Real Time Human Facial Expression Recognition System using Smartphone

Muhammad Hameed Siddiqi[†], Madallah Alruwaili[†], JaeHun Bang[††] and Sungyoung Lee[††],

[†]Department of Computer and Information Sciences, AlJouf University, Sakaka, AlJouf, Kingdom of Saudi Arabia
[††]Department of Computer Engineering, Kyung Hee University, Suwon, Rep. of Korea

Summary
Smartphone is a good indicator of mental status that has been employed in healthcare environments in order to assess the mental status of elderly patients. In such environments, the mental status of humans can be analyzed by expressions. Expressions have a significant role in improving the level of interaction between human-to-human communications. The features from human mouth, eyes, and eyebrow are considered the most informative features for expressions recognition. Commonly, these systems were tested on publicly available datasets which were collected using a fixed camera with static background. In this work, we proposed a real-time facial expression recognition (FER) system using smartphone camera. In order to make the system efficient and robust, we have extracted the features only from the contributing parts of the face, for which the angles and distances were measured. Then for classification, we used support vector machine (SVM) under 10$-$fold cross validation setting. The system was tested and validated in real time by 10 university students (which are not professional). The weighted average recognition rate for the proposed smartphone-based FER system is 85.6% for 5 basic expressions, which is a significant improvement in real-time domain.

Key words:
Expression Recognition, Face Recognition, Facial Feature Point Detection, Android Phone.

1. Introduction

Smartphone has been used widely by the human in real life. In healthcare, this technology may help to determine the mental status of a patient such as physically disabled people [1]. The body of a stroke patient is completely paralyzed due to which a normal human cannot understand his feelings. The facial expression can be one of the sources through which we can understand him. Such a system has been trained in order to identify accurately the moving parts of a patient face and determine the mood of the patient. Whenever the expressions have been identified for the stroke patient, then the doctors can easily produce strategies for their caution and security. This technology can be useful in care-donors and medical experts in order to efficiently employ the assets and delight the patient carefully. The heart failure patients can be monitored using facial expressions in telemedicine [2]. Likewise, for the fit people who cannot speak and share their feelings might be directed using expression systems, particularly in times of their aloneness. In healthcare, different types of communication are being used for medical purpose. Among them, verbal communication (such as speech) and non-verbal communication (such as facial expressions) [3] are commonly used.

Facial expressions recognition (FER) is an important concept for many applications, like image retrieval; human emotion analysis [4], neuroscience and psychology [5], access control and surveillance [6], and personality and child development [7]. Facial expressions not only represent emotions, but also reflect mental status, social interaction, and physiological signals [4]. Consistent with these ideas, psychophysiology studies using facial electromyography (EMG) have found that presented a system for facial expressions based on matching facial muscular activity in the viewer [8]. The analysis and development of automatic facial expressions recognition depends extensively on evolutions in the above-mentioned sciences.

General expression system has three modules, face recognition, feature extraction, and recognition modules. In the face recognition module, the faces have been detected and extracted [9]. Feature extraction module extracts informative features from different parts of the face. While, in recognition module, first a classifier has been trained and later has been then used to generate labels for the expressions using the training data.

In this work, we have proposed a real-time smartphone-based FER system. In this system, for the face recognition, well-known existing method such as Luxand face recognition [10] has been employed, which has the capability to track and detect the face in the expression frame. Then, the informative features are extracted from those parts of the face that have much contribution in expression making. Finally, for the expression, we utilized support vector machines (SVM) under the setting of 10□fold cross validation scheme. The proposed system

Manuscript received October 5, 2017
Manuscript revised October 20, 2017

achieved 85% recognition rate on 5 expressions in real time scenario.

The remaining paper is organized as follow. Section 2 presents the existing FER systems with their limitations. The overview of the proposed smartphone-based FER system is presented in Section 3. Section 4 shows the experimental results and discussion. Finally, the paper has been concluded with some future directions in Section 5.

2. Previous Works

A large number of facial expression recognition (FER) systems [9, 11–15], having varying success rates, have emerged over the past decade. However, these systems were assessed in control environments that are far away from real world circumstances which are common drawbacks for these systems. This is because most of the existing datasets are pose-based that were collected under predefined setup. The expressions in these datasets are collected using fix cameras with a continuous background and still ambient situations. In naturalistic environment, the expression systems are projected to deal with fluctuating ambient circumstances, vigorous background and camera angles, diverse size of face, and other human-related variations.

