International Journal of Innovative
Computing, Information and Control
Volume **9**, Number **4**, April **2013**

ICIC International ©2013 ISSN 1349-4198
pp. 1409–1421

AN OPTICAL FLOW FEATURE-BASED ROBUST FACIAL EXPRESSION RECOGNITION WITH HMM FROM VIDEO

Md. Zia Uddin[1,*], Tae-Seong Kim[2] and Byung Cheol Song[1,*]

[1]School of Electronic Engineering
Inha University
100 Inha-ro, Nam-gu, Incheon 402-751, Republic of Korea
*Corresponding authors: { ziauddin; bcsong }@inha.ac.kr

[2]Department of Biomedical Engineering
Kyung Hee University
26 Kyunghee-daero, Dongdaemun-gu, Seoul 130-701, Republic of Korea
tskim@khu.ac.kr

Received January 2012; revised May 2012

Abstract. *In this work, a novel method is proposed to recognize several facial expressions from time-sequential facial expression images. To produce robust facial expression features, optical flow extraction is utilized which are further improved by Principal Component Analysis (PCA) and Generalized Discriminant Analysis (GDA). Using these features, discrete Hidden Markov Models (HMMs) are utilized to model different facial expressions. Performance of our proposed FER system is compared against the conventional approaches and the proposed approach significantly improves the performance yielding the mean recognition rate of 99.16% whereas the conventional methods yield 82.92% at best.*

Keywords: Optical flows, Principal component analysis (PCA), Generalized discriminant analysis (GDA), Hidden Markov models (HMMs)

1. **Introduction.** Video cameras are being extensively used with many electronic devices for various surveillance applications. Among the applications, Facial Expression Recognition (FER) gets a considerable attention recently due to its potential use for Human Computer Interface (HCI). FER has been regarded as one of the key technologies for HCI, which enables computers to interrelate with humans in a way to human to human interactions [1,2]. As the FER technology provides computers a way of sensing a person's emotional information through video cameras, it is supposed to contribute to a HCI system by responding to the person's expressive conditions.

For non-motion feature extraction from the facial expression video images, most of the early FER researches extracted useful features using Principal Component Analysis (PCA). PCA is a second-order statistical method that derives the orthogonal bases containing the maximum variability in an unsupervised manner that provides global image features. Besides, it is most commonly used for dimension reduction. In [3], Padgett and Cottrell applied PCA on facial expression images to identify Facial Action Units (FAUs) and to recognize the facial expressions where their best analysis was performed on the separated face region such as eyes and mouth. In [4,5], the authors also employed PCA as one of the feature extractors for FER with the Facial Action Coding System (FACS). However, the recent research works such as [6] show that facial expressions do not take place in separate ways, but the whole spatial relationships of the facial features can be the additional source of information in the perception of facial emotions. Thus, most of the recent works such as [7-9] focused on the emotion specified feature extraction rather than focusing FAU. In [7,8], the authors utilized Linear Discriminant Analysis (LDA) to further classify the principal component (PC) features of the facial expression images. Lately, Independent Component Analysis (ICA) [9-15], a higher order statistical approach than PCA, has been extensively utilized for FER tasks due to its ability to extract local facial features. Basically, ICA is a generalization of PCA that seeks the independencies of the image features. In [9], Buciu et al. proposed ICA for the emotion-specified FER and in [13], Bartlett et al. extracted the local image representations for the facial expression coding using ICA to classify several facial expressions referred to FACS.

As all expressions contain different motion information, they can be separated easily utilizing motion features. Regarding the motion information-based FER, several works have been conducted to describe various expressions [16-19]. Optical flow methods are generally used for motion analysis of different expressions where using two consecutive frames of an image sequence, a two-dimensional vector field is computed to estimate the most likely displacement of image pixels from one frame to another. For example, in [18], Kruizinga and Petkov proposed to utilize optical flows in person identification. However, they considered the optical flow residue for expression classification. Optical flow can provide the visual motion information about facial expression images. Moghaddam et al. showed modeling of visual motion in [19]. They utilized pixel difference between images, and applied the Bayesian approach to modeling the pixel difference for all the subjects. A robust discriminant analysis called General Discriminant Analysis (GDA) has recently been used in different applications such as palmprint and face identification where GDA significantly outperforms the traditional feature extraction approaches such as PCA [20]. Thus, GDA can be used to obtain better discrimination among different facial expressions.

To train and recognize different facial expressions, several techniques have been applied so far [21-25]. In [21], Tian et al. employed neural network to extract FAUs applying the Gabor wavelet-based features. Support Vector Machine (SVM) classifiers were also utilized on facial expression image regions in FER [14,22]. Recently, Hidden Markov Model (HMM), a useful tool to encode time-sequential events was popularly tested for FER [23-25]. In [23], Otsuka and Ohya showed an HMM-based approach to recognize the facial expressions of multiple persons utilizing time-sequential velocity vectors extracted from consecutive facial expression images. In [24], Zhu et al. applied HMM to recognize the emotions utilizing moment invariants as features. In [25], Aleksic and Katsaggelos introduced a HMM-based FER approach using Facial Animation Parameters (FAPs) such as the movements of the eyebrows and outer-lip contours as observations. Due to its successful usage in pattern classification areas, we also utilize HMM to model and encode the time-sequential features in this work.

