Facial Expression Recognition using SVC Classification & INGI Method

Ashamol Joseph[1], P. Ramamoorthy[2]

[1]PG Scholar, Department of ECE, SNS College of Technology, Coimbatore, India
[2]Professor and Dean, Department of ECE, SNS College of Technology, Coimbatore, India
ashchalattu00@gmail.com
covairam33@ymail.com

Abstract: ***Facial expression analysis is the important area of Human Robot Interaction (HRI) because facial expressions represent human emotions. Most existing researches in facial expression analysis mainly focused on recognizing extreme facial expressions. In existing work they introduced a tensor perceptual color framework (TPCF) for facial expression recognition (FER), which is based on information contained in color facial images. Humans can perform expression recognition with a remarkable robustness without conscious effort even under a variety of adverse conditions such as partially unmarked faces, different appearances and poor illumination. To avoid this proposed face recognition system consists of a novel illumination-insensitive pre-processing method for eliminating the illumination. First, in the pre-processing stage, a face image is transformed into an illumination-insensitive image, called an "integral normalized gradient image," by normalizing and integrating the smoothed gradients of a facial image. The features are classified by using Support Vector clustering (SVC) classifier to avoid the complexity. The objective of this paper is to apply Support Vector Machines to the problem of classifying emotion on images of human faces.***

Keyword: Facial expression recognition, Log Gabor Filter, Support vector machine, perceptual color spaces

1. Introduction

The foundational studies on facial expressions that have formed the basis of today's research can be traced back to the 17th century. Darwin established the general principles of expression and the means of expressions in both humans and animals. The method of developing a facial expression recognition system consists of face detection, normalization, extraction, and feature classification. There are different techniques which we used for recognizing the facial expression. In this paper Viola and Jones describe the face detection technique using AdaBoost Haar classifier [5]. After performing the pre-processing operation the recognition is performed, the simplicity and robustness of the system is significant. Several approaches have been proposed for FER in the past several decades. The approaches mainly focus for facial expression classification on gray-scale image features [1] .In this we mainly focus on the colour scale images. Colour information may lead to more robust classification results. Current research reveals that colour information improves recognition rate and image retrieval performance. Liu and Liu proposed a new colour space for face recognition [9].

Seyed Mehdi Lajevardi and Hong Ren Wu explain the concept of facial expression recognition in perceptual colour space [7]. He describes that the perceptual colour spaces (CIELab and CIELuv) are better overall for FER than other colour spaces. This paper mainly explains the concept of tensor perceptual framework and the SVC classification. The advantage of tensors over a vector representation is their ability to encode local image features as orientation and orientation uncertainty. Tensors are of different types, first order tensor, and the second order tensor. First order tensor is a generalization of vector and second order tensor is generalization of matrix. Tensors are based on multilinear mappings over a set of vector spaces. The paper describes the concept of Facial expression based on the tensor representation and the support vector machine algorithm. There are two approaches, model based and the image based approaches. We consider the image based approaches in which Image-based methods extract features from images without relying on extensive knowledge about the object. This paper focuses on the static colour images and a holistic technique of the image-based method is used for feature extraction. The figure 1 shows the system level diagram. It consists of mainly face detection, feature extraction and the feature classification. The following sections will describe each module in detail.The facial components or facial feature points are extracted to form a feature vector that represents the face geometry. The appearance features present the changes of the face. The appearance features can be extracted from either the whole face or specific regions in a face image. This paper focuses on the static color images and a holistic technique of the image-based method is used for feature extraction.

The image based FER systems consist of several components or modules, including face detection and normalization, selection process and classification algorithm.

