EigenExpress Approach in Recognition of Facial Expression Using GPU

Qi Wu, Mingli Song, Jiajun Bu, and Chun Chen

College of Computer Science, Zhejiang University,
Hangzhou, PR China, 310027
{wuqi, brooksong}@ieee.org,
{bjj, chenc}@zju.edu.cn

Abstract. The automatic recognition of facial expression presents a significant challenge to the pattern analysis and man-machine interaction research community. In this paper, a novel system is proposed to recognize human facial expressions based on the expression sketch. Firstly, facial expression sketch is extracted by an GPU-based real-time edge detection and sharpening algorithm from original gray image. Then, a statistical method, which is called *Eigenexpress*, is introduced to obtain the expression feature vectors for sketches. Finally, Modified Hausdorff distance(MHD) was used to perform the expression classification. In contrast to performing feature vector extraction from the gray image directly, the sketch based expression recognition reduces the feature vector's dimension first, which leads to a concise representation of the facial expression. Experiment shows our method is appreciable and convincible.

1 Introduction

Facial expression is one of the most powerful, natural, and immediate means for human beings to communicate their emotions and intentions, Relevant studies have been performed on the relationship between emotion and facial expressions since 1970s. Ekman[2, 3] grouped the facial expressions into six "universality" ones. Those are so called happiness, sadness, anger, fear, surprise, and disgust.

Some approaches[4] extract the facial features from the image, and these features are used as inputs into a classification system to categorize different facial expressions automatically. Some approaches[5] estimate the facial expression and the gender of a person based on statistical data analysis and neural classifiers. Also, Ekman's[2] introduced Facial Action Coding System(FACS)which codes the facial expressions by these facial feature components. However, there is much useless information around the expression features in the gray face images which influence the precision and efficiency of recognition.

Therefore, enhancing the edge between the components in the expression face will decrease the ambiguity and noise in the image. Consequently, it will emphasize the maximum feature components and reduce the redundant information. However it is known that the time cost of such operation is heavy on CPU.

T.S. Huang et al. (Eds.): HCI/ECCV 2006, LNCS 3979, pp. 12–20, 2006.

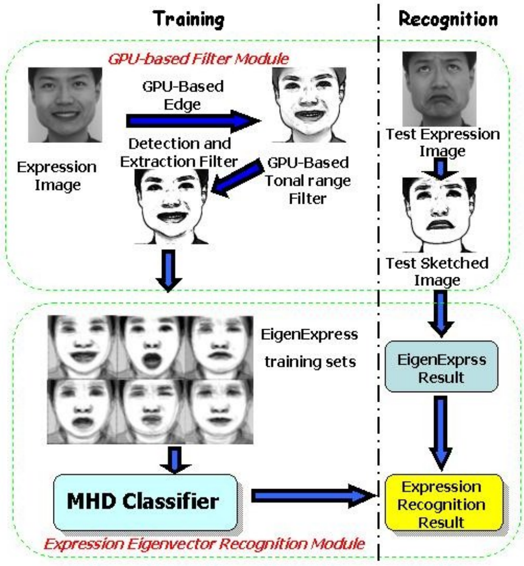


Fig. 1. System Overview

Recent development on computer graphics hardware makes GPU an extremely powerful graphics processing units (GPU). It can perform operations on very large amount of vectors very fast. It has more than one processing pipeline working in parallel. Also, the performance of GPU is growing at a high speed while Nvidia Corp. estimated that the Nvidia GeForce FX 7800 GPU performance peaks at 200 GigaFlops[10, 8]. Considering the parallel nature of implementing specific image processing with filters using vertex shader and pixel shader to process the image, this new graphics hardware is a suitable candidate for implementing them at an extremely high speed. This can be used to sharpen and detect the edge of the image of face expression to generate the corresponding sketch.

With the sketched images, the statistical based face representation, which we called *Eigenexpress*, is applied to represent them with serval eigenvectors that reserve the most energy of the images and set up the expression features training sets. Compared with the *Eigenfaces* [6], which plays a fundamental role and has been proved to be appreciable in face recognition, it is more robust and precise in facial expression recognition, especially for different people with different expressions. Finally, calculating the Modified Hausdorff distance between the test image's *Eigenexpress* representation and the seven basic expression training sets, the test expression eigenvectors, also generated from test expression images by GPU-base filter and Eigenexpress method, are classified into one of the seven basic categories. Figure 1. shows an overview of our facial expression recognition system with GPU-based filter which consists two parts: GPU-based facial feature extraction module, and expression recognition module.