In this work, we propose a novel approach for FER using optical flows, PCA, GDA, and HMM. We first extract the optical flows from the facial expression images and further extended by PCA and LDA. Discrete HMMs are used to train and test distinguished expressions containing time-sequential facial expression image sequences. Later on, we improve the approach by replacing LDA with GDA under the same HMM scheme to obtain superior recognition performance over others. To our best knowledge, this is the first attempt to address the time-sequential Optical Flows-PCA-LDA as well as Optical Flows-PCA-GDA feature information for FER via HMM. To compare the performance of our proposed approach, six comparison studies have been conducted such as PCA, ICA, PCA-LDA, Optical Flows-PCA, Optical Flows-PCA-LDA, and Optical Flows-PCA-GDA as feature extractors in combination with HMM. The experimental results show that Optical Flows-PCA-GDA significantly outperforms the other approaches.

The rest of the paper is organized as follows. Section 2 illustrates the proposed system from facial expression image preprocessing to HMM via feature extraction and vector quantization. Sections 3 and 4 represent the experimental setups and results respectively utilizing different feature extraction techniques including the proposed one. Finally, Sections 5 and 6 show discussion as well as concluding remarks respectively based on the experimental results obtained applying different approaches.

2. **Methods.** Our proposed FER system consists of preprocessing of sequential facial expression images in video, optical flow-based feature extraction, codebook generation via vector quantization algorithm, and modeling as well as recognition via HMM. Figure 1(a) shows the basic steps of training a facial expression HMM where M indicates the number of video clips for an expression, T the number of frames in each video, and O the observation symbol sequence for each video. Figure 1(b) shows the testing process of a facial expression image sequence utilizing the likelihoods (i.e., L) of all trained facial expression HMMs (i.e., $\boldsymbol{H}$).

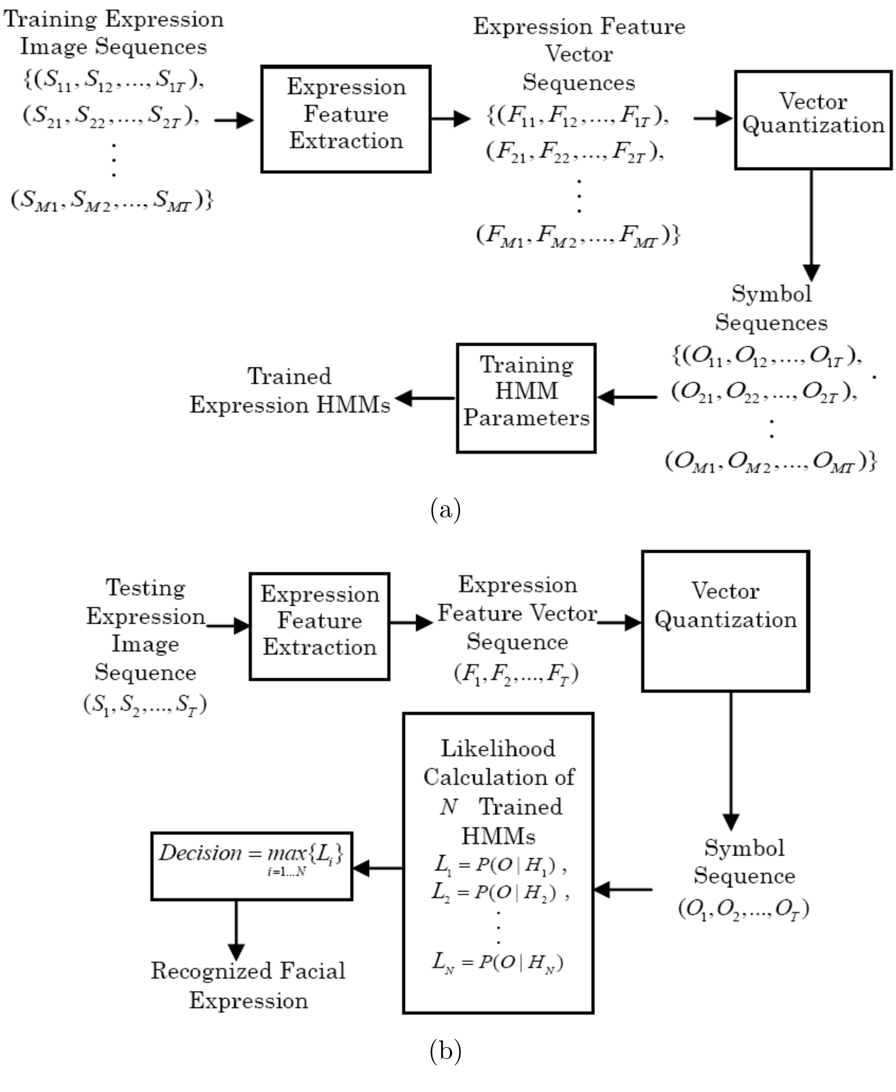
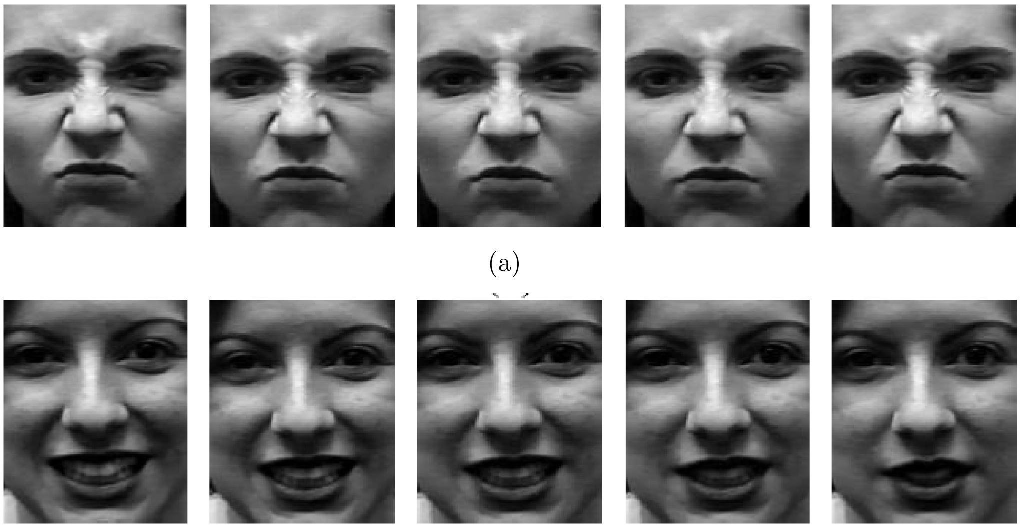


