Speech to Text Engine for Jawi Language

Zaini Othman[1], Nor Abdullah[2], Zaidi Razak[3], and Mohd-Yakub Mohd-Yusoff[4]
[1, 2, 3]Faculty of Computer Science and Information Technology, University of Malaya, Malaysia
[4]Academy of Islamic Studies, University of Malaya, Malaysia

Abstract: *This paper focused on the development of speech translation to special character that is Malay speech to Jawi text engine. Jawi is a unique character derived from Arabic but it is read in Malay language. There are not many research can be found on speech technology developed for Jawi and this research would be useful to researcher who wish to venture its benefit to many related ICT applications. The use of Zero Crossing Rate (ZCR) as a robust algorithm for accurate automatic detection of speech signal syllable boundary has been discussed. The combination of Linear Predictive Coding (LPC) and Artificial Neural Network (ANN) are used in this research to extract and classify the speech signals with backpropagation training method. This paper also, discussed on the use of Jawi Unicode in the final character tagging process to represent each of the Jawi character existed in the spoken word. As there are no standard lists of Jawi Unicode published, in this research, the existing of Jawi Unicode table produced by previous research is further investigated and enhanced in order to have better accuracy in Jawi character-phoneme representation. This list is based on the combination of Traditional Arabic and other scripts. A prototype educational learning tool was also, developed to enable school children to recognize and read Jawi text, check their pronunciation, and learn from their mistakes independently.*

Keywords: *Speech-to-text, jawi unicode, LPC, ANN, ZCR.*

Received May 24, 2012; accepted April 11, 2013; published online February 26, 2014

1. Introduction

Jawi language is an old Malay writing system which was first introduced during British colonial era [11], particularly for religious and cultural traditions in the South-east Asia region [19]. However, literature shows that since, the introduction of modern Roman-based writing system, Jawi has become less popular [4, 5, 11, 13]. In fact, many Malaysian has forgotten on how to read Jawi. To preserve the language, the Malaysian government has started the Jawi, Quran, Arabic, Fardu Ain (JQAF) initiative [13]. However, studied has shown that Malaysian students are weak in reading Jawi especially in reading as current primary school children only learn Jawi at school [1, 13].

The authors' related paper has investigated the combination of Linear Predictive Coding (LPC) and Artificial Neural Network (ANN) for extracting and classifying speech features into Jawi text [19]. While previous paper only discussed the early stage of this research, this paper focused more on the needs to enable students to learn how to read and write in Jawi language independently [13]. This has extended our research to develop a robust Jawi Speech-To-Text application, as a supplementary educational aid.

This paper introduces the creation and use of Jawi Unicode as suggested by [18] in order to display the recognized signal into respective Jawi text. In this research we have enhanced the Jawi Unicode list based on the research done earlier by Toru [16] for a better accuracy. The list of Jawi Unicode is formed with combination of Traditional Arabic, Persian and Urdu Unicode. This paper also, focused on the use of Zero Crossing Rate (ZCR) as a robust algorithm for accurate automatic detection of syllable boundary. The process of segmenting the speech signal into smaller parts is very important, in order to indentify consonant-vowel phonemes sounds.

2. Linguistic Properties of Jawi Script

In this section, we described the linguistic properties of Jawi script that is relevant to this study. Malay is a phonetic language which can be written in two types of characters; Roman and Jawi characters. Jawi consists of six vowel phonemes (*/a, e, é, i, o, u/*), three diphthong phonemes (*/ai, au, oi/*) and 25 consonant phonemes (*/b, c, d, f , g, ğ, h, dZ, k, kh , l, m, n, ŋ, ń, p , q, r, s, ś, t, v, w, y, z/*) as proposed by Dewan [3], an organization that is responsible for monitoring the use of Malay language in Malaysia.

In order to recognize spoken words, the system analyses speech signal in a small segments and this segment is compared to a Jawi database to identify the phonemes it represents. Once the phoneme has been identified, the next task is to display the correct characters respective to the actual output by using the Unicode. Unicode is a way to represent characters and symbols in a system by providing a unique sequence of numbers for every character. In Jawi characters, there is no standard version of Jawi Unicode created by Unicode Consortium. Therefore the Jawi Unicode list

Table 1. Unicode listing for jawi character ([16, 18]).

