Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics
San Antonio, TX, USA - October 2009

Facial-component-based Bag of Words and PHOG Descriptor for Facial Expression Recognition

Zisheng LI, Jun-ichi IMAI, and Masahide KANEKO
Department of Electronic Engineering
The University of Electro-Communications
Tokyo, Japan
lizisheng@radish.ee.uec.ac.jp

***Abstract*—A novel framework of facial appearance and shape information extraction for facial expression recognition is proposed. For appearance extraction, a facial-component-based bag of words method is presented. We segment face images into 4 component regions, and sub-divide them into 4×4 sub-regions. Dense SIFT (Scale-Invariant Feature Transform) features are calculated over the sub-regions and vector quantized into 4×4 sets of codeword distributions. For shape extraction, PHOG (Pyramid Histogram of Orientated Gradient) descriptors are computed on the 4 facial component regions to obtain the spatial distribution of edges. Our framework provides holistic characteristics for the local texture and shape features by enhancing the structure-based spatial information, and makes the local descriptors be possible to be used in facial expression recognition for the first time. The recognition rate achieved by the fusion of appearance and shape features at decision level using the Cohn-Kanade database is 96.33%, which outperforms the state of the arts.**

***Keywords*—facial expression recognition, bag of words, SIFT, PHOG, appearance extraction, shape extraction**

I. INTRODUCTION

Facial expression is one of the most powerful, natural and immediate means for human beings to communicate their emotions and intensions. Automatic facial expression recognition has many potential applications in areas such as human-computer interaction (HCI), emotion analysis, interactive video, indexing and retrieval of image and video databases, image understanding, and synthetic face animation. Many research efforts have been performed on facial expression analysis during the past two decades. The facial expressions under examination are defined by psychologists as a set of six basic facial expressions: anger, disgust, fear, happiness, sadness and surprise [1]. A survey on the research concerning facial expression recognition can be found in [2-3].

Deriving an effective facial representation from original face images is a vital step for successful facial expression recognition. There are two main approaches to extract facial features: appearance-based methods and shape-based methods [4]. For appearance information, the major works have focused on using Gabor wavelets to extract the facial appearance changes as a set of multi-scale and multi-orientation coefficients [5-7] on facial sub-regions or the whole face image. Accordingly, Boosted methods and subspace based methods have been applied to extract lower dimensional features from such texture information [5, 8]. Recently, Local Binary Pattern (LBP) features have been used for facial expression analysis [9-10] since the success of LBP in texture recognition [11]. For shape information, the facial motion and facial deformation are extracted by tracking a small set of predefined feature points [12-14] that are located usually around the permanent facial components. To measure the displacement between the current face image and the referenced neutral face, the positions of all tracked facial feature points should be normalized by mapping them into a standard face. When only texture or shape information is used, the recognition of facial expressions has certain drawbacks [15]. The fusion of both features has been proved to have better performance [4, 15].

In our work, a novel framework for extracting both appearance and shape information is proposed. A new method called facial-component-based bag of words is presented to extract facial appearance variations, and a facial-component-based PHOG [16] (Pyramid Histogram of Orientated Gradient) descriptor is proposed to represent facial shape features. Bag of words (BoW) methods [17-18], which represent an image as an orderless collection of local features, have recently demonstrated impressive levels of performance for scene categorization. However, the original bag of words has difficulties in facial expression recognition, because the object images belong to the same category (face images), histograms of orderless local features from face images with different facial expressions do not have large enough between-class variations. In this paper, a novel facial-component-based bag of words method is proposed to extract facial texture information for facial expression recognition. We firstly segment face images into 4 regions which contain different facial components, then equally divide each region into 4 sub-regions and calculate SIFT [19] (Scale-Invariant Feature Transform) descriptors on a sliding grid over each sub-region. Finally the SIFT features of each sub-region are vector quantized (VQ) into codewords to represent facial appearance features.

The shape information extraction is also implemented under the facial-component-based framework. Differing from the previous works, we use the spatial distribution of edges to represent facial shape features. We compute PHOG [16] descriptors over the 4 component regions of each face image respectively and concatenate the resulting histograms as shape representation. The PHOG descriptor is a pyramid

978-1-4244-2794-9/09/$25.00 ©2009 IEEE

representation of HOG [20] (histogram of orientated gradient) descriptor and does not need to track facial feature points. The facial component regions and the pyramid structure can provide enough geometric information for the local shape features.