2 GPU-Based Facial Feature Extraction

Modern graphics hardware give GPU a powerful capability to carry out Matrix-Vector operations. A series of filtering operations on image such as edge detection and sharpening, tone mapping, etc., can be performed at an extremely high speed. So it is suitable for GPU to implement such operations to obtain qualified sketched image at a very high speed.

In our GPU-based Edge Filter, each pixel in Gray-level Image can be operated concurrently under the stream pipeline of GPU.

Firstly, we convert it to luminance value $P_{i,j}$, $i \in Weight$, $j \in Height$ and sample all four around texture stages $P_{i,j}$, $P_{i+1,j}$, $P_{i,j+1}$, $P_{i+1,j+1}$.

Secondly, the two diagonal luminance differences of all four samples are computed, which Fig.2 shows to us,and we square each differences (it is easier and faster than obtaining its absolute value in GPU) ,then sum up them.

$$P'_{i,j} = (P_{i,j} - P_{i+1,j+1})^2 + (P_{i+1,j} - P_{i,j+1})^2 \tag{1}$$

Thirdly, we multiply $P'_{i,j}$ with a large number δ to make the values visible.

$$P'_{i,j} = P'_{i,j} \cdot \delta \tag{2}$$

Then, the result is subtract form 1 to invert edges black on white for we have normalize the gray range to $(0, 1)$.

$$P'_{i,j} = 1 - P'_{i,j} \tag{3}$$

Finally, we multiply edge-image with the luminance values and obtain the final result $P''_{i,j}$.

$$P''_{i,j} = P'_{i,j} \cdot P_{i,j} \tag{4}$$

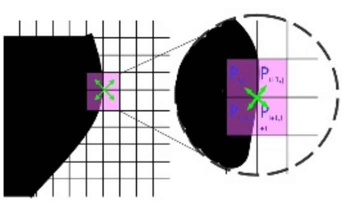


Fig. 2. Compute The Square Diagonal Luminance Differences

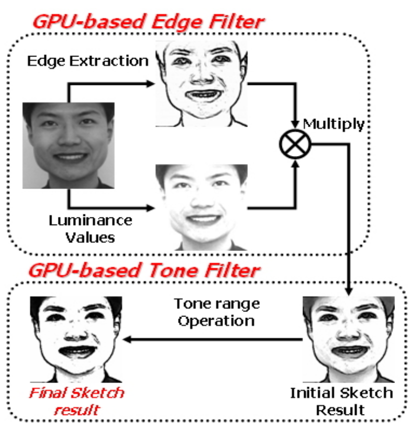


Fig. 3. Framework of GPU-based Filter

However, the result obtained from the method mentioned above still contains much noise make the image blurred and unreliable. So in our approach, a further step is taken to eliminate these high frequency noise by changing the tonal range, which can also be run extremely fast on GPU. For the world luminance L, it can be defined corresponding with the equation below [9].

$$TM(L) = LDMAX \frac{C(L) - C(L_{min})}{C(L_{max}) - C(L_{min})} \tag{5}$$

In our case, the mid-tone and the highlight are 0.25 and 175 respectively to compute L_{max} and L_{min}. As shown in Figure 3, we can see the luminance filtered sketch is cleaner than the initial one while keeping the detail of the face.

The following Figure 4. is one of our examples of creating the sketches corresponding to six basic expressions by the GPU-based filter which just takes 0.00153 second for each on average.

Fig. 4. Examples of sketched expressions

3 Facial Expression Recognition

As an important human behavior for conveying psychological information, facial expression has been studied for some ten's of years in different modal: visual, speech, etc. But there is still much work need to be done to get a higher accuracy. Similar to the previous work, we also categorize given face images into seven classifications: neutral, happy, anger, surprise, disgust, sad and fear.

3.1 EigenExpress Method

The way of EigenExpress uses the Karhunen-Loeve Transform (KLT)[12] for the representation of face expressions. Once a set of eigenvectors is computed from the ensemble face covariance matrix, a facial expression image can be approximately reconstructed using a weighted combination of the eigenvectors, called EigenExpress, The weights that characterize the expansion of the given image in terms of eigenvectors constitute the feature vector. When a new test image is given, the weights are computed by projecting the image onto the vectors. The classification is then carried out by comparing the distances between the weight vectors of the test image and the images from the training sets.

