Facial expression recognition system

Abstract

Human-computer interaction will be much more effective if a computer know the emotional state of human. Facial expression contains much information about emotion. So if we can recognize facial expressions, we will know something about the emotion. However, it is difficult to categorize facial expressions from images. Neural network may be suitable in this problem because it can improve its performance. Moreover, we do not need to know much about the features of the facial expressions to build the system. The input to the system is a 96x72 pixels image containing a human face. The outputs are 6 numbers. Each number represents a facial expression (smile, angry, fear, disgust, sadness, surprise). The number is 1 if that facial expression is present and 0 otherwise.

1. Introduction

Faces are accessible "windows" into the mechanisms which govern our emotional and social lives. About 70% of human communication is based on non-verbal communication such as facial expressions and body movements.

Facial expressions are generated by contraction or relaxation of facial muscles or by other physiological processes such as coloring of the skin, tears in the eyes or sweat on the skin. We restrict ourselves to the contraction or relaxation of the muscles category. When tackling the problem of Facial Expression recognition system, one generally decomposes it in three sub problems:

• Face Detection
• Facial Features Detection
• Expression
  Recognition/Classification

The first thing to do when one wants to design a facial expression recognition system is to select the Face under which experiments will have to be run. In our case, as we want the system to be fully automatic, we have to start by detecting the user's face inside the scene. Camera captures the image, which includes the human face. Then human face will be detects from image. Features, which are required for facial expression recognition will be, detect and then next module of the system detects the expression.

2. Face Detection

First we have to find the Face from the image. Although it seems like an easy problem at first glance, we quickly realized that the high variability in the types of faces encountered makes the automatic detection of the face a tricky problem.

Two steps are there to detect the face 1.Detect the region which likely to contain the human skin in the color image.

2.Then extract information from these regions, which might indicate the location of face in the image. Now the skin detection is performed using a skin filter, which relies on color and texture information.

Skin Filter:-

 Skin filter uses hue and saturation as the main arguments. Using Constraints on hue and saturation region of skin can be marked. Face can be detected by many different methods, such as

1.Face can be detected by Open CV face tracker

•It find the face localization on the image
•and surround it with bounding box.

2. Boosted Cascade Face Detector:- This algorithm utilises a boosting method known as AdaBoost to select and combine a set of features, which can discriminate between face and nonface image regions. etc

3. Feature extraction

Our second module is Feature Extraction. Now we require Features, which are essential for expression. Human face consists of different number of features, which are used to detect the expression.

Following Feature plays important role to identify the human expressions. • Eyebrows
•• Lips
•• Nose

FEATURE REGIONS DENTIFICATION

Edge projection analysis can use for extraction features. Using integral projections of the edge map of the face image facial features can be extraction.

A typical human face follows a set of anthropometrics standards, which have been utilized to narrow the search of a particular facial feature to smaller regions of the face.

Facial feature can be extracted by using following steps from the localized face image—

1.   An approximate bounding box for the feature is obtained using the anthropometrics standards.

2.   Sobel edge map is computed to obtain edges along the boundary of the feature.

3.   The integral projections V (x) and H (y) are calculated on the edge map.

4.   Median filtering followed by Gaussian smoothing smooths the projection vectors so obtained. Higher value of projection vector at a particular point indicates higher probability of occurrence of the feature. The relative probability E(i) of the ith region containing the feature is calculated.   5. The bounding box so obtained is processed further to get an exact binary mask of the   feature.

Eyebrow

The approximate bounding box is the top half of the face. The generic steps uses horizontal sobel edges to compute bounding box containing *eye* and *eyebrow*. The segmentation algorithm cannot

give bounding box for the eyebrow exclusively because the edges due to eye also appear in the chosen bounding box. Brunelli suggests use of template matching for extracting the eye.

Eyebrow is segmented from eye using the fact that the eye occurs below eyebrow and its edges form closed contours (fig),obtained by applying
Laplacian of Gaussian
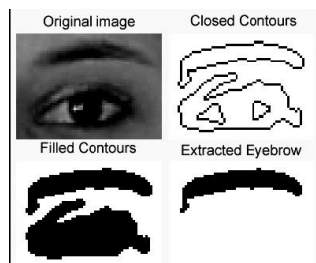
operator at zero threshold. These contours are filled and the resulting image containing masks of eyebrow and eye is morphologically filtered by horizontally stretched elliptic structuring elements. From the two largest filled regions, the region with higher centroid is chosen to be the mask of eyebrow.

