Photogenic Expression Recognition using Gabor Filters and Support Vector Machines*

Luana Bezerra Batista and Herman Martins Gomes
Universidade Federal de Campina Grande, Departamento de Sistemas e Computação
{luana, hmg}@dsc.ufcg.edu.br

Abstract

This paper presents a new view of the facial expression recognition problem, by addressing the question of whether or not is possible to classify previously labeled photogenic and non-photogenic face images, based on their appearance. In the proposed approach, a Support Vector Machine (SVM) is trained with Gabor representations of the face images to learn the relationships between facial expressions and the concept of a good photography of the face of a person. The obtained results showed that Gabor filters and SVM are promising, having presented good recognition rates.

1. Introduction

Facial Expression Recognition Systems (FERS) are generally applied to human-machine interfaces [7]. Such interfaces enable the automated provision of services that require a good appreciation of the emotional state of the user, as it would be the case in transactions that involve negotiation [2].

The main goal of this work is to give a new focus to the problem of facial expression recognition, by addressing the *photogeny* question. That is, to investigate the relationship between facial expressions of a person and the concept of a good photography of that person. Within this context, we propose an approach to classify facial expressions in two classes: photogenic and non-photogenic. This approach combines Gabor filters [4] together with Support Vector Machines [5].

This paper is organized as follows. Sections 2 and 3 briefly introduce, respectively, Support Vector Machines and Gabor filters. After that, in Section 4, the new approach is proposed. In Section 5, the performed experiments are presented. Finally, Section 6 presents the conclusions of this work.

2. Support Vector Machines

Support Vector Machines (SVM) is a kernel-based learning machine that has shown successful results in applications of various domains. It has been successfully used for pattern recognition, regression estimation, density estimation, novelty detection, and others [5].

Pattern recognition is usually accomplished through functions that divide the space of features into regions, separated by an optimal hyperplane, according to the number of classes. In the simplest case, an optimal separation hyperplane can be defined as the one with maximal margin of separation between a pair of classes. The margin can be seen as the sum of the distances of the hyperplane to pairs of opposite-class training points.

3. Gabor Filters

Gabor filters, also known as Gabor wavelets, are receiving great attention in the area of Facial Expression Recognition [6].

The two-dimensional Gabor function describes a sinusoid of frequency W modulated by a Gaussian. The Gabor functions form a complete, but non-orthogonal, basis set and any given function $f(x,y)$ can be decomposed in terms of these basis functions [4]. Such a decomposition results in a family of Gabor filters, making possible to detect features at various scales and orientations. The convolution of the image with a family of Gabor filters produces salient local features, such as eyes, nose and mouth, that are suitable for visual recognition.

4. Proposed Approach

The goal of the proposed approach is to train a classifier to learn the relationships between face expressions and the concept of a photogenic picture of a person. We used the common sense idea that when people are asked to pose for a picture, they usually make a neutral or happy face, rarely they use expressions such as anger, sadness etc.

Only the left side of the faces (see Figure 1), that is moved more extensively during facial expressions [1], was considered. Gabor filters were applied to these images in order to extract vertical and horizontal features.

After that, a SVM is trained with expressions that were labeled as photogenic and non-photogenic.

* This work was developed in collaboration with Hewlett-Packard.

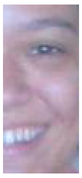

Figure 1. Area of the face used in the proposed approach

5. Experiments and Results

To perform the experiments, we selected a set of 324 images from the Cohn-Kanade Facial Expression Database (http://vasc.ri.cmu.edu/idb/html/face/facial_expression/). A total of 162 pictures were labeled as photogenic ones, corresponding to neutral and happiness expressions; whereas 162 pictures, corresponding to the others expressions, were labeled as non-photogenic.

After the left sides of the faces were cropped, they were resized to 50x25 pixels and transformed to 256 gray levels. Next, two Gabor filters, tuned to extract vertical and horizontal features, were applied to these images. A set of SVMs were trained with feature vectors (training patterns) of 2500 components each (50x25x2).

