Facial Expression Recognition in Video using Adaboost and SVM

Surabhi Prabhakar
Department of CSE
Amity University
Noida, India

Jaya Sharma
Department of CSE
Amity University
Noida, India

Shilpi Gupta
Department of CSE
Amity University
Noida, India

ABSTRACT
In human-computer interaction facial expression is the characteristic and proficient method for correspondence, and has been acknowledged as essential input of such interface. In this paper, we present an enhancement in facial expression recognition for image sequence. The most important step is to extract essential features from face to efficiently determine facial expression. Experimentation shows that LBP method performs well while extracting facial features. We further found that Boosted-LBP extracts most distinct features and the best recognition is calculated by SVM classifier.,

General Terms
Pattern Recognition, Facial expression recognition.

Keywords
Facial expression, Adaboost, Support vector machines.

1. INTRODUCTION
In this fast advancement of technology period, automatic facial expression recognition has picked up an expanding enthusiasm toward building intelligent systems. The facial expression recognition strategy could be embraced as the important component in sensing the feelings of people. To recreate virtual ideas to real actions, the framework must be able handle the physiological indicators instantly and react to them. Henceforth, making a powerful facial expression recognition system is the essential objective of this study. The facial expression recognition systems might be partitioned into two primary classifications dependent upon information sources: image sequence [1]. In image sequence based identification, one expression is portrayed with a sequence of images. The image based strategies uses just the current image to identify the expression of the image.

Facial expression is nearly identified with face detection where a lot of examination has been carried out and an immeasurable exhibit of algorithms has been presented. Facial expression recognition can additionally be acknowledged as special case of pattern identification issue and numerous methods are accessible. In the planning of a Facial expression recognition framework, we can exploit these assets and furthermore utilize existing algorithms as building pieces of framework. So a real challenge of this work is to focus on the ideal combination of calculations. To do this, we first divide the framework into three modules, i.e. Preprocessing, Feature Extraction and Classification, then for each of them some methods are executed, and inevitably the ideal setup is found by analyzing the execution of diverse combinations. In this work, we observe facial representation based on features extracted from Local Binary Pattern (LBP) [3, 4] for facial expression recognition. LBP features were proposed initially for surface dissection, and recently have been presented for face representation in facial expression identification. The tolerance against illumination changes and effortless computation are the most important properties of LBP. In this experiment we have used Adaboost with LBP to determine the most discriminative LBP characteristics. Executions of diverse classifiers are enhanced by utilizing the Boosted-LBP characteristics. Finally we use Support Vector Machine classifier for training and classifying, this combination enhances both speed and accuracy of the system.

2. FACIAL EXPRESSION DATA
The facial expression interpretation framework was prepared and tested on Cohn and Kanade's DFAT-504 dataset [2]. This dataset comprises of 100 university students extending in age from 18 to 30 year. 65% were female, 15% were African-American, and 3% were Asian or Latino. Video sequences were recorded with simple analog S-feature utilizing a camera placed straight to the frontal face of the subject. The students were instructed by the experimenter to perform an arrangement of 23 facial expressions. Subjects started and finished each expression with a natural face. When performing each expression, the experimenter depicted and demonstrated the wanted expression. Each image sequence starts from neutral face to target expression and were digitized into 640 by 480 pixel array with 8-bit exactness for grayscale values. For our study, we chose Cohn-Kanade dataset that were marked as one of the 6 fundamental expressions. The arrangements hailed from 90 subjects, with 1–6 feelings for every subject. Cohn-Kanade dataset sample is shown in fig.1.

Figure 1: Sample from Cohn-Kanade dataset

3. DESIGN AND IMPLEMENTATION
3.1 Face detection
Emulating Tian [5], to maintain a fixed distance between two eyes we normalized the faces. Face representation might be attained by face recognition [6] which is more confined to eyes area. Facial images of 110 x 150 pixels were cropped from original edges dependent upon the two eyes area. No further enrollment such as arrangement of mouth was performed in our approach. As the facial images in the

database are frontal perspective, we did not consider head posture changes. There were illumination variations exist in the database, because of LBP's gray scale variance we did not made any calculations to evaluate illumination changes in our work, Fig. 2 shows a sample of the frontal face image of cropped image.