To implement the EigenExpress, assuming F_i is a column vector of dimension N^2 representation of a sample facial expression image with the mean face computed as $\boldsymbol{m_p} = 1/M \sum_{i=1}^{M} F_i$, where M is the number of training samples, Removing the mean face from each training image, we have $P_i = F_i - \boldsymbol{m_p}$. The training set the form an N^2 by M matrix $A_p = [P_1, P_2, \cdots, P_M]$. We can find the orthogonal eigenvectors u_n corresponding to the C-largest eigenvalues λ_n of matrix $A_P^T A_P$

$$(A_P^T A_P)u_n = \lambda_n u_n \quad (n = 0, 1, \cdots, C-1) \tag{6}$$

In order to reduce the dimension of the matrix $A_P^T A_P$ which is N^2 by N^2 to decrease the computational complexity, we want to use M instead of N^2. Therefore, We use dominant eigenvector estimation method [11] multiplying both sides by A_P and we have

$$(A_P A_P^T)A_P u_n = A_P \lambda_n u_n \tag{7}$$

The orthonormal eigenvector matrix of the covariance matrix is

$$u_n = A_P u_n \lambda_n^{-1/2} \tag{8}$$

For a new facial expression image P_k, its projection coefficients in the eigenvector space form the vector $b_P = u_P^T P_k$ which is used as a feature vector for the classification. We call this feature vector *Eigenexpress*.

Because of the structural similarity across all face images, strong correlation exists among facial expression images. The *Eigenexpress* method takes advantage of such a high correlation to produce a highly compressed representation of facial expression images, thus improves expression recognition more efficiently.

3.2 Modified Hausdorff Distance

Hausdorff distance[14, 13] is a robust method to describe the similarity between two point sets. Given two finite point sets $A = \{a_1, a_2, \cdots, a_p\}$ and $B = \{b_1, b_2, \cdots, b_q\}$, The Hausdorff distance is defined as

$$H(A, B) = max(h(A, B), h(B, A)) \tag{9}$$

where

$$\begin{cases} h(A, B) = \max\limits_{a \in A} \min\limits_{b \in B} \parallel a - b \parallel \\ h(B, A) = \max\limits_{b \in B} \min\limits_{a \in A} \parallel b - a \parallel \end{cases} \tag{10}$$

And $\|\|$ is some underlying norm on the points of A and B (e.g., the L_2 or Euclideam norm).

The function $h(A, B)$ is called the directed Hausdorff distance from A to B. That is, $h(A, B)$ in effect ranks each point a_i of A based on its distance $\parallel a_i - b_j \parallel$ to the nearest point b_j of B and then used the largest distance as the value of $h(A, B)$. The Hausdorff distance $H(A, B)$ is the maximun of $h(A, B)$ and $h(B, A)$. Thus, it measures the degree of maximum mismatch between two sets.

In our approach, For both the gray based and sketch based facial expression recognition, we compute the the basic expression mean face's *Eigenexpress* first,which is represented as b_{P_m}, $m = [1, .., 7]$.Then we compare b_P with b_{P_m} by Modified Hausdorff Distance(MHD) which defines:

$$H_{LK}(b_P, b_{P_m}) = \max(h_L(b_P, b_{P_m}), h_K(b_{P_m}, b_P)) \tag{11}$$

where,

$$\begin{cases} h_K(b_{P_m}, b_P) & = & \frac{1}{N_{b_{P_m}}} \sum_{b \in b_{P_m}} \min\limits_{a \in b_P} \parallel a - b \parallel \\ h_L(b_P, b_{P_m}) & = & \frac{1}{N_{b_P}} \sum_{a \in b_P} \min\limits_{b \in b_{P_m}} \parallel b - a \parallel \end{cases} \tag{12}$$

where $N_{b_{P_m}}$ and N_{b_P} are the number of eigenvectors in Eigenexpress sets b_{P_m} and b_P.

The Hausdorff distance for recognition is different for the seven basic *Eigenexpress*. It determines which basic expression the test image should belong to.