The authors of [16] proposed simple and effective system for the identification of expressions from the video frames. In this system, a set of SVM classifiers and active shape model (ASM) features were utilized for the classification of different expressions. However, in this system, the author assumed the frame rate (i.e., 3 frames per second), which is not the case in real world. Similarly, in [17], the authors proposed a robust FER system using smartphone, which has two strategies. In the first strategy, they used active appearance model (AAM) with some edge-detector algorithms. While, in the second strategy, the back propagation neural network was used for the expression classification. However, the system was not tested in real world, but just validated on existing standard dataset of facial expressions (such as Yale B face dataset) which is far away from real world scenario. Likewise, developed a new approach based on haar classifiers, skin detection, feature extraction, feature points tracking and SVM for the real time FER. The system achieved 60% accuracy in different scenarios. However, this approach was tested in lab environment under a controlled camera setting. Moreover, computational wise, it is much expensive when used in smartphone.

A frame-based FER system has been proposed by [19] based on several geometrical features. The system showed better performance on existing standard dataset of facial expressions. However, prior knowledge was required for this system. Also, the system was tested and validated on extended Cohn-Kanade (CK+) dataset (that has been collected using fix camera with static background and light), due to which the system is far away from real world scenarios. In [20], the authors proposed a FER system based on multi-layer perceptron (MLP) neural network under a constructive training algorithm. The system showed better performance on Cohn-Kanade datasets; however, the system is unable to achieve higher recognition rate in real world environments due to the usage of multi-layer perceptron. Moreover, lots of training is required for this system. A new frame-based FER system was proposed by [21] based on artificial neural network (ANN), linear discriminant analysis (LDA), and K-nearest neighbor (KNN). The system was tested on two standard datasets such as multimedia understanding group (MUG) and FEEDTUM datasets of facial expression. However, this is a pose-based FER system which showed better performance in controlled lab environment with fix camera and light. The performance of this system degrades when we move from lab environment to the real world scenarios.

A fully automatic FER system was designed by [22, 23] based on geometric features and appearance features. In this system, features are represented by three different forms such as point, line, and triangle respectively. The system showed better performance on existing datasets. However, it is pose-based FER system that does not have the capability to maintain the same performance in real time scenarios because of the variations in lighting condition and viewpoint [13]. Similarly, for appearance-based approaches, a prior knowledge is required, i.e., at the time of implementation for these techniques, it is compulsory to decide randomly which intensity information will be important [24].

A very recent smartphone-based FER system has been proposed by [25] using deep learning. They claimed that system has shown better performances for various face datasets compared to a classifier based on hand-crafted features. However, the system has been trained on GPU using existing facial expression datasets such as Cohn-Kanade, and then the system has been tested in real time. However, the system used a complex classifier like deep learning which requires lots of time for training that may not applicable in real world.

3. Paragraphs and Itemizations

The flow diagram for the proposed smartphone-based FER system is shown in Fig. 1.

3.1 Face Recognition

In this work, we utilized well-known existing method such as Luxand face application for the purpose of face recognition. This application is a kind of library that can easily integrate into smartphone that has the capability to detect and recognize the face on still images and videos. Moreover, it also tracks the facial features in order to recognize gender. This application provides a Tracker API (Application Programming Interface) that helps to track and recognize faces in live video. This API simply works with video streams, contribution the functions to tag subjects with names that further can be recognized. This application provides the coordinates of 66 facial feature points (such as eyes, eyebrows, mouth, nose and face contours). This application can use multiple processor cores to speed up the recognition process. For more information, please refer [10].

