modeling, Face Detection using Viola Jones Algorithm, Feature Extraction using Template Matching and Classification using Gini Index. The basic flow of each of these steps has been outlined in detail in this paper and presents significant performance improvements over other related methods. Adaptive Mixture Modeling algorithm, used for background subtraction was a breakthrough in digital video analysis and is commonly used. Viola Jones is the most popular algorithm used for face detection and presents a robust way of detecting faces in various environmental conditions, although it is relatively complex to implement.

References

- [1] A New Algorithm to Classify Face Emotions through Eye and LipFeatures by Using Particle Swarm Optimization A. Habibizad Navin, Mir Kamal Mirnia, Tabriz Branch, Islamic Azad University, Tabriz, Iran. Department of Computer, Science and Research Branch, IEEE.
- [2] Chellappa, R.; Wilson, C. L.; and Sirohey, S. 1995. Human and Machine Recognition of Faces: A Survey. *Proceedings of the IEEE*, 83(5): 705-740.
- [3] Ekman, P., Friesen, W. V. *Facial Action Coding System: A Technique for Measurement of Facial Movement*. Consulting Psychologists Press Palo Alto, California, 1978.
- [4] Adaptive background mixture models for real-time tracking Chris Stauffer W.E.L. Grimson The Artificial Intelligence Laboratory Massachusetts Institute of Technology, Cambridge, MA 02139, USA
- [5] Haddadnia, J.; Ahmadi, M.; and Faez, K. 2002. An Efficient Method for Recognition of Human Faces using Higher Orders Pseudo Zernike Moment Invariant. In *Proceedings of the Fifth IEEE International Conference on Automatic Face and Gesture Recognition*, 315-320. Washington, USA.
- [6] Facial Expression Recognition Using Facial Characteristic Points and Gini Index Nazil Perveen, Shubhrata Gupta, and Kesari Verma.