Compared with basic Hausdorff distance, MHD is not sensitive to the noise and ambiguity of the image. It can decrease the influence of irrelevant eigenvectors and make more precise and robust classification.

4 Experiment and Evaluation

In our approach, GPU-based filter is used to preprocess facial images in performing the recognition. In this way, noise and redundant information in the image are reduced, and all the component such as eyes, mouth and chin which have obvious differences in seven basic expressions will be enhanced.

As for the *Eigenexpress*-based method, after expression sketching preprocessing, because the feature components have been enhanced, most of the energy concentrate on the top eigenvectors by which we can use much less eigenvectors to keep the same energy as before. Consequently, it reduces computing workload and makes the expression recognition more efficient.

In order to build a model that is flexible enough to cover most of the typical variation of face, we have collected more than 840 gray expression images form 40 people with different basic expressions and their corresponding sketched expression images processed by our GPU-based filter as the data sets.

Then we select 280 gray expression images which is $200 * 200$ pixels do the comparison. Table 1 shows the method's performance on the gray image, while Table 2 gives out a higher accuracy on the sketched image. We use Confusion matrix of the emotion recognition. (Columns represent the emotion elected by our method for samples belonging to the emotion of each row)

In our experiment, the most clearly recognizable expression is Surprise and Happy which have great difference form others, but Anger and Disgust are a bit lower with their difficulty for recognition not only by our method but also by human beings.

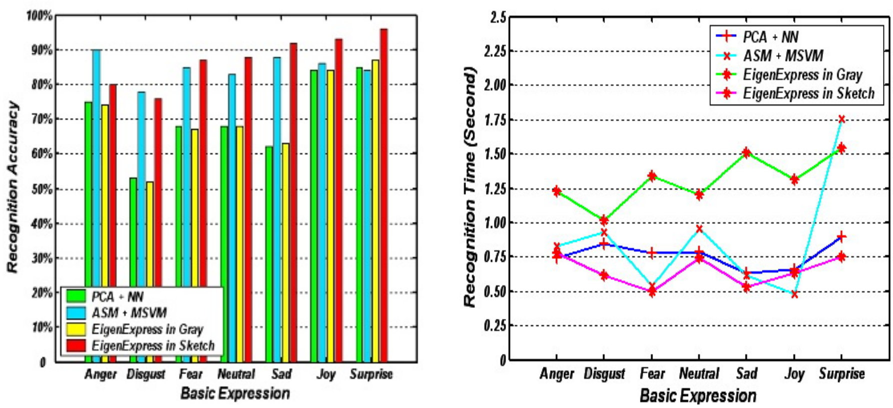
The comparison of recognition accuracy in different methods(PCA + Neural Network, ASM + SVM, Eigenexpress in Gray and Eigenexpress in Sketch)are

Table 1. Gray-level Image Based Expression Recognition with *Eigenexpress*

	Anger	Disgust	Fear	Neutral	Sad	Joy	Surprise	Tot	Ratio.
Anger	30	4	1	3	2	0	0	40	75%
Disgust	1	21	3	4	7	0	4	40	52.5%
Fear	1	4	27	3	0	2	3	40	67.5%
Neutral	1	3	3	27	4	0	2	40	67.5%
Sad	2	1	6	5	25	1	0	40	62.5%
Happy	0	0	1	3	0	34	2	40	85%
Surprise	0	1	3	0	0	2	34	40	85%
Total	35	34	44	45	38	39	45	280	70.71%

Table 2. Sketch Image Based Expression Recognition with *Eigenexpress*

	Anger	Disgust	Fear	Neutral	Sad	Joy	Surprise	Tot	Ratio.
Anger	32	2	3	2	1	0	0	40	80%
Disgust	2	31	2	2	2	0	1	40	77.5%
Fear	1	2	33	2	0	0	2	40	82.5%
Neutral	2	0	2	34	1	1	0	40	85%
Sad	0	1	1	1	37	0	0	40	92.5%
Happy	0	0	0	1	0	38	1	40	95%
Surprise	0	0	0	0	0	1	39	40	97.5%
Total	37	36	41	42	41	40	43	280	87.14%



1) Comparison of The Accuracy 2) Comparison of The Cost Time

Fig. 5. Comparison between previous methods and our method

listed in Fig.5(1). It is obvious from the figures that our method has the highest accuracy than any others, and the significant improvements (almost 27.13% compared to PCA+NN, 4.7% compared to ASM+MSVM and 28.34% compared to Eigenexpress in Gray) are obtained by using Eigenexpress in Sketch generated by GPU. However, it is also clear from the comparison of the recognition cost time in Fig.5(2) that the using of GPU in our method improves the recognition speed greatly(almost 11.52% compared to PCA+NN, 34.2% compared to ASM+MSVM and 15.43% compared to Eigenexpress in Gray).