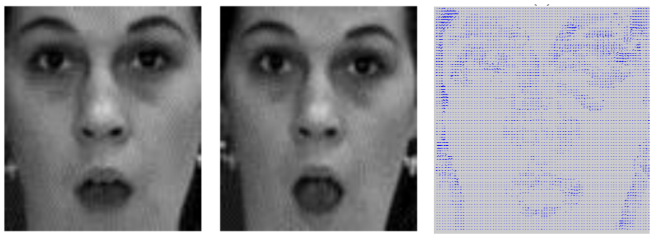
FIGURE 1. (a) Training HMM for a facial expression and (b) testing an expression in an image sequence



(a)

(b)

FIGURE 2. A sequential facial expression images of (a) anger and (b) joy



(a) (b)

FIGURE 3. (a) Two consecutive surprise image and (b) corresponding optical flows

2.1. **Preprocessing.** In preprocessing of sequential images of facial expressions, first image alignment is performed to realign the common regions of the face. A face alignment approach in [26] is applied by matching the eyes and mouth of the faces in the designated coordinates. The realigned image consists of 60 by 80 pixels. Histogram equalization is then applied on the realigned images for lighting correction. Figure 2 shows an exemplar set of facial expression sequence of (a) anger and (b) joy respectively. After extracting the facial expression images, the optical flows are obtained from the consecutive images. For optical flow computation, the Lucas-Kanade method has been utilized [27]. Figure 3 shows two consecutive surprise images as well as corresponding optical flows.

2.2. **PCA on optical flows.** Once optical flows of each pixel in an image of the facial expression are obtained, the vectors are then converted to a single row vector as $\widetilde{K} = \left[\bar{K}_1, \bar{K}_2, \ldots, \bar{K}_{4800}\right]$ where $\overline{K}$ consists of two components along x- as well as y-axis. Furthermore, PCA is applied on the optical flow row vectors obtained from the facial expression images of all the expressions to reduce the long dimension.

PCA is the most popular method to approximate original data in the lower dimensional feature space. Basically, it computes the eigenvectors of the covariance data matrix C and then does the approximation using the linear combination of top eigenvectors. The covariance matrix of the sample training image vectors and the PCs of the covariance matrix can be calculated respectively as

$$C = \frac{1}{T} \sum_{i=1}^{T} (\widetilde{K}_i \widetilde{K}_i^T) \tag{1}$$

$$E^T C E = \Lambda \tag{2}$$

where $\widetilde{K}_i$ represents the zero-mean processed optical flow row vector for i^{th} frame, E the matrix of orthonormal eigenvectors (i.e., PCs), and Λ diagonal matrix of the eigenvalues. E reflects the original coordinate system onto the PCs where the eigenvector corresponding to the largest eigenvalue indicates the axis of largest variance and the next largest one is the orthogonal axis of largest one indicating second largest variance and so on. Usually, the eigenvalues close to zero carry negligible variance and hence can be excluded from further processing. So, the top m eigenvectors corresponding to the largest eigenvalues can be used to define the subspace. Thus, the full dimensional expression image vectors can be easily represented in the reduced dimension.

