Spectral Coefficients System for Osteoarthritis Detection

Gan Hong Seng, Tan Tian Swee

Abstract— According to our knowledge, VAG serves great interest for the detection of Osteoarthritis. However, there is no scientific research being carried out to study the effect of LPCC and MFCC on vibroarthrographic (VAG) signals. Hence, the objective of this project is to evaluate the effectiveness of LPCC and MFCC in extracting features from VAG signals of 30 subjects. We have carried out quantitative analysis on the VAG cepstral model from inter-subject and intra-subject perspective. Our study exhibits high recognition rates of 90.36% for LPCC and 88.64% in the intra-subject analysis of the VAG signal. In conclusion, the cepstral analysis of VAG signal has been showing great potential for future research given the high intra-subject analysis. Nevertheless, we strongly suggest a larger research population with the inclusion of OA patients.

Keywords—Linear Predictive Cepstral Coefficient, Mel frequency Cepstral Coefficient, Signal Processing, Vibroarthrography

I. Introduction

Knee arthritis is caused by the degeneration of the knee joint cartilages [1]. With more than 100 types of arthritis being identified until recently, the three most common types of arthritis are Osteoarthritis (OA), Rheumatoid arthritis (RA) and Post traumatic arthritis. OA, identified as degenerative arthritis, is caused by the degeneration of articular cartilage due to family inheritance, injuries and aging [2]. The clinical symptoms of OA includes knee joint pain, swelling and bony spur.

Meanwhile, RA is classified as inflammatory arthritis [3]. Although the major cause of RA is unknown, early symptoms such as infiltration of the synovial membrane and morning stiffness could serve as possible warning signs indicating development of RA.

The post traumatic arthritis occurs when patient sustains serious injury to the knee joint section or undergoes invasive surgical procedures. The deterioration to the articular cartilage and subchrondral bone after trauma encourage degeneration to the surrounding cartilage, resulting in the development of Osteoarthritis.

In 2009, the population of the American arthritis population has been standing at 46.3 million, engulfing $128 billion of direct and indirect medical cost. A survey conducted by Centre for Disease Control and Prevention (CDC) estimated that in 2030 [4], there will be 63 million Americans affected by arthritis. Although there is still no cure to knee arthritis, diagnosis techniques have been able to provide early warning to patients for early treatment and prevention. These detection methods are arthroscopy, Magnetic Resonance Imaging (MRI), Computed Tomography scanning (CT scan) and X-ray.

Arthroscopy is an invasive surgical procedure to provide direct visualization of cartilage quality. The surgical procedures involve inserting the arthroscopic through a small incursion on the knee joint to monitor the knee joint cartilages. As the golden standard to monitor cartilage, this direct knee cartilage visualization approach ensures reliable diagnostic outcome [5]. Nevertheless, arthroscopy is more suitable for assessing mild cartilage degeneration [6] since applying this technique on highly degenerated knee poses high risk.

MRI, CT scan and X-Ray represent the indirect cartilage visualization approaches. These imaging based techniques are non-invasive and able to provide excellent depiction of the knee joint degeneration level. Furthermore, MRI is non-radiation while CT scan could be completed in a short duration. However, imaging modalities do not indicate integration function of knee joint cartilage, providing only gross picture of the knee joint [7]. Furthermore, MRI and CT scan are expansive. Although various image processing techniques have been introduced [8, 9], X-ray fails to clearly delineate the cartilage tissue of knee joint [10, 11].

Vibroarthrography (VAG) represents the analysis of patelloferomal joint vibration signals produced by leg movement. This method was first proposed by Robert Hooke in 17th century [12] and its potential as a possible arthritis diagnosis method has been explored since 1902 [13]. The possible detection technique is non-invasive and low cost while able to provide important information on the physiological condition of the cartilage. At the early stage of the scientific experiment on VAG signals, Steindler [14]

Gan Hong Seng is with the Department of Biotechnology and Medical Engineering, Medical Implant Technology Group, Material and Manufacturing Research Alliance (MMRA), Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Malaysia (phone: +60-17-655-1561; e-mail: meditagutm@gmail.com).

Tan Tian Swee is with Department of Biotechnology and Medical Engineering, Medical Implant Technology Group, Material and Manufacturing Research Alliance (MMRA), Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia, 81310 Skudai, Malaysia (e-mail: tantswee@biomedical.utm.my).

attempted the use of cardiophone, oscilloscope and recorder to record knee joint sound signals. He was reported to be the first to improve the VAG signals by using filter and recorded the knee angle. In 1978, Chu [12] concluded that the energy parameter of the VAG signal was directly proportional to the level of degeneration of the knee joint.

Aside of fundamental researches on VAG signals, analyses have been conducted to improve the pre-processing and feature extraction method. Initial study has been using the fixed segmentation to segment the VAG signals and analysed the signals by Short Time Fourier Transform (STFT). However, this signal analysis method produced large number of signal segments [12].

In 1996, Zahra [16] reported the use of Recursive Least Squares (RLS) modelling technique to segment the data samples into fewer model parameters. Dominant pole was then extracted from the processed signals by using forward backward linear prediction method and was classified by using Statistical Package for the Social Sciences (SPSS). Her analysis method has been further improved by Krishnan [13] who shifted to Recursive Least Squares lattice (RLSL) to segment VAG signals and introduced the Burg Lattice method in Auto Regressive modelling (AR) to extract dominant poles from VAG signals. Krishnan's methodology retained the SPSS as classification tool.

In 2000, Sridhar [7] proposed the decomposition of VAG signal by using Matching Pursuit (MP) algorithm in adaptive Time Frequency Distribution (TFD). After that, energy, energy spread, frequency and frequency spread are extracted from the TFD. A statistical pattern classification system based on stepwise logistic regression analysis was applied to classify the signals.