We apply multi-class SVM classifiers to classify the six basic facial expressions using the facial-component-based bag of words and PHOG descriptors respectively. Then we fuse the appearance and shape information at decision level using different combination rules [21]. The recognition rate achieved using the Cohn-Kanade database [23] is 96.33%, which is better than the state-of-the-art research works [5, 9-10, 15, 22].

The main contributions of our study are:

- An extension of bag of words, facial-component-based bag of words method, is presented to extract facial appearance features for facial expression recognition. Structure-based spatial information is provided for the bag of words to maintain both holistic and local characteristics. It makes the highly distinctive local descriptors such as SIFT possible to be used in facial expression recognition for the first time.
- Facial-component-based PHOG descriptor is proposed to extract facial shape information. The method can represent local shape features using the spatial distribution of edges, and concatenate the features in a holistic way.
- The decision level fusion of the extracted appearance and shape features for facial expression recognition covers drawbacks of either feature, and achieves a promising result comparing with the state of the arts.
- The novel framework of facial feature extraction should also be suitable for facial action unit classification, face recognition and other facial analysis researches.

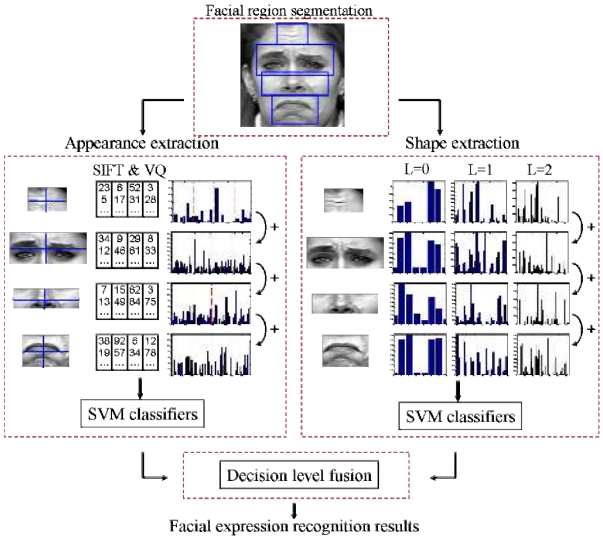


Figure 1. System architecture.

Section II describes the database used in our research, and Section III gives the system framework of our works. Section IV and V illustrate the proposed facial-component-based bag of words and PHOG descriptor respectively. Section VI and VII describe the experimental results and conclusions.

II. FACIAL EXPRESSION DATABASE

Our facial expression recognition system is trained and tested on the Cohn-Kanade database [23]. The database consists of 97 subjects. Each subject was instructed to display six basic facial expressions, each expression sequence started with neutral and ended with peak expressive frames. For our study, we selected 300 expressive frames from 90 subjects. The images were digitized into 640 by 480 pixel arrays with 8-bit precision for grayscale values.

III. SYSTEM FRAMEWORK

The proposed framework illustrated in Figure 1 consists of four sub-systems: facial region segmentation, appearance information extraction, shape information extraction and their fusion for final classification.

Face images are segmented into 4 regions which contain different facial components according to the detected positions of eyes and mouth [24]. The regions are with sizes 100×60, 200×80, 186×56 and 141×88 pixels. For appearance extraction, we equally sub-divide each region into 4 ROIs (Region of Interest), and calculate SIFT features on a sliding grid with spacing 2 pixels over each ROI. The SIFT features of each ROI from different images are vector quantized as a set of codewords respectively. As a result, appearance representation of a face image becomes 4×4 sets codewords of SIFT features. For shape extraction, PHOG descriptors are computed on the 4 facial component regions respectively. Histograms representing the spatial distribution of edges at different pyramid level are concatenated according to the geometric positions of the corresponding facial regions. Implementation details are given in Section IV and V respectively. The appearance and shape features are then input into multi-class SVM classifiers respectively, and fused at decision level finally.

