Project Report for CS9840 Machine Learning and Computer Vision

# Facial Expression Recognition

## Name: Qijie Chen

## Jan 22th, 2016

# Introduction

Expression is humans’ natural abilities, it can reflect humans’ internal emotions and feelings. For example, If someone hang a smile on his face, we can know that he is happy; and if someone hang with tears on his face, we know that he is in some pessimistic emotion right now. So it is important for us to identify the expression, it is an effective way to better understand others in our daily lives.

For this reason, facial expression recognition is one of popular research fields in machine learning. It is somehow similar to the face recognition techniques. With this technique, the computer can smoothly identify different expressions among a large number of people, which can significantly important for human activity analysis and human psychology analysis.

In this project, I download some facial expression image database from the Internet, using gabor filter, pca and svm method to implement an effective approach for facial expression recognition, run some test to validate its performance. And this report is the description of my working process and research result of my project.

# Data Preparation

The data I choose for this project is called pain expression database, which provided by the British academic intitude. It contains seven expressions from twelve different women, which makes the database has totally eighty-four images. The most significant features of this database is that all seven expressions from each women own same eye locations. I devide the database into training data set and testing data set, and specific proporation is 75% and 25%.

Figure samples from the image database (from left to right is angry, happy and surprised expressions)

# Feature Extraction

The features I choose for this project is called Gabor filter, a popular and effective technique created by Gabor on 1946 and widely use among face recognition.When we process face recognition,Some factors like facial expression, illumination characteristics and various effects geometric transformation may affect pattern features, so that it is difficult to obtain a high recognition accuracy. And Gabor filter can capture the corresponding spatial location, spatial frequency and local structure information by orientation selectivity, and its transformation coefficients describes the given gray feature near the specific area on the image which makes it is fit for showing local features of the face. Basically, Gabor filter is is a linear filter used for edge detection. Frequency and orientation representations of Gabor filters are similar to those of the human visual system, and they have been found to be particularly appropriate for texture representation and discrimination. The initial and general gabor filter definition shows in following pictures

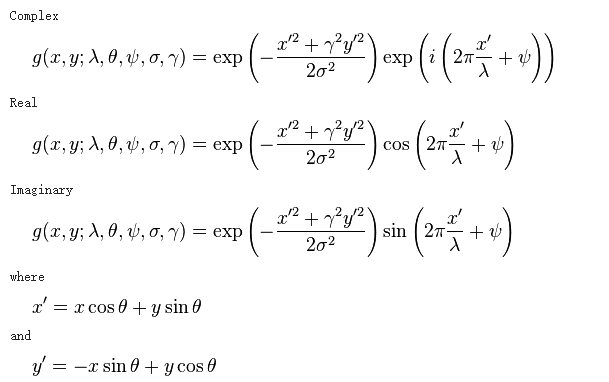


Figure Gabor Filter definition (<https://en.wikipedia.org/wiki/Gabor_filter>)

Since the images are 2-dimension objects which use a matrix of pixels to represent, we can use a 2-D gabor filter function to extract gabor features. In 1985, Daugman[2] proves that in a spatial domain, a 2D Gabor filter is a Gaussian kernel function modulated by a sinusoidal plane wave. Based on this theory, 2-D Gabor fuction can be defined as the following equation:

In this fomula, , ;

i is an imaginary number;

γand η are the standard deviation of the Gaussian envelope projected on the x axis and y axis respectively (depends on the f);

f is the frequency of gabor filter;

And there is some relationship between f, γand η. Based on the sampling theory, the frequency of gabor filter(f) can be ranged from 0~0.5, andγand η are related to the peak frequency fp and scale factor of adjacent filter k, here is the equation:

So in my project, I use 5 scales and 8 orientation gabor filter sets, which totally contains 40 gabor filters and it can retreive most characteristics from original images[2]. And each gabor filter size is 15\*15, which provides the best performance among the filter size range from 1\*1 to 50\*50. We should notice that scales influence the fluence of each gabor filter, we define s as the scale, and it should satisfy the following equation

And the orientation is the rotation angle of gabor filter set, each orientation o satisfy the following equation

The following picture is the wavelet representation of gabor filter sets

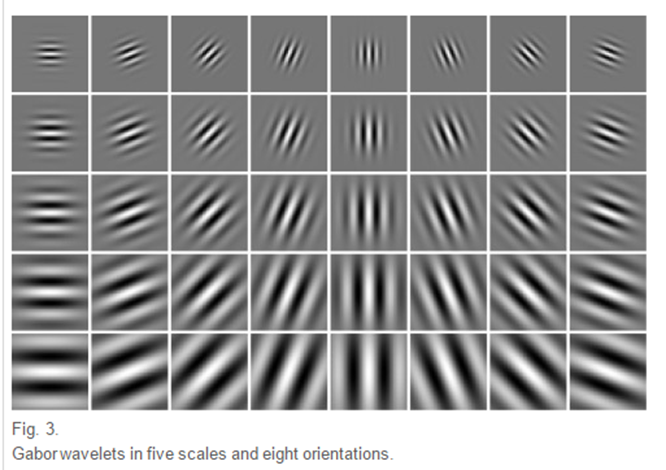


Figure Gabor wavelets in five scales and eight orientations[3]

Since we create the gabor filter sets, we just filter every pixels from every images of training set and testing set, then we can obtain the feature matrix of training set and testing set. All images should be transformed into grayscale representation before this feature extration process.