Keo [17] segmented and normalised the VAG signals using Dynamic Time Warping (DTW), transformed the VAG signals by Wigner Ville Distribution (WVD), reduced the noise using Singular Value Decomposition (SVD) and extracted energy, energy spread, frequency and frequency spread from VAG signals. The study classified the signals using BNPP classification system.

In 2008, Rangayyan [25] has conducted study on the possibility of using statistical parameters of VAG signals such as form factors for certain duration of the signals, skewness, kurtosis and, entropy to investigate the nature of the VAG signals. The study implemented the Fisher linear discriminant analysis (FLDA) to identify the pattern of the signals, where the best discriminant result was used to derive a receiver operating characteristics (ROC) curve for calculating the associated area under the curve using ROCKIT. The method was able to produce screening efficiency of up to 0.82 in terms of the area under the receiver operating characteristics curve.

In latter attempt, Rangayyan [33] has introduced a novel statistical modelling method to analyse the VAG signals statistical characteristics by studying their probability density functions (PDFs). Parzen windows were applied to derive the PDF models and Kullback-Leibler distances between the signals' PDF models were interpolated. In assessing the viability of this method, the study has included mean, standard deviation, coefficient of variation, skewness, kurtosis and entropy into the analysis. An overall accuracy of 77.53%, sensitivity of 71.05%, and specificity of 82.35% were achieved in the PDF analysis.

After reviewing various efforts being tested to improve the VAG technique, a novel attempt to evaluate the effectiveness of LPCC and MFCC on VAG signals analysis have been made the objective in this project. The effectiveness is gauged by evaluating the mean recognition rate and standard deviation of standard deviation of LPCC and MFCC.

II. Establishment of the Experiment

There are four major stages in this experiment i.e. data collection, pre-processing, feature extraction and classification, which is shown in Fig. 1. The VAG signals are collected using vibroarthrography intended recording stethoscope. A total of 30 normal subjects have been recruited for the recording process with their consent obtained. Other important pre-recording elements in VAG signal analysis are taken into consideration prior to the analysis and will be discussed in section *A*.

After that, desired features will be extracted from the signals by using LPCC and MFCC. We have introduced the application of Hidden Markov Model (HMM) as the classifier to produce the recognition rates for analysis.

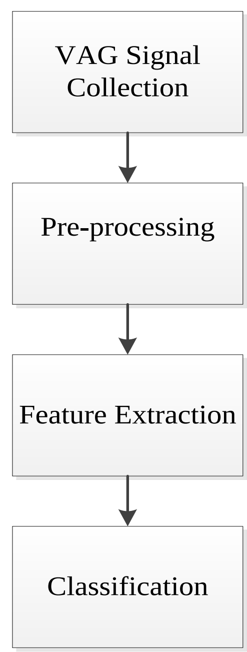


Fig. 1 The flow process of analyzing VAG signals using cepstral coefficient

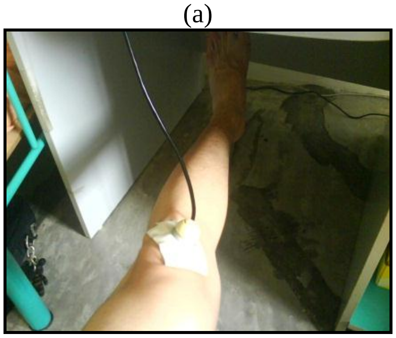
A. Data Collection

A total of 900 VAG signals are collected from 30 subjects in this experiment. We use a built in microphone recording stethoscope because its' recording head can fit nicely to the uneven surface of knee joint. The illustration of our recording stethoscope is shown in Fig. 2. Double side tape attaches the recording stethoscope to the subject knee joint and reduces the sensitivity of noise.

Medial condyle of patella is identified as the ideal auscultation location after considering the muscle contraction interference (MCI) factor. The subjects are required to sit straight before the swing. A full swing cycle in this experiment is comprised of a 90 degree full flexion and 90 degree full extension, in a 4 second recording period. The recording procedures are illustrated in Fig. 3 (a) and (b) Only one full swing cycle will be recorded in each recording period. The VAG signals are digitalized at a sampling rate of 8 KHz. The VAG signal sample is illustrated in Fig. 4

Fig. 2 Recording stethoscope

(a)



(b)

Fig. 3 knee position at (a) full flexion and (b) full extension

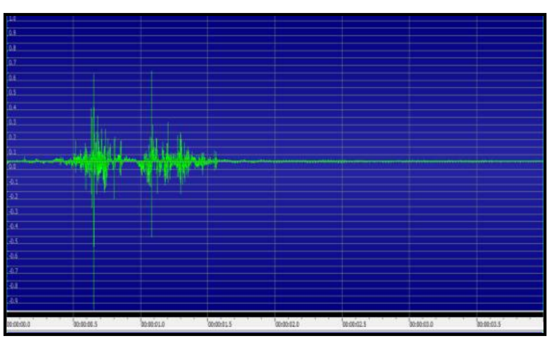


Fig. 4 Sample of normal VAG signal

B. Pre-processing

In the pre-processing stage, non-stationary VAG signals will be divided into frames to make the signal "stationary". The pre-emphasis filtering will flatten the frames to evaluate the high frequencies of the spectral reflectance signal since the high frequency bands is always weaker than that in lower frequency bands. (1) defines the pre-emphasis filter in mathematical terms and (2) reflects the relationship between the input $S(n)$ with its output $S_S^{'}(n)$.

