

# Article

DOI: 10.1111/j.1468-0394.2008.00483.x

## Electromyography signal analysis using wavelet transform and higher order statistics to determine muscle contraction

M.S. Hussain,<sup>1</sup> M.B.I. Reaz,<sup>2</sup> F. Mohd-Yasin<sup>3</sup> and M.I. Ibrahimy<sup>1</sup>

(1) Department of Electrical and Computer Engineering, International Islamic University Malaysia, Gombak, 53100 Kuala Lumpur, Malaysia

(2) Department of Electrical, Electronic and Systems Engineering, National University of Malaysia, 43600 UKM, Bangi, Selangor, Malaysia  
Email: mamun.reaz@gmail.com

(3) Faculty of Engineering, Multimedia University, 63100 Cyberjaya, Selangor, Malaysia

**Abstract:** Electromyography gives an electrical representation of neuromuscular activation associated with a contracting muscle. The electromyography signal acquires noise while travelling through different media. The wavelet transform is employed for removing noise from surface electromyography (SEMG) and higher order statistics are applied for analysing the signal. With the appropriate choice of wavelet, it is possible to remove interference noise (denoise) effectively in order to analyse the SEMG. Daubechies wavelets (db2, db4, db5, db6, db8), symmlet (sym4, sym5) and the orthogonal Meyer (dmey) wavelet can efficiently remove noise from the recorded SEMG signals. However, the most effective wavelet for SEMG denoising is chosen by calculating the root mean square difference and signal-to-noise ratio values. Results for both root mean square difference and signal-to-noise ratio show that wavelet db2 performs denoising best out of the wavelets. Furthermore, the higher order statistics method is applied for SEMG signal analysis because of its unique properties when applied to random time series, such as parameter estimation, testing of Gaussianity and linearity, deterministic and non-deterministic signal detection etc. Gaussianity and linearity tests as part of higher order statistics are conducted to understand changes in muscle contraction and to quantify the effectiveness of the noise removal process. According to the results, the SEMG signal becomes less Gaussian and more linear with increased force.

**Keywords:** electromyography, motor unit, muscle contraction, wavelet, denoising, higher order statistics

### 1. Introduction

The electromyography (EMG) signal represents the electrical activity of muscles. A muscle is composed of many motor units. EMG signals are usually detected via surface electrodes attached on the skin and show a train of motor unit action potentials (MUAPs) corrupted with noise. The simplest method of removing narrow

bandwidth interference from recorded signals is to use a linear recursive digital notch filter. But the disadvantage of the notch filter is that it distorts the signal (Mewette *et al.*, 2001). Wavelet-based noise removal (denoising) is performed in this research as the initial element of EMG signal processing and analysis. Wavelet denoising has already been used in denoising a number of physiological signals (Carre *et al.*,

1998; Mark *et al.*, 2000). The method is preferred over signal frequency domain filtering because it can maintain signal characteristics even while reducing noise. This is because a number of threshold strategies are available, allowing reconstruction based on selected coefficients. In this research, wavelet functions (WFs) Daubechies (db2, db4, db5, db6, db8), symmlet (sym4, sym5) and orthogonal Meyer (dmey) are used for the wavelet transform (WT). The wavelet denoising technique is applied to remove interference noise (inherent noise, motion artefact) caused while recording surface EMG from the rectus femoris muscle of three subjects during walking trials. Both root mean square (RMS) difference and signal-to-noise ratio (SNR) values of the denoised signals are calculated to quantify the most effective WF.

The technique of studying muscle function by using a surface electrode is normally known as surface electromyography (SEMG). With increasing muscle force, the raw EMG signal shows an increase in the number of MUAPs recruited at increasing firing rates. The firing pulses are normally considered a random function of time, which is non-Gaussian in nature (Nikias & Mendel, 1993; Kaplanis *et al.*, 2000; Reaz *et al.*, 2006b). Reaz *et al.* (2006b) proposed a method for SEMG signal analysis using the WT with various WFs. But the selection of the WF was not appropriate for SEMG signal processing and analysis because the choice of the mother wavelet is an important issue. Furthermore, the experimental conditions were not explained well and the conclusions were of limited interest due to the limitations to one subject only. The proposed method in this paper gives a firm algorithm and conclusions supported by sound evidence.

Furthermore, this study exploits the use of higher order statistics (HOS) in SEMG signal analysis in order to extract new parameters that could enhance the diagnostic character of SEMG. The conventional techniques of signal processing are generally based on the first- and second-order moments and cumulants and their spectral analysis. These techniques do not provide the full information available for the

signal. Higher order moments, cumulants and their frequency representation can offer advantages and access more information about the signal.

Higher order spectra defined in terms of the higher order cumulants of a random signal serve as useful tools for investigating non-Gaussianity and non-linearities. Moreover, the HOS method is applied for suppressing Gaussian white noise. A fundamental means by which probability theory can describe a random process is by using the different order moments and cumulants. The first-order moment or cumulant of any process indicates the mean of that process. The second-order moment of the process indicates the autocorrelation, while its second-order cumulant indicates the variance. HOS start from the third-order to  $n$ th-order moment (Nikias & Mendel, 1993). Along with other HOS measures, the bispectrum, a frequency-domain measure of third-order cumulants, has been developed as a tool of signal analysis to provide information about signals' Gaussianity and linearity.

After removing the interference noise, bispectrum analysis is introduced in this paper for analysing the SEMG signals. The Gaussianity test and linearity tests of the normalized bispectrum show changes in muscle contraction during the walking trials. The analysis also determines the effectiveness of the wavelet based denoising method. Results in this study show that the wavelet based noise removal technique using WF db2 works best to remove interference noise from the SEMG signals. The bispectrum analysis shows that the SEMG becomes less Gaussian and more linear with increased walking speed/force (increase in mean voluntary contraction). The signal after denoising is free from interference noise, which enhances the bispectrum analysis of the SEMG signals.