5 Conclusion

The automatic recognition of facial expression presents a significant challenge to the pattern analysis research community because the expression is generated by nonrigid object deformations vary from person to person. The expression recognition from a static image is, particularly, a more difficult problem compared to the recognition from an image sequence[15] due to lack of information during expression actions. In this paper, we propose a novel GPU enhanced facial expression recognition which use GPU-based filter to preprocess and convert the gray expression image to sketched facial expression at extremely high speed. *Eigenexpress* is introduced to analyze the features of different basic expressions to set up training sets. Finally, we use Modified Hausdorff distance to classify the expression images into seven basic expressions. Experiment result shows sketch based approach can obtain better result not only on the recognition accuracy but also the time cost.

In future work, more detailed work and complex action units will be considered in our system. Our goal is to find a way which can acquire a better effect with getting higher recognition accuracy and speeding less recognition time. we will employ new way to track the facial expression features more precisely and quickly. In addition, the combination of facial expression recognition and speech recognition will make our system more efficient and accurate.

Acknowledgments

This work is supported by National Science Foundation China 60203013 and Key Technologies R&D Program of Zhejiang Province (2004C11052 & 2005C23047).

References

1. Maja.Pantic and Leon J.M.Rothkrantz. Automatic analysis of facial expression: *The state of the art.*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(12):1424-1445,2000
2. P. Ekman, W.V. Friesen, *Facial Action Coding System: Investigator's Guide*. Consulting Psychologists Press, 1978.
3. P.Ekman, *Strong Evidence for Universals in Facial Expressions: A Reply to Russel's Mistaken Critique.* Psychological Bulletin, 115(2):268-287, 1994.
4. Ira Cohen, Nicu Sebe, Fabio G. Cozman, Marcelo C. Cirelo, Thomas S. Huang, *Learning Bayesian Network Classifiers for Facial Expression Recognition with both Labeled and Unlabeled data*IEEE Conference on Computer Vision and Pattern Recognition, pp.595-602, 2003.

5. T.Wilhelm, A.Backhaus, *Statistical and Neural Methods for Vision-based Analysis of Facial Expressions and Gender*. IEEE International Conference on Systems, Man and Cybernetics, Vol. 3, pp.2203-2208. 2004.
6. M. Turk and A. Pentland, *Eigenfaces for Recognition*, Journal of Cognitive Neuroscience, 3(1):71-86, 1991
7. M. Macedonia, *The Gpu Enters Computing's Mainstream* IEEE Computer, October 2003.
8. Jens Krüger and Rüdiger Westermann, *Linear Algebra Operators for GPU Implementation of Numerical Algorithms* ACM Siggraph 2003, pp. 908-916.
9. Michael Ashikhmin: *A Tone Mapping Algorithm for High Contrast Images.* Eurographics Workshop on Rendering (2002) 1–11
10. NVIDIA Corporation, *Industry's Premiere Game Developers Proclaim NVIDIA GeForce 7800 GTX GPUs Platform of Choice for Next-Generation Game Development*. 2005.
11. K.Fukunaga, *Introduction to Statistical Pattern Recognition*, New York: Academic, 1972.
12. M.Kirby and L.Sirovich, *Application of the karhunen-loeve procedure for the characterization of human faces*, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.12, pp.103-108, 1990.
13. D.P. Huttenlocher, G. A. Klanderman and W.J. Rucklidge. *Comparing images using the Hausdorff distance*, In IEEE Trans. PAMI, Vol. 15, No. 9, 1993 9.
14. W.J. Rucklidge. *Efficient Visual Recognition Using the Hausdorff Distance.* Springer-Verlag, 1996.
15. Mingli Song, Jiajun Bu, Chun Chen, *Expression Recognition from Video using A Coupled Hidden Markov Model*, IEEE TENCON'2004, 2004.