Facial Expression Recognition Based on Facial Region Segmentation and Modal Value Approach

Gibran BENITEZ-GARCIA[†,††], Gabriel SANCHEZ-PEREZ[††], *Nonmembers*, Hector PEREZ-MEANA[††], Keita TAKAHASHI[†], *Members*, *and* Masahide KANEKO[†a)], *Fellow*

SUMMARY This paper presents a facial expression recognition algorithm based on segmentation of a face image into four facial regions (eyes-eyebrows, forehead, mouth and nose). In order to unify the different results obtained from facial region combinations, a modal value approach that employs the most frequent decision of the classifiers is proposed. The robustness of the algorithm is also evaluated under partial occlusion, using four different types of occlusion (half left/right, eyes and mouth occlusion). The proposed method employs sub-block eigenphases algorithm that uses the phase spectrum and principal component analysis (PCA) for feature vector estimation which is fed to a support vector machine (SVM) for classification. Experimental results show that using modal value approach improves the average recognition rate achieving more than 90% and the performance can be kept high even in the case of partial occlusion by excluding occluded parts in the feature extraction process.
key words: *facial expression recognition, partial occlusion, facial segmentation, modal value*

1. Introduction

Since the appearance of the first computers and robots, one of the main targets has been to attain a complex human-computer interface [1], [2]. According to reports in the field of human communication, facial expressions and other non-verbal information are the most important elements to show human emotions [3], [4]. For that reason, in the past two decades many research efforts have been performed regarding facial expression recognition (FER) [5]–[12].

Psychologists have defined the human facial expressions as a set of six basic facial expressions: anger, disgust, fear, happiness, sadness and surprise, plus the neutral state [4]. Based on this division of facial expressions many systems of FER use this set of basic expressions. There exist many applications such as virtual reality, smart environments, video-conferencing, user profiling and customer satisfaction studies that require efficient methods of FER. However many problems concern this topic such as illumination changes, pose, angle of the input camera, partial occlusion, and so on [5], [6].

Especially, most of the existing approaches on FER just use non occluded facial images taken under controlled lab-oratory conditions. Nevertheless, this is not the case in real life, because one of the most frequently encountered problems in FER is partial occlusion [7], [8].

Partial occlusion can be seen as noise that disturbs facial expression feature extraction or it will cause information loss in FER. There are two types of partial occlusion: temporary and systematic; temporary occlusion is when a part of the face is obscured momentarily as a result of a person moving or an external factor, systematic occlusion results when the person is wearing something which covers part of his face [9]. Therefore to develop a FER algorithm, robustness under occlusion conditions has become an important research topic [10]–[12].

This paper proposes a robust FER algorithm which takes into account the problem of partial occlusion. The proposal is based on the segmentation of a face image into several regions using the sub-block eigenphases in each of them, instead of working with a whole image. It uses, specifically, the facial regions that are not occluded. In addition, the use of facial region segmentation also helps to improve the performance of FER for non-occluded faces because it allows the possibility to get several decisions from one facial image. Therefore, this paper also proposes a method to unify these results, using the most frequent decision of the classifiers (modal value approach).

The proposed algorithm was evaluated using leave-one-subject-out method with 300 frames of the Cohn-Kanade database [21] that includes face images of 97 subjects, where each one was instructed to display the six basic facial expressions. In order to evaluate the effectiveness of the proposed method for facial expression recognition under partial occlusion four types of occlusion were adopted: half left, half right, mouth and eyes occlusion. The performance of the proposed method was tested with non-occluded faces as well as partially occluded faces and compared with two recent approaches which use sub-block eigenphases [17] and linear binary pattern (LBP) [18] respectively.

The rest of the paper is organized as follows. Section 2 describes the system framework. Section 3 describes the modal value approach in detail. Section 4 shows the specifications of the database as well as the partial occlusion database. Section 5 presents the experimental results. Finally, discussion and conclusions are drawn in Sect. 6.

2. System Framework

The block diagram of the proposed system is shown in Fig. 1. In the stage of facial region segmentation, the face images are segmented into 4 facial regions: eyes-eyebrows, forehead, mouth and nose. The feature extraction stage is based on the sub-block eigenphases algorithm, which is applied independently to each facial region. From each facial region, the phase spectrum is firstly obtained, and then, principal component analysis (PCA) is applied in order to estimate a feature vector. Finally, all the feature vectors obtained from the different facial regions are concatenated to yield the final feature vector that represents the entire face. The number of concatenated features, N, depends on the number of the facial regions that are used in the process. In the classification stage the SVM is applied so as to make the recognition of facial expressions, using the multi-class mode specifically employing 6 classes, one for each expression (anger, disgust, fear, happiness, sadness, and surprise). Afterwards based on the decision from different classifiers, the modal value approach is applied.