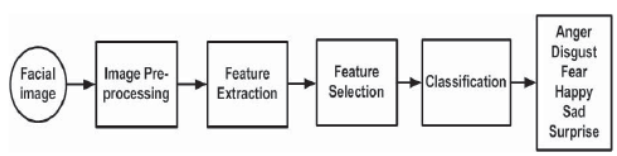

Figure 1: System Level Diagram

1.1 Face Detection and Normalization

The face area of an image is detected using the Viola–Jones method based on the Haar-like features and the AdaBoost learning algorithm [5]. In Viola-Jones method, the pixels are selected from rectangular areas. The value of a two-rectangle feature is the difference between the sums of the pixels within two rectangular regions. The size and shape of regions are the same and they are horizontally or vertically adjacent. After face detection stage, the colour values of face images are then normalized with respect to RGB values of the image. The first step of Viola-Jones face detection algorithm is to turn the input image into an integral image. It allows for the calculation of the sum of all pixels inside any given rectangle. These values are the pixels in the integral image that coincide with the corners of the rectangle in the input image. Haar Classifier is also a machine learning algorithm for the visual object detection. The variance of contrast between the pixel groups are used to determine relative light and dark areas.

The set of basic Haar-like-feature is rotating which the other features can be generated. The Haar-like feature is the difference between the sums of the pixel gray level values within the black and white rectangular regions. The illumination-insensitive image integral normalized gradient image (INGI) method is proposed to overcome the unexpected illumination changes in face recognition with limited side effects such as image noise and the halo effect.

1.2 Feature Extraction

Various methods of feature extraction have been studied including principal component analysis, (PCA), linear discriminant analysis (LDA), the Gabor filter bank, etc. But the Gabor filter has some disadvantages, the maximum bandwidth of Gabor filters is limited, and the Gabor filters are not optimal to achieve broad spectral information. The response of the Log-Gabor filter is Gaussian when viewed on a logarithmic frequency scale instead of a linear one. A bank of 24 Log-Gabor filters is used to extract the facial features. Six scales and four orientations are implemented to extract features from face images. This leads to 24 filter transfer functions representing different scales and orientations. Also tensors are used to extract the features and compare the performances of log Gabor filter and the tensor based extraction.

Implementing the filter for each component of the colour image is complicated, so to avoid this complication, a tensor of the colour image is generated and the filtering operation is directly applied. The image filtering takes place in frequency domain .It should be faster compared with spatial domain. In log Gabor filter the information to be captured in high frequency areas.

1.3 Feature Selection

The features are selected by using minimum redundancy maximum relevance algorithm based on mutual information (MI).The mutual information means that mutual dependence between the two random variables. In MIQ ,if a feature vector has expressions randomly or uniformly distributed in different classes, C denotes a set of classes C = {c1, c2, . . . , ck }.It is mainly selected based on the probability distribution function(PDF).This paper mainly deals with the mutual information quotient algorithm for feature selection. The mutual information between the features is selected based on the probability function. The mutual information between the two features is given by

$$I(\mathrm{V_t};\mathrm{V_s}) = \sum_{\mathrm{v_t}\in \mathrm{V_t}} \sum_{\mathrm{v_s}\in \mathrm{V_s}} p(\mathrm{v_t},\mathrm{v_s}) \log \frac{p(\mathrm{v_t},\mathrm{v_s})}{p(\mathrm{v_t})p(\mathrm{v_s})} \quad \textbf{(1)}$$

Where $p(\mathbf{vt}, \mathbf{vs})$ is the joint probability distribution function and $p(\mathbf{vt})$ and $p(\mathbf{vs})$ are the marginal PDFs of two features.

1.4 Feature Classification

In the previous paper, features are classified by using the multiclass LDA classifier. In this paper, we are doing the two classification algorithms, Linear Discriminant Analyser and Support Vector Clustering algorithm and compare their performances. This method should be a combination of support vector machine and the fuzzy clustering method. The features are selected by using genetic algorithm. Clustering is the classification of similar objects into different groups. Using clustering method to decrease the complexity of SVM classification, In SVM, Kernel functions are used to map input data which may not be linearly separable to a high dimensional feature space where linear methods can then be applied. The memberships of the fuzzy logic system are tuned by genetic algorithms to generate the optimal fuzzy logic system.