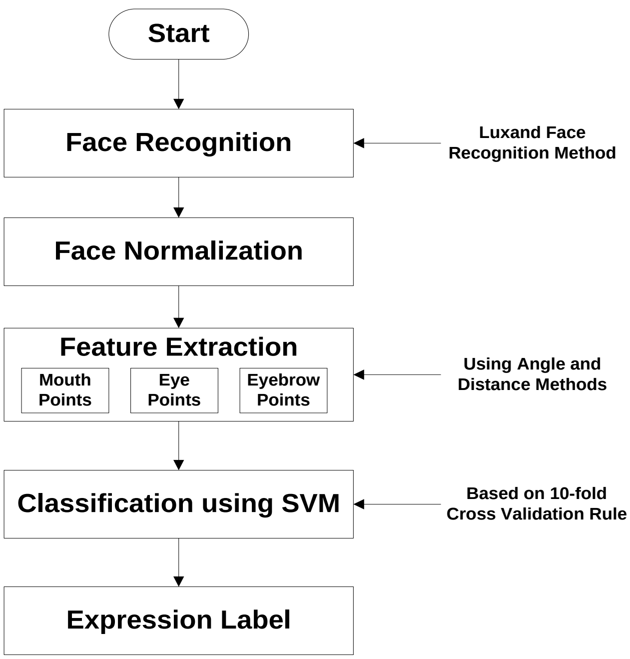

Fig. 1 Flow diagram for the proposed smartphone-based FER system.

3.2 Feature Extraction

Once the face has been recognized; then in the next step, we extract the features from the contributing parts of the face. For this particular purpose, we extracted the movable points as done by [26].

Consider a sliding window that has been moved pixel by pixel in order to find the motion of the pixels. By using this, we can compare pixels before and after using sum of square differences (SSD) that means "error" (*E(u, v)*). The mathematical form of this process is given in Eq. 1.

$$E(u,v)=\sum_{(i,j)\in Window}\left[\mathrm{Im}\,g(i+u,j+v)-\mathrm{Im}\,g(i,j)\right]^2 \qquad (1)$$

where Window represents the sliding window, *Img* indicates the corresponding image (recognized face), and *x* and *y* are the respective coordinates. If there are some small motions in the pixels, then Taylor Series expansion of *Img* can be utilized as shown in Eq. 2.

$$\mathrm{Im}\,g(i+u,j+v)=\mathrm{Im}\,g(i,j)+\frac{\partial\,\mathrm{Im}\,g}{\partial i}u+\frac{\partial\,\mathrm{Im}\,g}{\partial j}v + higher\ order\ terms \qquad (2)$$

If there is small motion in (*u*, *v*), then first order approximation is good as given in Eq. 3.

$$\begin{aligned}\mathrm{Im}\,g(i+u,j+v)&\approx \mathrm{Im}\,g(i,j)+\frac{\partial\,\mathrm{Im}\,g}{\partial i}u+\frac{\partial\,\mathrm{Im}\,g}{\partial j}v\\ &\approx \mathrm{Im}\,g(i,j)+\left[\mathrm{Im}\,g_i \quad \mathrm{Im}\,g_j\right]\begin{bmatrix}u\\v\end{bmatrix}\end{aligned} \qquad (3)$$

Img_i and Img_j represent two consecutive video frames. Combining Eq. 1 and Eq. 3, we can get

$$\begin{aligned}E(u,v)&=\sum_{(i,j)\in Window}\left[\mathrm{Im}\,g(i+u,j+v)-\mathrm{Im}\,g(i,j)\right]^2\\ &\approx\sum_{(i,j)\in Window}\left[\mathrm{Im}\,g(i,j)+\left[\mathrm{Im}\,g_i\ \ \mathrm{Im}\,g_j\right]\begin{bmatrix}u\\v\end{bmatrix}-\mathrm{Im}\,g(i,j)\right]^2\\ &\approx\sum_{(i,j)\in Window}\left[\left[\mathrm{Im}\,g_i\ \ \mathrm{Im}\,g_j\right]\begin{bmatrix}u\\v\end{bmatrix}\right]^2\end{aligned} \qquad (4)$$

Eq. 4 can be written as

$$E(u,v)\approx[u,v]\underbrace{\left(\sum_{(i,j)\in Window}\begin{bmatrix}\mathrm{Im}\,g_i^2 & \mathrm{Im}\,g_i\,\mathrm{Im}\,g_j\\ \mathrm{Im}\,g_i\,\mathrm{Im}\,g_j & \mathrm{Im}\,g_j^2\end{bmatrix}\right)\begin{bmatrix}u\\v\end{bmatrix}}_{H} \qquad (5)$$