Jawi Character	Name	16 Bits	Jawi Character	Name	16 Bits	Jawi Character	Name	16 Bits	Jawi Character	Name	16 Bits
ا	Alif	0627	ذ	Dzal	0630	ع	Ain	0639	ه	Ha	0647
ب	Ba	0628	ر	Ra	0631	غ	Ghain	063A	ء	Hamzah	0621
ت	Ta	062A	ز	Zai	0632	ف	Fa	06A7	چ	Ca	0686
ة	Ta Marbutah	0629	س	Sin	0633	ق	Qaf	0642	ڠ	Nga	06A0
ث	Tha	062B	ش	Syin	0634	ک	Kaf	06A9	ڤ	Pa	06A8
ج	Jim	062C	ص	Sad	0635	ل	Lam	0644	ݢ	Ga	0762
ح	Ha	062D	ض	Dhad	0636	م	Mim	0645	ڽ	Nya	06BD
خ	Kha	062E	ط	Tho	0637	ن	Nun	0646	ۏ	Va	06CF
د	Dal	062F	ظ	Zho	0638	و	Wau	0648	ی	Ye	06CC
						لا	Lam Alif	FEFB	ي	Ya	064A

grouped by Toru [16] has been updated and enhanced to get a more accurate Unicode representation as shown in Table 1 below. This is because some of the former Unicode number chosen is not representing the correct character or symbol for Malaysian Jawi. For an example, the character 'Ga' in Malaysian Jawi is more accurate to be represented by this character ݢ (Unicode:U+0762) instead of گ (Unicode:U+06AF) as suggested by [16].

3. Related Works

As mentioned before, our previous paper discussed the early stage of this research and investigated the combination of LPC and ANN for extracting and classifying speech into Jawi text [19]. In this paper, 15 words from Jawi character are chosen for training process by 15 speakers (10 males, 5 females). In controlled environment, the result of accuracy exceeded 91%.

3.1. Automatic Speech Recognition (ASR)

Salam *et al.* [15] in their paper explained works in Malay isolated digit speech recognition using neural network. The network used in the experiment is feed forward multilayer perceptron trained with backpropagation scheme. Speech data is analyzed using LPC to represent a fixed overlapped window. In this paper, digit utterances were recorded from 4 speakers (2 males, 2 females) without any dialect in silent environment. Best recognition rate achieved was 95% using 320 input nodes, 45 hidden nodes and 4 output nodes.

Lori *et al.* [7] developed a speech-to-text system or transcription of broadcast data in Arabic. The development of a speech recognition system for the Arabic language faced few challenges, rising from the differences in its spoken and written forms. This is because Arabic is a strongly consonantal language and has a large variety of vowelization for each written form. It is also, spoken in many different dialects and pronunciation which only affect the spoken form and not the written forms. This research focuses on the morphological decomposition method to address the challenge of dealing with huge Arabic lexical database.

Moaz and Rasheed [8] in their research have explored phonetic recognition of Arabic letters using neural networks. The main features of the voice signal are extracted using PCA technique and classified using Neural Network. PCA coefficients corresponding to each alphabet are used to train feed-forward neural networks to produce recognized letter. About 96% detection rate has been achieved over defined dataset. In this research, 2016 sound samples taken for analysis from 6 different speakers involving 28 letters.

3.2. Arabic Speech Segmentation

The study on related Arabic language speech recognition system has given good guidance and examples to our research as Jawi script consists of characters similar to Arabic in its written forms.

Mohammed *et al.* [9] have discussed the comparison between automatic and manual approach of speech segmentation using energy measurement method.

The phonemes classification has been divided into two parts; high energy phonemes and low energy phonemes corresponding to voiced and unvoiced phonemes.

Muhammad *et al.* [10] have reported a basic speech segmentation application developed for Arabic language with the aim to further develop it as a language tutor. The focus is on the Quranic Arabic as there are standards available which help in obtaining better accuracy. The need to identify the phoneme boundaries have cues them to include ZCR and Power Spectral Density (PSD) in their research as segmentation methods. Their system has demonstrated up to 89% accuracy on Arabic speech spoken words.