IV. FACIAL-COMPONENT-BASED BAG OF WORDS FOR APPEARANCE EXTRACTION

Recently, the bag of words method has been successfully used in object recognition [17-18]. However, since the object category (face image) is the same in facial expression recognition, the orderless collection of local patches can not provide strongly distinctive information for different classes of expressions. In our work, we segment the face image into 4 component regions: forehead, eyes-eyebrows, nose and mouth regions. Moreover, each region is equally divided into 4 sub-regions to generate 4×4 ROIs for each face image. Over each ROI, we calculate 128-dimensional SIFT features on a regular grid with spacing 2 pixels. As a result, 4×4 sets of SIFT features are obtained for each face image. At training stage, we use *k*-means algorithm [25] to learn one codebook for each set of SIFT features respectively. The number of cluster K for the appearance features of each ROI is in proportion to the size of the ROI. Codewords are then defined as the centers of the learnt clusters. Finally, 4×4 codebooks are formed using the

training images. At testing stage, 4×4 sets of SIFT features of each image are converted into 4×4 histograms of codeword distribution using the trained codebooks, and these histograms are concatenated together to represent the facial appearance information. The spatial information of different facial component regions and ROIs provides the local appearance features for holistic characteristics.

Figure 2 shows the implementation of the appearance extraction for an eyes-eyebrows region. Four sets of SIFT features are converted into different codewords using the four trained codebooks and *k*-means algorithm. The horizontal axis of histograms (a)~(d) represents different codewords of the corresponding sub-regions, while the vertical axis represents codeword distributions. The four histograms are then concatenated into one, and finally concatenated with those of other facial component regions.

We compare different implementing approaches of bag of words for appearance extraction in facial expression recognition. Four methods are tested using the Cohn-Kanade database:

- Method A): The original bag of words [18]: We consider the whole faces as the ROIs, and crop the whole face region out of the images with size of 270 by 270 pixels. SIFT features are computed on a dense grid over the face ROIs, and the features are learned to form a codebook.
- Method B): Bag of words with sub-regions of face: Each face ROI in A) is equally divided into 4 sub-regions, and 4 codebooks are learned using the SIFT features from the 4 sub-regions respectively.
- Method C): Bag of words with facial component regions: The segmented facial component regions (without being divided into sub-regions) mentioned above are used as ROIs. Similarly, 4 codebooks are learned based on the dense SIFT features of the 4 facial component regions.
- Method D): The proposed method: 4×4 codebooks are formed using the SIFT features of the 4×4 ROIs.

The average recognition rates of these four approaches on the Cohn-Kanade database using a linear SVM classifier are shown in Figure 3, the total size of codebook is 268, the radius of the sliding grid is 20 pixels, and the sampling interval is 2 pixels. We can see that the proposed method has the best performance since the spatial information is maintained to the largest extent.

We test different radii of the sampling grid for computing SIFT features using the proposed method. The sampling interval is 2 pixels, total size of 4×4 codebooks is 268. We also test different sizes of codebooks, the radius of the dense grid is 20 pixels. Results are shown in Figure 4. We can see that when the radius is 20 pixels and the size of codebook is 268, the best performance is achieved. Detailed recognition results are illustrated in Section VI.

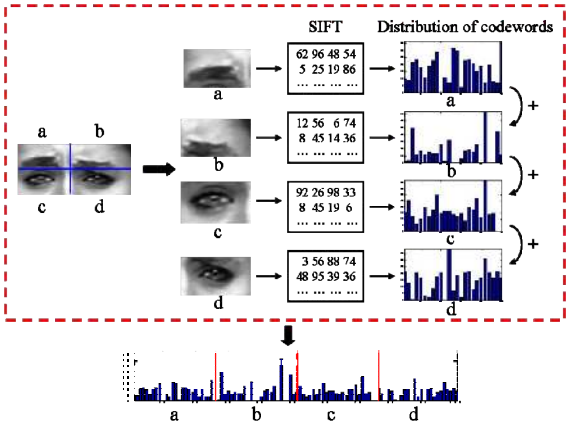


Figure 2. Appearance extraction of a facial component region.