The baseline algorithm of sub-block eigenphases method has already been proposed in [17], but the original algorithm works with the whole facial image. The proposed method in the present paper handles each of the facial regions individually and the main focus is how to combine the available facial regions to overcome partial occlusions as well as to achieve better FER performance for non-occluded faces.

2.1 Facial Region Segmentation

Face images are segmented into 4 regions that contain information of eyes-eyebrows, forehead, mouth and nose. This segmentation enables us not only to exclude some facial parts in the case of partial occlusion, but also to evaluate the contribution that each facial region has on the FER, which results in a robust modal value approach.

First of all a robust algorithm of face detection proposed by Vukadinovic et al. [13] is used to detect the eyes position of a face image. Then based on the distance between the irises (ED) and accordingly with the relation proposed in [14] the mouth region is defined. This relation proposes that the top of the mouth region is 0.85ED and the
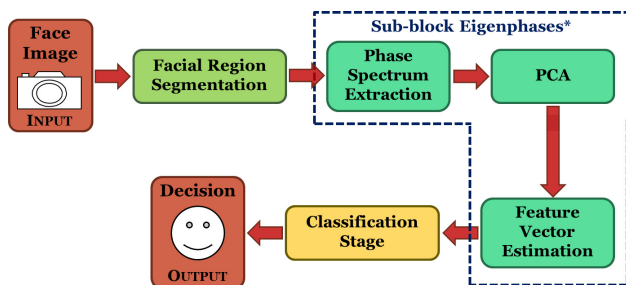
bottom 1.5ED from the irises position. This approach is shown in Fig. 2(a). Based on this issue, the bottom of the nose region is 0.85ED and the top 0.35ED, the bottom of the eyes-eyebrows region is 0.35ED and the top -0.4ED, finally the bottom of the forehead region is -0.4ED and the top -ED, where "-" means the distance in the upward direction from irises position.

Subsequently the mouth, eyes-eyebrows, nose and forehead regions are segmented which have size of 1.2ED(width)x0.65ED(height), 2EDx0.75ED, 1.7EDx0.5ED and EDx0.6ED respectively as shown in Fig. 2(b), (c), (d) and (e). In this paper we assumed that the segmentation described above is correctly achieved.

2.2 Phase Spectrum Extraction

Each of the face segments is further divided into sub-blocks of 2x2 pixels, and the phase spectrum is extracted for each facial region by using fast Fourier transform. The phase spectrum is employed in the eigenphases algorithm [15], because the Oppenheim's study [16] proved that the most important information of an image is contained in the phase instead of its magnitude.

2.3 PCA

The phase spectrum matrix is converted into a column vector, and subsequently the column vectors of all training images will form a matrix in order to apply the PCA algorithm to finally get the principal matrix (PM), which consists of the principal component vectors. It is important to notice that in contrast to the original sub-block eigenphases algorithm [17] which yields only one principal matrix from one face image, we calculate 4 principal matrices from one face image, one matrix from each of the 4 facial regions, respectively.

2.4 Feature Vector Estimation

The feature vectors are the product of the principal matrix by each column vector of the training images. Similarly to the previous stage this process is applied to each facial region independently. The final step is to concatenate the individual feature vectors in order to create the final feature vector that represents the entire face. As shown in Fig. 3, (a), (b), (c) and (d) represent the process to get the individual feature vector related to eyes-eyebrows, forehead, mouth and nose
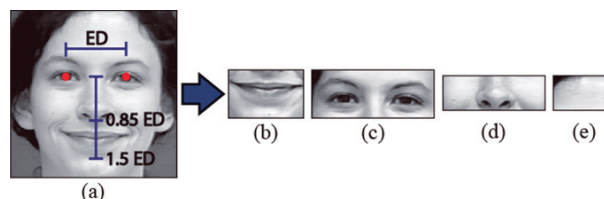

Fig. 1 Block diagram of proposed system.