Kirchhoff and Vergyri [6] have done research on the varieties appeared in Arabic dialects which has makes the development of Arabic automatic speech recognition a challenging task. The paper focused on automatic vowelization Modern Standard Arabic (MSA) data to address this problem. The process is done using a combination of syntactic, morphological and acoustic information sources.

3.3. Speech for Language Learning

Apart from the research done in the Arabic language, we also, studied several researches involving speech recognition developed for education. Neri *et al.* [12] studied on how automatic speech recognition is able to train pronunciation in second language learning. A system known as Computer Assisted Pronunciation Training (CAPT) has been reviewed to be able to detect student's mispronunciation, making the learning process more realistic.

Rebecca [14] in her paper investigated on the effectiveness of speech recognition system to a group of students who enrolled for 200-hours, ten-week English course in Sweden. In this course, five hours have been allocated for students to use computer-assisted pronunciation tutoring program. Eleven students tried the program Talk to Me by French company Auralog and were encouraged to practice at home. Talk to Me uses speech recognition technique to provide conversational practice, visual feedback and scoring of pronunciation. In this evaluation, students have to keep track of how many hours they used the program. The resulting hours spend throughout this course ranged widely from 2 to 48 hours with an average of 12.5 hours spend by each student. Students have reported high satisfaction with the software and thought the program has improved their English pronunciation. However, as the program used a lot of British English model in their database, it seems to be beneficial for those students who have a strong foreign accent (e. g., British and American) only.

Speech technology provides an alternative method for teaching and learning language. These researches proved that with proper adaptation into language course at school, speech technology allows students to practice spoken language outside the classroom.

4. Proposed Development of Malay Speech to Jawi Text Engine

Figure 1 below shows the experimental framework. In this section, the speech segmentation process, which is an important phase in feature extraction, will be discussed in subsection 4.1. The subsection 4.2 will discuss on the application of LPC and ANN as the chosen algorithm for extraction and classification.

4.1. Boundary Estimation Based on Zero Crossings

Basically, the speech signal that has been recorded by speakers is often too long to make the extraction process possible for the whole signal. The properties of speech are also, time-varying and we cannot analyse the whole signal to extract its features. To solve this problem, we should divide or segment the signal into small frames and prepare them for short term analysis. The process is known as speech segmentation which involves dividing speech utterances into different chunks which are recognizable and meaningful.

This paper focuses on the implementation of ZCR in segmenting the recorded voice based on the syllables boundary. These segments will determine the segment's type; voiced-unvoiced-silent group of syllables. ZCR is the rates of changes in amplitude that occur throughout the signal waves, everytime the signal passes through the value of zero.

The reason why ZCR is important to trace syllable boundary in Jawi spoken word is because ZCR has been proven to be effective in Arabic speech recognition system [9, 10].

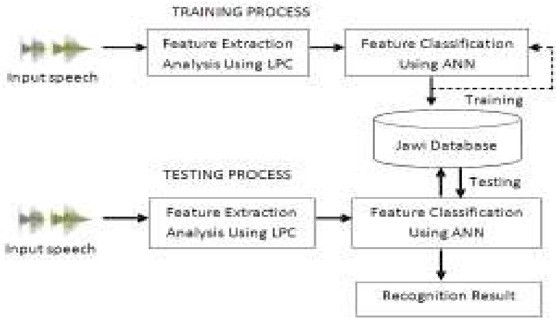


Figure 1. An experimental framework for malay speech to jawitext engine.

