❏     875

# Distinctive features for normal and crackles respiratory sounds using cepstral coefficients

**N. H. Mohd Johari, Noreha Abdul Malik, K. A. Sidek**
Department of Electrical and Computer Engineering, International Islamic University Malaysia, Malaysia

| Article Info | ABSTRACT |
|---|---|
| | Classification of respiratory sounds between normal and abnormal is very crucial for screening and diagnosis purposes. Lung associated diseases can be detected through this technique. With the advancement of computerized auscultation technology, the adventitious sounds such as crackles can be detected and therefore diagnostic test can be performed earlier. In this paper, Linear Predictive Cepstral Coefficient (LPCC) and Mel-frequency Cepstral Coefficient (MFCC) are used to extract features from normal and crackles respiratory sounds. By using statistical computation such as mean and standard deviation (SD) of cepstral based coefficients it can differentiate between crackles and normal sounds. The statistical computations of the cepstral coefficient of LPCC and MFCC show that the mean LPCC except for the third coefficient and first three statistical coefficient values of MFCC's SD provide distinctive feature between normal and crackles respiratory sounds. Hence, LPCCs and MFCCs can be used as feature extraction method of respiratory sounds to classify between normal and crackles as screening and diagnostic tool.<br><br> |

*Corresponding Author:*

Noreha Abdul Malik,
Department of Electrical and Computer Engineering,
International Islamic University Malaysia,
Jalan Gombak, 53100 Kuala Lumpur, Malaysia.
Email: norehaa@iium.edu.my

## 1.    INTRODUCTION

Traditionally, one of the methods used by physician to diagnose respiratory diseases is by chest auscultation using a stethoscope [1]. However, it is difficult to diagnose the lung condition by using only the stethoscope. Hence, modern computerized auscultation, CT scan and X-Ray are used by doctors to capture various distinctive parameters of the lung.

Respiratory sounds can be classified as normal and abnormal or adventitious. There are many types of adventitious sounds, such as crackles, pleural rubs, stridor, and wheezes (ronchi) where the abnormality of the pulmonary system can be the cause of these sounds [2]. It is analyzed and used to detect respiratory system related diseases like Chronic Obstructive Pulmonary Disease (COPD), asthma and bronchitis.

Crackles can be detected in lung or heart auscultation of COPD, pneumonia, heart failure and asbestosis patients. The presence of crackles helps doctor to diagnose these patients. Crackles are heard mostly during inspiration and sometimes it is overheard during expiration [3]. Crackle sounds are short, explosive nonmusical sounds that normally lasts less than 20 ms [4]. Explosive opening of the small airways caused lung and heart to produce this crackle sounds ranging from 100 to 2000 Hz or even higher [5].

There are many methods to extract the features of lung auscultation, for example, Discrete Wavelet Transform (DWT), LPCC, MFCC and others. In this paper, the LPCC and MFCC will be used to extract the features of crackles and normal respiratory sound and statistical computation will be performed to evaluate the features extracted.

## 2.    FEATURE EXTRACTION OF RESPIRATORY SOUNDS

In speech signal processing, MFCC is considered one of the most highly effective feature extraction method [6]. This is because in MFCC analysis, mel scale is used to wrap the frequency and it is approximately close to the human auditory perception [7]. Chin et. al. used MFCC as the feature extraction of the lung sounds to classify between normal and abnormal sounds including crackles, wheezes and ronchi. K-Nearest Neighbor Classifier (k-NN) was used as the classifier which differentiate between normal and abnormal sounds [8].

Meanwhile, Nandini et al. [9] used MFCC for the feature extraction and Artificial Neural Network (ANN) as the classifier and yield different results. Nandini et. al. in their study used statistical features extracted from MFCC and other feature extraction methods such as Linear Frequency Cepstral Coefficient (LFCC), Perceptual Linear Prediction Coefficient (PLPCC) and several others. The classification using features extracted from MFCC yield a better result as compared to using other feature extraction in terms of the accuracy, specificity and sensitivity [9].