$$H(z)=1-\alpha z^{-1} \tag{1}$$

$$S_n^{'}=S_n-\alpha S_{n-1} \tag{2}$$

During the pre-processing stage, leakage occurs during the flame blocking process of non-stationary signals, leading to faulty information about the spectral amplitude and frequency of the signal. In order to minimise spectral leakage, VAG frames will pass through the windowing process [20]. The mathematical expression of Hamming window is illustrated in (3).

$$W(n)=0.54-0.46\cos(2\pi n/N-1) \tag{3}$$

Hamming window can be applied in any type of signals without compromising the frequency resolution of the signals. The size of window in this experiment is 20 ms. Overlapping occurs during the windowing to avoid discontinuities on both side lobes of each window. The overlapping in this experiment is 33.33%. After that, the stationary frames will undergo DFT and change from time domain to frequency domain [21]. The mathematical expression for DFT is shown in (4):

$$S(n)=R(n)+jI(n) \tag{4}$$

III. Cepstral Analysis

Homomorphic system is a nonlinear signal processing system that abides by the generalized principal of superposition [31]. The superposition-type multiplication, exponentiation as well as addition operations are one of the vital properties of the homomorphic system whereby the mathematical expression of the combined superposition operations shown in (5). In traditional speech processing, the cepstral analysis is derived from a simple linear acoustic model. According to the model, a speech signal is produced by convoluting an excitation waveform, $e(n)$ with the vocal filter, $h(n)$ as illustrated in (6). The convolution is then converted into frequency domain in (7) and expressed in logarithmic terms to linearly combine the excitation signal and vocal tract filter in linear expression in (8).

$$D\left[\{x_1(n)\}^{\alpha}\cdot\{x_2(n)\}^{\beta}\right]=\alpha D[x_2(n)]+\beta D[x_2(n)] \tag{5}$$

$$s(n)=e(n)\otimes h(n) \tag{6}$$

$$S(f) = E(f) \cdot H(f) \tag{7}$$

$$\log(|S(f)|) = \log(|E(f)|) + \log(|H(f)|) \tag{8}$$

In order to analyze the characteristic of the vocal tract response, homomorphic system can decouple the excitation signal, $e(n)$ from the vocal tract response $h(n)$ in a process known as cepstral deconvolution. Through liftering operation, the excitation part is removed from the logarithm speech signal in (7). Hence, inverse Fourier transform can be applied onto the vocal tract filter to obtain the cepstrum, which is divided into real and complex cepstrum. The definitions of the real and complex cepstrum are illustrated in (8) and (9) respectively. Since the logarithmic function obeys the generalized superposition property, the cepstral analysis produces output that is a linear superposition of the input signals under a nonlinear transformation.

$$c[n] = FT^{-1}\{\log|FT(x[n])|\} \tag{9}$$

$$\bar{x}[n] = FT^{-1}\{\log(FT\{FT[x[n]]\})\} \tag{10}$$

In previous studies, cepstral analysis has been showing better accuracy in analyzing VAG signals compared to other modeling parameters. The cepstral coefficient is capable of disclosing higher and more prominent discriminant information by weighting on lower band frequencies and exhibit superior separabilty in feature [38]. Moreover, cepstral coefficients are uncorrelated, thus, these coefficients are suitable for building machine learning models. The MFCC has its real cepstral coefficients derived from the Discrete Fourier Transform (DFT) of the signal. Nonetheless, the MFCC is carried out based on Mel-frequency scale instead of the nonlinear frequency scale. On the other hand, LPC coefficients can be converted to complex cepstrum by using iterative technique.

A. Mel Ferquency Cepstrum Coefficinet (MFCC)

MFCC is a popular feature extraction technique in the human speech analysis field due to its high sensitivity to the low order cepstral coefficients in overall spectral slop [23] and its ability to capture phonetically important characteristics of human voice [24,42]. To perform MFCC, we convert the DFT into Mel-frequency scale using a set of triangular bandpass filters shown in Fig. 5. These filters act linearly at the range between 0 to 1000 Hz but increases logarithmically for frequency above 1000 Hz [36]. The mapping of linear frequency to Mel frequency or Mel-frequency wrapping is shown in (11) and its critical bandwidth function could be calculated by applying (12).

.

$$Mel(f) = 2595\log_{10}\left[1 + \left(f/700\right)\right] \tag{11}$$

$$BW_{critical} = 25 + 75\left[1 + 1.4\left(f/1000\right)^2\right]^{0.69} \tag{12}$$

As a result, the Mel frequency scale is linearly represented in low range but logarithmically represented in high range. This representation gives more weight to the low frequency components. The Mel parameters are transformed back into time domain by using Discrete Cosine Transform (DCT) and the result of the transformation is recognized as MFCC or acoustic vectors [26]. We transform the Mel parameters back into time domain by using the mathematical equation expressed in (13).

$$MFCC_i = \sqrt{(2/N)}\sum_{j=1}^{N} m_j \cos\left[\left(\pi i/N\right)(j - 0.5)\right] \tag{13}$$

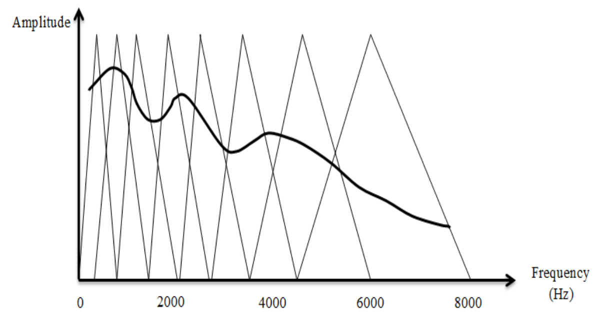


Fig. 5 Triangular bandpass filters with band limit between 0 to 8000 Hz for filtering the spectral envelop

B. Linear Predictive Cepstrum Coefficient (LPCC)

LPC arose from the concept of analyzing speech signal by estimating the resonance (formant) in the signal [30, 37, 40], remove their effect from the signal (inverse filtering) and estimate the intensity and frequency of the remaining signal, recognized as residue. LPC synthesizes the VAG signals by using the signal residue to create a source signal, and use the formants to create a filter, and run the source through the filter.