In the first experiment, the 10-fold cross-validation method was employed. The LIBSVM Software (http://www.csie.ntu.edu.tw/~cjlin/libsvm) was used in this phase. Table 1 shows the recognition rates using different types of SVM algorithms combined with different kernel functions. From this table, we can see that the best recognition rate is associated with the epsilon-SVR algorithm using a RBF kernel.

Table 1. Recognition Rates using Gabor Filters (%)

Algorithm/ Kernel	RBF	Poli-nomial	Linear	Sigmoid
C-SVC	79.93	73.76	81.79	76.85
nu-SVC	80.25	82.72	78.08	79.01
epsilon-SVR	86.32	83.66	81.98	85.01
nu-SVR	85.23	78.59	78.03	82.20

After that, the same experiment was performed using the Haar wavelet [3]. The recognition rates using Haar are presented in Table 2. From these results, it is possible to observe that, in general, the Gabor filters have a better performance.

Table 2. Recognition Rates using Haar Wavelets (%)

Algorithm/ Kernel	RBF	Poli-nomial	Linear	Sigmoid
C-SVC	59.26	50.00	79.94	41.97
nu-SVC	84.26	80.25	81.48	81.79
epsilon-SVR	83.10	54.45	85.74	77.77
nu-SVR	77.71	75.03	84.80	76.38

In the last experiment, the image database was separated in training (75%; 244 images) and test (25%; 80 images), so that the people contained into training set are not contained into test set. 75% of recognition rate was obtained with C-SVC algorithm using a *Polinomial Kernel*. Table 3 presents the confusion matrix for this experiment. The 11 images that caused confusion in the non-photogenic class (classified as photogenic) are related with the expressions anger and fear that have some visual resemblance to smiles. On the other hand, the images with other expressions considered as non-photogenic, as surprise and sadness, were correctly classified. The 9 images misclassified in the photogenic class resemble to non-photogenic expressions, as sadness or fear.

Table 3. Confusion Matrix

	Photogenic	Non-Photogenic
Photogenic	**31**	9
Non-Photogenic	11	**29**

6. Conclusions

In this paper, we presented a new approach that linked facial expressions with the concept of a photogenic picture of the face of a person.

Different types of SVM algorithms were combined with various kernel functions. Moreover, a comparison between Gabor filters with Haar wavelets was performed. The experiments have shown that Gabor filters and SVM are promising, having presented good recognition.

As future work, we intend to incorporate to the experiments images containing facial expressions of people with the eyes closed. Another future work is to create a custom-built larger image database to further examine the problem.

References

[1] Borod, J, Koff, E., Yecker, S., Santschi-Haywood, C. and Schmidt, J. Facial asymmetry during emotional expression: Gender, valence, and measurement technique. *Neuropsychologia*, vol. 36, no. 11, pp. 1209-1215, 1998.

[2] Chibelushi, C. and F. Bourel. "Facial Expression Recognition: A Brief Tutorial Overview". *On-Line Compendium of Computer Vision*, 2003.

[3] Gonzalez, R. and R. Woods. Digital Image Processing. Addison-Wesley Publishing Company, Inc, Massachusetts, 1992.

[4] Manjunath, B. and R. Chellappa. "A unified approach to boundary perception: Edges, textures, and illusory contours". *IEEE Transactions on Neural Networks*, vol. 4, no. 1, pp. 96-108, 1993.

[5] Vapnik, V. *The Nature of Statistical Learning Theory*. Springer-Verlag, New York, 2nd Ed., 1999.

[6] Tian, Y. "Evaluation of Face Resolution for Expression Analysis". *Workshop on Face Processing in Video*, 2004.

[7] Zue, V. and J. Glass. "Conversational Interfaces: Advances and Challenges". *IEEE*, vol. 88, no. 8, pp. 1166-1180, 2000.