In a study by Fatma et al. [10], the normal and the asthmatic breath sounds were classified using wavelet transform as the feature extraction. The wavelet transform is used for analyzing the sounds segment and to characterize the local regularity of the signals by decomposing the signals into elementary building blocks with well localized time and frequency.The computational burden of the wavelet transform is reduced using DWT. Wavelet packet transform (WPT) is used as the wavelet transform extension. The DWT decomposed the signal into lower frequency band and higher frequency band. Meanwhile, the wavelet packet transform gives a balanced binary tree structure by decomposing both the lower frequency and higher frequency into two sub bands [10]. Fatma et al. compared the result of DWT and WPT as the analyzer of the respiratory sounds. The result showed that the use of DWT as the analyzer gaves slightly better accuracy as compared to WPT. This is because DWT is more effective in processing load and computational time [10].

Grønnesby et al. [11] used 5-dimensional feature vector as the feature extraction. They extracted four features from the time domain such as range, variance, sum of simple moving average (coarse) and sum of simple moving average (fine) and another one feature from frequency domain which is the spectrum mean. As reported in the study, the advantage of using simple summary statistic is that it is easy to relate to actual data whereas the disadvantage is many information are lost [11].

In a study by Abdul Malik et al. [12], fifteen different features are extracted from each segment of the respiratory sounds and Artificial Neural Network (ANN) was used as the classifier. DWT was used to decompose the respiratory sounds into seven different frequency band based on Daubechies (db7) and Haar mother wavelet. Mean, standard deviation and maximum power spectral density were calculated from five frequency band (D3, D4, D5, D6, and D7) and these features value were used as the input of the ANN. Result of the study showed that db7 outperform Haar with perfect 100% sensitivity, accuracy and specificity in both testing and validation stage by using 15 nodes at the hidden layer. Meanwhile, using 10 nodes at the hidden layer, Haar was able to obtain perfect 100% sensitivity, specificity and accuracy for testing stage only [12].

A study to analyze the performance of the Automatic Speech Recognition using MFCC and LPCC by M Rana et al. showed that MFCC gave better performance to the system [13]. Automatic Speech Recognition System using MFCC was 80 percent accurate, while LPCC gave only 60 percent accuracy. Meanwhile, a study by Azmy [14] to classify abnormal lung sounds (stridor and polyphonic) using DWT and LPCC resulted in high recognition percentage, 95.24 percent. The LPCCs were calculated from the 3-level coefficients of DWT. LPCC, delta LPCC and delta-delta LPCC were extracted and the variance and kurtosis were calculated. Next, the study used Support Vector Machine (SVM) to classify the features extracted [14].

Previous study by Johari et al. [15] analyzed the statistical value of the features extracted using MFCC and showed that the first three coefficients were able to distinguish between normal and crackles sounds. This paper is the extended paper of the study that will analyze the statistical value of cepstral coefficient of LPCC.

## 3.    STATISTICAL CEPSTRAL COEFFICIENT USING LPCC AND MFCC

### 3.1.  Data acquisition and pre-processing

The respiratory sound signals were collected from 20 healthy subjects and 23 lung cancer patients at the University Malaya Medical Centre (UMMC) [7]. The study was approved by medical ethics committee of University Malaya Medical Centre (UMMC) with reference number (MREC ID NO: 201698-4242). The signals were recorded using One Thinklabs stethoscope with sampling frequency 11025 Hz. The data were recorded, saved and pre-processed using Thinklabs Phonocardiography by Audicity. For the pre-processing, the signals were filtered from unwanted noise, outliers, and artifacts. The lung sounds were enhanced and the heart sounds were filtered out. Next, 30 crackles and 30 normal segment sounds were

extracted from the pre-processed signals. The segment includes inhale and exhale sounds. The methodology of this research is shown in Figure 1.
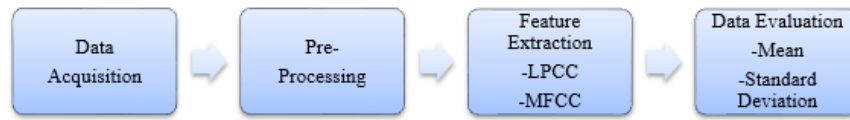


Figure 1. The proposed methodology of the study

### 3.2. Features extraction using LPCC
The features of crackles and normal sounds were extracted using LPCC by using MATLAB. The flow of the LPCC feature extraction is shown in Figure 2. Finite Impulse Response (FIR) digital filter is used to pre-emphasis the signal and hamming window is used for windowing the signal. In the next step, the filter coefficients are find using autocorrelation method of autoregressive (AR) modeling. The signals are applied to Levinson-Durbin Recursion to find the Linear Predictive Coding (LPC) and next, the Cepstral Coefficient of the linear predictive coding (LPCCs) is calculated using (1). The output cepstral coefficient is set to 10 coefficients.