However, the LPC faces limitation in determining the formants from the given signals, which can be resolved by a difference equation shown in (14). An all pole mode is used to model the VAG signal $S(n)$ that expresses each sample of the VAG signal as a linear combination of its previous samples. The coefficients of the difference equation, $\alpha_k^{(p)}$ characterize the VAG formants. In autocorrelation method, we could estimate these coefficients and their gain factor through minimizing the mean-square error (MSE) between the predicted VAG signal and the actual VAG signal. We have set the modeling order p at 12.

$$s[n] = \sum_{k=1}^{p} \alpha_k^{(p)} s[n-k] + Gx[n] \tag{14}$$

LPC can be computed through the transfer function of the all pole model filtering in z transform can be represented by $H(z)$ [41]. The function's mathematical expression of all pole models in z transform and their gains is shown in (15) and (16) [35].

$$H(z) = S(z)/X(z) \tag{15}$$

$$H(z) = G/[1 - \sum_{K=1}^{p} \alpha_{\kappa}^{p} z^{-k}] \tag{16}$$

The best all-pole model can be obtained when the linear predictor coefficients, α_k that provides the best all pole model to the power spectral density function of a particular random process $S(n)$ are also the coefficients that minimize the mean error between the current value of $S(n)$ and its linear estimate $\bar{s}[n]$, shown in (17). After that, we can find the linear predictor coefficient α_{κ} that minimizes the MSE ξ^2 through (18). By considering (15) and (16), we can obtain a resultant MSE which is illustrated in (19).

$$e[n] = \bar{s}[n] - s[n] \tag{17}$$

$$\xi^2 = E\left|(e^2)\right| = E\left[(\bar{s}[n] - s[n])^2\right] \tag{18}$$

$$\xi^2 = E\left[\left(\sum_{k=1}^{p} \alpha_k^{(p)} s[n-k] - s[n]\right)^2\right] \tag{19}$$

In order to express the optimal LPC coefficients implicitly, we differentiate the MSE with respect to linear predictor coefficient and set the result to zero. (20) and (21) illustrate the differentiation of the MSE and the result of the differentiation.

$$d\xi^2 \Big/ d\alpha_k = E\left[\left(\sum_{l=1}^{p} \alpha_l^{(p)} s[n-l] - s[n]\right) s[n-k]\right] = 0 \tag{20}$$

$$E\left[\sum_{l=1}^{p} \alpha_l^{(p)} s[n-l] s[n-k]\right] = E\big[s[n]s[n-k]\big] \tag{21}$$

To solve the LPC equation illustrated in (21), we have implemented Levinson-Durbin recursion attributable to its relative simplicity and fast computational ability. The equations of the Levinson-Durbin recursion that have been used to compute the corresponding reflection coefficient, k_i and LPC parameter are illustrated from (22) to (26) to subsequently acquire the linear predictor coefficients. The reflection coefficients comprise of an alternate specification of the random process $S(n)$ that is unique and complete as the LPC predictor coefficient.

$$E^{(0)} = \phi[0] \tag{22}$$

$$k_i = \left[\left\{ \phi[i] - \sum_{j=1}^{i-1} \alpha_j^{(i-1)} \phi[i-j] \right\} \Big/ E^{(i-1)} \right] \tag{23}$$

$$\alpha_i^{(i)} = k_i \tag{24}$$

$$\alpha_j^{(i)} = \alpha_j^{(i-1)} - k\alpha_{i-j}^{(i-1)} \tag{25}$$

$$E^{(i)} = (1 - k_i^2) E^{(i-1)} \tag{26}$$

By reiterating the (22) through (26) for i = 1, 2… p, the reflection coefficient can be obtained. (27) illustrates the final computational of LPC through Levinson-Durbin recursion method where the recursion ranged between 1 and p.

$$\alpha_j = \alpha_j^{(p)} \; for 1 \le j \le p \tag{27}$$

LPCC is the LPC representation in cepstral domain [28] and resembles the equivalents of to the smoothed envelop of the log spectrum of the speech in MFCC. LPCC decomposes the smoothed spectrum envelop to extract desired features while MFCC extracts desired features directly from the FFT power spectrum. LPCC could be derived from LPC through recursive conversion method shown in (28).

$$LPCC_i = \begin{cases} In(G); i = 0 \\ LPC_i; i = 1 \\ LPC_i + \sum_{j=1}^{i-1} \left(i-j \big/ i \right) LPCC_{i-j} LPC_j; 1 < i \le p \\ \sum_{j=1}^{p} \left(i-j \big/ i \right) LPCC_{i-j} LPC_j; i > p \end{cases} \tag{28}$$

IV. Classification model

With the desired cepstral features extracted through the LPCC and MFCC, we would analyze the performance of these cepstral features through intra-subject and inter-subject approaches. The intra-subject approach is carried out by analyzing every subject's knee joints signals separately to test the applicable of the cepstral features on VAG signals. The inter-subject analysis is intended to test the capability of cepstral coefficients in representing VAG signals in large scale study.