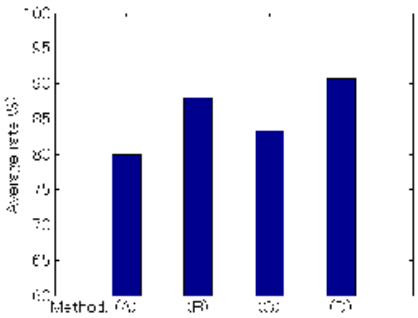
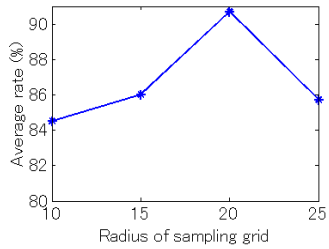
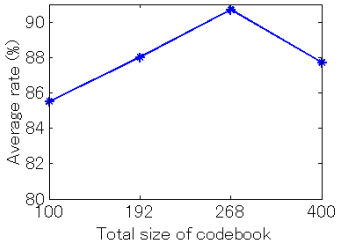


Figure 3. Comparison of different approaches of bag of words for appearance extraction.



(a) Different radii of sampling grid



(b) Different size of codebook

Figure 4. Different parameters of appearance extraction.

V. FACIAL-COMPONENT-BASED PHOG DESCRIPTOR FOR SHAPE EXTRACTION

In our work, we use the spatial distribution of local edges to represent facial shape information instead of tracking a set of predefined facial feature points. HOG (Histogram of orientated gradient) descriptor [20] counts occurrences of gradient orientation in localized portions of an image. Pyramid HOG (PHOG) descriptor [16] is a spatial pyramid representation of HOG descriptor, and achieved promising performance in object recognition. For facial expression recognition, we enhance the spatial information by computing PHOG features on the segmented facial component regions mentioned in Section III and IV.

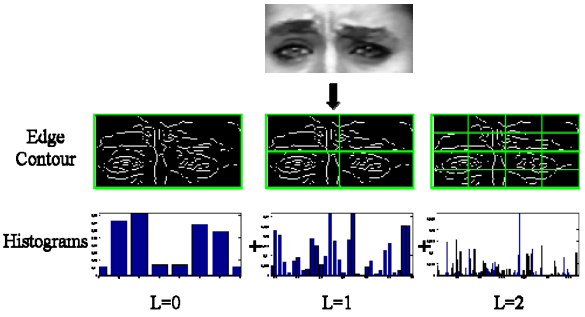
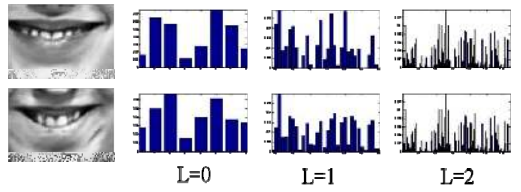
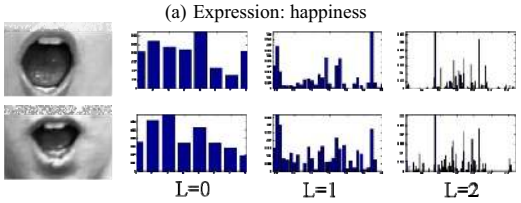


Figure 5. PHOG descriptor of a facial region.



(a) Expression: happiness



(b) Expression: surprise

Figure 6. PHOG features of mouth region of different expressions.

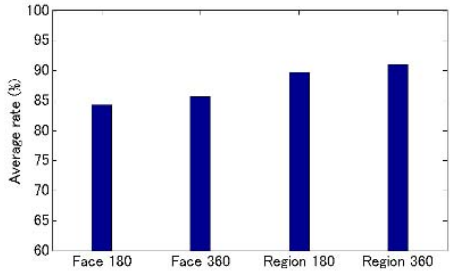


Figure 7. Comparison of different approaches for shape extraction.

As illustrated in Figure 5, edge contours are extracted using the Canny edge detector for each facial component region. Then each region is divided into a sequence of increasingly finer spatial grids by repeatedly doubling the number of divisions in each axis direction. The grid at resolution level $L = l$ has 2^l cells along each dimension. The orientation gradients are then computed using a 3×3 Sobel mask without Gaussian smoothing [20]. Histogram of edge orientations within each cell is quantized into N bins, and histograms of the same level are concatenated into one sequence. In Figure 5, horizontal axis represents different bins, and the vertical axis represents the numbers of orientations counts in corresponding bins. Here, two shape descriptors are used: one with orientations in the range [0-180] (where the contrast sign of the gradient is ignored) and the other with range [0-360] using all orientations [16]. PHOG descriptor shown in Figure 5 is calculated using $L = 2$ levels, $N = 8$ bins and range of [0-360]. The shape representation of a face image is the concatenations of PHOG features of the 4 segmented facial component regions. Figure 6 shows PHOG features of the mouth region of two different facial expressions. These PHOG descriptors are computed from two different subjects. We can see that shape descriptors of the same expression from different subjects have similarities and those of different classes from the same subject still have discriminative differences.