## 2. Methodology

The sample raw EMG signals were collected from three normal subjects aged 22–40 at Universiti Kebangsaan Malaysia. For this experiment, six separate EMG data files were used

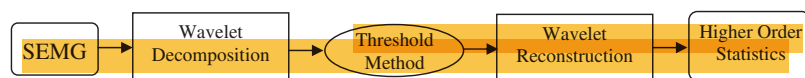
(three subjects were considered for the experiment and SEMG signals were recorded from two channels). The SEMG signals were recorded from the right rectus femoris muscle and all analogue channels were recorded at 1 kHz. Two surface electrodes were used to capture the raw EMG signals and the distance between the electrodes was 3.5 inches. SEMG signals were captured during the subjects' walking trial in which all the subjects increased their walking speed/force with time. Recording was performed for 3 minutes for each subject. The subjects were asked to increase their walking speed in every 45 second interval (categorized into slow walking speed, medium walking speed, fast walking speed, very fast walking speed). During slow walking speed there was least muscle contraction from the rectus femoris muscle. The muscle contraction was higher from slow walking speed to very fast walking speed. The speed of walk directly affects the rectus femoris muscle and the more force given to the muscle generates higher muscle contraction (Kaplanis *et al.*, 2000; Reaz *et al.*, 2006b). The recorded SEMG signals were denoised using the discrete wavelet transform (DWT) and a threshold method. The DWT and threshold based denoising was implemented with MATLAB Wavelet toolbox. The algorithm for bispectrum estimation was implemented with MATLAB Signal Processing Toolbox. Figure 1 gives the flow of the algorithm which is able to remove interference noise using the WT and analyse the SEMG signals using HOS (bispectrum analysis). The details of the block diagram are explained in the following sections.

### 2.1. Wavelet denoising

Wavelet theory stems from a field of study known as multiresolution analysis. Wavelet techniques can localize both time and frequency

components, as signals are processed and analysed at various scales, or resolutions. Moreover, wavelet processing provides good frequency resolution at high frequencies, so the noise components in a signal can be isolated while important high-frequency transients can also be preserved. Since noise is easily isolated in the wavelet domain, it can be removed while leaving important components intact. Additionally, there are a large number of WFs from which to select the one that most closely fits the specific application (Mark *et al.*, 2000). Wavelets commonly used for denoising biomedical signals include the Daubechies db2, db4, db5, db6 and db8 wavelets and the orthogonal Meyer wavelet. Wavelets for SEMG are generally chosen with shapes similar to those of the MUAP (Laterza & Olmo, 1997; Mark *et al.*, 2000; Ren *et al.*, 2006).

From the signal processing point of view EMG signals are multicomponent signals, made up of the superimposition of MUAPs characterized by individual time and frequency localizations. Moreover, it can be shown that under certain conditions the various MUAPs can be considered as scaled versions of a single wave (Gabor, 1946). MUAPs due to deeper motor units are dilated with respect to those due to more superficial ones. Therefore, the global signal can be modelled as the superimposition of delayed and scaled versions of a basic component. The WT represents a very suitable analysis method for this class of signals. In particular, if the analysis wavelet is chosen to match the shape of the MUAP, the resulting WT yields the best possible energy localization in the time-scale plane (Guglielminotti & Merletti, 1992). The ability of the WT to extract features from the signal is dependent on the appropriate choice of the mother WF. Even though there is no well-defined rule for selecting a wavelet basis function in a particular applica-



**Figure 1:** Block diagram of wavelet based denoising and HOS analysis of SEMG signals.

tion or analysis, some properties of the wavelets make a specific mother wavelet more suitable for a given application and signal type. There are a number of basis functions that can be used as the mother wavelet for WTs. Since the mother wavelet produces all WFs used in the transformation through translation and scaling, it determines the characteristics of the resulting WT. Therefore, the details of the particular application should be taken into account and the appropriate mother wavelet should be chosen in order to use the WT effectively. The wavelets are chosen based on their shape and their ability to process the signal in a particular application. For this research, the wavelets are chosen for EMG signal processing. The wavelet based denoising method consists of three steps. First, the WT of a signal is computed. Next, a threshold for the resulting wavelet coefficients is determined. Coefficients whose absolute values are below this threshold are set to zero. Finally, the inverse WT is computed on the modified set of coefficients, resulting in a cleaned signal. The following sections explain the wavelet decomposition, the thresholding process and the wavelet reconstruction process.

**2.1.1. Wavelet decomposition** The DWT is computed by successive low-pass and high-pass filtering in the discrete-time domain. The DWT of a signal  $x[n]$  is calculated by passing it through a series of filters. First the samples  $x[n]$  are passed through a low-pass filter with impulse response  $g[n]$ , resulting in a convolution  $y[n]$  of the two. The signal also goes simultaneously through a high-pass filter with impulse response  $h[n]$ . The outputs give the detail coefficients (from the high-pass filter,  $y_{\text{high}}[n]$ ) and the approximation coefficients (from the low-pass filter,  $y_{\text{low}}[n]$ ). The filter outputs are then down-sampled (or subsampled) as given by

$$y_{\text{low}}[n] = \sum_{k=-\infty}^{\infty} x[k].g[2n-k] \quad (1)$$

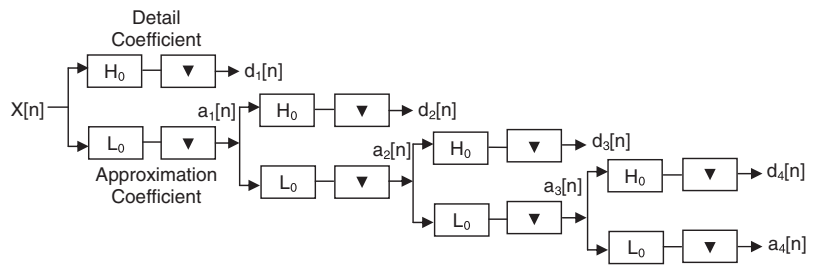
$$y_{\text{high}}[n] = \sum_{k=-\infty}^{\infty} x[k].h[2n-k] \quad (2)$$

respectively. For this research WFs Daubechies (db2, db4, db5, db6, db8), symmlet (sym4, sym5) and orthogonal Meyer (dmey) are used for the decomposition. Wavelet based noise removal exploits the time-scale characteristics of the DWT. Normally, noise consists of high-frequency components and is thus localized in the finer, detailed level of the transform. In these levels, the important coefficients, corresponding to true signal information, have a relatively high magnitude, and the lower magnitude coefficients represent the noise. Previous research shows that four to five levels of decomposition are suitable for EMG signals (Moshou *et al.*, 2000). However, four levels of decomposition are considered for the SEMG signals in this paper. Figure 2 gives the four levels of wavelet decomposition.