A. Vector Quantization

In vector quantization, signal data is encoded statistically into a set of k-dimensional of data vector or codeword, C*i*, where i = 1, 2, 3, … These set of data vectors are represented by a finite set of M symbols. A codebook is a set of complete set of M codewords, which is constructed through a training process by training large amount of vector data [27].

During the encoding process, each encoded k-dimension vector X_i is compared to each of the M codewords in the codebook. The distortion $D(X_i, C_i)$ where i = 1, 2, … , M between the input vector and the codeword is computed to find out the minimum distortion. The receiver, which is assumed to have a copy of the codebook, uses this index to look up the corresponding codeword, C_i. Codeword, C_j is then used as the encoded value of the vector X [31].

The average quantization distortion is an important element which is used to match score during the identification process [32]. The average quantization distortion with N frames is defined in (29).

$$D(X,C) = \left(1/N\right)\sum_{i=1}^{N} \min_{cj\in C}\left\|X_i - C_j\right\| \qquad (29)$$

The codebook used in this experiment is 64 bit per vector dimension. Typically, size of the codebook, M is equal to 2^L and the rate of the vector quantization is L/K bits per vector dimension. The 64 bit per vector dimension is appropriate because the primary computational burden during the encoding process is that of computing the distortion between the input vector and each of the M codewords [33].

B. Hidden Markov Model (HMM)

The non-stationary nature of VAG signals give rise to frequency components within a wide range at a given period of time, making the temporal structure of the VAG signals unpredictable. Hence, we have opted for HMM, with the assistance of MFCC and LPCC, to serve as an effective tool in capturing the non-linear variability in auscultatory signals [39]. The training and testing of VAG signals using HMM is shown in Fig. 6

A HMM is a Markov chain where the output observation is a random variable generated according to an output probabilistic function associated with each state [22, 29, 34]. The fundamental of HMM constitutes of five imperative elements:

i. The number of hidden states in the model, S
ii. The number of distinct observation symbols, K
iii. The state transition probability matrix, A = $\{a(i|j)\}$ where $a(i|j)$ is the probability of taking a transition from state i to state j
iv. The set of state output probability distribution, B = $\{b_i(o_t)\}$ where the $\{b_i(o_t)\}$ is the probability of emitting O_t when the state i is entered
v. The initial state distribution, $\pi = \{\pi_i\}$

Two important assumptions, i.e. first order process assumption and independent observation assumption have been applied in the HMM to reduce the number of parameter that need to be estimated as the model complexity without significantly affecting the VAG signal analysis performance.

The first order Markov assumption states that the state transition at time t depends only on the previous state at time $t-1$. (30) expresses the time invariant state transition probability.

$$\begin{aligned} P(Q|\lambda) &= P(q_1, q_2, \ldots, q_T|\lambda) = \\ &\pi_{q_1} a(q_2|q_1) a(q_3|q_2) \ldots a(q_T|q_{T-1}) \end{aligned} \qquad (30)$$

The output independence assumption states that the present observation depends only on the current state and neither chain evolution nor past observations influence if the last chain transition is specified. The mathematical expression of the second assumption is stated in (31).

$$\begin{aligned} P(O|Q,\lambda) &= P(o_1, o_2, \ldots, o_T|q_1, q_2, \ldots, q_T, \lambda) = \\ &P(q_1|\lambda)\prod_{t=2}^{T} P(q_t|q_{t-1}, \lambda) \end{aligned} \qquad (31)$$

V. RESULT

Although no previous study has been carried on the application of LPCC and MFCC on VAG signals, these two feature extraction methods have proved to be successful in extracting desired features from many types of sound signals aside from speech recognition application.

For instance, research on the application of LPCC on milling sound signals has been carried out by Ai [18]. The milling sound has been playing an important role in monitoring the tool wear as high level manufacturing process which emphasizes on precision, efficiency and cost is emphasized. LPCC was applied and the results showed that LPCC was able to extract related features from milling sound signals and concluded the LPPC was appropriate to be utilized to extract parameters from milling sound signals for monitoring purpose.

A feature extraction monitoring analysis has been conducted by M. Hariharan [19] on LPCC for infant cry signal parameters. His experiment enhanced the LPPC method by giving out suitable weighting to the cepstrum coefficients which aimed at reducing the noise sensitivity of the baby