We test different implementations of shape extraction using PHOG descriptor: PHOG features on the whole face ROIs mentioned in Section IV with range [0-180] and [0-360], which refer to 'Face 180' and 'Face 360' in Figure 7 respectively; the proposed concatenations of PHOG descriptors of the 4 facial component regions with range [0-180] and [0-360], which refer to 'Region 180' and 'Region 360' respectively. The number of resolution level for the whole face ROI method is $L = 3$, while that of the proposed method is $L = 2$ to make the lengths of feature vectors the same. Bin number is $N = 8$ for both approaches. Results illustrated in Figure 7 show that the proposed facial-component-based PHOG method has better performance because the spatial information of local shapes is enhanced. Detailed facial recognition results of shape extraction are shown in Section VI.

VI. EXPERIMENTAL RESULTS

We use multi-class SVM classifiers with RBF kernels to classify the six basic facial expressions using both appearance and shape information on the Cohn-Kanade database which is described in Section II. The evaluation method is leave-one-subject-out. The average recognition rates and the confusion matrices have been computed to represent the accuracy of facial expression recognition. The confusion matrix is a $n \times n$ matrix ($n = 6$ in our case) containing information about the actual classification results (in its columns) and different category labels through the classification (in its rows). The diagonal entries of the confusion matrix are the percentage number of facial expressions that are correctly classified, while the off-diagonal entries correspond to misclassifications.

A. Recognition Results of Appearance Extraction

We concatenate the facial-component-based bag of words features with sampling grid radii $r = 15, 20, 25$ pixels as the

final appearance features. The sampling interval is 2 pixels and the total size of 4×4 codebooks is 268. The average recognition rate is 93.33% and the confusion matrix is shown in Table I.

B. *Recognition Results of Shape Extraction*

We concatenate the facial-component-based PHOG features of range [0-180] and [0-360] with the resolution level $L = 2$, and concatenate the PHOG features of the whole face ROIs of range [0-360] with level $L = 3$ as the final shape features. The bin number is $N = 8$ for all of the features. The average recognition rate is 94.33% and the confusion matrix is shown in Table II.

C. *Fusion of Appearance and Shape Information*

In order to achieve more robust and accurate results, we fuse the appearance and shape information at decision level by six combination rules [21]. Table III shows the average recognition rates of these rules. The best result is obtained by the sum rule, the average rate reaches 96.33%, which is better than using either appearance or shape information alone. The corresponding confusion matrix is shown in Table IV. By comparing with Table I and II, we can see that the appearance feature works better in anger and happiness while the shape feature outperforms in fear and sadness (the results of disgust and surprise are the same, 100%). Therefore, it can be seen that the fusion of appearance and shape can take advantages of both features while covers drawbacks of either one.

Figure 8 shows the comparison of the achieved recognition rate of every expression with the state of the arts [5, 9-10, 15, 22] using the same database (the Cohn-Kanade database), and Table V illustrates the comparison of average rates. It should be noted that the results are not directly comparable due to different protocols, preprocessing methods, and so on, but they still give an indication of the discriminative power of each approach. It can be seen that, except anger, our method works best in all of the classes, especially in disgust and fear. Our method also achieves the best average rate. Although the average rate difference between ours and [9]'s is imperceptible, our method clearly outperforms in most of the expression categories, and has better robustness. Our framework not only successfully takes advantages of the highly distinctive local descriptors such as SIFT and PHOG descriptors to capture the subtle changes of facial motions, but also enhances the holistic characteristics by the structure-based spatial information. The experimental results show that the proposed method outperforms the other recent methods.