Jawi is a phonetic language where its character derived from Arabic but it is read in Malay. Furthermore, some of the Jawi words are pronounce in Arabic styles and dialects, for example the presence of 'makhraj'. The term 'makhraj' in Arabic refers to the articulation stress sound which comes out of the mouth based on the position or shape of articulation muscle [17]. Considering this pronunciation styles and dialect, we can say that Arabic language has a huge influence for Jawi character. Literally, we have divided the Jawi phonemes into two main class, voiced and unvoiced phonemes. Voiced class of signal usually represents vowel and contains higher energy as compared to the unvoiced [2]. In Jawi, voiced phonemes refers to sound produced by vowel characters, / ا, ي, و / and unvoiced phonemes refers to sound produced by consonant characters, / ب ، ث ,ت، ج، ح، خ، د، ذ، ر، ز، س، ش، ص، ض، ط، ظ، غ، ف، ق، ک، ل، م، ن، و، ه، ي، چ، ڠ، ڤ، ݢ، ۏ، ڽ /. Automated segmentation aims to segment the speech signal according to this voiced-unvoiced class before they can be recognized.

Figure 2 shows the ZCR value detected from the word 'توليس', which is phonetically read as /*t-u-l-i-s*/. From the figure, we can see that the unvoiced frames are relatively marked as 0 and voiced frames are marked as 1. If the value of the zero crossing rates is higher than the threshold value, the speech segment is considered as unvoiced (marked as 0) and if the value of the zero crossing rates is lower than threshold value, the speech segment is considered as voiced (marked as 1).

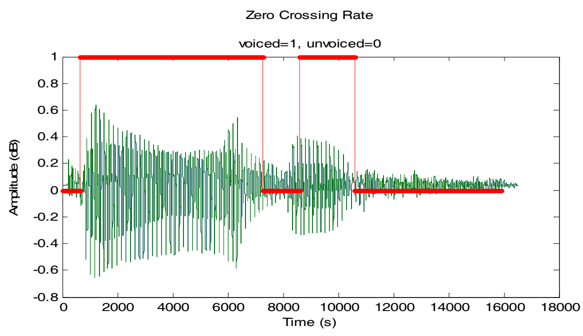


Figure 2. Value of zero crossing rate for the word 'توليس'.

The energy level of that particular speech segment also, determines the speech segment type as shown Figure 3.

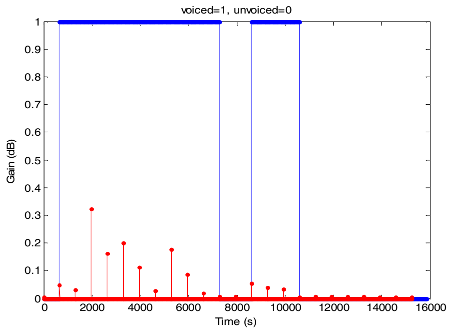
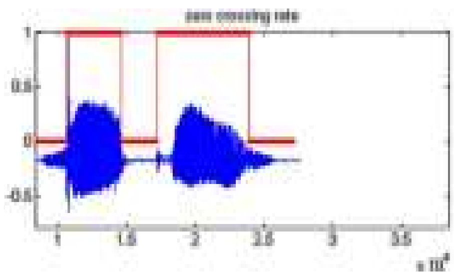
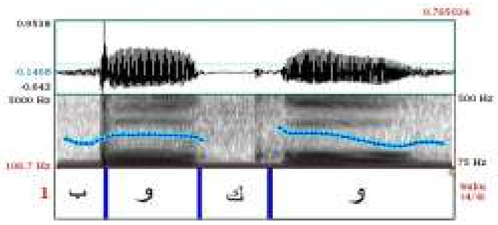


Figure 3. Energy content for the word 'توليس'.

The result obtained is consistent with the threshold theory as the ZCR value are low and energy level is high, the speech segment can be considered as voiced signal. In order to ensure that the segmentation is accurate, we have make a comparison between the automated and manual segmentation from the Praat software, a time-frequency tools and also, through repeated listening by user to identify the targeted phoneme as shown in Figure 4 below.



a) Automated Segmentation for the word 'بوكو'.



b) Manual Segmentation using Praat for the word 'بوكو'.

Figure 4. Segmentation comparison.

In the automated segmentation test as shown in Figure 4-a, the word 'بوكو' is recorded from speaker and extracted using LPC algorithm. ZCR threshold is set to segment the processed signal based on phonemes boundary. The segmented boundary is then compared to the manual segmentation to check its accuracy level. For example, the character 'و' in Figure 4-b is classified as a vowel phoneme in Jawi and a vowel phoneme is referred as voiced signal. Therefore, we can conclude that automated segmentation using ZCR has successfully detected the voiced segment and syllable boundary as those obtained from the manual segmentation.