# Data Reduction

Data reduction process is necessary for the feature matrix. From the process we mention on the previous section, if the image size of our dataset is 100\*100, and we have 40 gabor filters on our gabor filter sets, since every pixel need to be filtered, and the dimension of feature vecter for each image will be 40\*100\*100=400000. It is a very large feature vector and it may affect the efficiency and performance of the SVM classifier which I will use later. So, here we should figure out how to reduce the dimension of feature vectors. Normally, there is two way to solve this problem, downsample the features and PCA.

Downsample the features is an intuitive method to reduce the dimensions , it is based on the fact that adjacent pixels in the image are usually correlated[4]. Using this method,first we need to define a reduction factor, it cannot be very large because if we set up a large reduction factor, it may lose too many features from original feature vector. Normally the reduction factor is varied from 2~5.

The downsampling procedures include 5 steps, we assume the downsample factor is k:

1. transform every number in the feature matrix into its absolute value representation
2. downsample the rows: From top to bottom, in every k rows, we only reserve the first row and abandon the rest k-1 rows
3. downsample the columns: From left to right, in every k columns, we only reserve the first column and abandon the rest k-1 columns
4. resize the fecture matrix into a fecture vector
5. Normalize the feature vector into zero mean and unit variance

For example, if we define the downsample factor is 4, and the new feature vector will be 400000/(4\*4)=25000, it is only 1/16 compared to the original size.  
 However, the downsampling approach cannot give us an appropriate size of feature vector, it is still relatively large. So we need to apply PCA to further decrease the dimensions. The procedures of PCA was introduced in the lecture slides. I follow this procedures and inplement a PCA function, which can significantly reduce the dimensions of feature vector.

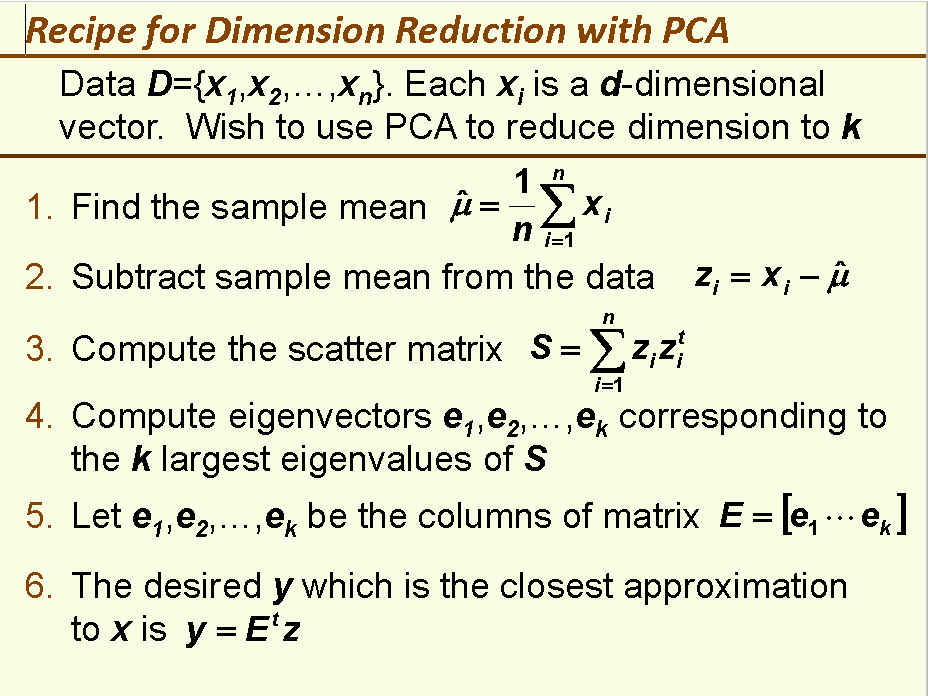


Figure Major procedures of PCA implementation (Cource lecture slid 6, page 28)

# Classifier Selection

From the previous steps, we have the appropriate feature matrix for training set and testing set. And now I should select a classifier which have better performance. After serious consideration and paper reading, I choose SVM as my classifier. SVM classifier are widely used in face recognition project and have better performance integrated with Gabor filter. the original reason is that SVM has better performance to handle pattern recognition problem with small samples and nonlinear features.