The scalar product of the image and a Haar layout provides a Haar-like feature in Viola-Jones algorithm. Suppose an image I and a Haar layout H, the layout H has background gray to support the pattern, they should be of same size M x M (Fig 3). Black and white portion in pattern symbolizes the feature that needs to be extracted. The gray region around the pattern changes according to the dimension of image. To remunerate the effect of different illumination conditions, we have to normalize mean and variance of all the images in advance.

Figure 2: Cropped fontal face of image sequence

Figure 3: Image I and Haar layout H

Four patterns are recognized when we use the algorithm to recognize face, these patterns are shown in fig 4. The inferred characteristics are expected to hold all the data required to determine a face. There is an alternate vital component which lets this set of features outweigh everything else: the integral image which permits to evaluate them at very low computational expense. The strategy is to use the cumulative approach rather than summing up all the pixels within a rectangle.

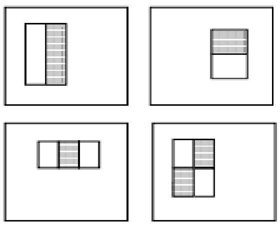

Figure 4: Rectangular features for detecting face

3.2 Extracting data for expression analysis

Local binary patterns (LBP) were originally introduced by Ojala et al.[6]. Basically LBP was used for texture analysis. It can be used in Facial expression feature extraction by exchanging pixels of an image by 3 x 3 neighborhood of each pixel with its central value and converting the result into a binary number Fig.5. The binary numbers found are called LBP codes, these codes help in defining various primitives of an image edges, flat surface, any spot in image etc. By analyzing circular neighbors pixels can allow us to use any radius and number of neighborhood pixels. Suppose P are points on circle drawn with radius R. the points are spaced equally. When a binary number of LBP circular pattern has two transitions from 0 to 1 or 1 to 0 then it is called uniform pattern. Almost 90% of all LBP patterns of eight neighborhood circular representation are uniform patterns. Histogram of labeled image with LBP operator is determined. This histogram contains information related to patterns in an image like edges, flat surfaces, any spot in image. Using this histogram we can identify various features of an image. The extracted histogram represents the features and shape of face in an image of image sequence. We perform these steps repeatedly for each image in a particular image sequence of video.

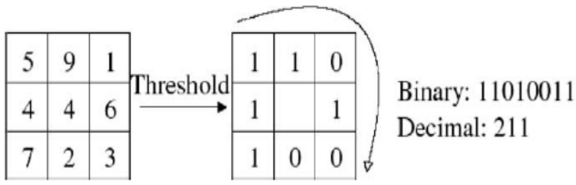

Figure 5: LBP operator

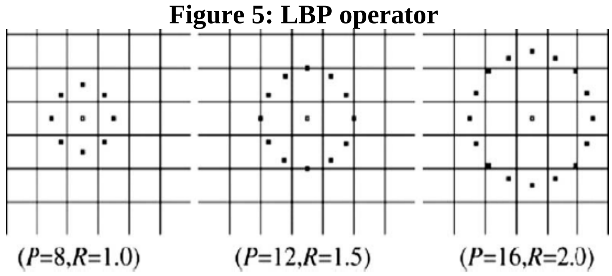

Figure 6: Circular neighborhood with different radius and different points

3.3 Adaboost with LBP

The above experimentation show that the LBP features are powerful for facial expression recognition, and performed in the same way or superior to existing strategies. In the above experiment, facial images are divided into equal sub-areas. LBP histogram of each sub region is calculated and linked into a single feature vector. Clearly the concentrated LBP features depend on the partitioned sub-areas, so this LBP feature extraction strategy is based on equal sub-area size and positions. Furthermore by scaling and shifting a sub-window over facial image, we can identify various other sub- areas which helps in identification of more LBP histogram features. To minimize an expansive number of LBP histograms presented by moving and scaling a sub-window, boosted learning approach could be used to identify the most discriminate LBP histograms that holding much discriminative data. Adaboost [7, 8] gives a basic yet successful methodology for stage-wise learning of a non-linear classifier.

