Differential Qiraat Processing Applications using Spectrogram Voice Analysis

Muhammad Saiful Ridhwan, Akram M. Zeki and Akeem Olowolayemo
International Islamic University Malaysia, Malaysia
akramzeki@iium.edu.my

ABSTRACT

This study provides a review of efforts made so far to develop applications that assist in proper Quranic recitation based on the spectrogram voice analysis. The voice recorded from users' recitation is used to compare with the sample recitation features collected from expert reciters that have been stored in the database. These comparison is utilized to evaluate the performance result of the various algorithms that have been applied in Quranic recitation processing. Before reviewing how to develop the actual applications, a study was made to determine the effectiveness of spectrogram voice analysis. Samples are often collected to measure the articulation of Arabic letters. The collected samples are analyzed using Fourier analysis technique. The sound features samples' waveforms are transformed into spectrums that are basically frequency representations of the signals. The spectrogram is used to determine the formant frequency and are observed for each subject to determine its mean formant frequency. Subsequently after features extraction, a classification algorithm is employed to compare the features extracted in real-time with that available in the knowledge base. Based on the results obtained, the processes are often developed into mobile applications for portability.

KEYWORDS

Fast Fourier Transform, formant frequency, Signal analysis, spectrogram, voice recognition.

1 INTRODUCTION

Quran is written in Arabic language and for the general observance of religious duties, it is read in Arabic by majority of the Muslims as well as non-Muslims to learn and understand the content. People who are intent to learn the Quran must possess basic knowledge of Arabic in order to read, memorize and understand the Quran. In recent times, the use of technology has been studied by researchers to assist the process of learning and memorization of the Quran. One of the methods is the adoption of speech recognition techniques that have been attracting attention since the 1990s [1]. Speech is considered as a medium for communication between human being. A look at speech recognition technology nowadays, would reveal that many algorithms and methods have been developed and tested to detect the content of user's speech and have been found to be very useful in numerous voice recognition applications. Previous authors such as [2] has outlined two categories of speech sound - vowels and consonants. It is understood that consonants are produced by compressing the air inside the lungs whereas the vowel sound is produced according to the tongue and lips' movements.

As mention earlier, the research effort in testing and developing speech applications has been growing considerably, however, often the real concern on those application is not the implementation or use of modern technology but the processing capabilities of the algorithms for generating and interpreting speech, that cannot be said to be perfect or completely effective at the moment.

In recent development, researchers have found the use of spectrogram to analyze sound. The spectrogram analysis uses Fast Fourier Transform (FFT) to decompose the speech for spectrogram to process the features of the speech [3] [4]. This technique in voice processing is to analyze the power spectrum of a speech sample. If the power spectrum is

expressed in terms of intensity, one is able to produce an image called spectrogram, showing how the power distribution changes with time. The spectrogram reflects the location of the formants [4]. Formant is "the spectral peaks of the sound spectrum of the voice". It is often measured as an amplitude or peak in the frequency spectrum of the sound, when using a spectrogram or a spectrum analyzer. For different speakers, due to the frequency differences in their speech, even when uttering the same word, their major resonance frequencies will most likely be different when processed with spectrum analyzer [3] [4]. The purpose of this write up is to review previous breakthrough that have been made in applying spectrogram voice analysis to Quranic recitation to aid learners in recitation where there may not be an expert teacher to guide such learners.

The remaining parts of this paper are organized thus; the succeeding section reviews the existing literatures related to the topic of the discuss. It basically outlines a review of voice recognition related researches, algorithms and extraction techniques as well as details such as algorithm information, application programming interface, data collection and features extraction procedures as well as data analysis techniques employed. The final section concludes the treatise by identifying other areas for improvement.

2 PREVIOUS WORK

Quran is written in Arabic language and as a result, over the years, modern Arabic speakers have used Arabic letters in different ways in writing and pronunciations of conversational speech. In order to properly recite the Quran, people have to practice until precision is achieved. This proper pronunciation will take into consideration the rules of *Tajweed*. *Tajweed*, by definition is enhancing, improving, and becoming excellent in Quran recitation or Qiraat. Functionally, it means articulating every letter in the Quranic verses in its correct timing and from its proper point of articulation [5]. For people who just started to learn the Quran, they have to understand letters used in the Quran, the ways to pronounce those letters and practice often before they can reach fluentness in recitation. When people start to learn the Quran they will normally consult a *hafiz* or *hafizah* (who can be understood as teachers). A *hafiz* (male) or *hafizah* (female) is a person that memorizes the whole Quran and have vast knowledge in *Tajweed* (rules and regulations when reciting the Quran). A *hafiz* or *hafizah* has to practice every day to memorize and to improve themselves such that they can recite the whole sacred book without seeing it. Memorizing verses in the Quran is encouraged and people who are trying to memorize verses continues to practice recitation under the guidance and presence of an expert. Even though there may be challenges encountered in learning the Quran, people are still encouraged to learn as well as to memorize as much as they can, even a few verses from the Quran. This shows that an automated aid may be handy to assist the disparate learners. Most of the time, people would try to practice independently by listening to recitation by famous reciters. This independent learning is to proactively try to learn with or without the presence of people [6]. Self-determining learning helps learners to make informal choices and to take responsibility for deciding what learners need to do in order to learn [6]. One of the strategies used as mentioned earlier is by using technology such as media content. Using this strategy can attract people or ease learning to recite the Quran. Nowadays, the advancement of mobile technology has opened up a new door for researchers and developers to create new type of contents in mobile platform. This encourages researchers and developers to develop and test mobile applications such as Quran speech recognition, Quran reader, and Quran book in digital format. Those applications linked with recitation of samples of the Quran that can be easily retrieved from the Internet. Looking at the speech recognition