As for the recognition of anger, Table IV shows that 13.33% of anger faces are misclassified as sadness. The reason

TABLE I. CONFUSION MATRIX OF APPEARANCE FEATURES

	Anger	Disgust	Fear	Happiness	Sadness	Surprise
Anger	83. 33%	0%	0%	0%	5. 56%	0%
Disgust	6. 67%	100%	2. 13%	0%	0%	0%
Fear	0%	0%	85. 11%	2. 86%	5. 56%	0%
Happiness	0%	0%	8. 51%	97. 14%	0%	0%
Sadness	10. 00%	0%	2. 13%	0%	88. 89%	0%
Surprise	0%	0%	2. 13%	0%	0%	100%

TABLE II. CONFUSION MATRIX OF SHAPE FEATURES

	Anger	Disgust	Fear	Happiness	Sadness	Surprise
Anger	70. 00%	0%	0%	0%	3. 70%	0%
Disgust	3. 33%	100%	2. 13%	0%	0%	0%
Fear	6. 67%	0%	93. 62%	4. 29%	0%	0%
Happiness	0%	0%	4. 25%	95. 71%	0%	0%
Sadness	20. 00%	0%	0%	0%	96. 70%	0%
Surprise	0%	0%	0%	0%	0%	100%

TABLE III. FUSION RESULTS OF DIFFERENT RULES

	Combination Rules					
	Sum	***Product***	***Max***	***Min***	***Median***	***Majority rote***
Rates	96. 33%	96. 00%	94. 67%	95. 33%	96. 00%	94. 00%

TABLE IV. CONFUSION MATRIX OF THE FUSION BY THE SUM RULE

	Anger	Disgust	Fear	Happiness	Sadness	Surprise
Anger	83. 33%	0%	0%	0%	3. 70%	0%
Disgust	3. 33 %	100%	2. 13%	0%	0%	0%
Fear	0%	0%	93. 62%	1. 43%	0%	0%
Happiness	0%	0%	4. 25%	98. 57%	0%	0%
Sadness	13. 33%	0%	0%	0%	96. 70%	0%
Surprise	0%	0%	0%	0%	0%	100%

TABLE V. COMPARISON WITH DIFFERENT METHODS

	Subject Num	**Sequence/ Frame Num**	**Class Num**	**Measure**	**Recognition Rate (%)**
[5]	90	313	7	leave-one-subject-out	93. 3
[9]	97	374	6	two-fold	95. 19
[9]	97	374	6	ten-fold	96. 26
[10]	96	320	7	ten-fold	91. 4
[10]	96	320	6	ten-fold	95. 1
[15]	97	374	7	five-fold	92. 3
[15]	97	374	6	five-fold	94. 5
[22]	90	284	6	-	93. 66
Ours	90	300	6	leave-one-subject-out	96. 33

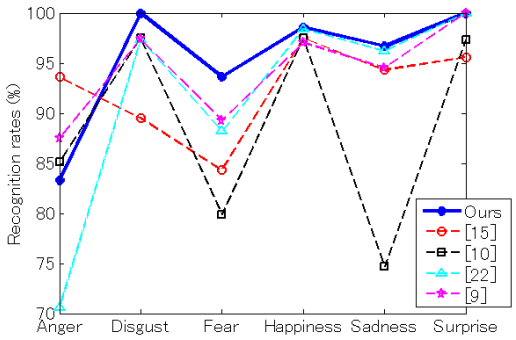


Figure 8. Comparison of recognition rates by different methods.

(a) Anger

(b) Sadness

Figure 9. Misclassified examples

can be seen from Figure 9, here, (a) gives some anger examples which are misclassified as sadness, while (b) shows the ground truth of sadness. It is difficult even for a human to recognize them accurately, which is also reported in [9].

VII. Conclusions and Future Works

In this paper, a novel framework of appearance and shape information extraction for facial expression recognition is proposed. Facial-component-based bag of words method is presented to extract local facial appearance changes while maintaining the holistic characteristics; similarly, facial-component-based PHOG descriptor is proposed to extract face local shape while enhancing the spatial information. Our method makes the bag of words methods and local descriptors be possible to be used in facial expression recognition for the first time. The decision level fusion of the extracted appearance and shape information achieved the average recognition rate as 96.33%, which outperforms the state-of-the-art research works.

The proposed method should also be valid in other facial image analysis works, such as face recognition, facial action unit classification and so on. In the future, we will focus on applying our method to dynamic expressive sequences and the recognition of facial actions.

References

[1] P. Ekman and W.V. Friesen, "Emotion in the Human Face," Prentice-Hall, New Jersey, 1975.

[2] M. Pantic and L.J.M. Rothkrantz, "Automatic analysis of facial expressions: the state of the art," IEEE Trans. Pattern Anal. Mach. Intell., vol. 22, no. 12, pp. 1424-1445, 2000.