I use the LIBSVM library to support my svm classifier. Because we need to find the best value of parameter ‘beta’ which is the cost variable in the LIBSVM library, I use the coarse grid search to find this value.[6] Besides, I use linear kernel SVM instead of Gaussian kernel SVM or polynomial kernel SVM because my project has a best performance on linear kernel SVM.

# Test Result

Now, we have all a comple machine learning approach and we should run a test to validate its performance. The first test we use the original pain expression subset, there are 84 images in this dataset, and we devided the dataset into training set and testing set on 75/25 ratio, which means 63 images for training and 21 images for testing. And here is the result from the matlab.

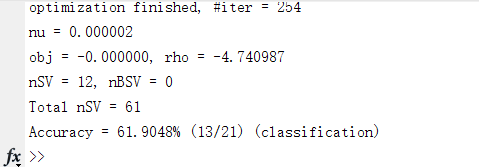
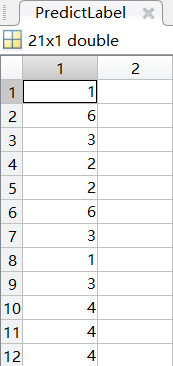
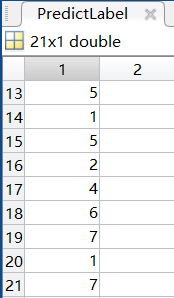


Figure The Accuracy rate of the test (using pain expression subset as a database)

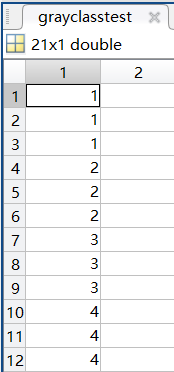
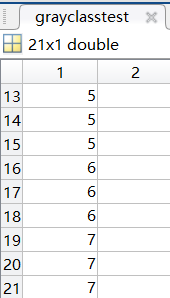
 

Figure The test result (The upper images are predition labels, and the lower images are the original testing label

From the testing result we can know that, for the pain expression subset, the classifier has 61.9048% accuracy rate. In total 21 samples, it can correctly predict 13 samples. The result shows that this classifier have good performance for this database.

Now, I will choose another database to run the test. The new database called nott-faces-originals, It contains 4 expressions from 68 people, including women and men. I also devided this database into training set and testing set, 51 for training and 17 for testing for each expressions.



Figure The test result of nott-faces-originals dataset

However, The classifier has a poorer performance, and the accuracy rate is only 35.2941%. After analysis and consideration, the reason for this poor accuracy rate is that the illumination variation, uncertain face location and uncertain background may affect the performance of this approach.

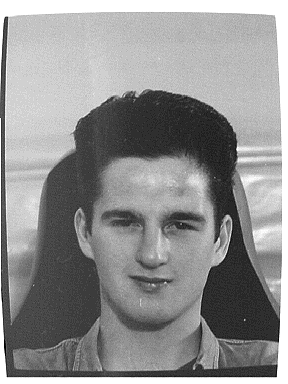
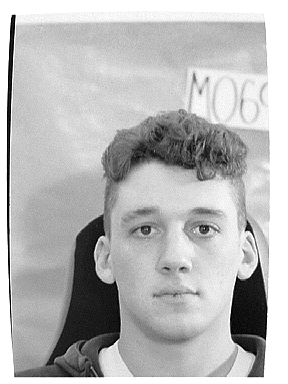
 

Figure two samples from the nott-faces-originals dataset, we can see that these images have different illumination variation, face location and background

# Summary

In this project, I implement an effective machine learning approach for facial expression recognition. Basic tools I applied for my project is Gabor filter+PCA+SVM, and the test result shows that this approach has nice performance on our datasets. Besides this coding work, in this project I read many papers and research reports, which also expand my horizon on machine learning and methemetic knowledge. Although the project is finished, There is some shortcomings on my work. For example, I haven’t solve the problem of accuracy declining on the large facial database. Another shortcoming is that I still cannot find the reason for the different performance when I apply different SVM kernel on my project. Nevertheless, this project is my first personal project on machine learning aspect, and I feel really happy to finsh this project!

And thank you, Profeesor Veksler, for your teaching and tutoring on this course!

# Reference

[1] JG Daugman. Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression, IEEE Transactions on Acoustics Speech & Signal Processing, 1988, 36(7):1169-1179

[2] TS Lee. Image representation using 2D Gabor wavelets. Pattern Analysis & Machine Intelligence IEEE Transactions. 1996, 18(10):959-971

[3] M Haghighat, S Zonouz, M Abdel-Mottaleb. CloudID: Trustworthy cloud-based and cross-enterprise biometric identification. Expert Systems with Applications, 2015, 42(21):7905-7916

[4] V Kyrki, JK Kamarainen, H KälviäinenSimple. Gabor feature space for invariant object recognition. Pattern Recognition Letters, 2004, 25(3):311-318

[5] L Chengjun, W Harry. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. IEEE Transactions on Image Processing A Publication of the IEEE Signal Processing Society, 2002, 11(4):467-476