In this work, 150 principal components (i.e., E_{opt}) were considered from the whole PC feature space. Thus, the projection of the zero-mean optical flow row vectors (i.e., $\tilde{K}$) on the PC feature space can be represented as

$$R = \tilde{K} E_{opt} \tag{3}$$

In this way, the size of the PC-based reduced motion feature representation of each frame becomes 1×150. To make the motion features more robust, PC-based optical flow features of the facial expression images of the six classes (i.e., anger, joy, sad, disgust, fear and surprise) are further classified by LDA as well as GDA.

2.3. **GDA on PC-features vs. LDA on PC-features of optical flows.** To produce more robust features, we do discriminant analysis of the PCA features: namely LDA and GDA respectively. Basically, LDA produces an optimal linear discriminant function which maps the input into the classification space based on which the class identification of the samples can be decided. The within scatter matrix, S_W and the between scatter matrix, S_B. The optimal discrimination matrix D_{LDA} is chosen from the maximization of ratio of the determinant of S_B and S_W as

$$D_{LDA} = ARG\ \underset{D}{MAX} \frac{|D^T S_B D|}{|D^T S_W D|} \tag{4}$$

where D_{LDA} is the set of discriminant vectors of S_W and S_B.

Figures 4(a) and 4(b) demonstrate the 3-D plot of the motion features of the facial expression images of the six classes obtained from applying Optical Flows-PCA and Optical Flows-PCA-LDA respectively. In Figure 4(b), the prototypes of each class are clustered and separated well using Optical Flows-PCA-LDA indicating the improved facial expression feature representation.

As LDA tries to find a classification space to classify the samples of different classes linearly, a more robust discriminant analysis called General Discriminant Analysis (GDA), a nonlinear approach to classify patterns from different classes can be applied to outperform LDA. Basically, GDA produces an optimal discriminant function, which maps the input into a classification space. GDA maps the training data into a high dimensional feature space Z by a nonlinear function as

$$Y : F \to Z, \quad R_i \to Y(R_i). \tag{5}$$

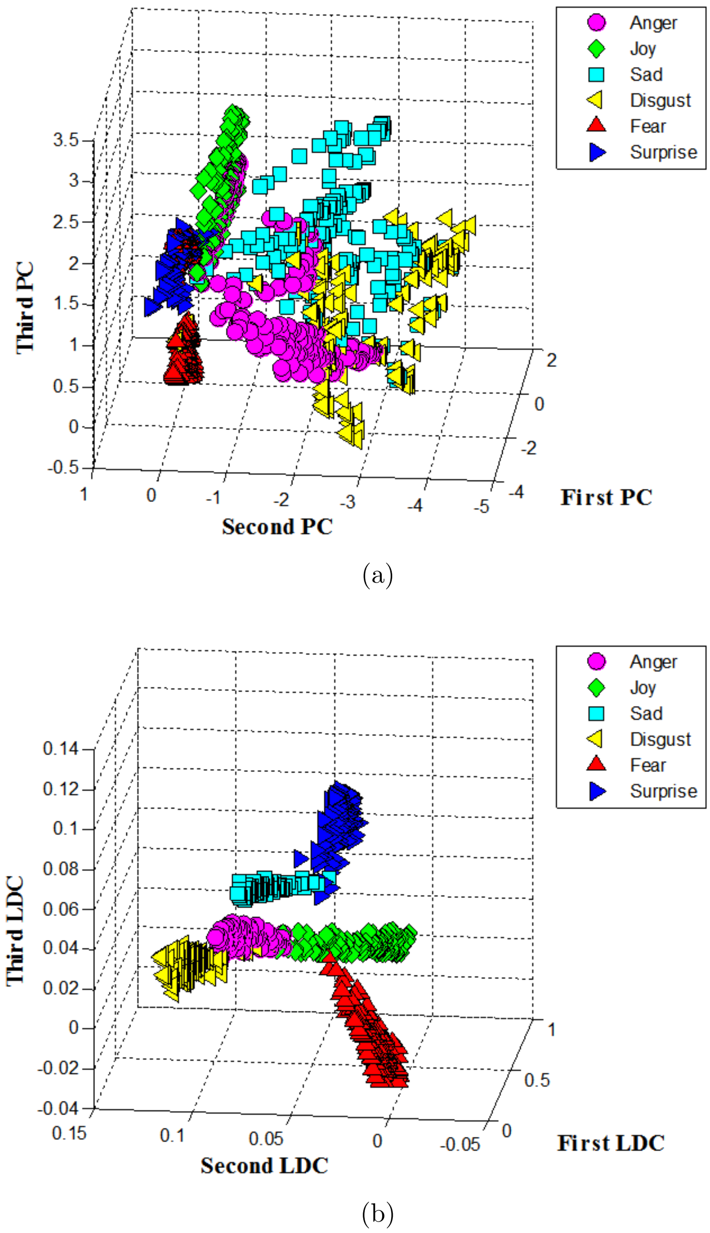


FIGURE 4. 3-D plot of the features from (a) PCA and (b) LDA on the PC features of the optical flows of the six classes (i.e., anger, joy, sad, disgust, fear, and surprise)

The advantage of GDA is building a nonlinear feature space for classification that cannot be separable by linear methods. A Gaussian kernel function $K(R_i, R_j)$ and a map Y is used that satisfy $Y(R_i)Y(R_j) = K(R_i, R_j)$. GDA maximizes the following as

$$\Lambda = \frac{|\theta^T B^Y \theta|}{|\theta^T T^Y \theta|} \tag{6}$$

where B^Y and T^Y are the between-class and total sample scatter matrices respectively.