$$Cepstrum = ifft(log(PowerSpectrum), 1024) \tag{1}$$



Figure 2. Flow of LPCC to extract features from crackles and normal sounds

### 3.3. Features extraction using MFCC
Another feature extraction method used is MFCC. The flow of the MFCC feature extraction is shown in Figure 3. (2) which is the first order difference equation is applied to the sample of the signal at each window to pre-emphasis the signal:

$$s'_n = s_n - ks_{n-1} \tag{2}$$

where $\{s_n, n=1, N\}$ and $k$ is the pre-emphasis coefficient which should be in the range $0 \leq k < 1$.
The crackles duration is around 20 ms and therefore, the signal is divided into 20 ms frames. Meanwhile, frame shift is set to 10 ms (50% overlap) to allow some overlap to the frame as to not miss any signal. Hamming window is applied after the pre-emphasis process and next, the power spectrum is calculated using fast fourier transform (FFT).
Mel-filterbank is applied to the periodogram power spectral from the previous step to compute the mel-spaced filterbank. The number of filterbank channels is set at 26 where every channel is to indicate the amount of energy in each filterbank, called filterbank energies (FBEs). Equation (3) and (4) is the formula to convert the frequency to Mel Scale frequency and vice versa in order to obtain the filterbank. The formula to convert the signal from frequency to Mel-Scale frequency is shown in (3):

$$M(f) = 2595 \, log_{10}(1 + \frac{f}{100}) \tag{3}$$

The formula to convert Mel-Scale frequency to frequency in Hz is shown in (4):

$$M^{-1}(m) = 700(100^{\frac{m}{2595}} - 1) \tag{4}$$

Discrete Cosine Transform (DCT) is applied to the logarithm of FBEs to get the 26 cepstral coefficients of filterbank channels. However, only the first 13 coefficients are kept to give a better performance as the higher DCT coefficients show fast changes in the FBES.

*Distinctive features for normal and crackles respiratory sounds using cepstral… (N. H. Mohd Johari)*

Figure 3. Flow of MFCC to extract features from crackles and normal sounds

### 3.4. Mean and standard deviation of LPCCs

There are 30 crackles and normal segments used in the feature extraction and 10 Linear Predictive Cepstral Coefficients (LPCCs) are extracted for each segment. The mean and standard deviation of the first coefficient for all segments are calculated and analyzed using (5) and (6). This process is continued for the next 9 coefficients.

### 3.5. Mean and standard deviation of MFCCs

There are 13 coefficients extracted from each frames of a segment. Thus, there are 13 mean values ($\mu 1$- $\mu 13$) and SD values ($\sigma 1$- $\sigma 13$) coefficients calculated for all frames of a segment and for all segments. The mean of MFCCs and SD are calculated using (5) and (6), respectively. The average of mean and SD for all the segments are also calculated and analyzed. Mean/average of the coefficients is calculated as in (5):

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} x_{ij} \tag{5}$$

SD of the coefficients is calculated as in (6):

$$\sigma_i = \sqrt[2]{\frac{1}{N}\sum_{j=1}^{N}(x_{ij} - \mu_i)^2} \tag{6}$$

where $i$ ($i=1,2,\ldots 13$) is the coefficient and $j$ ($j=1,2,\ldots N$) is the frame of the segment. $N$ is the number of frames.

Mean and SD calculation are used in this study because the presence of a few abnormally high values of MFCCs coefficients have effect to the mean value and the best method to measure the variation of these MFCCs coefficients is by using SD. These statistical values evaluate the pattern of the MFCCs coefficients in every segment and analyze which statistical value will show distinct outcome between crackles and normal lung sounds.

## 4.    RESULTS AND ANALYSIS

MFCC and LPCC are widely used in speaker identification as well as in speech researches. In this study, MFCC and LPCC are used to extract the features of the respiratory sounds. The mean and SD calculated from the features extracted are plotted. Figure 4 and Figure 5 show the mean and SD values of the LPCCs for each segment.