Fig. 2 Example of the facial region segmentation.
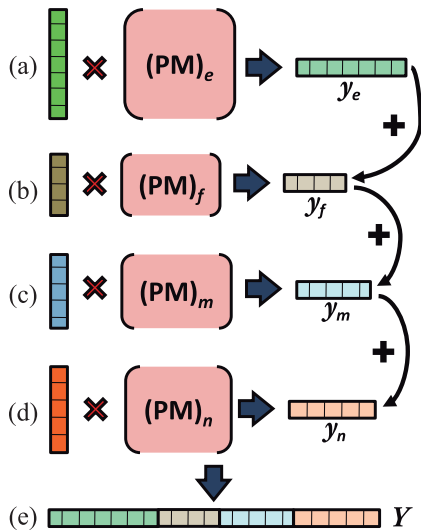
Fig. 3 Example of feature vector estimation.

Fig. 4 Diagram of modal value approach.

Table 1 Example of modal value approach.

Sample	1	2	3	4	Res.
Exp.	Anger	Anger	Fear	Fear	–
Frame					–
SVM1	Anger	Anger	Fear	Fear	3/4
SVM2	Anger	Anger	Fear	Disg	3/4
SVM3	Fear	Anger	Happy	Fear	2/4
Mod	**Anger**	**Anger**	**Fear**	**Fear**	**4/4**

Table 2 Special case when two or more classes have the same decision.

Expression	Anger		
Frame			
SVM1	Anger	0.82	**mean**
SVM2	Anger	0.89	**0.855**
SVM3	Fear	0.75	mean
SVM4	Fear	0.80	0.775
Final Decision	**Anger**		

respectively, finally (e) represents the concatenation of the 4 individual feature vectors which results in the final feature vector. The final feature vector depends on the number of facial regions used; up to 4 feature vectors from one facial image are concatenated, although in the case of occlusions, the number could be less than 4. To show the contribution of each facial region, all possible combinations of the 4 facial regions are presented in the experimental results section.

2.5 Classification Stage

For the classification of the six basic expressions, we use LIBSVM [20], which provides an implementation of multi-class support vector machine (SVM) [19], employing RBF kernels. The SVM output values provide similarity scores for each class whose range is 0.0 and 1.0, denoting 1.0 the exact match. The classes correspond to 6 basic expressions, which are anger, disgust, fear, happiness, sadness, and surprise.

3. Modal Value Approach

Our feature extraction method mentioned in Sect. 2.4 is capable of producing several feature representations from one facial image, because several combinations are possible by concatenating up to 4 facial regions. For each feature representation, we train a SVM individually, so that we have several independent decisions from one face image. We propose to unify these decisions using modal value approach, which helps to improve the recognition performance.

Modal value approach consists in selecting as a final decision the most frequently (modal value) class gotten from the classifiers, each of which is associated with the final feature vector representation that depends on the combinations of facial regions.

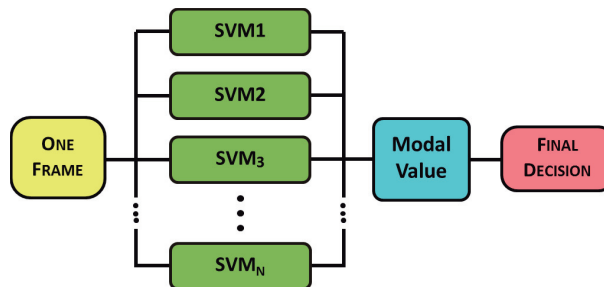Figure 4 shows the procedure to apply the modal value approach, as we can see more than one SVM are applied to one frame of facial images, where N depends on the number of classifiers used in the approach (it depends on the combinations of facial regions used in feature vector estimation). The modal value is selected from the decision of these classifiers in order to take the final decision. It is important to mention that in order to apply this approach at least 3 classifiers are necessary, because with only two decisions it is not possible to calculate correctly the modal value.

Table 1 shows an example of modal value approach using 4 sample faces. In this example two frames are displaying the expression of anger and the other two of fear. None of the 3 classifiers used here, SVM1, SVM2, and SVM3 produced the perfect results; the recognition accuracies are 3/4, 3/4 and 2/4 respectively. However, unifying the decisions from these classifiers by taking the modal values leads to 100% of accuracy, as shown in the bottom row of the Table 1.

As shown in Table 2, when two or more classes have the same number of positive decisions, the output values of such classifiers are averaged and taking as the final decision the result with the highest numerical value. In this example the decision of the classifiers SVM1 and SVM2 was anger while the decision of SVM3 and SVM4 was fear. How-

ever taking into account the average among them the highest value was provided by the average among SVM1 and SVM2 therefore the final decision was anger.