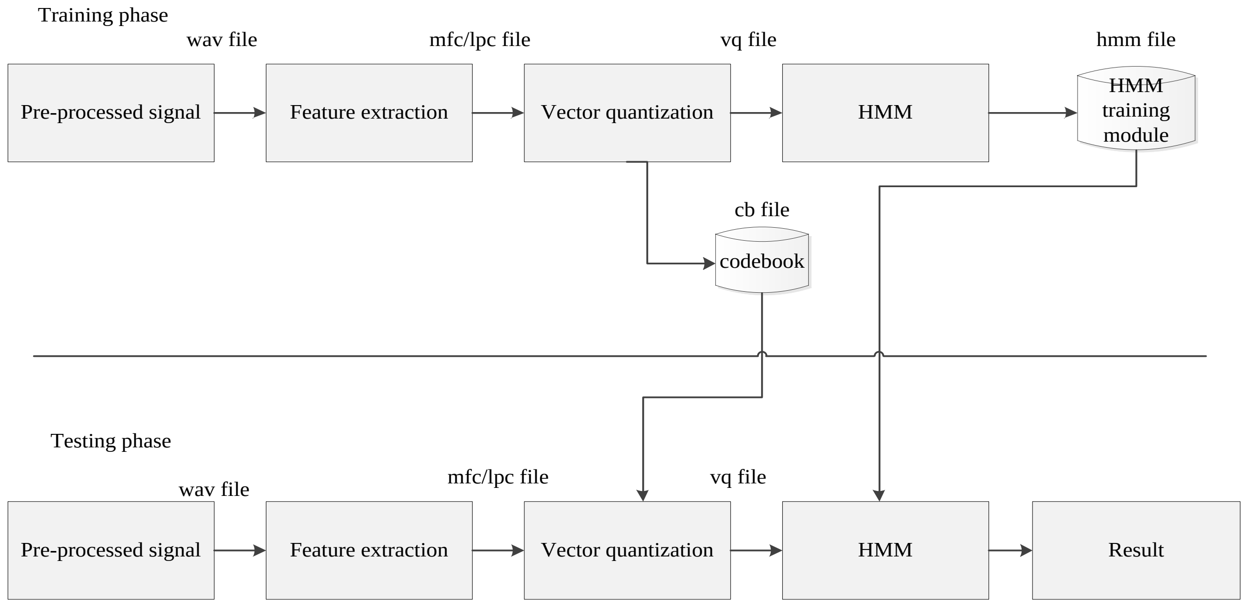


Fig. 6: Training and testing of VAG signal using HMM for the cepstral analysis model

Table 1: Recognition Rates of generated by using LPCC and MFCC. The recognition rate is classified by two elements i.e. intra-subject and inter-subject. Intra-subject measures the accuracy of the experiment by training the VAG signals by subject while inter-subject measures the accuracy of the experiment by training the VAG signals by number of training signal.

	Scale of Analysis	
	Intra-subject	Inter-subject
LPCC Recognition Rate, %	90.17	60.33
MFCC Recognition Rate, %	88.83	56.83

Table 2: Quantitative analysis on the reproducibility of cepstral analysis on VAG signals using mean, variance and standard deviation. In order to ensure the specificity of the study, we have conducted the analysis from the intra-subject and inter-subject perspective.

	Intra-subject		Inter-subject	
	LPCC	MFCC	LPCC	MFCC
Mean, %	90.36	88.64	61.11	57.33
Variance, %	0.23	0.99	3.92	0.49
Standard Deviation, %	0.48	0.99	1.98	0.70

crying signals. The LPCC yielded a classification accuracy of more than 98%.

A study by Gao [15] has been extracting spectral and temporal parameters such as pitch and spectral envelop from the music signal segments by using LPCC and MFCC. The result of the study was promising as an accuracy of 87.92% was achieved for the broad class LPCC and an accuracy of 88.37% was achieved broad class MFCC of the experiment.

In this study, we have quantified the effect of cepstral analysis application by evaluating the accuracy of the study. Table 1 shows the recognition rates generated by using LPCC and MFCC in the VAG signals. The intra-subject analysis exhibited high recognition rates for both LPCC and MFCC, producing accuracy of 90.17% and 88.83% respectively.

However, feature extraction using LPCC and MFCC in the inter-subject analysis rendered unimpressive recognition rates, at 60.33% and 56.83% respectively. This observation reflects the inverse relationship between accuracy of the representation of VAG signal using cepstral coefficient with database size. The VAG signals may contain highly discriminant information that require additive features to harness an effective VAG signal extraction or a weighted cepstral analysis specially designed for VAG signals.

Second important indicator contributed by this paper shows that the LPCC outperforms MFCC in both dependent and independent class of database. The mean recognition rates of LPCC for both inter-subject and intra-subject analyses are obviously higher than those of MFCC. The better overall performances shown by the LPCC may suggest the LPCC is more capable of extracting salient information from the VAG signals, thus, enable the construction of a more comprehensive training database.

Thirdly, the increasing standard deviation as the scale of analysis increases and the large variation indicate the LPCC analysis of VAG signals are inconsistent and need further optimization for better VAG signal parametric representation.

Although the MFCC is not showing good recognition rates, the consistency of MFCC features are obviously better than LPCC as the scale of analysis increase.

VI. Conclusion

We have proposed the use of cepstrum coefficients to extract spectral features from the VAG signals in order to represent the VAG signal for digital signal processing. We concluded that LPCC and MFCC render accurate recognition rates in the case of intra-subject analysis, where the mean recognition rates are 90.36% for LPCC and 88.64% for MFCC. The standard deviations of no more than 1.0% in both methods indicate LPCC and MFCC are able to produce consistent results.

On the other hand, inter-subject analysis requires further improvement and revision before it could produce satisfactory results. The claim is justified by the inaccurate performance displayed by inter-scale analysis. The mean recognition rate of both approaches achieved only 61.11% for LPCC and 57.33% for MFCC.

Despite high accuracies observed by LPCC and MFCC in other applications, the relative inaccuracy of inter-subject analysis demonstrated in this experiment indicates more optimizations are needed and provide further room for future researches on this technology.

Acknowledgment

The authors gratefully acknowledge the research grant provided to this study by Research Management Centre sponsored by Ministry of Higher Education, Malaysia. Vot: Q.J130000.2545.04H41, GUP Universiti Teknologi Malaysia, Johor Bahru, Malaysia

References

[1] Frederic H. Martini, M. J. T., Robert B. Tallitsch, *Human Anatomy 7th Edition*, Pearson Education, 2011.

[2] Ahmad Fadzil Mohd Hani, A. S. M., Dileep Kumar, Raja Kamil, Ruslan Razak, Azman Kiflie, Features And Modalities For Accesing Early Knee Osteoarthritis, *International Conference on Electrical Engineering and Informatics*, Bandung, Indonesia, 2011.

[3] Franssila, R., K. Hedman, Viral Causes Of Arthritis, *Best Practice & Research Clinical, Rheumatology*, Vol.20, No.6, pp. 1139-1157, 2006.