4.2. Feature Extraction and Feature Classification

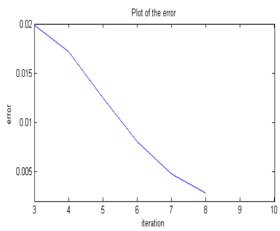
In order to recognize voice signal, basically a speech sample need to be extracted to retrieve the speeches important features for example, the pitch and energy values. The speech first needs to be segmented and the segment will be represented in coefficient values that will be the input for feature classification process. The coefficients are then will be classified in order to recognize the phoneme or word represented by the segment. This is usually done by comparing the parameter coefficients with the sample parameter from database. In this research, LPC has been applied as feature extraction technique and ANN has been applied as feature classification technique.

Speech is a non-stationary signal but it is assumed that vocal tract is stationary for at least 10-30 ms [15]. Hence, in this research the speech signal recorded from speaker will be divided into small frames of equal size range from 10 to 30 ms so that stationary operation can be performed. ZCR algorithm is used to determine the segmentation boundary which has been discussed in section 4.0. Based on pitch and energy values, each of speech segments will produce a feature vector which is called as LPC coefficient.

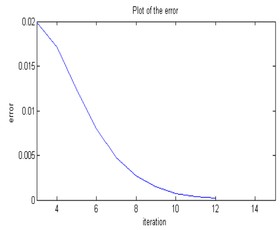
The LPC coefficient will be the input for neural network to classify the phoneme represented. In this research, the combination of feed-forward and back-propagation network training algorithm is used to get the final output since it has been successfully applied to Arabic speech recognition [9, 10]. Each neuron in a network is interconnected by a weight value. Accurate weight value will help recognized the correct word more efficiently. As the first loop of feed-forward algorithm assigned random weight value from input layer towards the output layer, the backpropagation algorithm goes through the network iteratively, update the weight in each layer backwards from the output layer towards the input layer so, that the error gradually becomes smaller.

Training process is done to get the correct weight value for database reference. In this research, 10 data of LPC coefficients are used as neural network training input while 30 iteration loops is applied to train the database. This means that the training is terminated when it reached the correct targeted word or reached 30, depends on which comes first. This maximum iteration, which is also, known as epochs, is decided after several testing on neural network using 5, 10, 20, 30, 40 and 50 epochs. In this research, 30 epochs are chosen due to the results of neural network computation that produce the most minimal error reduced with shorter time taken to complete.

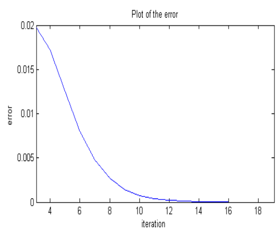
Figure 5 shows the training graph for the word 'جاوي'. The goal of Backpropagation training is to minimize the error so, that the actual output matches the target output. As we can see, the Backpropagation algorithm has gradually reduces the error values from 0.0267 to 0.0001, which makes the actual output closer to the target output. Five neuron output has been set as the final output in the network to produce binary results for phoneme's reference in Jawi database which indicates all phonemes in binary representation. For example, the binary target output for Jawi consonants starts from 001101 until 100101 for a total of 25 Jawi consonant phonemes. Each of this target output represents a phoneme and thus represents a Jawi character. This character is later combined at the end of recognition/testing process to form Jawi word.



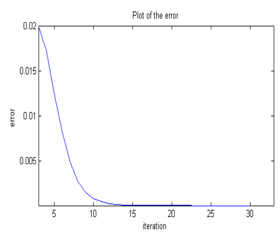
a) ANN training graph at 10 epochs.



b) ANN training graph at 14 epochs.



c) ANN training graph at 18 epochs.



d) ANN training graph at 30 epochs.

Figure 5. ANN training graph progress.