In order to find the directions of the pixels, we used the eigenvectors of H (that is 2_2 matrix). This will result in the largest and smallest E values as shown in Eq. 6.

$$\det\begin{bmatrix}h_{11}-\lambda & h_{12}\\ h_{21} & h_{22}-\lambda\end{bmatrix}=0 \qquad (6)$$

The solution is given as

$$\lambda_{\pm} = \frac{1}{2}\left[\left(h_{11} + h_{22}\right) \pm \sqrt{4h_{12}h_{21} + \left(h_{11} - h_{22}\right)^2}\right] \quad (7)$$

Once we know the □, then we can find *i* by solving

$$\begin{bmatrix} h_{11} - \lambda & h_{12} \\ h_{21} & h_{22} - \lambda \end{bmatrix}\begin{bmatrix} i_1 \\ i_2 \end{bmatrix} = 0 \quad (8)$$

The detected points are presented in Fig. 2.

Fig. 2 Laxuad Face SDK with 66 facial feature points.

Once the points are extracted and detected, then the distance and angle are calculated between the pixels as shown in Eq. 9 and 10 respectively

$$Dist = \sqrt{\left(i_2 - i_1\right)^2 + \left(j_2 - j_1\right)^2} \quad (9)$$

$$\theta = \tan^{-1}\left(\frac{j}{i}\right) \quad (10)$$

where *Dist* indicates the distance, i_1, i_2 and j_1, j_2 represent x and y coordinates of the two points, and represents the angle between two points. The resultant image (after the angle and distance) is shown in Fig. 3.

3.3 Classification via Support Vector Machine

Support vector machine (SVM) is one of the famous statistical techniques that can be employed in machine vision, computer vision, and pattern recognition [27]. SVM have been exploited widely for linear and binary classification purposes. It is dependent on the optimum splitting decision hyper-plane within two or more classes with the maximum boundary amongst the patterns of individual class. SVM uses the so-called function that projects data from the original feature space to another higher dimensional space due to which the linear classification in the new space is equivalent to non-linear classification in the original space.

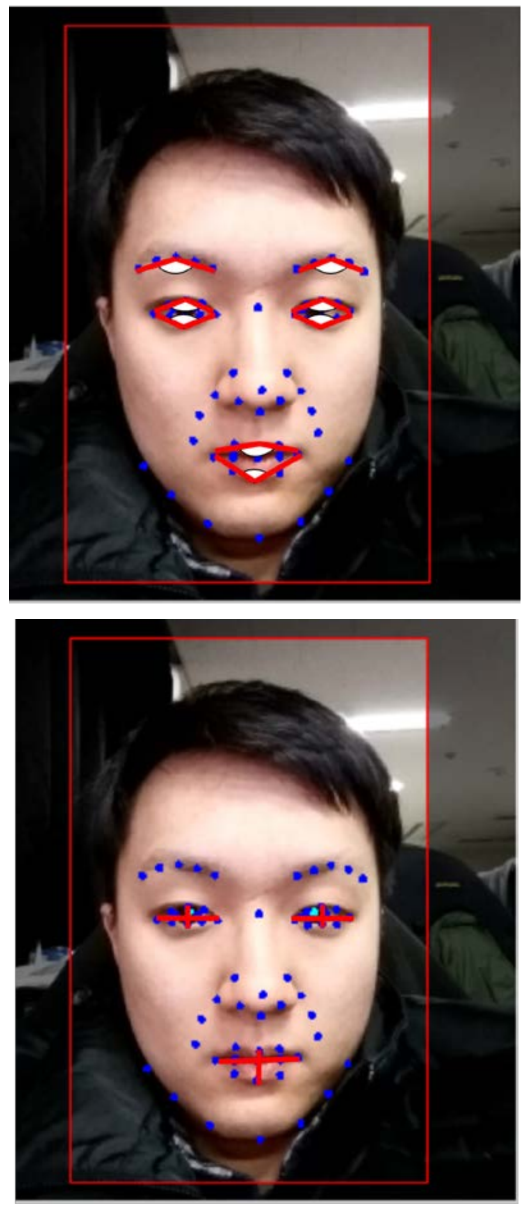

Fig. 3. Angle (top image) and distance (bottom image) features that are utilized in this work.