4. Database

For this paper 300 peak expressive frames of the Cohn-Kanade database [21] were used. The database contains face images of 97 subjects ranging in age from 18 to 30 years old, 65 percent were female and 45 male. The images were taken under a controlled environment and digitized into 640x480 pixels in grayscale values. For the experiments the face part was cropped with 280x280 pixels. For the regions of eyes-eyebrows, forehead, mouth and nose their sizes were normalized at 200x80, 100x60, 140x80 and 175x50 pixels respectively. Table 3 shows the number of frames for each of the 6 expressions.

In order to evaluate the effectiveness of the proposed method under partial occlusion, four different types of occlusion were adopted, as mentioned below.

4.1 Simulation of Partial Occlusion

There is not public available facial expression database that contains different types, position or size of partial occlusions. Therefore, four different types of partial occlusion were simulated in this work: occluded half left, occluded half right, occluded eyes and occluded mouth. The motivation for applying partial occlusions on these regions comes from real situations of daily life. For example, the use of sunglasses, scarves, medical masks, some hair styles and shadows. Figure 5 shows the four different occlusions applied to one subject who displays the 6 basic expressions. These samples are generated by superimposing graphically black mask regions on the non-occluded 300 frames selected from Cohn-Kanade database.

It is worth noting that the occlusions introduced in the face images are more critical than real life occlusions. For example in the left and right side occlusions, which emulate occlusion due to hair styles and shadows, half of the face is completely occluded with a black mask which does not happen in the situations mentioned above. Also to emulate the use of sunglasses, a black mask completely occludes the eyes-eyebrows part of the face which is larger than the real life occlusion. Finally the mouth occlusion used is similar to the occlusion produced by scarves or medical masks. These kinds of occlusions, shown in Fig. 5, which are more critical than real life ones were used because several real life occlusions due to sunglasses and shadows are efficiently solved by sub-block eigenphases algorithm [17]. Therefore because the occlusion used in this paper are more critical than real life ones, we can expect that if the proposed algorithm

performs well with these it will be able to perform fairly well with real life occlusions.

4.2 Handling Half Face Occlusion

For the cases of half left/right face occlusions, the proposed method is not applicable directly because facial regions used in the method depend on both sides of the face. Therefore, we generate a mirror image based on the half side that is not occluded in order to work with a whole face instead of a half. Then the facial region segmentation is applied to mirror images as described above. In this way we intended to solve the half face occlusion problem to increase the recognition rate. Figure 6 shows two subjects of the database with half left/right occlusion and its consequent mirror images.

Fig. 5 Example of partial occlusion simulation of database. From top to bottom: no occlusion, half left, half right, eyes and mouth occlusion.
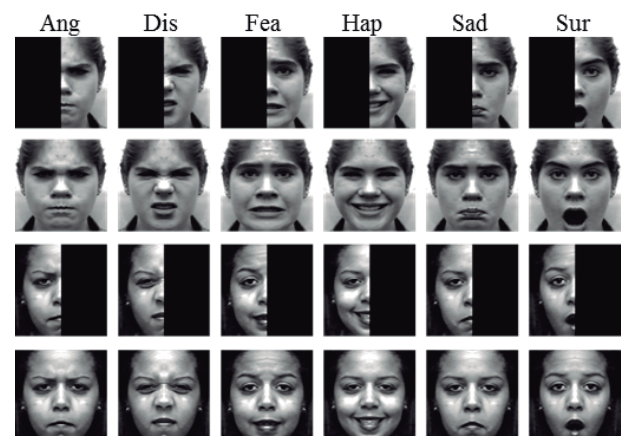

Fig. 6 Example of half occluded and mirror images. From top to bottom: half left occlusion, mirror right, half right occlusion, and mirror left.
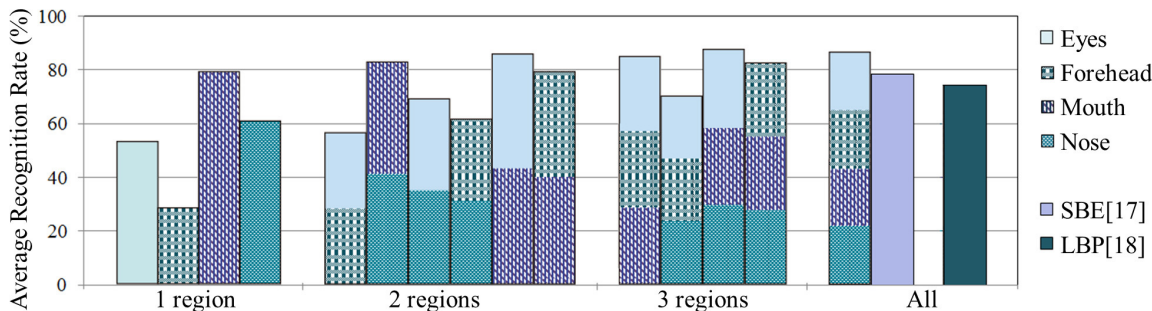
Table 3 Frame numbers of each expression used in the experiments.