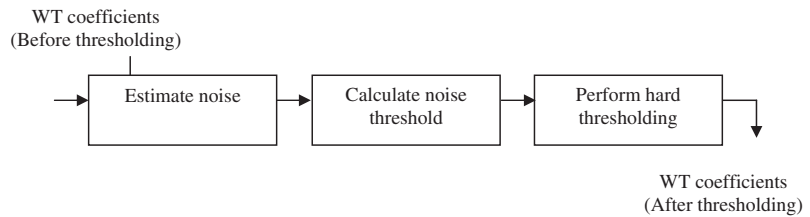
According to Figure 2, the raw EMG signal  $x[n]$  is decomposed to get the detail coefficient ( $H_0$ ) and approximation coefficient ( $L_0$ ) during the first level of decomposition. The detail coefficients  $d_1[n]$  are stored and the approximation coefficients  $a_1[n]$  are further decomposed into the second-level detail coefficients and approximation coefficients. As mentioned earlier, the decomposition process continues until the fourth level when using the SEMG signal. After the four levels of WT decomposition,  $d_1[n]$ ,  $d_2[n]$ ,  $d_3[n]$  and  $d_4[n]$  detail coefficients and  $a_4[n]$  approximation coefficients are the resulting coefficients of the transform. The threshold can then be chosen such that most of the coefficients are set to zero, while the few important coefficients are untouched.

**2.1.2. Threshold method** A threshold is determined for the raw EMG signals which is applied on the wavelet coefficients ( $d_1[n]$ ,  $d_2[n]$ ,  $d_3[n]$ ,  $d_4[n]$ ,  $a_4[n]$ ) after the WT. The thresholding process of the wavelet coefficients is illustrated in Figure 3.

According to Figure 3, the WT coefficients are used to estimate the noise and calculate the threshold. Hard thresholding is performed on the WT coefficients based on the calculated threshold. The detail of the process is given in the following paragraphs.



**Figure 2:** Four levels of wavelet decomposition.



**Figure 3:** Thresholding process on wavelet coefficients.

*Estimation of noise level:* Calculate the median absolute deviation  $\delta_{\text{mad}}$  on the largest coefficient spectrum by

$$\delta_{\text{mad}} = \frac{\text{median}\{|c_0|, |c_1|, \dots, |c_n - 1|\}}{0.6745} \quad (3)$$

where  $|c_0|, |c_1|, \dots, |c_n - 1|$  are the wavelet coefficients and 0.6745 in the denominator rescales the numerator so that it is a suitable estimator for the standard deviation for Gaussian white noise.

*Calculation of noise threshold:* The noise threshold  $\tau$  is determined by

$$\tau = \delta_{\text{mad}} \sqrt{\ln(N)} \quad (4)$$

where  $\delta$  is the estimated noise and  $N$  is the total number of samples.

*Hard thresholding:* Suppose that the contaminated signal  $f$  equals the original EMG signal  $s$  plus the noise signal  $n$ , i.e.  $f = s + n$ . The threshold method works under the following assumptions, where  $T_s$  is the signal threshold and  $T_n$  (threshold  $\tau$ ) is the noise threshold (Walker, 1999).

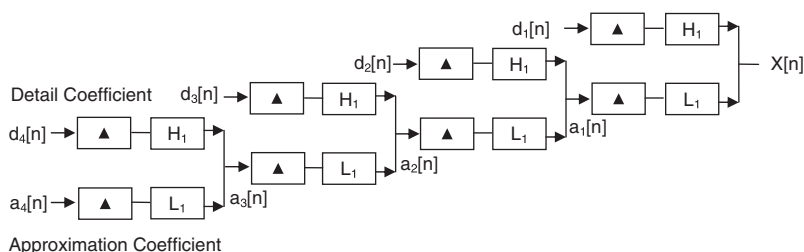
- (1) The energy of the original signal  $s$  is effectively captured, to a high percentage, by transform values whose magnitudes are all greater than some threshold  $T_s > 0$ .

- (2) The noise signal's transform values all have magnitudes that lie below some noise threshold  $T_n$  satisfying  $T_n < T_s$ .

Then the noise in  $f$  can be removed by thresholding its transform where all values of the transform whose magnitude lies below the noise threshold  $T_n$  are set equal to 0. This way of thresholding is called hard thresholding.

**2.1.3. Wavelet reconstruction** An inverse transform is performed on the coefficients after thresholding, providing a good approximation of the EMG signal. The reconstruction is the reverse process of wavelet decomposition. The approximation ( $a_4[n]$ ) and detail coefficients ( $d_1[n], d_2[n], d_3[n], d_4[n]$ ) at every level are up-sampled by two, passed through the low-pass and high-pass synthesis filters and then added. This process is continued through the same number of levels as in the decomposition process to obtain the original SEMG signal (Walker, 1999). For this research, four levels of wavelet decomposition/reconstruction are applied as mentioned earlier. Figure 4 gives the four levels of wavelet reconstruction based on Figure 2.

According to Figure 4, the approximation coefficient  $a_4[n]$  and detail coefficient  $d_4[n]$  are passed through the low-pass and high-pass



**Figure 4:** Four levels of wavelet reconstruction.

filters and then added to get  $a_3[n]$ .  $a_3[n]$  and  $d_3[n]$  are added to get  $a_2[n]$  and the process continues until the original SEMG signal  $x[n]$  is achieved. Here  $x[n]$  is the denoised EMG signal. As mentioned earlier, various WFs are used in this research and each WF is suitable for EMG signals but the most efficient WF is quantified by calculating the RMS difference and SNR values.

## 2.2. RMS difference calculation

The RMS difference of the contaminated signal  $f[n]$  compared with the denoised signal  $s[n]$  is defined by

RMS difference =

$$\sqrt{\frac{(f_{[1]} - s_{[1]})^2 + (f_{[2]} - s_{[2]})^2 + \dots + (f_{[N]} - s_{[N]})^2}{N}} \quad (5)$$

where  $f$  is the raw SEMG signal and  $s$  is the signal after denoising.  $N$  is the total number of samples (length of data). The RMS difference is calculated for various WFs. According to equation (5), the greater the RMS difference, the better the denoising performance of the WF.

## 2.3. SNR calculation

The SNR is calculated using

$$\text{SNR} = 10 \log_{10}(X_n/X_s) \quad (6)$$

where  $X_n$  is the variance for the noisy signal and  $X_s$  is the variance for the denoised signal. According to equation (6), the higher the value of negative SNR (–db), the better the performance of the WF.