However, the optimal discriminant vectors θ_{GDA}^T can be achieved through solving the following eigenvalue problem.

$$B^Y \theta = \Lambda T^Y \theta \tag{7}$$

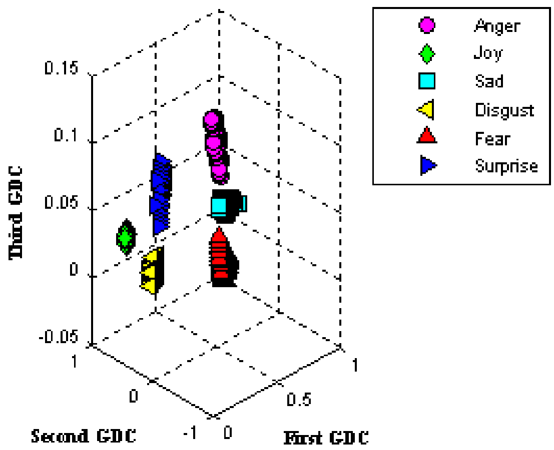


FIGURE 5. 3-D plot of the features from GDA on the PC features of the optical flows of the six classes

Thus, the Optical Flows-PCA features of the expression images can be extended by GDA as

$$G = \theta_{GDA}^T P. \tag{8}$$

Figure 5 shows exemplar plot of 3-D using GDA on PC features of the optical flows of six classes where it shows a better separation among the expression prototypes of different classes than other approaches such as LDA on PC features as shown in Figure 4(b). Hence, GDA seems to be a better choice over LDA to generate robust features in this regard. Thus, we prefer Optical Flows-PCA-GDA features in this work.

2.4. **Vector quantization.** To decode the sequential variations of the facial expression features, discrete HMMs have been employed. As discrete HMMs are usually trained and tested with symbol sequences, the feature vectors are symbolized by means of comparing with the codeword vectors of a codebook. To obtain the codebook, vector quantization algorithm is performed on the feature vectors from the training datasets. In our work, the Linde, Buzo and Gray (LBG) algorithm has been utilized for codebook generation [28]. The LBG approach selects the initial centroids and splits the centroids of the whole

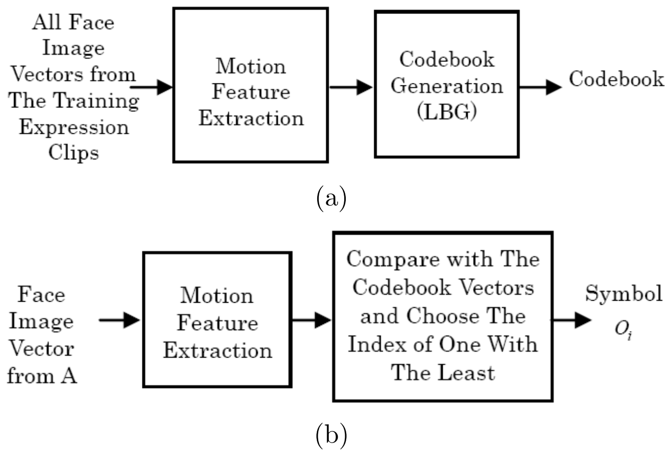


FIGURE 6. Steps for (a) codebook generation and (b) symbol selection

dataset. Then, it continues to split the dataset according to the codeword size, and the optimization is performed to reduce the distortion.

Once the codebook is obtained, the index numbers of the codewords are regarded as symbols to be used with discrete HMMs. As each facial expression image is converted to a symbol, an expression video clip of T consecutive images will result in T symbols after the vector quantization. Figure 6 demonstrates the basic steps for codebook generation from all the facial expression image vectors and symbol selection for a sample facial expression image vector.

2.5. **Modeling and recognition via HMM.** The basic theory of HMM was developed by Baum et al. and it has been applied extensively to solve a large number of problems [23-25]. Due to its successful usage in pattern classification to decode the time-sequential events, HMM has been adopted as a model and recognizer for expression recognition. A HMM denoted as $H = \{\Xi, \pi, A, B\}$ can be expressed as follows.