Ange	Disg	Fear	Happ	Sad	Surp
30	34	47	70	54	65

Fig. 7 Average recognition rate of all possible combinations using different number of facial regions.

5. Experimental Results

The recognition accuracy was measured using leave-one-subject-out. For all experiments the training was performed with non-occluded database. The average recognition rates and the confusion matrices have been computed to represent the accuracy of facial expression recognition. The confusion matrix describes percentages of the predicted expressions in its columns against the actual expressions in its rows. The diagonal entries are the percentage of facial expressions that are correctly classified, while the off-diagonal entries correspond to misclassifications.

5.1 Experiments with Different Number of Facial Regions

Since our method is based on facial region segmentation into 4 regions and the combination of the features individually obtained from those regions, all possible combinations of the 4 facial regions had been analyzed. Our method was also compared with two basic methods for reference that use the whole face image at once to extract a feature vector, without facial region segmentation [17], [18]. For the case of reference methods, the whole image is used for calculating one principal matrix (PM), and that matrix was used to obtain a feature vector that describes the entire face. In this case, recognition was performed with only one classifier, since there are not multiple facial regions.

Figure 7 shows the average recognition rates of the combinations divided in one, two, three, and all facial regions, compared with the result of reference methods, subblock eigenphases (SBE) [17] and LBP method (LBP) [18]. The best result is achieved using 3 regions, by Eyes-Mouth-Nose (EMN) with 87.7% which is better than employing the 4 facial regions (All) with 86.7%. EMN case also outperformes by about 9% the recognition rate obtained using the SBE (with 78.3%) and by about 13% the result provided by the LBP (with 74.3%). Moreover we can notice that Mouth (M) with 79.3% and Eyes-Mouth (EM) with 86% are the best result achieved by using 1 and 2 regions respectively. Another important point that we can see from Figure 7 is that the combinations which use mouth region provide the highest average recognition rate independently of the number of facial regions used in the process. Meanwhile, when

Table 4 Average recognition rate of the best results using modal value approach.

No. SVMs	Combinations	Result (%)
4	EM - FM - All - SBE	92.00
4	M - FM - EMN - All	91.67
6	M - N - EM - FM - EMN - All	90.00
4	M - EM - EMN - All	89.33
3	FM - EMN - SBE	88.00

Table 5 Confusion matrix of the best result.

	Ang	Disg	Fear	Happ	Sad	Surp
Ang	**83.3**	0.0	3.3	3.3	10.0	0.0
Disg	0.0	**88.2**	5.9	2.9	0.0	2.9
Fear	0.0	0.0	**76.6**	17.0	6.4	0.0
Happ	0.0	0.0	2.9	**97.1**	0.0	0.0
Sad	0.0	0.0	0.0	1.9	**98.1**	0.0
Surp	0.0	0.0	1.5	0.0	0.0	**98.5**

the mouth region is not employed, the combinations of other regions do not provide competing recognition rates.

Next, all possible combinations using modal value approach described in Sect. 3 were tested. In this approach, several SVMs associated with different combinations of facial regions were executed in parallel, obtaining a recognition result from each classifier, and taking the modal value to unify the recognition results. The number of classifiers used in this test was 16, 15 from the combinations of 4 subregions and one from the whole image (which is SBE). It is important to mention that at least 3 classifiers are required to find the modal value.

Average recognition rates from the best results using modal value approach are shown in Table 4. It is possible to see that the best result obtained by modal value approach achieves 92% of average recognition rate using 4 classifiers, the combinations used in this case were: Eyes-Mouth (EM), Forehead-Mouth (FM), the four regions (All) and without region segmentation (SBE). This result provides the highest recognition rate obtained in this paper, outperforming by almost 15% the average recognition rate of the SBE and by almost 19% when LBP is used. On the other hand around 5% is the improvement compared with the best result obtained when only one classifier is used (EMN).