2. Methodology

2.1 Adaboost Algorithm

Face Detection was done by using Adaboost algorithm, it was mainly based on viola-jones Method [5].A new image representation called the “Integral Image”, and it can be computed from an image using a few operations per pixel. An efficient classifier which is using the AdaBoost learning algorithm to select a small number of features from a very large set of potential features. The third method is used for combining classifiers in a “cascade” which allows background regions of the image to be quickly discarded while spending more computation on promising face-like regions. Within any image sub-window the total number of Haar-like features is very large, far larger than the number of pixels. In order to ensure fast classification, the learning process must exclude a large majority of the available features, and focus on a small set of critical features. More specifically, we use three kinds of features. The two-rectangle feature is the difference between the sums of the pixels within two rectangular regions. The size and shape of regions are the same and are horizontally or vertically adjacent. A three-rectangle feature is the sum within two outside rectangles subtracted from the sum in a center rectangle. The four-rectangle feature computes the difference between diagonal pairs of rectangles. The steps are given below

1. Evaluate the rectangle features
2. Compute the weak classifier for each feature
3. Combine the weak classifiers

2.2 Basic Steps of LDA Algorithm

There are many possible techniques for classification of data. Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) are two commonly used techniques for data classification and dimensionality reduction This method maximizes the ratio of between-class variance to the within-class variance in any particular data set by maximal separability. To implement an algorithm for LDA providing better classification compared to Principal

Components Analysis. The difference between the two classification algorithms is that LDA and PCA does more of feature classification and LDA does data classification. Linear discriminant analysis (LDA) is a known technique used for dimensionality reduction.

LDA uses subspace as input data, *i.e.*, matrix. The advantage is cutting the eigenvectors in matrix that are not important for face recognition (this significantly improves computing performance). LDA considers between and also within class correspondence of data. It means that training images create a class for each subject, *i.e.*, one class = one subject (all his/her training images).

Determine LDA subspace, determining the line from training data. Calculate the within class scatter matrix

$$\mathsf{S}_w = \sum_{i=1}^{C} \mathsf{S}_i\,, \quad \mathsf{S}_i = \sum_{x \in X_i} (x - m_i)(x - m_i)^\top \quad \textbf{(2)}$$

Where C is the number of classes. m_i is the mean of the images in the class, Calculate the class scatter matrix using this equation is give by

$$\mathsf{S}_B = \sum_{i=1}^{N} n_i (m_i - m)(m_i - m)^\top \quad \textbf{(3)}$$

Where m is the mean of all the images, n_i is the number of images in the class; m_i is the mean and solves the generalized eigenvalue problem.

The following steps are performed by both methods

1. All training images are projected onto particular method's subspace
2. Each test image is also projected to the same subspace and compared by distance metrics between the image and training images (distance metrics are different for both methods).

2.3 SVC Classification

Features are classified by using SVC (Support Vector Clustering). Support vector machines (SVMs) are a set of related supervised learning methods used for classification and regression. In another terms, Support Vector Machine (SVM) is a classification and prediction tool that uses machine learning theory to maximize predictive accuracy. Support Vector machines based on fuzzy clustering can be defined as systems which use hypothesis space of a linear functions in a high dimensional feature space, trained with a learning algorithm from optimization theory. The aim of clustering is to separate a finite unlabeled data set into a finite and discrete set of hidden data structures.SVM using pixel maps as input; it gives maximum accuracy compared to sophisticated neural networks. The general form of objective function is given by

$$J(u_{ij}, \mathrm{v}_k) = \sum_{i=1}^{k} \sum_{j=1}^{n} \sum_{l=1}^{k} g\left[w(x_i), u_{ij}\right]\ d\left(\mathrm{x}_j, \mathrm{v}_i\right) \quad \textbf{(4)}$$

Uij is a numeric value in [0, 1], *w* (*xi*) is the a priori weight for each **x***i*, *d* (**x***j*, **v***i*) is the degree of dissimilarity between the data **x***j* and the supplemental element **v***i*, which can be considered as the central vector of *i*-th cluster.