5. An Application of Malay Speech to Jawi Text

Apart from the significant value in Jawi-Arabic speech recognition field, this project also, aims to enhance the Jawi reading ability among students. Multimedia elements in computer application are fascinating, not only because of their power to influence cognitive and motivational processes, but especially because of their capability to attract children's attention to continue learning [4]. Therefore, in order to allow users to access the system in a more user-friendly and attractive way, an application for learning Jawi has been developed. The application basically displays Jawi text for users to learn to read, and other functions for users to record, play, analyze and recognize the recorded speech. Figure 6 shows the screenshots of this application for a correct pronunciation of Jawi text.

This application allow user to record his/her voice by clicking the button 'Record' and recognized the pronounced word by clicking the button 'Recognize'. The result will be shown at the recognition area. From the figure, the recognition area produces result of the correctly recognized word ('ساتو ') in black colour. For incorrect pronunciation, the resulted Jawi characters are formed in red colour, which indicates that the spoken output did not match the selected word. Users are also, be able to listen to the correct pronunciation by clicking the 'Correction' button and try to record their reading again.

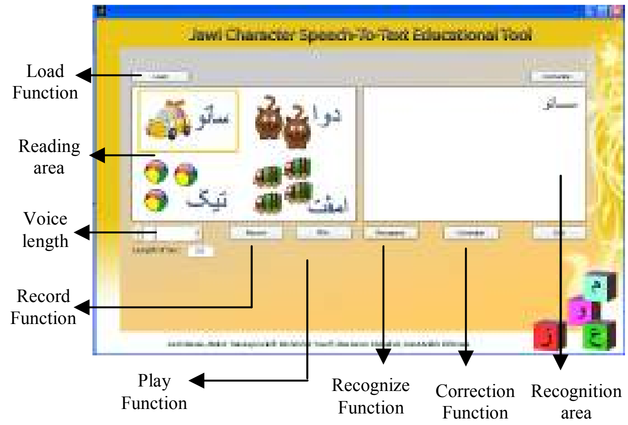


Figure 6. Screenshots of malay speech to jawi text application.

6. Experimentation and Results

In this section we described experimental results and discussion of this research. In the training phase, the system is actually trained to map the spoken word with the correct pronunciation from the database. The correct coefficients used for correct mapping is saved.

6.1. Speech Samples

Experimentation process begins with speech collection by recording samples of voice using MATLAB® wavrecord and wavwrite function. The numbers of speech samples are 320 samples, which consists of 64 words spoken by 5 speakers. In Jawi writing system, there are some Jawi characters that produce similar phoneme. For example, vowel phoneme /a/ can be represented by character ا or ع. This is because some of the Malay words are derived from the Arabic, which influence the pronunciation in a more Arabic-like sound with the additional of 'makhraj'. There are six phonemes which match this similarity; /a, d, k, t, s, z/. Therefore, the selected words are divided in 6 sets based on the six phonemes. The purpose of having dataset for the six phonemes is to test the pronunciation accuracy of Arabic-derived words by Malay students.

6.2. Experimental Results

LPC coefficients values are the output from segmentation process with a total of 10 vectors for each segment. In testing process, the recognition of speech samples are considered success if the final outputs of each phoneme are correctly recognized and tagged according to its Jawi character. This is done by manually comparing the final output with the correct word displayed at the recognition area. Table 2 shows the overall accuracy percentage from testing set 1-6.

Based on the recognition accuracy percentage shown in Table 2, we can see that testing set 4 get the highest score with 100% accuracy.

Table 2. Testing set 1-6: recognition accuracy percentage.

Test Set	Phoneme	Num. of words	Num. of samples	Accuracy Percentage (%)
1	/a/	10	50	88
2	/d/	7	35	86
3	/k/	10	50	86
4	/t/	7	35	100
5	/s/	12	60	88
6	/z/	8	40	83

In set 4, both characters 'ت' and ' ط' produce the similar /t/ phoneme sound. However, the word with character 'ط' is supposed to be pronounced differently with the presence of 'makhraj'. In this testing set, most of the Arabic words listed are well known and widely used. Therefore, they are pronounced well in the correct Arabic style. This makes the accuracy of the words to be detected higher.