## 2.4. Bispectrum analysis (HOS)

The two-dimensional discrete-time Fourier transform of the third-order cumulant gives the bispectrum. Knowing the frequency components  $X(k)$  and  $X(l)$  of the output signal  $x(k)$ , the bispectrum  $B_x(k, l)$  can be estimated by

$$B_x(k, l) = E\{X(k)X(l)X^*(k + l)\} \quad (7)$$

where  $E\{\cdot\}$  donates the statistical expression,  $k$  and  $l$  are the discrete frequency components and  $*$  denotes the complex conjugate.

**2.4.1. Bispectrum estimation process** Consider a set of real observations (here the EMG signal)  $\{x(n)\}$  for  $n = 0, 1, 2, \dots, N - 1$ , where it is assumed that the data set is stationary (Nikias & Raghuvver, 1987; Shahid *et al.*, 2005; Reaz *et al.*, 2006a). The algorithm for the bispectrum estimation is given in detail as follows.

1. Segment the EMG data into  $G$  segments of  $M$  samples each ( $N = G \times M$ ). The new denotation of  $\{x(n)\}$  is  $x^g(m)$  where  $m = 0, 1, 2, \dots, M - 1$  and  $g = 0, 1, 2, \dots, G - 1$ . For good estimation, the sample size  $M$  is selected as the number of frequency points of the fast Fourier transform.
2. Subtract the average (mean) value of each segment from each record of that segment. If necessary, add zeroes to each segment to obtain a length  $M$  for the fast Fourier transform such that  $M$  is a power of 2. The subtracted value  $x_a^g(m)$  is given by

$$x_a^g(m) = x^g(m) - \mu_g \quad (8)$$

where

$$\mu_g = \frac{1}{M} \sum_{m=0}^{M-1} x^g(m)$$



3. Multiply each segment by a suitable data window  $w(m)$  to control the effect of spectral leakage. A one-dimensional Hamming window is used.
4. Within each segment compute the discrete Fourier transform (DFT) coefficients  $X^g(k)$ . The DFT is computed by

$$X^g(k) = \frac{1}{M} \sum_{m=0}^{M-1} x^g(m) w(m) \exp\left(\frac{j2\pi mk}{M}\right) \quad (9)$$

where  $k = 0, 1, 2, \dots, M/2$  is the frequency index. The frequency resolution is  $(2\pi f_s/N_f)k$  where  $f_s$  is the sampling frequency and  $N_f$  is the total number of frequency samples.

5. Compute the estimated bispectrum by using DFT coefficients for the principal domain  $0 \leq l \leq k$ ,  $0 \leq k \leq M/2$ ,  $2k+1 \leq M/2$ . The bispectrum  $\hat{b}^g(k, l)$  is given by

$$\hat{b}^g(k, l) = X^g(k) X^g(l) X^{g*}(k+l) \quad (10)$$

where the sign  $*$  indicates the complex conjugate.

6. The bispectrum estimate  $\hat{B}(k, l)$  of the given data is averaged over the  $G$  pieces as

$$\hat{B}(k, l) = \frac{1}{G} \sum_{g=1}^G \hat{b}^g(k, l) \quad (11)$$

7. The power spectrum  $\hat{P}^g(k)$  can also be estimated during the process and is given by

$$\hat{P}^g(k) = X^g(k) X^{g*}(k) \quad (12)$$

where the  $*$  sign is the complex conjugate. The average over  $G$  pieces is given by

$$\hat{P}(k) = \frac{1}{G} \sum_{g=1}^G \hat{P}^g(k) \quad (13)$$

**2.4.2. Gaussianity and linearity tests** To quantify the non-Gaussianity of the EMG signal, the normalized bispectrum gives the bicoherence  $B_n(k, l)$  which is estimated by

$$B_n(k, l) = \frac{B_x(k, l)}{\sqrt{P(k)P(l)P(k+l)}} \quad (14)$$

where  $B_x(k, l)$  is the bispectrum and  $P(\cdot)$  is the power spectrum. Bicoherence is a mixed function of second- and third-order statistics that is the power spectrum and bispectrum. The third-order and fourth-order cumulant gives the skewness. The bicoherence and skewness function are very useful in the detection and characterization of non-linearities in deterministic and non-deterministic signals and also for discriminating non-linear processes from linear ones. The skewness function of a linearly filtered non-Gaussian signal is flat (Hinich, 1982) but the skewness function of the output signal of a non-linear system is not flat. If a signal is Gaussian then its bispectrum is zero. The theoretical bispectrum of any Gaussian signal is identically zero. Fackrell (1996) reported that the bispectrum is zero for any independent and identically distributed signal with a symmetric probability density function (pdf) whereas skewed signals have an asymmetric pdf whose bispectrum is non-zero (Raghuveer, 1995). It is also reported that any signal which is not white has a non-zero bispectrum. The power spectrum and bispectrum of a non-Gaussian random signal always carry a constant value. This property is very useful to identify a process from a time series whether it is Gaussian or not.

A common problem in signal processing is that the observed signal consists of a non-Gaussian stationary signal in additive Gaussian noise. Let us consider a mixture process which is the sum of two processes (Gaussian and non-Gaussian), as follows:

$$x(n) = e(n) + w(n) \quad (15)$$

where  $e(n)$  is non-Gaussian zero mean and  $w(n)$  is the Gaussian white noise. From the property of the power spectrum and the bispectrum, equation (15) can be represented in terms of the power spectrum and bispectrum by equations (16) and (17) respectively.

$$P^x(k) = P^e(k) + P^w(k) \quad (16)$$

$$\begin{aligned} B^x(k, l) &= B^e(k, l) + B^w(k, l) \\ &= B^e(k, l) = \gamma_3^e \end{aligned} \quad (17)$$

**Table 1:** Average (from three subjects) RMS differences of various WFs

WF	Walk style				
	Slow	Medium	Fast	Very fast	Average
db2	0.023394	0.025761	0.036118	0.035513	0.030196
db4	0.023216	0.025692	0.035767	0.034624	0.029825
db5	0.023209	0.025558	0.035648	0.033597	0.029503
db6	0.023029	0.025532	0.035666	0.033161	0.029347
db8	0.023024	0.025156	0.036438	0.035398	0.030004
sym4	0.022986	0.025331	0.03521	0.032953	0.02912
sym5	0.022983	0.025398	0.033601	0.032748	0.028682
dmey	0.022282	0.024669	0.033562	0.03215	0.028165