[4] Batra, Abhinandan, Comparative Study of the Effects of Tai Chi and Strength Training on Osteoarthritis in Older Adults, *Gerontology Theses*, Georgia State University, Georgia, United States, 2011.

[5] Karim, Z., R. J. Wakefield, et al., Validation and reproducibility of ultrasonography in the detection of synovitis in the knee: A comparison with arthroscopy and clinical examination, *Arthritis & Rheumatism*, Vol.50, No.2, , pp. 387-394, 2004.

[6] K.S. Kim, C. G. S., J.H. Seo, K.H. Park, J.Y. La, and J.C. Kim, An Efficient Algorithm to Improve Feature Extraction and Classification of Knee Joint Sound, *ICME International Conference on Complex Medical Engineering, 2009. CME*, pp. 1-6, 9-11 April 2009 2009.

[7] Sridhar Krishnan, C. B. F., Rangaraj M. Rangayyan, G. Douglas Bell, Adaptive Time-Frequency Analysis of Knee Joint Vibroarthrographic Signals for Non-Invasive Screening of Articular Cartilage Pathology, *IEEE Transactions on Biomedical Engineering*, Vol.47, No.6, pp. 773-783, 2000.

[8] Y.C. Hum, K.W. Lai, T.S. Tan, S. Sh-Husssain, "Adaptive Crossed Reconstructed (ACR) K-mean Clustering Segmentation for Computer-aided Bone Age Assessment System," *International Journal of Mathematical Models and Methods in Applied Sciences*, Vol. 5, No. 3, pp. 628-635, 2011.

[9] Y.C. Hum, K.W. Lai, T.S. Tan, S. Sh-Husssain, Y. C. Lim, "An Artifacts Removal Post-processing for Epiphyseal Region of Interest (EROI) Localization in Automated Bone Age Assessment (BAA)," *Biomedical Engineering Online*, Vol. 10, No. 87, pp. 1-22, 2011.

[10] Y.C. Hum, K.W. Lai, T.S. Tan, S. Sh-Husssain, "GLCM based Adaptive Crossed Reconstruction (ACR) K-mean Clustering Hand Bone Segmentation," *Proceeding of 10th WSEAS international conference on electronics, hardware, wireless and optical communications, and 10th WSEAS international conference on signal processing, robotics and automation, and 3rd WSEAS international conference on nanotechnology, and 2nd WSEAS international conference on Plasma-fusion-nuclear physics*, pp. 192-197, 2011.

[11] Y.C. Hum, K.W. Lai, T.S. Tan, S. Sh-Husssain, Grey-Level Co-occurance Matrix Bone Fracture Detection, *WSEAS Transactions on Systems*, Vol. 10, No. 1, pp. 7-16, 2011.

[12] Sanjeev Tavathia, R. M. R., Cyril Basil Frank, Gordon Douglas Bell, Katherine O. Ladly, Yuan-Ting Zhang, Analysis of Knee Vibration Signals Using Linear Prediction. *IEEE Transactions on Biomedical Engineering*, Vol.39, No.9, pp. 959-970, 1992.

[13] S. Krishnan, R. M. R., G.D. Bell, C.B. Frank, K. O, Ladly, Adaptive Filtering, Modelling and Classification of Knee Joint Vibroarthrographic Signals for Non-Invasive Diagnosis of Articular Cartilage Pathology. *Med. Bio. Eng. Comput*., Vol.35, pp. 677-684, 1997.

[14] A. Steindler, Auscultation of Joints. *The Journal of Bone and Joint Surgery*, Vol.21, No.1, pp. 121-136, 1937.

[15] Sheng Gao, Namunu chiothaka Maddage, Chi Hui Lee, A Hidden Markov Model Based Approach To Music Segmentation And Identification, *ICICS-PCM 2003*, Singapore, pp. 1576-1580, 2003.

[16] Zahra M. K. Moussavi, Rangaraj M. Rangayyan, G. Douglas Bell, Cyril B. Frank, K. O. Ladly, Yuan Ting Zhang, Feature Extraction of Knee Joint Sound for Non-Invasive Diagnosis of Articular Pathology, *IEEE Transaction on Biomedical Engineering*, Vol.43, No.1, pp. 349-352, 2008.

[17] Kim, K. S., J. H. Seo, et al, Feature Extraction of Knee Joint Sound for Non-Invasive Diagnosis of Articular Pathology, 2008, pp. 349 - 352.

[18] C.S. Aia, Y.J. Sun, G.W. He, Tool Wear Monitoring Based on Milling Acoustic Spectrum LPCC, *Key Engineering Materials*, pp. 353-358, 2012.

[19] M. Hariharan, Lim Sin Chee, Sazali Yaacob, Analysis of Infant Cry Through Weighted Linear Prediction Cepstral Coefficients and Probabilistic Neural Network, *J Med Syst*, Vol.36, No. 3, 2012, pp. 1309-1315.

[20] Juhani Saastamoinen, Evgeny Karpov, Ville Hautamäki, Pasi Fränti, Accuracy of MFCC-Based Speaker Recognition in Series 60 Device, *EURASIP Journal on Applied Signal Processing*, Vol.17, pp. 2816–2827, 2005.

[21] Nizamettin Aydin, Hugh S. Markus, Optimization of Processing Parameters for the Analysis and Detection of Embolic Signals, *European Journal of Ultrasound*, Vol.12, pp. 69–79, 2000.

[22] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE*, Vol. 77, pp. 257-286, 1989.

[23] S.A.R. Al-Haddad, Speech Disorder Malay Speech Recognition System, Sensors, Signals, Visualization, Imaging, Simulation and Materials.