Table 5 shows the confusion matrix of the best result of our system (EM-FM-All-SBE). We can see that for the pro-

posed method, surprise, sadness, happiness and disgust expressions are easy to recognize, while anger and fear expressions are not. Also, the average recognition rate of surprise is the highest and fear is the lowest with 98.5% and 76.6% respectively. The problem to recognize fear is related with happiness, because in 17% of the cases our system misrecognizes fear expression as happiness.

5.2 Experiments Despite Partial Occlusion

The proposed algorithm was evaluated using four types of partial occlusion: occluded half left, occluded half right, occluded eyes-eyebrows and occluded mouth, which are described in Sect. 4.1 and shown in Fig. 5

To solve the above mentioned occlusions, in addition to the 4 regions shown in Fig. 2, 4 additional regions were included in order to have more possible combinations when modal value approach is used, such as: left eye, right eye, half left face and half right face. Here left/right eye is the half part of eyes-eyebrows region shown in Fig. 2(c) and half left/right face is the half part of the whole image shown in Fig. 2(a). To determine the contribution of these additional regions to the facial expression recognition each of them is used independently and the evaluation results are shown in Table 6.

For each type of partial occlusion this paper proposes a different solution. For the specific case of occluded half face a solution using mirror images is employed in which the reconstructed image is segmented into 8 regions (E, F, M, N, LE, RE, LF and RF) described above, which are used together with the whole mirror image in the modal value approach. Otherwise the problem of eyes-eyebrows occlusion is overcome using the facial regions which are not occluded: forehead, mouth and nose. The same solution is employed for mouth occlusion using only eyes-eyebrows, left eye, right eye, forehead and nose regions.

The recognition rates of the best results obtained from

Table 6   Average recognition rate of additional regions.

Region	Result (%)
Left Eye (LE)	61.00
Right Eye (RE)	50.33
Half Left Face (LF)	75.00
Half Right Face (RF)	79.33

each type of partial occlusion are shown in Table 7. Note that hyphens "-" indicates that the modal value approach was adopted to unify the outputs of multiple classifiers. Therefore we can see that the best solutions for the partial occlusion problems were obtained using 4 classifiers.

In order to compare the results of the proposed system despite occlusion, recognition was performed without mirroring nor facial region segmentation, using only whole image (SBE). The comparison is shown in Fig. 8, where none occluded recognition serving as baseline is also presented. In all cases the proposed method improves the results of the approach using whole image. The maximum improvement is obtained for half left occlusion, thus proposed method using modal value approach improves around 55% the average recognition rate of SBE approach that provides only 33%. Moreover it is possible to see that the results among half left/right occlusion provide almost the same recognition rate, and the recognition rate with occluded eyes is higher than the result when the mouth is occluded.

Figure 9 shows, by each facial expression, the recognition performance despite the four types of partial occlusion employing the proposed method. The effect of partial occlusion differs for different expressions. For example, the

Table 7   Average recognition rate of the best results from each type of partial occlusion.

Occluded Region	Best Solution	Result (%)
Half Left	EMN - All - LF - RF	87.00
Half Right	EMN - All - LF - RF	83.33
Eyes-Eyebrows	M - FM - MN - FMN	87.67
Mouth	N - EN - FN - EFN	75.33

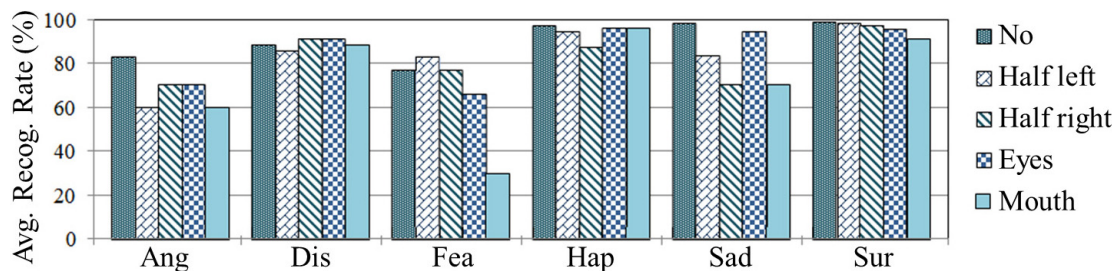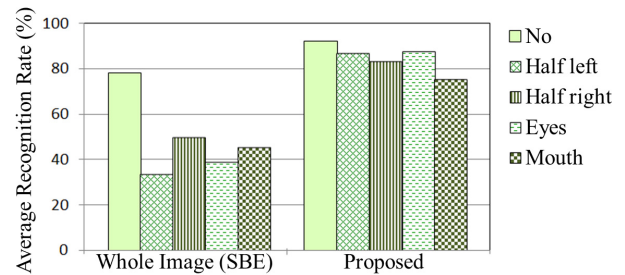**Fig. 8**   Comparison between approach using whole image and proposed methods despite partial occlusion.