**Table 2:** Average (from three subjects) SNR values (−db) of various WFs

WF	Walk style				
	Slow	Medium	Fast	Very fast	Average
db2	−3.8504	−2.3722	−1.2662	−1.7506	−2.3098
db4	−3.2413	−2.3739	−1.288	−1.6904	−2.1484
db5	−3.7259	−2.3608	−1.2791	−1.5977	−2.2408
db6	−3.6403	−2.3506	−1.2792	−1.6004	−2.2176
db8	−3.6132	−2.2804	−1.359	−1.7989	−2.2629
sym4	−3.6201	−2.302	−1.233	−1.5662	−2.1803
sym5	−3.6075	−2.2977	−1.1191	−1.5075	−2.1329
dmey	−3.2128	−2.174	−1.1185	−1.5312	−2.0091

**Table 3:** Average (from three subjects) RMS difference and SNR values (−db) for WF db45

Method	Walk style				
	Slow	Medium	Fast	Very fast	Average
RMS difference	0.24695	0.39401	0.76109	0.801975	0.551006
SNR value	20.005	21.445	22.185	22.4787	21.5284

where  $k$  and  $l$  are the frequency components. It is noted that the bispectrum of a Gaussian signal is zero and the resulting bispectrum of the modelled mixer signal gives the skewness value ( $\gamma_3^e$ ) only. Therefore, additive noise has been suppressed in the output signal. Therefore, the bispectrum offers robustness to additive Gaussian white noise. The test of Gaussianity  $S_g$  is based on the mean bicoherence power defined by

$$S_g = \sum |B_n(k, l)| \quad (18)$$

The Gaussianity test  $S_g$  (zero-skewness test) involves deciding whether or not the estimated

bicoherence is zero. The linearity test involves deciding whether or not the estimated bicoherence is constant in the bi-frequency domain, employing a measure of the difference ( $dR$ ) between a theoretical and an estimated inter-quartile range  $R$ . A window of 256 points was used for the Gaussianity test based on the size of the fast Fourier transform to estimate the bispectrum. In this project, the size of the fast Fourier transform was 256. Furthermore, the Gaussianity test was performed with 0.51 smoothing. If  $c$  is considered to be the resolution (smoothing) parameter, then  $0.5 < c < 1.0$ . Increasing  $c$  reduces the variance but increases the bias (window bandwidth). So



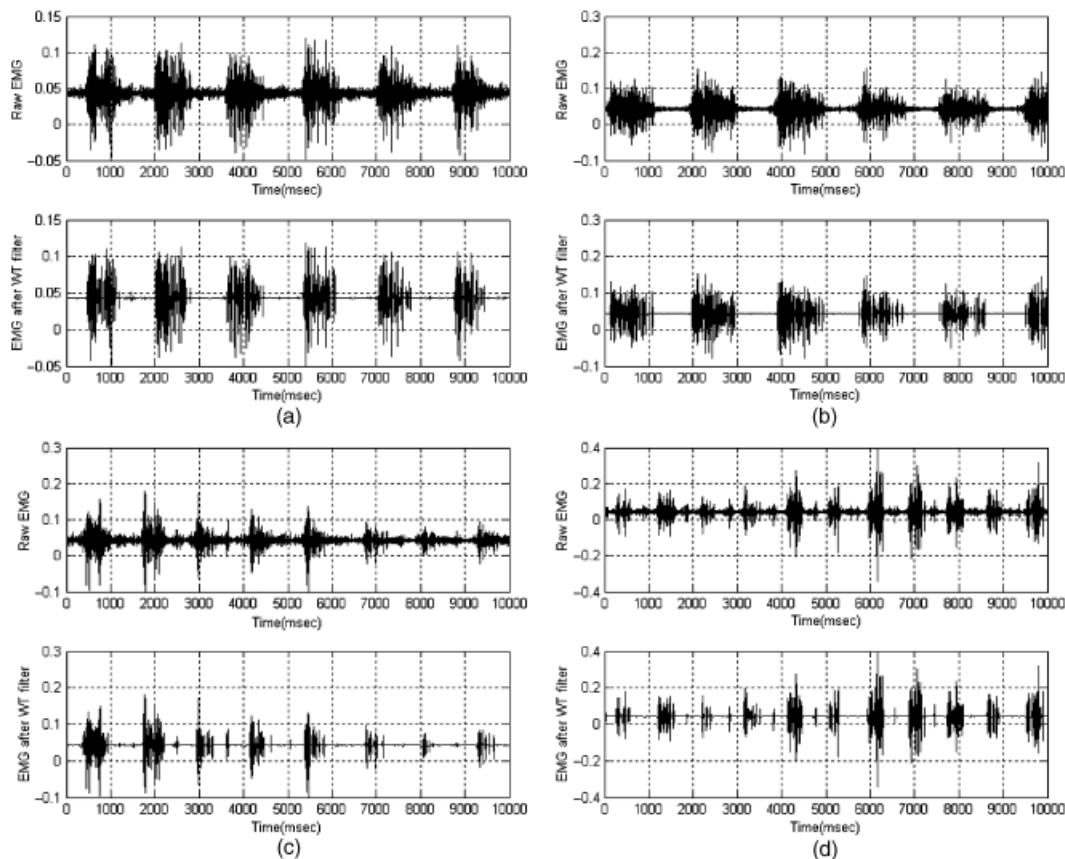
the default value 0.51 for smoothing is considered in this experiment for the Gaussianity.

### 3. Results and discussion

The wavelet denoising method was applied to the SEMG signal at various walk styles/force (slow walking style, medium walking style, fast walking style and very fast walking style) during the trial. The bispectrum was used to analyse the EMG signal to understand the muscle contractions through the Gaussianity and linearity tests.

Kumar *et al.* (2003) demonstrated that the differences between the EMG corresponding to fatigued muscles and non-fatigued muscles is

highlighted in the power of the wavelet coefficients when using WFs sym4 and sym5. Ren *et al.* (2006) used WF db5 to denoise EMG signals in their research. Reaz *et al.* (in press) used WF db45 to filter raw EMG signals and demonstrate muscle contraction through power spectrum and bispectrum analysis. In this experiment the appropriate WF for EMG is chosen on the basis of the calculated RMS difference and SNR values. The raw EMG signals were used to calculate the RMS difference and SNR values for all the WFs suitable for biomedical signal processing (db2, db4, db5, db6, db8 and dmev) including sym4, sym5 and db45 with four levels of decomposition. Table 1 gives the results of the average RMS difference and Table 2 gives