In the same table we can see that Testing Set 6 get the lowest score with 83% accuracy. In this set, the similar sound of phoneme /z/ is being pronounced by three characters 'ذ', 'ز' and 'ظ'. Based on the testing result, words for character 'ظ' is harder to be recognized compared to words for character 'ذ' and 'ز'. For example, the word 'zalim' in this set only achieve 40% accuracy. This is because the word 'zalim' is usually pronounced in local dialect but according to Jawi spelling, it is supposed to be pronounced in Arabic-like, for example, 'dzolim'. Therefore, the difference between these two pronunciations has affected the accuracy of the recognition. The overall 54 words recognition exceeded 88.5% accuracy. The highest accuracy percentage is 100% recognized (35 words) while the lowest percentage is 40% (2 words).

7. Conclusions

The significance of this research lies in the uniqueness of Jawi script as the characters are derived from Arabic characters but pronounced in Malay. The existing of Jawi Unicode table produced by Toru [16] is further enhanced for better accuracy. This list is based on the combination of Traditional Arabic, Persian and Urdu Unicode. This research has also, covered important aspects of speech technology applied for language with special characters. Based on experimental results, the advantage of LPC-ANN combination is the execution of LPC extraction processes which is simple and straight forward together with the used of ANN nonlinear speech modeling which is able to reach maximum accuracy upon controlled environment. Direct application of LPC also, allows us to focus more on the speech segmentation. The use of ZCR for accurate automatic detection of speech signal syllable boundary is experimented. This is the first attempt of using ZCR in Jawi research and it is proven to be accurate for both Jawi and Arabic language. The experimental results denote an average of 89% of recognition accuracy of the proposed engine.

References

[1] Amin A., "Tulisan Jawi Ke Arah Penggunaan Yang Lebih Meluas Dan Berkesan," *Journal Dewan Bahasa*, vol. 33, no. 10, pp. 937 - 941, 1989.

[2] Beritelli F., Casale S., Russo A., and Serrano S., "Adaptive V/UV Speech Detection Based on Characterization of Background Noise," *Journal on Audio, Speech, and Music Processing*, vol. 2009, no. 11, pp 1 - 12, 2009.

[3] Dewan B.,"Klik DBP Edisi Jun," Kuala Lumpur: Dewan Bahasa Pustaka, 2005.

[4] Diah N., Ismail M., Ahmad, S., and Syed A., "Jawi on Mobile Devices with Jawi Word Search Game Application," *in Proceedings of IEEE International Conference of Science and Social Research*, Kuala Lumpur, Malaysia, pp. 326 - 329, 2010.

[5] Hairul A., Abdul H., and Moahmud R., "Using an Edutainment Approach of a Snake and Ladder Game for Teaching Jawi Script," *in Proceedings of IEEE International Conference on Education and Management Technology*, New York, USA, pp. 228 - 232, 2010.

[6] Kirchhoff K. and Vergyri D., "Cross-Dialectal Data Sharing for Acoustic Modeling in Arabic Speech Recognition," *Journal of Speech Communication*, vol. 46, no. 1, pp. 37 - 51, 2005.

[7] Lori L., Abdelkhalek M., and Jean-LucG., "Automatic Speech-to-Text Transcription in Arabic," *ACM Transactions on Asian Language Information Processing*, vol. 8, no. 4, pp. 1 - 18, 2009.

[8] Moaz A. and Rasheed M., "Phonetic Recognition of Arabic Alphabet Letters Using Neural Networks," *International Journal of Electric & Computer Sciences IJECS-IJENS*, vol. 11, no. 1, pp. 44 - 49, 2011.

[9] Mohammed A., Mohammed I., and Mansour M., "Arabic Speech Segmentation: Automatic Verses Manual Method and Zero Crossing Measurements," *Indian Journal of Science and Technology*, vol. 3, no. 12, pp. 1134 - 1138, 2010.

[10] Muhammad J., Awais M., Shahid M., and Shafay S., "Automatic Arabic Speech Segmentation System," *International Journal of Information Technology*, vol. 12, no. 6, pp. 102 - 111, 2006.