Table 1: List of the features that are extracted by utilizing angle and distance methods.

No.	Extracted Features
1	Left Eye brow Angle
2	Right Eye brow Angle
3	Left Eye width
4	Right Eye width
5	Left Eye Height
6	Right Eye Height
7	Left Eye Upper
8	Left Eye Low Angle
9	Right Eye Upper Angle
10	Right Eye Low Angle
11	Mouth Upper Angle
12	Mouth Low Angle
13	Mouth Width
14	Mouth Height

SVM has the capability to classify two or more classes using hyper-planes and we used an optimization technique in order to determine the optimum separating hyper-plane among different class labels as shown in Fig. 4.

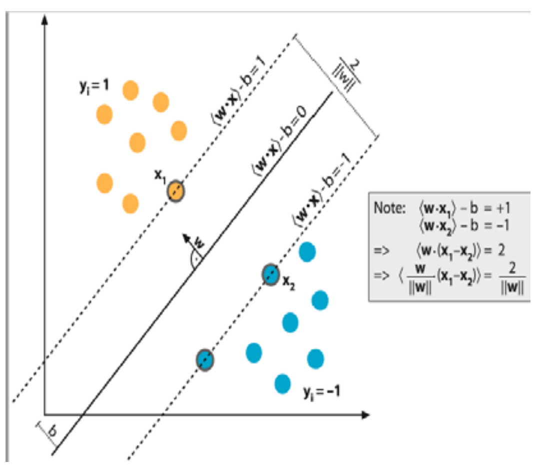

Fig. 4. A solid line indicates the optimal separating hyperplane [27].

Generally, SVM is given as below.

$$\langle N, \Phi \langle x \rangle \rangle + t = 0 \quad (11)$$

where N represents the normal vector to the hyper-plane that separates the two classes, Φ is the inherent function of the input data, x shows the data point, and t represents the training data. This corresponds to the resultant function as follows:

$$R(x) = sign(\langle N, \Phi \langle x \rangle \rangle + t) \quad (12)$$

where $R(x)$ is the resultant function that shows the training patterns; this is the so-called support vector that holds all the information about the classification problem.

4. Results and Discussion

The main purpose of this section is to evaluate abilities of the proposed smartphone-based FER system when operated on different settings of increasing complexity and practicality formulated in this work.

4.1 Dataset

In order to evaluate the performance of the proposed smartphone-based FER system in real time, 10 subjects (university students) were involved who performed five basic expressions like happy, anger, sad surprise, and normal. Each subject performed 173 instances for each expression in diverse scenarios. The age ranges of the subject were from 20 to 35 years and most of them were male. The tested dataset has been collected in realistic domain at different locations such as at University four court, bus stations, subway stations, and shopping malls at a speed of more than 15 fps (frames per second).

4.2 Setup

The proposed system has been tested and validated using Samsung Galaxy Note II withWeka and SVM using a Quad-Core^{TM} processor of speed 1.6 GHz Cortex–A9 and a RAM capacity of 2 GB. All the experiments were performed using 10-fold cross-validation rule in real time. In other words, each dataset was divided into 10 random subsets. Out of these 10 subsets, one subset was used as the validation data, whereas the remaining nine subsets were used as the training data, and this process (training and testing) was repeated 10 times.

4.3 Results

We performed the following three different experiments in order to assess the performance of the proposed FER system.

4.3.1 First Experiment

In this experiment, the proposed FER system was validated in real time. The overall results are shown in Table 2.
It can be seen from Table 2 that the proposed smartphone-based FER system showed better performance in real time

that is significant improvement in facial expression domain in real world scenarios.

Table 2: Recognition rates of the proposed of the smartphone–based FER system in real time (Unit: %).

Expressions	Happy	Anger	Sad	Surprise	Normal
Happy	**86**	3	5	4	2
Anger	2	**90**	3	4	1
Sad	6	3	**82**	4	5
Surprise	4	4	5	**81**	6
Normal	2	3	4	2	**89**
Average			**85.6**		

4.3.2 Second Experiment

In this experiment, we utilized one of the existing methods such as Different of Gaussian kernel [17] instead of using the proposed feature extraction method with angle and distance estimations. The results are shown in Table 3.