Figure 4 shows that the mean value of LPCCs is separable except at the third coefficient. Meanwhile, Figure 5 shows that LPCCs statistical features for SD shows undistinguishable result. These results are proved using t-test tabulated in Table 1. The t-test is calculated from the value of mean and SD of the LPCCs to define the hypotheses with hypotheses that the statistical value, mean and SD of LPCCs can distinguish between normal and crackles sounds. The t-score calculated is used to find the range of the p-value with the significance level is set at 0.05.
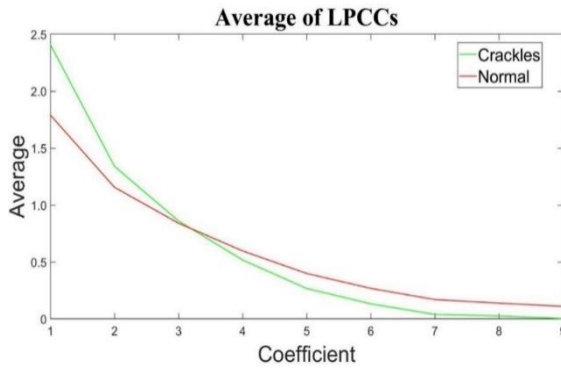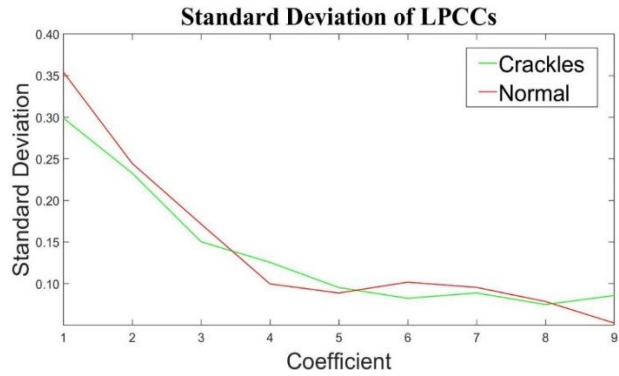
Figure 4. Mean value of LPCCs                  Figure 5. SD value of LPCCs

Table 1. T-Score and P-value calculated for average and standard deviation of the 9 LPCCs

| LPCC Coefficient | Mean | | Standard deviation | |
|---|---|---|---|---|
| | t-score | p-value | t-score | p-value |
| 1 | 7.322819 | The P-Value is <0.00001 | 0.653449 | The P-Value is 0.2580 |
| 2 | 2.981043 | The P-Value is 0.0021 | 0.189453 | The P-Value is 0.4252 |
| 3 | 0.467547 | The P-Value is 0.3209 | 0.511509 | The P-Value is 0.3055 |
| 4 | 2.710906 | The P-Value is 0.0044 | 0.883819 | The P-Value is 0.1902 |
| 5 | 5.561750 | The P-Value is <0.00001 | 0.275194 | The P-Value is 0.3921 |
| 6 | 5.688879 | The P-Value is <0.00001 | 0.812154 | The P-Value is 0.2100 |
| 7 | 5.477963 | The P-Value is <0.00001 | 0.278098 | The P-Value is 0.3910 |
| 8 | 5.676406 | The P-Value is <0.00001 | 0.184743 | The P-Value is 0.4270 |
| 9 | 5.851847 | The P-Value is <0.00001 | 1.808894 | The P-Value is 0.0378 |

Table 1 proved that mean of LPCCs are distinguishable except at coefficient three where the p-value does not fall in the significance range, 0.05. Meanwhile, calculated t-test of the SD of LPCC shows that most of the p-value does not fall in the significance range, thus proving that SD of the LPCCs are not separated. Next, average of mean and SD of the set of MFCCs for each frames of a segment are also calculated and analyzed. Figure 6 illustrates the average of the mean of MFCC values between frames in each 30 segments. It shows that the crackles and normal have similar values and are not separated. Meanwhile, Figure 7 shows the average of the SD of the MFCC values between frames in each segment. The first few coefficients show distinguishable result between crackles and normal respiratory sounds.
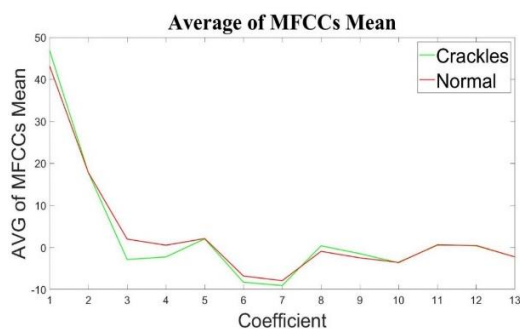