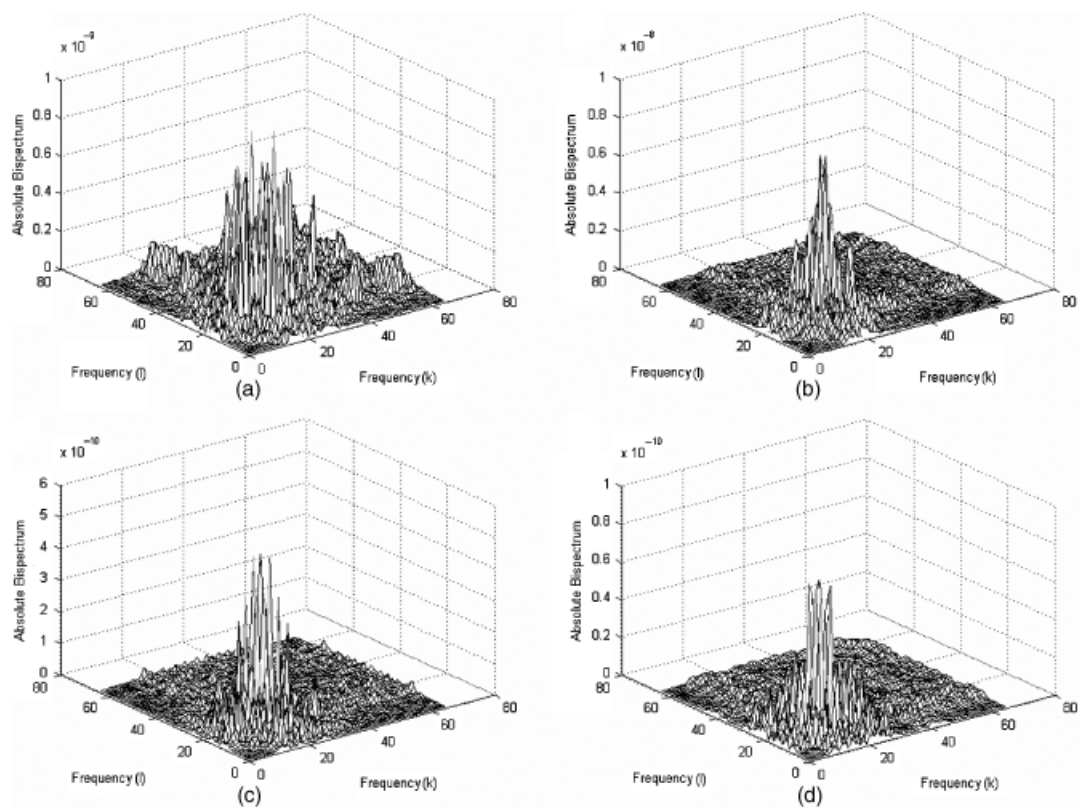
**Figure 5:** Raw SEMG signal and denoised SEMG signal at various walking styles for subject 1 using WF db2 at four levels of decomposition: (a) slow walking style; (b) medium walking style; (c) fast walking style; (d) very fast walking style.

the average SNR values of the chosen WFs for the three subjects at various muscle contraction positions.

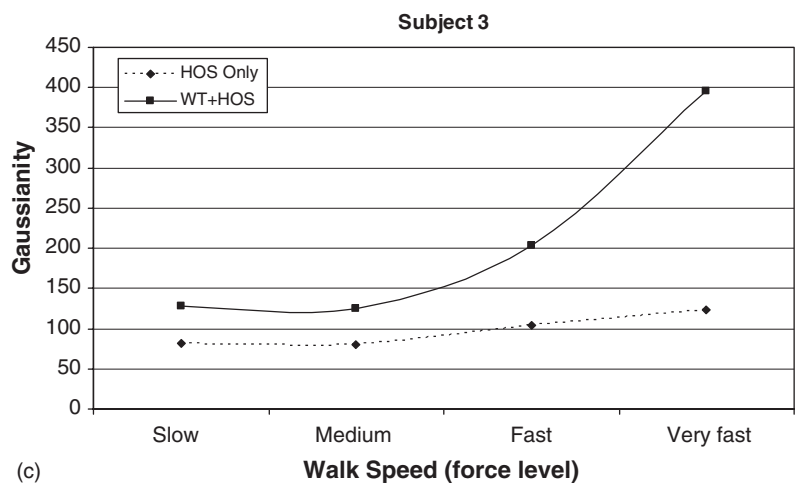
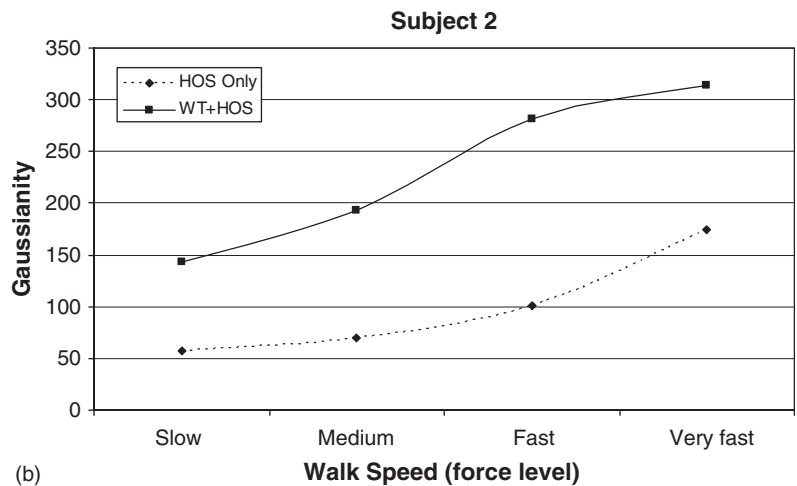
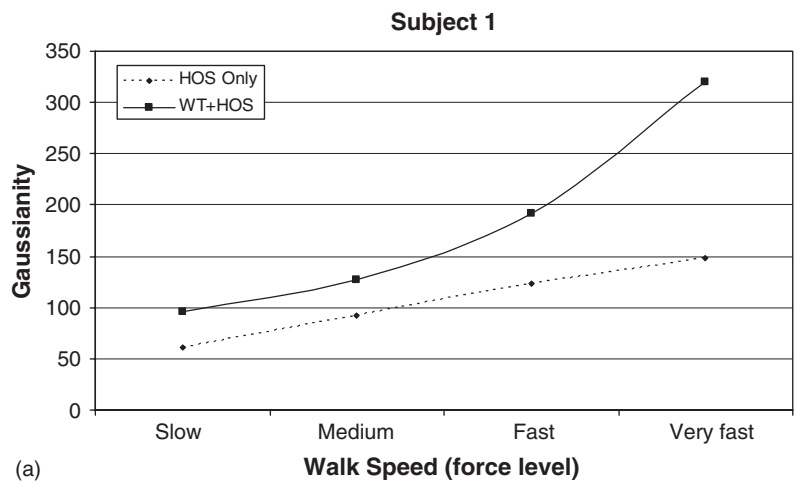
According to the results in Table 1, the WFs from the Daubechies family show better performance compared to the other families (symlet and orthogonal Meyer); however, the Daubechies WF db2 has a higher RMS difference than the other WFs. This means that WF db2 (from the average column in Table 1) is capable of denoising SEMG signals better than the other WFs and families. According to Table 2, the SNR values show similar results where the WF db2 shows better performance than the other WFs. It is clear that the WFs (sym4 and sym5) chosen by Kumar *et al.* (2003) are appropriate for wavelet based signal analysis (calculating the wavelet power) but are not appropriate for