Table 3: Recognition rates of the proposed smartphone–based FER system with different of Gaussian kernel as in [17], while removing the proposed feature extraction method (Unit: %).

Expressions	Happy	Anger	Sad	Surprise	Normal
Happy	**71**	9	7	5	8
Anger	10	**67**	8	6	9
Sad	6	5	**74**	7	8
Surprise	9	7	8	**70**	6
Normal	5	10	9	11	**65**
Average			**69.4**		

It is obvious from Table 3 that the proposed feature extraction has an important role in achieving the higher recognition rates. This is because the proposed method extracts the most important information from expression frames in the form of frequency, which compactly supported (like wavelet) on RGB images along with least irregularity and maximum number of disappearing instants for a given support width. The frequency-based supposition is considered in our experiments, and we measure the dependency statistics of the extracted facial points for all the facial frames. Combined likelihood of an RGB frame is calculated by accumulating the geometrically associated frames of the expression for each facial point. Joint information for these facial points is measured using these distributions, which is used to approximate the asset of the dependency numbers between the two expression frames. Moreover, the proposed method has the capability to extract the important features from RGB images with the help of neighborhood in frequency, orientation and in space as well. As mentioned before that during the expressions, the pixels have dynamic position; therefore, in the proposed method, we also find the motion related information of the pixels that improve the recognition rate.

4.3.3 Third Experiment

In this experiment, the proposed smartphone-based FER system has been compared with of the existing system. The comparison results along with the proposed system are presented in Table 4.

Table 4 Recognition rates of the proposed smartphone–based FER system with different of Gaussian kernel as in [17], while removing the proposed feature extraction method (Unit: %).

Existing FER System	Average Classification Rates	Standard Deviation
[17]	68.1	±6.9
[16]	74.2	±5.6
[25]	80.2	±1.2
[28]	77.8	±2.0
Proposed System	85.6	±4.1

It is clear from Table 4 that the proposed system showed better performance and achieved higher recognition rate than of the existing systems in real world scenarios. Thus, the proposed smartphone-based FER system shows significant improvement which means that the proposed system can easily recognize the expressions using smartphone in real world scenarios.

5. Conclusion

Nowadays, smartphone has widely been utilized in many telemedicine and healthcare domains in order to monitor the mental status of the stroke patients. In such environments, the mental status of a human can be analyzed by expressions. Expressions have important role in improving the quality of interaction between humans. In daily life, communication through facial expressions plays a significant role. There some parts in the face (mouth, eyes, and eyebrow) that has much involvement in expression creation. Most of the existing FER system either considered whole face or recognized the expressions in controlled environments that are far from real-life scenarios, due to which their performance degrades in real world environments. Therefore, in this work, we proposed a real-time facial expression recognition (FER) system using smartphone camera. In order to make the system efficient and robust, we have extracted the features only from the contributing parts of the face. We have extracted the facial points from these parts of the face. Once the points are extracted, then the distance and angle are calculated between these points. We utilized support vector machine (SVM) under 10□fold cross validation setting for expressions classification. The system was tested and validated in real time by 10 university students (which are not professional). The weighted average recognition rate for the proposed smartphone-based FER system is 85.6% for 5 basic expressions at a speed of more than 15 fps (frames per second), which is a significant improvement in real-time domain. The computational time for the proposed system is approximately 0.4 seconds in order to recognize an expression. Thus, the proposed FER system outperformed the existing FER systems.

As can be seen that the performance of the proposed system is better than of the existing systems; however, the accuracy is still not good. Therefore, it is desirable that in future, we will either tune the previous feature extraction, and recognition techniques, or propose new ones in order to improve the accuracy of the proposed smartphone-based FER system in real world.

Acknowledgments

This research was supported by the Industrial Core Technology Development Program 203 (100490790, Development of mining core technology exploiting personal big data) funded by theMinistry of Trade, 204 Industry and Energy (MOTIE, Korea), and by Basic Science Research Program through the National Research 205 Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (No. 2014–050807).

References

[1] N. Agrawal, R. Cosgriff, and R. Mudur, “Mood detection: Implementing a facial expression recognition system,” CS229 project, 2009.