Adaboost takes various weak classifiers and boost them with appropriate weights and try to make a strong classifier which is more accurate. The methodology of Adaboost takes different sets of training data. At every cycle, weights are minimized with each weak classifier and select an error rate accordingly. Adaboost has been effectively utilized within numerous issues for example, face recognition [6]. As every LBP histogram is figured from a sub-area, Adaboost help is really used to determine the sub-areas that hold more discriminative data for facial expression classification in term of the LBP histogram. For every sub-area, the LBP histograms in a given class are generated to create a layout for this class. The prepared weak classifier matches the input histogram with the closest layout, and yields the corresponding class.

We further consolidate features determination by Adaboost with Support vector machine classifier. Specifically, we prepare SVM network with the Boosted- LBP features. We connected Adaboost with LBP to take in the discriminative LBP histograms for each expression, and after that used the combination of selected LBP histogram to form a input to Support vector machine.

3.4 Expression classification

We initially analyzed classification of facial expressions by around Support Vector Machines (SVM). The framework performed a 7-way constrained decision between the emotion classifications: Happy, sad, surprise, disgust, fear, anger and neutral. There are a number of strategies for making a multiclass classifier using number of binary classifiers, [9]. Here, the seven ways constrained decision for six emotions in addition to neutral was prepared in two stages. In stage I, Support vector machines performed binary decisions. We investigated two methodologies to train SVM for binary decisions: one-versus-one and one-versus-all. Stage II converts the representation transformed by the first stage into a probabilistic dissemination over the seven expressions. The key characteristics of SVM are the utilization of kernels deficiency of local minima, the inadequacy of the result and the capacity control acquired by optimization of margin. Support vector machines utilize a specific type of class function called kernel, it creates a large margin in feature space. The kernel by introducing a factor of similarity index between data determines various possible patterns. Ordering information is a typical errand in machine learning. Assume some given information focuses each one have a place to one of two classes, and the objective is to choose which class a new information point will be in. On account of support vector machines, information is seen as a p-dimensional vector (list of p numbers) and we need to know whether we can separate such points with a one dimensional hyperplane (Fig.7.). This is known as a linear classifier. There are numerous hyperplanes that may arrange the information. One sensible decision as the best hyperplane is the particular case that speaks to the biggest division, or edge, between the two classes. So we pick the hyperplane so the separation from it to the closest data point on each side is maximized. In the event that such a hyperplane exists, it is known as the most maximum margin hyperplane and the linear classifier it characterizes is known as a maximum margin classifier.

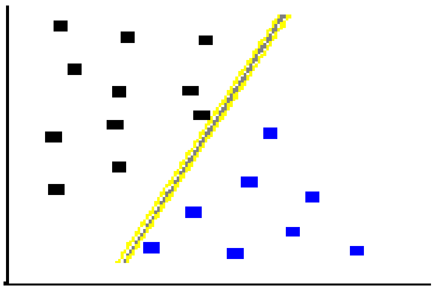

Figure 7: Hyperplane to separate 2-class

SVM's and Adaboost are both appropriate to the undertakings portrayed in this paper, in light of the fact that they can adapt to huge representation space, they sum up well, perform decisions in real time and are easy to train. Here we audit the likenesses between SVM's and Adaboost, and also indicate where the two algorthms diverge. SVM's and Adaboost are both substantial margin classifiers. The two methodologies might be considered maximizing the margin that relies upon the weights and hypothesis [10], in spite of the fact that Adaboost does not typically achieve the maximization.