- The states $\Xi = \{\Xi_1, \Xi_2, \ldots, \Xi_q\}$, where q is the number of states.
- The initial probability of the states π and can be obtained as

$$\pi = \{\pi_j\}, \quad \sum_{j=1}^{q} \pi_j = 1. \tag{9}$$

– The state transition probability matrix A where a_{ij} denotes the probability of a changing state from i to j, i.e.,

$$a_{ij} = P(\Omega_{t+1} = \Xi_j | q_t = \Xi_i), \quad 1 \leq i,\ j \leq q, \tag{10}$$

$$\sum_{j=1}^{q} a_{ij} = 1, \quad 1 \leq i \leq q. \tag{11}$$

– The observation symbols' probability matrix B where the probability $b_j(d)$ represents the probability of observing the symbol d from a state j and can be expressed as follows.

$$b_j(d) = P(O_t = d | \Omega_t = \Xi_j), \quad 1 \leq j \leq q. \tag{12}$$

To train each HMM, first vector quantization is performed on the training features from the facial expression image sequences to obtain discrete symbols. Those obtained sequential symbols are then trained with HMMs to learn the proper model for each expression. Four-state left to right HMMs are chosen in this work to model sequential events of facial expressions. Figure 7 shows the HMM structure and the transition probabilities of a trained sad expression.

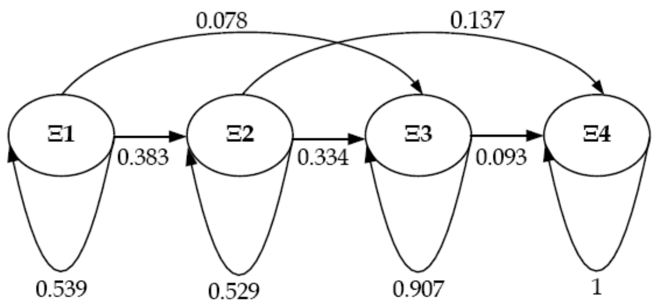


FIGURE 7. A sad expression HMM with the transition probabilities after training

When testing an expression image sequence, after vector quantization process, the obtained observation symbol sequence is used to determine the proper model by means of highest likelihood computation of all trained expression HMMs.

3. **Experimental Setups.** A set of comparison experiments were done using non-motion as well as motion feature extraction methods with HMMs including the proposed one. To report the recognition performance, the training and testing video clips were prepared utilizing Cohn-Kanade facial expression database obtained from [29]. Six facial expressions such as anger, joy, sad, disgust, fear, and surprise were tried to recognize in this work. From the facial expression database, we collected variable length consecutive frames from each video sequences. The collected frames are then realigned with the size of 60×80 pixels. A total of 90 sequences from all expressions were utilized to build the feature spaces. To evaluate the performance of the proposed system, we applied a total of 15 and 40 image sequences per expression for training and testing each expression respectively. Then, different feature extraction methods under the same HMM structure were compared. First of all, we went for non-motion feature-based traditional FER experiments using PCA, ICA, PCA-LDA with HMM. Then, we tried the motion feature-based experiments with Optical Flows-PCA, Optical Flows-PCA-LDA, and Optical Flows-PCA-GDA with HMM where Optical Flows-PCA-GDA showed superior recognition performance.

4. **Experimental Results.** To compare the proposed method with the conventional feature extraction methods, all the extraction methods were implemented with the HMMs for recognition of six different facial expressions.

4.1. **Results using non-motion features.** Regarding non-motion traditional feature-based FER experiments, we used PCA, PCA-LDA, and ICA features with HMM. As shown in Table 1, the recognition rate using PCA on the gray images is overall 55.83%, the lowest recognition rate. Then, we employed ICA to extract the ICs from the image dataset. The results using ICA method in Table 2 shows the improved recognition rate than the result of PCA. To apply the ICA method, we first chose the eigenvectors as used in the PCA method and processed ICA on the selected eigenvectors. Moreover, the conventional approach PCA-LDA was performed for the last comparison study and it achieved the recognition rate of 82.92% as shown in Table 3.

4.2. **Results using optical flow-based motion features.** Using the settings as above, we conducted the experiment utilizing Optical Flows-PCA with HMMs and achieved the total mean recognition rate of 92.50% as shown in Table 4. A better recognition rate is achieved (i.e., 96.67%) further using Optical Flows-PCA-LDA as shown in Table 5.

Table 1. Confusion matrix using PCA (unit: %)

Expression	Anger	Joy	Sad	Surprise	Fear	Disgust
Anger	**30**	0	20	0	10	40
Joy	5	**50**	10	10	30	5
Sad	0	7.5	**82.50**	12.50	0	0
Surprise	0	0	0	**70**	12.50	17.50
Fear	0	7.50	50	7.50	**35**	0
Disgust	0	7.50	25	0	0	**67.50**
Mean	**55.83**					

Table 2. Confusion matrix using ICA (unit: %)

Expression	Anger	Joy	Sad	Surprise	Fear	Disgust
Anger	**30**	0	10	30	10	20
Joy	5	**60**	0	0	35	0
Sad	0	5	**87.50**	7.50	0	0
Surprise	0	0	12.50	**82.50**	0	5
Fear	0	25	25	7.50	**35**	7.50
Disgust	0	7.50	25	0	0	**67.50**
Mean	**60.41**					