[2] I. H. Kraai, M. Luttik, R. M. de Jong, T. Jaarsma, and H. Hillege, “Heart failure patients monitored with telemedicine: patient satisfaction, a review of the literature,” Journal of cardiac failure, vol. 17, no. 8, pp. 684–690, 2011.

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

[4] A. Mehrabian, “Communication without words,” Psychology Today, vol. 2, no. 4, pp. 53–56, 1968.

[5] V. Bettadapura, “Face expression recognition and analysis: The state of the art,” Emotion, pp. 1–27, 2009.

[6] X. Wu and J. Zhao, “Curvelet feature extraction for face recognition and facial expression recognition,” in Natural Computation (ICNC), 2010 Sixth International Conference on, vol. 3, pp. 1212–1216, IEEE, 2010.

[7] M. Bartlett, J. Hager, P. Ekman, and T. Sejnowski, “Measuring facial expressions by computer image analysis,” Psychophysiology, vol. 36, no. 2, pp. 253–263, 1999.

[8] W. Sato, T. Fujimura, and N. Suzuki, “Enhanced facial emg activity in response to dynamic facial expressions,” International Journal of Psychophysiology, vol. 70, no. 1, pp. 70–74, 2008.

[9] M. H. Siddiqi, S. Lee, Y.-K. Lee, A. M. Khan, and P. T. H. Truc, “Hierarchical recognition scheme for human facial expression recognition systems,” Sensors, vol. 13, no. 12, pp. 16682–16713, 2013.

[10] “Luxand facesdk documentation.” http://www.luxand.com/download/Luxand FaceSDK Documentation.pdf. Accessed: 2016-03-18.

[11] M. H. Siddiqi, R. Ali, M. Idris, A. M. Khan, E. S. Kim, M. C. Whang, and S. Lee, “Human facial expression recognition using curvelet feature extraction and normalized mutual information feature selection,” Multimedia Tools and Applications, vol. 75, no. 2, pp. 935–959, 2016.

[12] M. H. Siddiqi, R. Ali, A. M. Khan, Y.-T. Park, and S. Lee, “Human facial expression recognition using stepwise linear discriminant analysis and hidden conditional random fields,” Image Processing, IEEE Transactions on, vol. 24, no. 4, pp. 1386–1398, 2015.

[13] M. H. Siddiqi, R. Ali, A. M. Khan, E. S. Kim, G. J. Kim, and S. Lee, “Facial expression recognition using active contour-based face detection, facial movement-based feature extraction, and non-linear feature selection,” Multimedia Systems, vol. 21, no. 6, pp. 541–555, 2015.

[14] M. H. Siddiqi, R. Ali, A. Sattar, A. M. Khan, and S. Lee, “Depth camera-based facial expression recognition system using multilayer scheme,” IETE Technical Review, vol. 31, no. 4, pp. 277–286, 2014.

[15] T. Ahsan, T. Jabid, and U.-P. Chong, “Facial expression recognition using local transitional pattern on gabor filtered facial images,” IETE Technical Review, vol. 30, no. 1, pp. 47–52, 2013.

[16] M. Suk and B. Prabhakaran, “Real-time mobile facial expression recognition system–a case study,” in Computer Vision and Pattern Recognition Workshops (CVPRW), 2014 IEEE Conference on, pp. 132–137, IEEE, 2014.

[17] K.-S. Cho, I.-H. Choi, and Y.-G. Kim, “Robust facial expression recognition using a smartphone working against illumination variation,” Appl. Math, vol. 6, no. 2S, pp. 403S–408S, 2012.

[18] S. Srivastava, “Real time facial expression recognition using a novel method,” arXiv preprint arXiv:1206.3559, vol. 4, no. 2, pp. 49–57, 2012.

[19] A. Saeed, A. Al-Hamadi, R. Niese, and M. Elzobi, “Frame-based facial expression recognition using geometrical features,” Advances in Human-Computer Interaction, vol. 2014, p. 4, 2014.

[20] H. Boughrara, M. Chtourou, C. B. Amar, and L. Chen, “Facial expression recognition based on a mlp neural network using constructive training algorithm,” Multimedia Tools and Applications, vol. 75, no. 2, pp. 709–731, 2016.