[11] Nafisah A., "Romanization of Multiscript/ Multilingual Materials: Experiences of Malaysia," *in Proceedings of the 65th IFLA Annual Conference Bangkok*, Bangkok, Thailand, pp., 1999.

[12] Neri A., Cucchiarini C., and Strik W., "Automatic Speech Recognition for Second Language Learning: How and Why it Actually Works," *in Proceedings of the 15th International Congress of Phonetic Sciences Barcelona*, Spain, pp. 1157 - 1160, 2003.
[13] Nik R., "Penguasaan Jawi Dan Hubungannya Dengan Minat Dan Pencapaian Pelajar Dalam Pendidikan Islam," *Jurnal Pendidik Dan Pendidikan*, vol. 22, pp. 161 - 172, 2007.
[14] Rebecca H., "Speech Recognition for Language Teaching and Evaluating: A Study of Existing Commercial Products," *in Proceedings of ICSLP*, Jeju Island, Korea, pp. 733 - 736, 2003.
[15] Salam M., Mohamad D., and Salleh S., "Malay Isolated Speech Recognition Using Neural Network : A Work in Finding Number of Hidden Nodes and Learning Parameters," *the International Arab Journal of Information Technology*, vol. 8, no. 4, pp. 364 - 371, 2011.
[16] Toru A., Jawi Study Group, "Annuals of Japan Association for Middle East Studies," *Annals of Japan Association for Middle East*, vol. 20, no. 20, pp. 399 - 404, 2005.
[17] Umm M., *A Brief Introduction to Tajweed*, Abul Qasim Publishing House, Saudi Arabia, 2000.
[18] Unicode Consortium., available at: http://www.unicode.org, last visited 2011.
[19] Zaidi R., Abdullah N., Othman Z., and Mohd Y., "Jawi Character Speech-to-Text Engine Using Linear Predictive and Neural Network for Effective Reading," *in Proceedings of the 3rd Asia International Conference on Modelling & Simulation*, Bali, Indonesia, pp. 348 - 352, 2009.

Zaini Othman obtained her Bachelor of Computer Science (Hons) degree, majoring in Software Engineering, and Master of Computer Science degree, specializing in the field of Signal Processing, both from the University Malaya, Malaysia in year 2008 and 2011 respectively. Her research interests revolve around speech and signal processing, biosignal and e-learning technology. She has published several research papers local and international, and also, won awards in Information Technology exhibitions.

Nor Abdullah is a senior lecturer at the Faculty of Computer Science and Information Technology in University of Malaya, Malaysia. She received her Bachelor of Computer Science (Hons) degree from the University Malaya, Malaysia in 1997. In the following year, she obtained her Master of Science degree in the field of interactive multimedia from the University of Westminster, London. In 2006 she received her Ph.D. degree in Computer Science from the Southampton University, United Kingdom, specializing in adaptive hypermedia research. Her research interests are e-learning technologies and tools, and context aware multimedia content retrieval. She has published several numbers of research papers such as in Australasian Journal of Educational Technology, Educational Technology Research and Development, Journal of Multimedia Tools and Applications, and Malaysian Journal of Computer Science.

Zaidi Razak obtained his Bachelor's degree in Computer Science and Master in Chip Design from University of Malaya in 2000. Currently, he is a senior lecturer at the Faculty of Computer Science and Information Technology, University of Malaya. His research area includes image processing, Jawi character recognition and System on Chip (SoC) design. He has published a number of papers related to these areas.

Mohd-Yakub Mohd-Yusoff obtained his Bachelor's degree, Master's degree and PhD from University of Malaya, University of Jordan and University of Wales, respectively. Currently, he is a professor and acting director at the Academy of Islamic Studies, University of Malaya, Kuala Lumpur, Malaysia. He also, actively involved as the Head/Director of Centre of Quranic Research (CQR) which has been a platform to venture research in the fields of Quranic and Islamic research field or multidisciplinary research field (scientific research/education/heritage etc.). His research areas include Quranic Exegesis, Quranic Studies and Methodology of Quranic Exegesis. A recognized leader in Quranic research field both in Malaysia and abroad, he has provided consultation and has been a research fellow, as well as visiting professor domestically and globally for more than 10 years.