Table 3. Confusion matrix using PCA-LDA (unit: %)

Expression	Anger	Joy	Sad	Surprise	Fear	Disgust
Anger	**60**	0	10	0	0	30
Joy	0	**87.50**	0	0	7.50	5
Sad	0	5	**87.50**	7.50	0	0
Surprise	0	0	0	**95**	5	0
Fear	0	7.50	7.50	10	**75**	0
Disgust	0	0	0	0	7.50	**92.50**
Mean	**82.92**					

Table 4. Confusion matrix using Optical-Flows-PCA (unit: %)

Expression	Anger	Joy	Sad	Surprise	Fear	Disgust
Anger	**82.50**	0	0	0	0	17.50
Joy	0	**95**	0	0	5	0
Sad	0	0	**90**	0	10	0
Surprise	0	0	0	**100**	0	0
Fear	0	7.50	0	0	**92.50**	0
Disgust	0	0	0	0	5	**95**
Mean	**92.50**					

Table 5. Confusion matrix using Optical-Flows-PCA-LDA (unit: %)

Expression	Anger	Joy	Sad	Surprise	Fear	Disgust
Anger	**95**	0	0	0	0	5
Joy	0	**97.50**	0	0	2.50	0
Sad	0	0	**95**	0	5	0
Surprise	0	0	0	**100**	0	0
Fear	0	5	0	0	**95**	0
Disgust	0	0	0	0	2.50	**97.50**
Mean	**96.67**					

Finally, Table 6 shows the recognition rates using the Optical Flows-PCA-GDA, i.e., approach that outperforms all the methods for all the expressions containing mean recognition rate of 99.16%.

Table 6. Confusion matrix using Optical-Flows-PCA-GDA (unit: %)

Expression	Anger	Joy	Sad	Surprise	Fear	Disgust
Anger	**100**	0	0	0	0	0
Joy	0	**97.50**	0	0	2.50	0
Sad	0	0	**97.50**	0	2.50	0
Surprise	0	0	0	**100**	0	0
Fear	0	0	0	0	**100**	0
Disgust	0	0	0	0	0	**100**
Mean	**99.16**					

5. **Discussion.** In this work, we have proposed a new method of FER by combining Optical Flows-PCA-GDA and HMM to model the time-sequential feature information from holistic facial expression images.

While executing the conventional non-motion feature extraction methods, the recognition rate achieved utilizing image representations derived from ICA was higher than the recognition rate obtained using PCA. The PCA representation is known for producing the global facial expression features. On the other hand, ICA highlights the local features of the faces. It seems that ICA is a better tool for capturing the more relevant features for determining the expressions than PCA from holistic facial expression images. So far, PCA-LDA produces the improved recognition rate in the non-motion feature-based FER system. However, PCA-LDA does not produce sufficient recognition rate in this regard.

On the contrary, Optical Flows-PCA approach showed higher recognition performance than conventional approaches representing time-sequential local motion features of each facial expression images. A far better recognition performance was achieved further using Optical Flows-PCA-LDA. Finally, Optical Flows-PCA-GDA showed its superiority over all the feature extraction methods by means of achieving the highest recognition rate.

6. **Conclusion.** We have proposed a video-based robust FER system using Optical Flows-PCA-GDA for facial expression motion feature extraction and HMM for recognition. We have illustrated the performance of our proposed method applied on sequential datasets for the six facial expression recognition problems. The experimental results show that Optical Flows-PCA-GDA improves the feature extraction task. Furthermore, HMM, dealing with the Optical Flows-PCA-GDA processed sequential facial expression images can provide superior recognition rate over the other feature extraction approaches, reaching up to the mean recognition rate of 99.16% where the conventional non-motion feature-based approach achieved 82.92% at best. In future, we plan to focus on more robust motion features for improved FER.

Acknowledgement. This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012000446). This research was also supported by the MKE (The Ministry of Knowledge Economy), Korea, under the ITRC (Information Technology Research Center) support program supervised by the NIPA (National IT Industry Promotion Agency) (NIPA-2012-(H0301-12-2001). This work was also supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No.2012-0000609). This research was also supported by the research grant of Inha University, Korea.