[21] C. Silva, A. Sobral, and R. T. Vieira, “An automatic facial expression recognition system evaluated by different classifiers,” pp. 208–212, 2014.

[22] D. Ghimire, J. Lee, Z.-N. Li, and S. Jeong, “Recognition of facial expressions based on salient geometric features and support vector machines,” Multimedia Tools and Applications, pp. 1–26, 2016.

[23] D. Ghimire, S. Jeong, J. Lee, and S. H. Park, “Facial expression recognition based on local region specific features and support vector machines,” Multimedia Tools and Applications, pp. 1–19, 2016.

[24] R. Cendrillon and B. Lovell, “Real-time face recognition using eigenfaces,” in Visual Communications and Image Processing 2000, pp. 269–276, International Society for Optics and Photonics, 2000.

[25] I. Song, H.-J. Kim, and P. B. Jeon, “Deep learning for real-time robust facial expression recognition on a smartphone,” in Consumer Electronics (ICCE), 2014 IEEE International Conference on, pp. 564–567, IEEE, 2014.
[26] “Feature points detection.” https://courses.cs.washington.edu/courses/cse455/09wi/Lects/lect6.pdf. Accessed: 2016-03-30.
[27] M. Weis, T. Rumpf, R. Gerhards, and L. Pl¨umer, “Comparison of different classification algorithms for weed detection from images based on shape parameters,” Bornimer Agrartechn. Ber, vol. 69, pp. 53–64, 2009.
[28] M. Suk and B. Prabhakaran, “Real-time facial expression recognition on smartphones,” in Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on, pp. 1054–1059, IEEE, 2015.

Muhammad Hameed Siddiqi is currently working as an Assistant Professor in Department of Computer and Information Sciences, Al-Jouf University, Sakaka, Kingdom of Saudi Arabia. He has completed his Bachelor of Computer Science (Hons) from Islamia College university of Peshawar, N-W.F.P, Pakistan in 2007, and Master and PhD from Ubiquitous Computing (UC) Lab, Department of Computer Engineering, Kyung Hee University, Suwon, Korea by 2012 and 2016, respectively. He was a Graduate Assistant at Universiti Teknologi PETRONAS, Malaysia from 2008 to 2009. His research interest is Image Processing, Pattern Recognition, Machine Intelligence and Activity Recognition, Facial Expression Recognition.

Madallah Alruwaili was born in Aljouf, KSA, in 1983. He received the B.E. (Hons) degree in computer science from the Aljouf University, KSA, in 2005, the M.S. degree in computer science from University of Science, Malaysia, in 2009, and the Ph.D degree in electrical and computer engineering from Southern Illinois University, Carbondale, Illinois, USA, in 2015. He is currently an Assistant Professor of Computer Engineering and Networks at Aljouf University, Sakaka, Aljouf, KSA. His research interests include image processing, image quality analysis, pattern recognition, computer vision, and biomedical imaging.

Jae Hun Bang was born in Korea. He has completed bachelor from the Department of Computer Engineering, Kyung Hee University, Suwon, South Korea. Currently, he is the student of MS leading to PhD in Ubiquitous Computing (UC) Lab, Department of Computer Engineering, Kyung Hee University, Suwon, Korea. His research interest is Signal Processing, Audio based Emotion Recognition, Sensor based Activity Recognition, Machine Learning.

Sungyoung Lee received his BS from Korea University, Seoul, Korea. He got his MS and PhD degrees in computer science from Illinois Institute of Technology (IIT), Chicago, USA in 1987 and 1991 respectively. He has been a professor in the Department of Computer Engineering, Kyung Hee University, Korea since 1993. He is a founding director of the Ubiquitous Computing Laboratory, and has been affiliated with a director of Neo Medical ubiquitous-Life Care Information Technology Research Center, Kyung Hee University since 2006. Before joining Kyung Hee University, he was an assistant professor in the Department of Computer Science, Governors State University, llinois, USA from 1992 to 1993. His current research focuses on ubiquitous computing and applications, wireless ad-hoc and sensor networks, context-aware middleware, sensor operating systems, real-time systems and embedded systems, activity and emotion recognition.
He is a member of ACM and IEEE.