One distinction between the classifiers is that the standards in the denominator of their equation. There are two norms for the standard manifestation of the SVM, while for Adaboost there is a one norm for alpha and infinite norm for hypothesis. We investigated preparing SVM classifiers on the features selected by Adaboost. Adaboost is not just a quick classifier; it is likewise a feature determination system. In feature selection by Adaboost, every Gabor filter is a treated as a weak classifier. Adaboost picks the best of those classifiers, and at that point supports the weights on the samples to weight the errors more. The next filter is chosen that gives the best execution on the errors of the previous filter. At each stage, the picked filter could be demonstrated to be independent with the result of the previous filters [10,11]. Training of SVM is done on the features extracted by Adaboost. When we prepared SVM's on the outputs yields of the chosen Gabor features, they performed no superior to Adaboost. In any case, we prepared SVM's on the consistent outputs of the selected filters. We casually call these joined classifiers as Ada-SVM.

4. CONCLUSION AND FUTURE WORK

In this paper, we exhibit a thorough experimental investigation of facial expression recognition focused around Local Binary Patterns features. The key issues of this work might be condensed as:

- Inferring a successful facial representation from unique face image is a crucial step for facial expression recognition. We experimentally evaluated LBP features to portray appearance progressions of representation image. Experimentation outline that LBP features are compelling and effective for facial expression recognition
- We used Adaboost to identify the most discriminative LBP features from an expansive LBP characteristic pool. We can get highest rate of expression classification by using SVM with boosted LBP features; this strategy has limitation on generalization to other datasets.

We will investigate temporal data in our future work. Volume LBP and LBP from 3-orthogonal planes have been presented for dynamic texture analysis [12]. Limitation of the current work is that we do not consider head position and impediments; we will consider this in our future work. We will also consider non-clear face images to detect facial expression.

5. ACKNOWLEDGMENTS

We would like to express our gratitude to Mr. Abhay Bansal, HOD Amity School of Engineering and Technology, Amity University, for his guidance, proofreading and help in completing this work successfully.

6. REFERENCES

[1] F. De la Torre and J. F. Cohn, "Guide to visual analysis of humans: looking at people, chapter Facial Expression Analysis, Springer, 2011.

[2] T. Kanade, J.F. Cohn, and Y. Tian. Comprehensive database for facial expression analysis. IEEE, pages 46–53, Grenoble, France,

[3] T. Ojala, M. Pietikäinen, T. Mäenpää, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence , 2002.

[4] Y. Tian, Evaluation of face resolution for expression analysis, in: CVPR Workshop on Face Processing in Video, 2004.Tavel, P. 2007 Modeling and Simulation Design. AK Peters Ltd.

[5] P. Viola, M. Jones, Rapid object detection using a boosted cascade of simple features, in: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2001.Forman, G. 2003. An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res. 3 (Mar. 2003), 1289-1305.

[6] T. Ojala, M. Pietikäinen, D. Harwood, A comparative study of texture measures with classification based on featured distribution, Pattern Recognition, 1996.

[7] Y. Freund, R.E. Schapire, A decision-theoretic generalization of on-line learning and an application to boosting, Journal of Computer and System Sciences ,1997.

[8] R.E. Schapire, Y. Singer, Improved boosting algorithms using confidence-rated predictions, Maching Learning, 1999.

[9] S. Har-Peled, D. Roth, D. Zimak, Constraint classification for multiclass classification and ranking, in: S. Becker, K. Obermayer (Eds.), Advances in Neural Information Processing Systems, vol. 15, MIT Press,Cambridge, MA, 2003.

[10] J. Freund, R.E. Schapire, A short introduction to boosting, Journal of Japanese Society for Artificial Intelligence 14 (5) (1999) 771–780.

[11] J. Friedman, T. Hastie, R. Tibshirani, Additive Logistic Regression: A Statistical View of Boosting, 1998.

[12] G. Zhao, M. Pietikäinen, Dynamic texture recognition using local binary patterns with an application to facial expressions, IEEE Transactions on Pattern Analysis and Machine Intelligence 29 (6) (2007) 915–928.