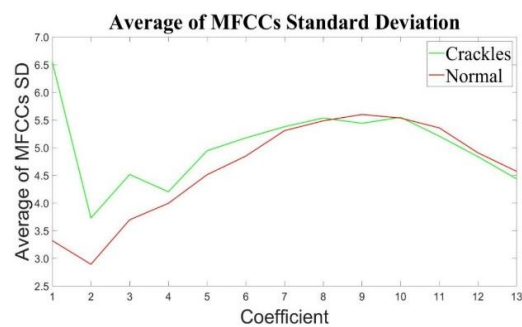Figure 6. Average value of MFCCs mean       Figure 7. Average value of MFCCs standard
between frames of a segment              deviation between frames of a segment

These results are also proven using the t-test calculation. T-test is calculated to define the hypotheses that the values of average mean and SD of MFCCs can distinguish between normal and crackles sounds. The results of the calculated t-test for average mean and average SD MFCCs are tabulated in Table 2. It shows that for the first three p-value of 13 SDs of the MFCCs' falls within the significance range, and most of the average MFCCs' mean is not separated.

Table 2. T-score and P-value calculated for 13 mean and standard deviation of MFCCs

| MFCC Coefficient | Average mean | | Average standard deviation | |
|---|---|---|---|---|
| | t-score | p-value | t-score | p-value |
| 1 | 3.199543707 | The p-Value is 0.0011. | 6.139165159 | The P-Value is <0 .00001 |
| 2 | 0.030327686 | The P-Value is 0.4880 | 3.852059716 | The P-Value is 0.0002 |
| 3 | 3.733222963 | The P-Value is 0.0002 | 3.483623261 | The P-Value is 0.0005 |
| 4 | 3.493882175 | The P-Value is 0.0005 | 1.574744534 | The P-Value is 0.0604 |
| 5 | 0.104253068 | The P-Value is 0.4587 | 3.250215986 | The P-Value is 0.0010 |
| 6 | 2.187802033 | The P-Value is 0.0164 | 3.267762884 | The P-Value is 0.0009 |
| 7 | 1.613127805 | The P-Value is 0.0561 | 0.65028139 | The P-Value is 0.2591 |
| 8 | 2.576424622 | The P-Value is 0.0063 | 0.52311489 | The P-Value is 0.3014 |
| 9 | 2.614309261 | The P-Value is 0.0057 | 1.714367015 | The P-Value is 0.0460 |
| 10 | 0.22179613 | The P-Value is 0.4127 | 0.161645305 | The P-Value is 0.4361 |
| 11 | 0.264513728 | The P-Value is 0.3962 | 1.415442366 | The P-Value is 0.0812 |
| 12 | 0.346124198 | The P-Value is 0.3653 | 0.787394324 | The P-Value is 0.2172 |
| 13 | 0.007517145 | The P-Value is 0.4970 | 1.580656933 | The P-Value is 0.0597 |

## 5. CONCLUSION

Based on the result of the experimentation, the statistical values of the mean LPCC's except for the third coefficient can indicate the existence of crackles in the respiratory sounds same as the first three SDs of MFCCs. The t-test calculation supports these results. On the other hand, the SD LPCC and average of MFCCs' mean were unable to distinguish between normal and crackles sounds. In order to classify the respiratory sounds, there are two major steps needed. First, feature extraction and second, classification. In this research, MFCC technique is studied and used as the feature extraction. The statistic cepstral coefficient values are analyzed and the output shows that the statistic cepstral values are able to distinguish the features between normal and crackles respiratory sounds. Further research needs to be done to find the most appropriate classifiers that is suitable with these feature extraction technique in order to distinguish respiratory sounds between normal sound in healthy subjects and crackles in lung cancer patients.

## ACKNOWLEDGEMENTS

## REFERENCES

[1] Melbye, H., Garcia-Marcos, L., Brand, P., Everard, M., Priftis, K. and Pasterkamp, H., "Wheezes, crackles and rhonchi: simplifying description of lung sounds increases the agreement on their classification: a study of 12 physicians' classification of lung sounds from video recordings." *BMJ open respiratory research* 3, no. 1 (2016): e000136.

[2] Malay, I. Madabhavi, N. Niranjan, and M. Dogra. "Auscultation of the respiratory system." *Annals of thoracic medicine*, *10*(3), p.158, 2015.