Fig. 9   Effect of four types of partial occlusion on the recognition performance by each facial expression.

performance of the system for fear degrades when mouth is occluded while for happiness the system performs fairly well with any kind of occlusion.

6. Discussion and Conclusion

This paper presented a facial expression recognition algorithm based on face image segmentation into four facial regions. Several combinations of facial regions are possible in this approach, resulting in different classifiers corresponding to the combination. We also proposed a modal value approach in order to unify the results obtained from different classifiers. Based on the experimental results we can conclude that the use of facial region segmentation improves the average recognition rate compared to the approaches using whole image all together (SBE and LBP). The best result obtained in this paper was provided by EM-FM-All-SBE combination in the modal value approach that achieves 92% of average recognition rate. Also in this paper we can conclude that the mouth is the most important part of the face for developing facial expression recognition.

Another advantage of the proposed method is that even with only one part of the face it is possible to make the recognition, achieving almost 80% of average recognition rate if the mouth region is available. This fact becomes very important when several regions of the face are invisible in the case of partial occlusion. If the left or right half of the face is occluded, we can employ mirroring of the non-occluded part, and improve the recognition result using facial region segmentation too.

The future work of this study will consider basically two lines. One is to increase the recognition rate, where we will focus on applying another way to use the modal value approach, specifically employing probability values of each class provided from the SVM process. The other is to automate the process applied for occluded images, including automatic detection of occluded regions.

Acknowledgements

References

[1] Y. Tian, T. Kanade, and J.F. Cohnr, Facial Expression Recognition, Handbook of Face Recognition, pp.487–519, 2011.

[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.13, pp.1424–1445, 2000.

[3] A. Mehrabian, Communication Without Words, Psychol. Today, vol.2, no.4, pp.53–56, 1968.

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

[5] Y. Tian, T. Kanade, and J.F. Cohn, Facial Expression Analysis, Handbook of Face Recognition, eds. Stan Z. Li and Anil K. Jain, Springer-Verlag, Dec. 2004.

[6] B. Fasel and J. Luettin, "Automatic facial expression analysis: A survey," Pattern Recognit., vol.36, no.1, pp.259–275, 2003.

[7] W. Gu, C. Xiang, Y.V. Venkatesh, D. Huang, and H. Lin, "Facial expression recognition using radial encoding of local Gabor features and classifier synthesis," Pattern Recognit., vol.45, no.1, pp.80–91, Jan. 2012.

[8] G. Song and R. Qiuqi, "Facial expression recognition using local binary covariance matrices," 4th IET International Conference on Wireless, Mobile & Multimedia Networks (ICWMMN 2011), IET, 2011.

[9] T. Howard and M. Slater, "Reconstruction and recognition of occluded facial expressions using PCA," Affective Computing and Intelligent Interaction 2007, pp.36–47, 2007.

[10] I. Kotsia, I. Buciu, and I. Pitas, "An analysis of facial expression recognition under partial face image occlusion," Image Visi. Comput., vol.26, no.7, pp.1052–1067, 2008.

[11] L. Zhang, D. Tjondronegoro, and V. Chandran, "Toward a more robust facial expression recognition in occluded images using randomly sampled Gabor based templates," IEEE International Conference on Multimedia and Expo (ICME) 2011, pp.1–6, July 2011.

[12] Y. Miyakoshi and S. Kato, "Facial emotion detection considering partial occlusion of face using Bayesian network," 2011 IEEE Symposium on Computers & Informatics (ISCI), pp.96–101, 2011.

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

[14] Z. Li, J. Imai, and M. Kaneko, "Facial expression recognition using facial-component-based bag of words and PHOG descriptors," J. ITE (The Institute of Image Information and Television Engineers), vol.64, no.2, pp.230–236, Feb. 2010.

[15] M. Savvides, B.V.K. Vijaya Kumar, and P.K. Khosla, "Eigenphases vs. Eigenfaces," Proc. 17th International Conference on Pattern Recognition, vol.3, pp.810–813, Aug. 2004.

[16] A.V. Oppenheim and J.S. Lim, "The importance of phase in signals," Proc. IEEE, vol.69, pp.529–541, May 1981.