Lip

The generic steps calculates edge maps on the transformed image. Edges for lips occur both in horizontal and vertical direction. In the bounding box computed by the generic algorithm, closed contours are obtained by applying *Laplacian of Gaussian* operator at zero thresholds. These contours are filled and morphologically filtered using elliptic structuring elements to get binary mask for the lips.

Nose

The approximate bounding box for the nose lies between the eyes and the mouth. The generic algorithm uses vertical sobel edges to compute the vertical position, which is required as a reference point on face.

4. FEATURE VECTOR AND CLASSIFICATION

A spatial-temporal representation of the face is created, based on geometrical relationships between features using Euclidean distance (figure). Such a representation allows robust handling of partial occlusion. Seven parameters form the feature vector **F—**

$$\mathbf{F} = \{H_e, W_e, H_m, W_m, R_{ul}, R_{ll},$$

N L }

Where He: Height of eyebrow We: brows distance
Hm: mouth height
Wm: mouth width
Rul: upper lip curvature
Rll: lower lip curvature

All components of the vector are normalized against the first frame to achieve scale independence. Radii of curvature of the upper and lower lips Rul and Rll are computed by approximating the binary mask of the lips with two parabolas. NL is the number of distinct peaks detected for upper and lower lips during edge projection analysis, indicating whether mouth was open or closed. The change in feature vector **F** when the face undergoes change from neutral state to some expressional state—

`$\Delta \mathbf{F} = \{ \Delta H e , \Delta W e , \Delta H m,$
$\Delta W m , \Delta R ul , \Delta R ll , \Delta N L \}$

Such dynamic characteristic of the feature vector provides shape independence. $\Delta \mathbf{F}$ serves as an input to the classifier. The classifier is a feedforward basis function net with one hidden layer (figure).
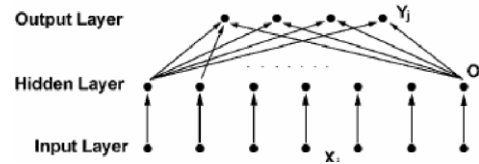
The basis function net for classification.

•Each image sequence depicted one of the expression classes (joy, surprise, sad and disgust against neutral). The first image in the sequence was a neutral image. Confidence level of each expression was calculated for each of the subsequent images against the neutral image. The calculated vector of confidence levels was added to give total confidence for each of the expressions. Expression having the highest total confidence level was declared as the expression of the sequence.

5. Applications

•Behavioral assessment of emotion and paralinguistic displays is important in a range of academic and applied fields including

- clinical psychology and psychiatry,

- child development, political science

- and advertising.

•Biomedical applications,
including
treatment of facial nerve disorders

•Computer systems that understand human behavior and respond appropriately.
•Speech recognition.
•Security systems.
•Lie detection.
•Facial animation.

6.Conclusion

In our day-to-day life Facial expression plays important role to perform non-verbal communication. For that we decided to build computer base system for Facial Expression recognition. When tackling the problem of Facial Expression recognition system, one generally decomposes it in three sub problems:
•   Automatic   Face
                Detection   and
Tracking
•   Automatic Facial Features
        Detection and Tracking
•   Automatic
                Expression
Recognition/Classification
        An efficient, local image-based approach for extraction of intransient facial features and recognition of six facial expressions present. The system requires no manual intervention (like initial manual assignment of feature points). The system, based on a local approach, is able to detect partial occlusions also.

7. References

[1]   M.J.T.   Reinders,   Model Adaptation for Image,Coding, Delft University Press

[2]   Analysis   of   Facial Expressions Based on Silhouettes by J.C. Wojdel, A.
Wojdel, L.J.M. Rothhkrantz

[3]   D. Cristinacce and T. Cootes. Facial feature detection using adaboost with shape constraints. In 14*th British Machine Vision Conference, Norwich,
England*

[4]   Measuring facial expressions by computer image analysis, Psychology, by Barlett, M. S., Hager, J. C., Ekman,
P., Sejnowski,

[5]http://www.cs.cmu.edu/~face/A_ange
r_tracksmall.avi

[6]   http://face-andemotion.com/dataface/nsfrept/nsf_contents.html

[7]   http://face-andemotion.com/dataface/visage/visage_forehead.jsp