[24] Sridharan, E. W. a. S., Comparison of Linear Prediction Cepstrum Coefficients, *Proceedings of 2001 lnternational Symposium on Intelligent Multimedia, Video and Speech Processing*, Hong Kong, pp. 95 -98, 2001.

[25] R. Rangayyan and Y. Wu, Screening of knee-joint vibroarthrographic signals using statistical parameters and radial basis functions, *Medical and Biological Engineering and Computing*, Vol.46, pp. 223-232, 2008.

[26] Anjali Bala, Abhijeet Kumar, Nidhika Birla, Voice Command Recognition System Based on MFCC and DTW, *International Journal of Engineering Science and Technology*, Vol.2, No.12, pp. 7335-7342, 2010.

[27] J. W. Picone, "Signal modeling techniques in speech recognition," Proceedings of the IEEE, Vol. 81, pp. 1215-1247, 1993.

[28] Biing-Hwang Juang, L. A. R. R. A. J. G. W., On the Use of Bandpass Liftering in Speech, *IEEE Transactions on Acoustics, Speech, and Signal Processing*, Vol.35, No.7, pp. 947-954, 1987.

[29] L. R. Rabiner, "A tutorial on hidden Markov models and selected applications in speech recognition," *Proceedings of the IEEE,* Vol. 77, pp. 257-286, 1989.

[30] D. Gargouri, M. A. Kammoun, and A. B. Hamida, "A comparative study of formant frequencies estimation techniques," presented at the *Proceedings of the 5th WSEAS international conference on Signal processing*, Istanbul, Turkey, 2006.

[31] M.Hassan Shirali-Shahreza, S. S.-S., Effect of MFCC Normalization on Vector Quantization Based Speaker Identification, *2010 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, pp. 250-253*),* 2011.

[32] Ashok K. Krishnamurthy, S. C. A., Douglas E. Melton, Prakoon Chen, Neural Networks for Vector Quantization of Speech, *IEEE Journal on Selected Areas in Communications*, Vol.8, No.8, pp. 1449-1457, 1990.

[33] R. M. Rangayyan and Y. Wu, Screening of knee-joint vibroarthrographic signals using probability density functions estimated with Parzen windows, *Biomedical Signal Processing and Control*, Vol.5, No.1, pp. 53-58, 2010.

[34] A. Cohen, "Hidden Markov models in biomedical signal processing," in Engineering in Medicine and Biology Society, *Proceedings of the 20th Annual International Conference of the IEEE, 1998*, Vol.3, pp. 1145-1150, 1998.

[35] S. Davis and P. Mermelstein, Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences, *IEEE Transactions on Acoustics, Speech and Signal Processing*, Vol.28, No.4, pp. 357-366, 1980.

[36] J. W. Picone, Signal modeling techniques in speech recognition, *Proceedings of the IEEE,* Vol.81, No.9, pp. 1215-1247, 1993.

[37] P. Bansal, A. Dev, and S. B. Jain, Role of different order ranges of autocorrelation sequence on the performance of speech recognition, *WTOS*, Vol 9, pp. 1-9, 2010.

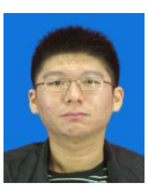
[38] K. Wen-Juh, S. Jiue-Rou, C. Cheng-Kung, L. Jin-Shin, T. Hen-Wai, and K. Te-Son, The application of cepstral coefficients and maximum likelihood method in EMG pattern recognition [movements classification], *IEEE Transactions on Biomedical Engineering*, Vol. 42, pp. 777-785, 1995.

[39] S. Chauhan, P. Wang, C. Sing Lim, and V. Anantharaman, "A computer-aided MFCC-based HMM system for automatic auscultation," *Computers in Biology and Medicine*, Vol. 38, pp. 221-233, 2008.

[40] O. Horak, “The voice segmentation type determination using the autocorrelation compared to cepstral method,” *WSEAS Transactions on Signal Processing*, Vol. 8, No. 1, PP. 11-20, 2012.

[41] H. Kuscu, K. Kahveci, U. Akyol, A. Cihan, “Robust autocorrelation testing in multiple linear regression,” *International Journal of Mathematics and Computers in Stimulation*, Vol. 6, No. 1, pp. 119-126, 2012.

[42] O. Grigore, C. Grigore, V. Velican, “Impaiered speech evaluation using Mel-Cepstrum analysis,” *International Journal of Circuits, Systems and Signal Processing,* Vol. 5, No. 1, pp. 70-77, 2011.



Gan Hong Seng, was born in Negri Sembilan, Malaysia on October 08, 1988. In 2012, he graduated with Bachelor of Engineering (Bioemdical) from Universiti Teknologi Malaysia.

Previously, he has been working as research assistance in Universiti Teknologi Malaysia. Now, he is a PhD (Biomedical Engineering) Candidate in Universiti Teknologi Malaysia (UTM) and a Research Member in Medical Implant Technology Group Materials and Manufacturing Research Alliance. His interests of research cover MR image processing, knee joint signal processing, artificial intelligence, and medical computing.

Mr. Gan is a member of Board of Engineering Malaysia (BEM).

Tan Tian Swee, was born in Johor, Malaysia on June 05, 1978. He received B.Eng (Electric-Mechatronic), M.Eng (Electrical Eng.), Ph.D (Electrical Eng) from Universiti Teknologi Malaysia.

He is a senior lecturer in Universiti Teknologi Malaysia. Prior to his appointment as senior lecturer, he was a tutor in Universiti Teknologi Malaysia and a researcher at Medical Implant Technology Group Materials. His current research interest includes signal processing for speech and medical image processing.

Dr. Tan is a senior member of Board of Engineering Malaysia (BEM).