denoising signals. The WF (db5) chosen by Ren *et al.* (2006) is suitable for denoising EMG signals since it is from the Daubechies family. The WF (db45) considered by Reaz *et al.* (in press) is suitable for EMG signal analysis but not at all appropriate for EMG signal denoising. WF db45 gives irrational results (from the same data used earlier) for both RMS difference and SNR values which is not right for SEMG signal denoising. Table 3 gives the results for the RMS difference and SNR values using WF db45.

Since WF db2 shows better performance from the results, it was considered for the SEMG signal denoising process and bispectrum analysis. Figure 5 gives the raw SEMG signal and denoised SEMG signal for one subject at various muscle walking stages. The denoising in the figures is obtained using WF db2. The result for



**Figure 6:** Bispectrum of the various walking styles for subject 1 after denoising: (a) slow walking style; (b) medium walking style; (c) fast walking style; (d) very fast walking style.



**Figure 7:** *Gaussianity tests for three subjects ((a), (b), (c)) during the walking trial using HOS only and WT plus HOS.*

bispectrum analysis is given in the following paragraphs.

For this experiment, the SEMG signal was captured from the rectus femoris muscle for the following force levels/walking styles: slow walk, medium walk, fast walk and very fast walk. Results for two subjects at the mentioned force level are presented in this paper using bispectrum analysis for two cases: the SEMG signal (a) with wavelet denoising; (b) without wavelet denoising. As stated earlier, the bispectrum of an additive Gaussian white noise signal is identically zero (Raghuveer, 1995; Fackrell, 1996). So, the bispectrum (HOS) is able to remove additive Gaussian white noise without performing the denoising technique. The denoising method is used for filtering the interference noise and by performing the filtration, the presence of noise in the signal is even less after applying the bispectrum. The results for SEMG signal analysis are obtained from the Gaussianity tests of the signal using only the bispectrum and the signal using the bispectrum after the noise removal technique. The bispectrum for one subject after denoising is given in Figure 6. The results for the Gaussianity tests are given in Figure 7 for the three subjects during their walking trial.

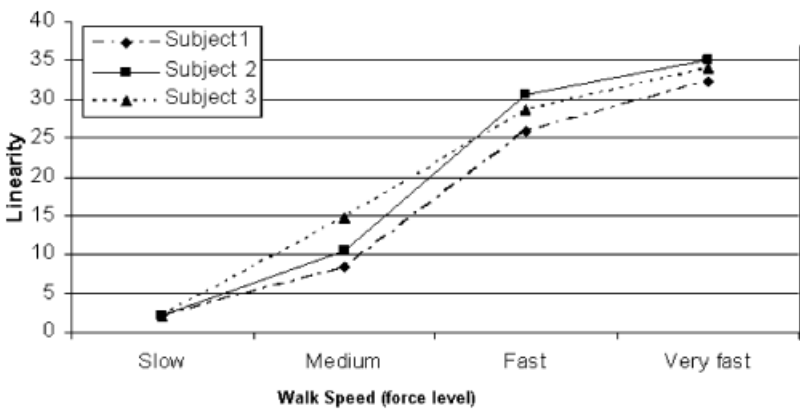
According to Figure 7, the SEMG signal becomes less Gaussian with increased force or walking speed (all subjects). Both the signals (using only HOS, using the WT plus HOS) show similar characteristics but the denoised signal is

less Gaussian compared to the signal using only HOS (bispectrum).

Bispectrum analysis was also used by Kaplanis *et al.* (2000) for analysing the biceps brachii muscle. It is reported that SEMG becomes less Gaussian and more linear on increasing mean voluntary contraction. Other research using HOS also showed similar results where SEMG becomes less Gaussian with increased muscle contraction due to load (Shahid *et al.*, 2005; Reaz *et al.*, 2006a, in press). The results for the linearity test for the three subjects are given in Figure 8.

According to Figure 8, it is demonstrated that the signals become more linear with increased walk speed/muscle force. The linearity tests show the same pattern as the Gaussianity tests which is the reverse pattern for the trial.

Results obtained by this research explain that the signal becomes less Gaussian as in Kaplanis *et al.* (2000), Shahid *et al.* (2005) and Reaz *et al.* (2006a, in press) and more linear as in Kaplanis *et al.* (2000) and Reaz *et al.* (2006a) with increased force. The dotted lines in Figure 7 represent the change in Gaussianity for the raw signals applying HOS only and the solid lines demonstrate the Gaussianity for the signal after noise removal applying the WT and HOS. The raw signal and denoised signal show similar characteristics where both SEMG signals become less Gaussian from the slow walking style to very fast walking style. The important thing to notice from Figure 7 is that the signal after



**Figure 8:** Linearity tests for three subjects during the walking trial using WT and HOS.

denoising is less Gaussian than the other signal using only HOS. This indicates that the wavelet based denoising method using WF db2 effectively filtered interference noise.

#### 4. Conclusion

The aim of this study was to investigate the usefulness of the wavelet based noise removal technique in SEMG signal processing and analysing the signals using HOS. The wavelet denoising method has already been successfully used in other biomedical signal processing. It provides a powerful complement to conventional noise-removal techniques like notch filters and frequency domain filtering methods. Results obtained through the research clearly show that WF db2 is the best choice among other wavelets to denoise SEMG signals. On the other hand HOS is preferable because it can suppress Gaussian white noise. The method used in this research applying WT and HOS proves to be a better and efficient technique for SEMG signal processing and analysis, where WT filters interference noise successfully and the bispectrum (HOS) analysis showed that the SEMG signal becomes less Gaussian and more linear (an exact reverse pattern of Gaussianity) with increased force.