The following steps are given below:

Step 1: Select a number of clusters and exponential weight and then choose initial partition matrix.
Step 2: Calculate the fuzzy cluster centers.
Step 3: Calculate the new partition matrix.

The fuzzy clustering is used to dividing the input data space into clusters. The fuzzy clustering obtains *k* clusters center of data and also the membership function of each data in the clusters minimizing the objective function.

2.4 Integral Normalized Gradient Image (INGI)

We can make the following assumptions: most of the intrinsic factor is in the high spatial frequency domain, and most of the extrinsic factor is in the low spatial frequency domain. Considering the first assumption, one might use a high-pass filter to extract the intrinsic factor, but it has been proved that this kind of filter is not robust to illumination variations. In addition, a high-pass filtering operation may have a risk of removing some of the useful intrinsic factor. Hence, we propose an alternative approach is used to remove the illuminations, namely, employing a gradient operation.

3. Tensor based Representation of Colour Facial Images

Each color image can be represented as 3-D; there is a technical challenge to proceed with applying a 2-D filtering process to a 3-D matrix, which represents the color image [7]. It can either process a single channel of the color image or perform the filtering operation on each color channel individually. The latter approach is to employ the 2-D filters three times over three component images, respectively. Instead of implementing the filter for each component of the color image, a tensor of the color image is generated and the filtering operation is directly applied to this tensor. The 3-D color image is unfolded to obtain 2-D tensors based on multilinear analysis criteria which are suitable for2-D feature extraction filters.

4. Results and Discussions

4.1 Input Images

The various expressions in RGB colour space are given below in Fig.2

Figure 2: Input Images

4.2 Face Detection

The face area of the image was detected based on Haar-like features using viola-Jones method and Adaboost algorithm. The detected output can be shown below in Figure 3.

Figure 3: Face Detection

The features used by Viola and Jones are derived from pixels selected from rectangular areas imposed over the picture, and exhibit high sensitivity to the vertical and the horizontal lines. After face detection stage, the face images are scaled to the same size.

4.3 Color Spaces

There are various color spaces which are used in image processing such as RGB, YCbCr, CIE Lab .The CIE Lab is shown as in Fig.4.

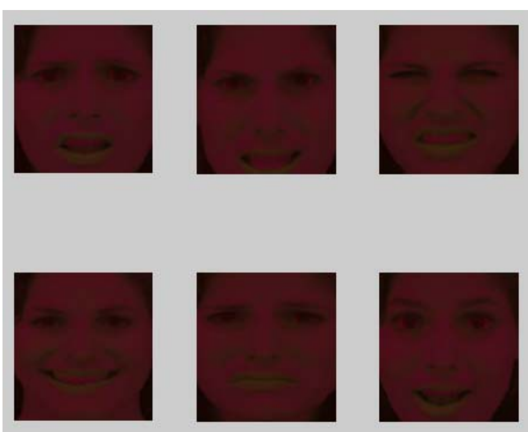

Figure 4: LAB Colour Space

4.4 Normalization

The color values of face images are then normalized with respect to RGB values of the image and its output is shown as in Fig.5. The purpose of color normalization is to reduce the lighting effect because mainly normalization process is the actually a brightness elimination process.

Figure 5: Normalization

4.5 Accuracy Comparison

The accuracy comparison is shown as below in Fig.6. The CIE Lab possesses the highest accuracy than the other colour spaces.