[17] G. Benitez-Garcia, J. Olivares-Mercado, G. Sanchez-Perez, M. Nakano-Miyatake, and H. Perez-Meana, "A sub-block-based eigenphases algorithm with optimum sub-block size," Knowl.-Based Syst., vol.37, pp.415–426, Jan. 2013.

[18] Y. Luo, C. Wu, and Y. Zhang, "Facial expression recognition based on fusion feature of PCA and LBP with SVM," Optik-International Journal for Light and Electron Optics, vol.124, pp.2767–2770, Sept. 2013.

[19] V. Vapnik, Statistical learning Theory, John Wiley & Sons, New York, 1998.

[20] C.-C. Chang and C.-J. Lin, "LIBSVM: a library for support vector machines," ACM Trans. Intelligent Systems and Technology, 2:27:1–27:27, 2011. Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm

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

Gibran Benitez-Garcia received the B.S. degree on Computer Science Engineer from National Polytechnic Institute, Mexico City in 2011. He is currently a Master of Science student at the Mechanical Engineering School of the National Polytechnic Institute of Mexico. He stayed in the University of Electro-Communications from April 2012 to March 2013 as a JUSST student. His research interests include image processing, facial information analysis and pattern recognition.

Gabriel Sanchez-Perez received the B.S. degree on Computer Science Engineer; and the Ph.D. degree on Electronic and Communications, in 1999 and 2005, respectively, from the National Polytechnic Institute, Mexico City. From January 2001 to October 2006 he joined the Computer Engineering Department, Electrical and Mechanical Engineering School at the National Polytechnic Institute as Assistant Professor. In October 2006 he joined the Graduate School of the National Polytechnic Institute of Mexico where he is now a professor. He is a member of the IEEE.

Hector Perez-Meana received the B.S. degree in Electronics Engineers from the Metropolitan Autonomous University (UAM), Mexico City in 1981, the M.S. degree from the University of Electro-Communications, Tokyo Japan in March 1986, and a Ph.D. degree in Electrical Engineering from Tokyo Institute of Technology,Tokyo, Japan, in 1989. In 1981 he joined the Electrical Engineering Department of the Metropolitan University where he was a Professor. From March 1989 to September 1991, he was a visiting researcher at Fujitsu Laboratories Ltd, Kawasaki, Japan. In February 1997, he joined the Graduate Department of the Mechanical and Electrical Engineering School on the National Polytechnic Institute of Mexico, where he is now a Professor. In 1991 Prof. Perez-Meana received the IEICE excellent Paper Award, and in 1999 and 2000 the IPN Research Award. In 1998 Prof. Perez-Meana was Co-Chair of the ISITA'98. His principal research interests are signal and image processing, pattern recognition, watermarking, steganography and related fields. Prof. Perez-Meana is a senior member of the IEEE, a member of the IEICE, the IET, the National Researchers System of Mexico and the Mexican Academy of Science.

Keita Takahashi received his B.E., M.S., and Ph.D. degrees in information and communication engineering from the University of Tokyo, Japan in 2001, 2003, and 2006, respectively. He was a research fellow of the Japan Society for the Promotion of Science in 2005–2006. He was a project assistant professor of the University of Tokyo in 2006–2011. From Oct. 2011 to April 2013 he was an assistant professor of the Graduate School of Informatics and Engineering, the University of Electro-Communications, Japan. He is currently an associate professor of Graduate School of Engineering, Nagoya University, Japan. His research interests include 3-D vision, image-based rendering, image coding, object recognition, and video segmentation. He is a member of the IEEE SPS&CS and IEICE.

Masahide Kaneko received his B.E., M.E., and D.E. degrees from the University of Tokyo, Japan in 1976, 1978, and 1981, respectively. From April 1981 to March 1994, he was with the Research and Development Laboratories of Kokusai Denshin Denwa Co., Ltd. (KDD). From April 1994 to March 1997, he was an associate professor of Department of Information and Communication Engineering, the University of Tokyo. In April 1997, he was reinstated in the Research and Development Laboratories of KDD. In April 1998, he joined the University of Electro-Communications as an associate professor. He is currently a professor of Department of Mechanical Engineering and Intelligent Systems. His research interests include the image coding, 3D image processing, processing of facial image information, and active interaction between humans and intelligent robots. He is a member of IEEE, the Institute of Image Information and Television Engineers, the Information Processing Society of Japan, the Robotics Society of Japan, and Japan Academy of Facial Studies.