#### References

- CARRE, P., H. LEMAN, C. FERNANDEZ and C. MARQUE (1998) Denoising of the uterine EHG by an undecimated wavelet transform, *IEEE Transactions on Biomedical Engineering*, **45** (9), 1104–1114.
- FACKRELL, J.W.A. (1996) Bispectral analysis of speech signals, PhD Thesis, University of Edinburgh.
- GABOR, D. (1946) Theory of communication, *Journal of the Institute of Electrical Engineering*, **93**, 429–457.
- GUGLIELMINOTTI, P. and R. MERLETTI (1992) Effect of electrode location on surface myoelectric signal variables: a simulation study, in *Proceedings of the 9th International Congress of ISEK*, 106.
- HINICH, M.J. (1982) Testing for Gaussian and linearity of a stationary time series, *Journal of Time Series Analysis*, **3** (3), 169–176.
- KAPLANIS, P.A., C.S. PATTICHIS, L.J. HADJILEONTIADIS and S.M. PANAS (2000) Bispectral analysis of surface EMG, in *Proceeding of the 10th Mediterranean Electrotechnical Conference*, Vol. 2, pp. 770–773.
- KUMAR, D.K., N.D. PAH and A. BRADLEY (2003) Wavelet analysis of surface electromyography to determine muscle fatigue, *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, **11** (4), 400–406.
- LATERZA, F. and G. OLMO (1997) Analysis of EMG signals by means of the matched wavelet transform, *Electronics Letters*, **33** (5), 357–359.
- MARK, P.W., G.S. RASH, P.M. QUESADA and A.H. DESOKY. (2000) Wavelet-based noise removal for biomechanical signals: a comparative study, *IEEE Transactions on Biomedical Engineering*, **47** (2), 360–360.
- MEWETTE, T.D., N. HOMER and J.R. KAREN (2001) Removing power line noise from recorded EMG, *Proceedings of the 23rd Annual EMBS International Conference*, 2190–2193.
- MOSHOU, D., I. HOSTENS, G. PAPAIOANNOU and H. RAMON (2000) Wavelets and self-organising maps in electromyogram (EMG) analysis, in *Proceedings of the European Symposium on Intelligent Techniques (ESIT)*, 186–191.
- NIKIAS, C.L. and J.M. MENDEL (1993) Signal processing with higher-order spectra, *IEEE Signal Processing Magazine*, **10**, 10–37.
- NIKIAS, C.L. and M.R. RAGHUVEER (1987) Bispectrum estimation: a digital signal processing framework, *Proceedings of the IEEE*, **75** (7), 869–891.
- RAGHUVEER, M.R. (1995) Third-order statistics: issue of PDF symmetry, *IEEE Transactions on Signal Processing*, **43** (7), 1736–1738.
- REAZ, M.B.I., M.S. HUSSAIN and F. MOHD-YASIN (2006a) Kinesiological surface EMG analysis using higher order statistics, *Proceedings of the 4th APT Telemedicine Workshop*, 167–171.
- REAZ, M.B.I., M.S. HUSSAIN and F. MOHD-YASIN (2006b) Techniques of EMG signal analysis: detection, processing, classification and applications, *Biological Procedures Online*, **8** (1), 11–35.
- REAZ, M.B.I., M.S. HUSSAIN and F. MOHD-YASIN (in press) EMG signal analysis to determine muscle fatigue, *Journal of Medical Engineering and Technology*, forthcoming.
- REN, X., X. HU and Z. WANG (2006) MUAP extraction and classification based on wavelet transform and ICA for EMG decomposition, *Medical and Biological Engineering and Computing*, **44**, 371–382.
- SHAHID, S., J. WALKER, G.M. LYONS, C.A. BYRNE and A.V. NENE (2005) Application of higher order statistics techniques to EMG signals to characterize the motor unit action potential, *IEEE Transactions on Biomedical Engineering*, **52** (7), 1195–1209.

WALKER, J.S. (1999) *A Premier on Wavelets and their Scientific Applications*, London: Chapman and Hall/CRC.

## The authors

### M.S. Hussain

M.S. Hussain was born in Bangladesh. He received the BSc in information technology from Multimedia University, Malaysia, in 2005. He is presently pursuing his Master of Engineering degree in the Department of Electrical and Computer Engineering, International Islamic University Malaysia, Malaysia. He is author and co-author of more than five papers in biomedical signal processing.

### Mamun Bin Ibne Reaz

Mamun Bin Ibne Reaz was born in Bangladesh in December 1963. He received the BSc and MSc in applied physics and electronics, both from the University of Rajshahi, Bangladesh, in 1985 and 1986, respectively. He received the DEng from Ibaraki University, Japan, in 2007. He has vast research experience in Norway, Ireland and Malaysia. Currently he is a Senior Lecturer at the Department of Electrical, Electronic and Systems Engineering, The National University of Malaysia, Malaysia, involved in teaching, research and industrial consultation. Dr Mamun has published extensively in the area of integrated circuit (IC) design and biomedical application IC. He is author and co-author of more than 100 papers in design automation and IC design for biomedical applications.

### Faisal Mohd-Yasin

Faisal Mohd-Yasin was born in Malaysia. He received the BSc in electrical engineering and MSc in telecommunications and computers, both from the George Washington University, Washington, DC, in 1999 and 2001, respectively. He was also a Graduate Fellow working with the AOL-GWU collaboration in the area of a distributed sensors network for smart homes. Currently he is a senior lecturer at the Faculty of Engineering, Multimedia University, Malaysia. Mr Faisal is author and co-author of more than 80 papers in very large scale integration system design, microelectromechanical systems and design automation. His current research interests are in the area of mixed-signal and radio frequency identification design and microelectromechanical systems.

### Muhammad Ibn Ibrahimy

Muhammad Ibn Ibrahimy was born in Bangladesh in August 1962. He received the BSc and MSc in applied physics and electronics, both from University of Rajshahi, Bangladesh, in 1985 and 1986, respectively. He received the PhD from Universiti Kebangsaan Malaysia, Malaysia, in 2002. He has vast research experience in Japan where he has done his post-doctorate. Currently he is an assistant professor in the Department of Electrical and Computer Engineering, International Islamic University Malaysia, Malaysia, involved in teaching, research and industrial consultation. Dr Ibrahimy has published extensively in the area of biomedical signal processing. He is author and co-author of more than 10 papers in biomedical signal processing.