[3] Pramono, R.X.A., Bowyer, S. and Rodriguez-Villegas, E. "Automatic adventitious respiratory sound analysis: A systematic review." *PloS one* 12, no. 5 (2017): e0177926

[4] Reichert, S., Gass, R., Brandt, C. and Andrès, E., "Analysis of respiratory sounds: state of the art." *Clinical medicine. Circulatory, respiratory and pulmonary medicine* 2 (2008): CCRPM-S530.

[5] Zhang, Kexin, Xuefeng Wang, Fangfang Han, and Hong Zhao. "The detection of crackles based on mathematical morphology in spectrogram analysis." *Technology and Health Care* 23, no. s2 (2015): S489-S494.

[6] P., R., K. Sundaraj, and S. Sundaraj. "A comparative study of the svm and k-nn machine learning algorithms for the diagnosis of respiratory pathologies using pulmonary acoustic signals." *BMC bioinformatics*, *15*(1), p.223, 2014.

[7] Tiwari, V., "MFCC and its applications in speaker recognition." *International journal on emerging technologies* 1, no. 1 (2010): 19-22.

[8] Chen, Chin-Hsing, W.T. Huang, T.H. Tan, C.C. Chang, and Y.J. Chang. "Using k-nearest neighbor classification to diagnose abnormal lung sounds." *Sensors*, *15*(6), pp.13132-13158, 2015.

[9] Sengupta, Nandini, M. Sahidullah, and G. Saha. "Lung sound classification using cepstral-based statistical features." *Computers in biology and medicine*, *75*, pp.118-129, 2016.

[10] F. Z. Göğüş, B. Karlik, and Güneş Harman. "Classification of asthmatic breath sounds by using wavelet transforms and neural networks." International Journal of Signal Processing Systems, 3(2), pp.106-111, 2015.

[11] M. Grønnesby, J. C. A. Solis, E. Holsbø, H. Melbye, and L. A. Bongo. "Feature Extraction for Machine Learning Based Crackle Detection in Lung Sounds from a Health Survey."arXiv preprint arXiv:1706.00005(2017).

[12] Malik NA, Idris W, Gunawan TS, Olanrewaju RF, Ibrahim SN. "Classification of Normal and Crackles Respiratory Sounds into Healthy and Lung Cancer Groups". *International Journal of Electrical and Computer Engineering (IJECE)*. 2018 Jun ;8(3):1530-1538.

[13] Rana, M., & Miglani, S. "Performance Analysis of MFCC and LPCC Techniques in Automatic Speech Recognition". *International Journal Of Engineering And Computer Science*. 3(08). 2014.

[14] M. M. Azmy, "Classification of lung sounds based on linear prediction cepstral coefficients and support vector machine," *2015 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*, Amman, 2015, pp. 1-5.

[15] N. H. M. Johari, N. A. Malik and K. A. Sidek, "Distinctive Features for Classification of Respiratory Sounds Between Normal and Crackles Using Cepstral Coefficients," *2018 7th International Conference on Computer and Communication Engineering (ICCCE)*, Kuala Lumpur, 2018, pp. 476-479.

## BIOGRAPHIES OF AUTHORS

Nabila Husna Mohd Johari received her BEng in Communication Engineering (2017) and currently pursuing her Meng in Computer and Information Engineering in International Islamic University Malaysia (IIUM). She is currently a research assistant at IIUM. Her research interest is in Biomedical Signal Processing.

Noreha Abdul Malik received her BEng in Medical Electronics from University of Technology Malaysia (2001) and later pursued her MEng in Communication and Computer Engineering at National University of Malaysia (2004). She later received her PhD in Electronics and Electrical Engineering from University of Southampton, United Kingdom (2011). She is currently an assistant professor at International Islamic University Malaysia (IIUM). Her research interests are in biomedical signal processing and biomedical applications. She is a member of Institute of Engineers Malaysia (IEM) and Board of Engineer Malaysia (BEM).

Khairul Azami Sidek received a B.Eng. degree in Computer and Information Engineering from the International Islamic University Malaysia (IIUM) in 2004, and a M.Eng. degree in Computer and Communication Engineering from Universiti Kebangsaan Malaysia in 2006. He obtained a Ph.D. degree in Computer Science from the Royal Melbourne Institute of Technology, Melbourne, Australia in 2014. He is currently an associate professor at the Kulliyyah of Engineering, IIUM and his research interests include biometric, biomedical signal processing, and pattern recognition.