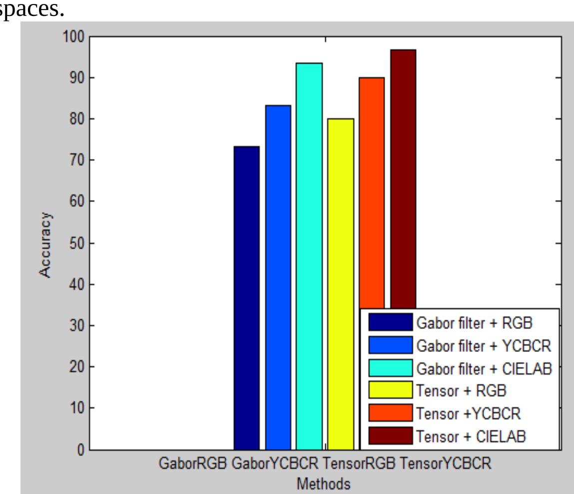

Figure 6: Accuracy Comparison

5. Conclusion

The study of facial expressions will present continued opportunities for the study of emotion-relevant experience and autonomic and central nervous system physiology. It presented a facial expression with pre-processing, feature extraction for uncontrolled illumination situations. This paper presented an approach for face detection using Viola-Jones method which minimizes computation time while achieving high detection accuracy. It also proposed a SVC method to classify the features. It also exhibit high accuracy. The contribution of this paper is a simple and efficient classifier built from computationally efficient features using AdaBoost for face detection.

References

[1] B. Fasel and J. Luettin, “Automatic facial expression analysis”: A survey Pattern Recognit.,vol. 36, no. 1, pp. 259–275, 2003

[2] C. J. Young, R. Y. Man, and K. N. Plataniotis, “Color face recognition for degraded face images,” IEEE Trans. Syst., Man, Cybern. B, Cybern., vol. 39, pp. 1217–1230, Oct. 2009
[3] J. Wang, L. Yin, X. Wei, and Y. Sun, “3-D facial expression recognition based on primitive surface feature distribution,” in Proc. IEEE Conf. Comput. Vis.PatternRecognit.,Jun.2006,pp.1399–1406
[4] L. Torres, J. Y. Reutter, and L. Lorente, “The importance of the color information in face recognition,” in Proc. Int. Conf. Image. Process, vol. 2. Kobe,Japan,Oct.1999,pp.627–631
[5] P.Viola and M. J. Jones, “Robust real-time face detection,” Int. J.Comput. Vis.,vol.57,no.2,pp.137–154,2004
[6] P.Ekman, E. T. Rolls, D. I. Perrett, and H. D. Ellis, “Facial expressions of emotion: An old controversy and new findings discussion,” Phil.Trans. Royal Soc. London Ser. B, Biol. Sci., vol. 335, no. 1273, pp.63–69, 1992
[7] Seyed Mehdi Lajevardi and Hong Ren Wu “Facial Expression Recognition in Perceptual Color Space” vol. 21, no. 8, august 2012
[8] S. M. Lajevardi and Z. M. Hussain, “Emotion recognition from color facial images based on multilinear image analysis and Log-Gabor filters,” in Proc. 25th Int. Conf. Image. Vis. Comput., New Zealand, Dec. 2010, pp. 10–14
[9] T. S. Lee, “Image representation using 2-D Gabor wavelets,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 18, no. 10, pp. 959–971, Oct.1996
[10] X. Xie and K.-M. Lam, “An efficient illumination normalization method for face recognition,” Pattern Recognition.Lett., vol. 27, no. 6, pp. 609–617,2006
[11] Y. Lijun, C. Xiaochen, S. Yi, T. Worm, and M. Reale, “A high-resolution 3-D dynamic facial expression database,” in Proc. 3rd Int. Conf. Face Gesture Recognition., Amsterdam, The Netherlands, Sep. 2008, pp. 1–6
[12] Z. Liu and C. Liu, “A hybrid color and frequency features method for face recognition,” IEEE Trans. Image Process., vol. 17, no. 10, pp. 1975–1980, Oct. 2008