REFERENCES

[1] Y.-K. Chang, C.-C. Lien and L.-Y. Lin, A new appearance-based facial expression recognition system with expression transition matrices, *International Journal of Innovative Computing, Information and Control*, vol.5, no.11(B), pp.4089-5104, 2009.
[2] R. Zhi and Q. Ruan, Facial expression recognition using fuzzy laplacianfaces, *International Journal of Innovative Computing, Information and Control*, vol.6, no.5, pp.1999-2011, 2009.
[3] C. Padgett and G. Cottrell, Representation face images for emotion classification, *Advances in Neural Information Processing Systems*, vol.9, Cambridge, MA, USA, 1997.
[4] G. Donato, M. S. Bartlett, J. C. Hagar, P. Ekman and T. J. Sejnowski, Classifying facial actions, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.21, no.10, pp.974-989, 1999.
[5] P. Ekman and W. V. Priesen, Facial action coding system: A technique for the measurement of facial movement, *Consulting Psychologists Press*, Palo Alto, CA, USA, 1978.
[6] M. Meulders, P. D. Boeck, I. V. Mechelen and A. Gelman, Probabilistic feature analysis of facial perception of emotions, *Applied Statistics*, vol.54, no.4, pp.781-793, 2005.
[7] A. J. Calder, A. M. Burton, P. Miller, A. W. Young and S. Akamatsu, A principal component analysis of facial expressions, *Vision Research*, vol.41, pp.1179-1208, 2001.
[8] S. Dubuisson, F. Davoine and M. Masson, A solution for facial expression representation and recognition, *Signal Proc. of Image Communication*, vol.17, pp.657-673, 2002.
[9] I. Buciu, C. Kotropoulos and I. Pitas, ICA and Gabor representation for facial expression recognition, *Proc. of the IEEE*, pp.855-858, 2003.
[10] F. Chen and K. Kotani, Facial expression recognition by supervised independent component analysis using MAP estimation, *IEICE Transactions on Information and Systems*, vol.E91-D, no.2, pp.341-350, 2008.
[11] A. Hyvarinen, J. Karhunen and E. Oja, *Independent Component Analysis*, John Wiley & Sons, 2001.
[12] Y. Karklin and M. S Lewicki, Learning higher-order structures in natural images, *Network: Computation in Neural Systems*, vol.14, pp.483-499, 2003.
[13] M. S. Bartlett, G. Donato, J. R. Movellan, J. C. Hager, P. Ekman and T. J. Sejnowski, Face image analysis for expression measurement and detection of deceit, *Proc. of the 6th Joint Symposium on Neural Computation*, pp.8-15, 1999.
[14] C.-F. Chuang and F. Y. Shin, Recognizing facial action units using independent component analysis and support vector machine, *Pattern Recognition*, vol.39, pp.1795-1798, 2006.
[15] A. J. Calder, A. W. Young and J. Keane, Configural information in facial expression perception, *Journal of Experimental Psychology: Human Perception and Performance*, vol.26, no.2, pp.527-551, 2000.
[16] J. J. Lien, A. Zlochower, J. F. Cohn and T. Kanade, Automated facial expression recognition, *Proc. of the 3rd IEEE International Conference on Automatic Face and Gesture Recognition*, pp.390-395, 1998.
[17] Y. Yacoob and L. S. Davis, Recognizing human facial expressions from long image sequences using optical flow, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol.18, no.6, pp.632-646, 1996.
[18] P. Kruizinga and N. Petkov, Optical flow applied to person identification, *Proc. of the 1994 EUROSIM Conference on Massively Parallel Processing Applications and Development*, pp.871-878, 1994.
[19] B. Moghaddam, T. Jebara and A. Pentland, Bayesian face recognition, *Pattern Recognition*, vol.33, pp.1771-1782, 2000.
[20] L. Shen, L. Bai and M. Fairhurst, Gabor wavelets and general discriminant analysis for face identification and verification, *Image and Vision Computing*, vol.25, pp.553-563, 2007.
[21] Y. Tian, T. Kanade and J. Cohn, Evaluation of Gabor wavelet based facial action unit recognition in image sequences of increasing complexity, *Proc. of the 5th IEEE International Conference on Automatic Face and Gesture Recognition*, pp.229-234, 2002.
[22] I. Kotsia and I. Pitas, Facial expression recognition in image sequences using geometric deformation features and support vector machine, *IEEE Transactions on Image Processing*, vol.16, pp.172-187, 2007.
[23] T. Otsuka and J. Ohya, Recognizing multiple person's facial expressions using HMM based on automatic extraction of significant frames from image sequences, *Proc. of the International Conference on Image Processing*, pp.546-549, 1997.

[24] Y. Zhu, L. C. De Silva and C. C. Ko, Using moment invariants and HMM in facial expression recognition, *Pattern Recognition Letters*, vol.23, no.1-3, pp.83-91, 2002.

[25] P. S. Aleksic and A. K. Katsaggelos, Automatic facial expression recognition using facial animation parameters and multistream HMMs, *IEEE Transactions on Information and Security*, vol.1, pp.3-11, 2006.

[26] L. Zhang and G. W. Cottrell, When holistic processing is not enough: Local features save the day, *Proc. of the 26th Annual Cognitive Science Society Conference*, 2004.

[27] J. L. Barron, D. J. Fleet and S. Beauchemin, Performance of optical flow techniques, *International Journal of Computer Vision*, vol.12, no.1, pp.43-77, 1994.

[28] Y. Linde, A. Buzo and R. Gray, An algorithm for vector quantizer design, *IEEE Transactions on Communications*, vol.28, no.1, pp.84-94, 1980.

[29] J. F. Cohn, A. Zlochower, J. Lien and T. Kanade, Automated face analysis by feature point tracking has high concurrent validity with manual FACS coding, *Psychophysiology*, vol.36, pp.35-43, 1999.