[3] B. Fasel and J. Luettin, "Automatic facial expression analysis: a survey," Pattern Recognition, vol. 36, no. 1, pp. 259-275, 2003.

[4] Y. Tian, T. Kanade and J. Cohn, "Handbook of Face Recognition," Springer, 2005 (Chapter 11. Facial Expression Analysis).

[5] G. Littlewort, M.S. Bartlett, I. Fasel, J. Susskind, and J. Movellan, "Dynamics of facial expression extracted automatically from video," Trans. on Image and Vision Computing, vol. 24, pp. 615-625, 2006.

[6] Y. Tian, "Evaluation of face resolution for expression analysis," IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'04), pp. 82-88, 2004.

[7] Z. Zhang, M.J. Lyons, M. Schuster, and S. Akamatsu, "Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perception," IEEE Conference on Automatic Face & Gesture Recognition (FG), pp. 454-459, 1998.

[8] Z. Li, J. Imai, and M. Kaneko, "Comparisons of facial expression recognition in image sequences with and without speech," Procedings of the 13th Image Media Processing Symposium (IMPS2008), pp. 47-48, 2008.

[9] G. Zhao and M. Pietikainen, "Dynamic texture recognition using Local Binary Patterns with an application to facial expressions," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 29, no. 6, pp. 915-928, 2007.

[10] C. Shan, S. Gong, and P.W. McOwan, "Facial expression recognition based on Local Binary Patterns: A comprehensive study," Image and Vision Computing (2008), doi: 10.1016/j.imavis.2008.08.005.

[11] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971-987, 2002.

[12] I. Cohen, N. Sebe, A. Garg, L. Chen, and T. Huang, "Facial expression recognition from video sequences: Temporal and static modeling," Computer Vision and Image Understanding, vol. 91, no. 1-2, pp. 160-187, 2003.

[13] Y. Tian, T. Kanade, and J.F. Cohn, "Recognizing action units for facial expression analysis," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 23, no. 2, pp. 97-115, 2001.

[14] Y. Zhang and Q. Ji, "Active and dynamic information fusion for facial expression understanding from image sequences," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 27, no. 5, pp.699-714, 2005.

[15] I. Kotsia, S. Zafeiriou, and I. Pitas, "Texture and shape information fusion for facial expression and facial action unit recognition," Pattern Recognition, vol. 41, no. 3, pp. 833-851, 2008.

[16] A. Bosch, A. Zisserman, and X. Munoz, "Representing shape with a spatial pyramid kernel," Proc. of the 6th ACM International Conference on Image and Video Retrieval (CIVR'07), pp. 401-408, 2007.

[17] G. Csurka, C.R. Dance, L.Fan, and C. Bray, "Visual categorization with bags of keypoints," Proc. of European Comference on Computer Vision, pp. 1-22, 2004.

[18] F. Li and P. Perona, "A bayesian hierarchical model for learning natural scene categories," Proc. of IEEE Conference on Computer Vision and Pattern Recgnition, vol. 2, pp. 524-531, 2005.

[19] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, 2004.

[20] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," Proc. of IEEE Conference on Computer Vision and Pattern Recgnition (CVPR'05), pp. 886-893, 2005.

[21] J. Kittler, M. Hatef, R.P.W. Duin, and J. Matas, "On Combining Classifiers," IEEE Trans. on Pattern Analysis and Machine Intellegence, vol. 20, no. 3, pp. 226-239, 1998.

[22] P.S. Aleksic and A.K. Katsaggelos, "Automatic facial expression recognition using facial animation parameters and multistream HMMs," IEEE Trans. on Information Forensics and Security, vol. 1, no. 1, pp. 3-11, 2006.

[23] T. Kanade, J. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," Proc. of IEEE International Conference on Face and Gesture Recognition (FG), pp. 46-53, 2000.

[24] D. Vukadinovic and M. Pantic, "Fully automatic facial feature point detection using Gabor feature based Boosted classifiers," Proc. of IEEE International Conference on Systems, Man and Cybernetics, vol. 2, pp. 1692-1697, 2005.

[25] O. Duda, P.E. Hart, and D. G. Stork, "Pattern classification," John Wiley & Sons, 2000.