technology, many algorithms and methods have been developed and tested for differential Qiraat processing. However, there is lack of adequate efforts that have been done related to mobile applications in the context of Qiraat voice analysis. Efforts have largely been concentrated on non-automated applications that are capable to evaluate the recitation accuracy of Quranic verses in comparison to recorded user recitation.

The literature provides information on the area that research should be focused and conducted and other necessary issues to achieve an acceptable level of accuracy. The usual research objectives of most of the previous work centers around selection of appropriate features extraction techniques to build efficient knowledge base, adoption of feature classification algorithms to compare features extracted in real time with features previously stored in the knowledge base, as well as making efforts to implement the feature extraction and classification processes on a mobile application that will use the technology of spectrogram for voice analysis suitable for learners to correct their mistakes during recitation or practice.

The final part is often to study the performance of every stage in the process of voice analysis as highlighted. This survey is significant in the sense that it can be viewed through its contribution to two major areas, namely, theory and practice. Theoretically, this study contributes in evaluating the progress that has been achieved so far in developing and testing phases of Qiraat voice recognition. Apart from that, the challenges and progress made so far in presenting the process on mobile application using the tools and techniques available to provide benefits and ease the practice of all the people who are relentlessly trying to read correctly or memorize the Quran.

The previous section outlined the background of important matters related to conducting research in video recognition for Qiraat differentiation. In this section, a study based on previous research work has been done related to this research topic. This review is a compilation of report in written forms (journal) on the research topic which have been written by researchers who are authorities in this field. Before a detail review on voice recognition technology is presented, first, discussion on the concept of the Quran and *Tajweed* were presented. Then, experimental studies from the literature are presented, which shows the gap and differences between normal voice analysis and voice analysis in Quran. As a final point, a summary is presented in the form of table for normal voice analysis and voice analysis in Quran context. Three tables are presented in the next frames to provide a background of letters that depict different sounds in Quranic recitation as well as relevant voice recognition techniques that have been applied previously. Table 1 presents basic summary of the letters in Quran. In Table 2 likewise, the different symbols and their respective meanings in Quran recitation are presented. Table 3 on the other hand provides a summary list of speech recognition techniques that have been famously used by researchers. In this table, points that were outlined are related to numerous types of voice analysis extractions, classifications and matching techniques.

2.1 Reciting and Memorizing the Quran.

Quran is continuously been learnt by many people around the world which include Muslim and non-Muslim alike. In trying to learn the Quran, they have to understand the fundamentals of Quranic alphabets. They have to pronounce correctly so that they guide against mistakes that are critical to Quran recitations. In the process of learning Quran, people normally consult a *hafiz* or *hafizah* to learn the proper *Tajweed* of reciting the Quran. Table 1 below contains the list of letters used in the Quran. People who beginners in the study of the Quran have to understand these letters, the correct ways to pronounce those letters and practice often before they can perfect the act of recitation of the Quran.

2.2 Voice recognition techniques

Voice recognition technologies such as automatic speech recognition and text-to-speech have been under development since the early days of computing technology. Voice recognition development has been carried out by many researchers since early 1990s [7]. Today, users have access to very powerful, large-vocabulary speech recognition techniques for the creation of new analysis methods and experiment.

Table 1. Letters in Quran [8].

Letter	Name	Consonant (sokoon)	Short vowel (fatha)	Short vowel (damma)	Short vowel (kasra)	Long vowel (ا)	Long vowel (و)	Long vowel (ي)
ب	Baa	بْ	بَ	بُ	بِ	با	بُو	بِي
ت	Taa	تْ	تَ	تُ	تِ	تا	تُو	تِي
ث	Thaa	ثْ	ثَ	ثُ	ثِ	ثا	ثُو	ثِي
ج	Gym	جْ	جَ	جُ	جِ	جا	جُو	جِي
ح	Haa	حْ	حَ	حُ	حِ	حا	حُو	حِي
خ	Khaa	خْ	خَ	خُ	خِ	خا	خُو	خِي
د	Daal	دْ	دَ	دُ	دِ	دَا	دُو	دِي
ذ	Thal	ذْ	ذَ	ذُ	ذِ	ذا	ذُو	ذِي
ز	Zayn	زْ	زَ	زُ	زِ	زا	زُو	زِي
ر	Raa	رْ	رَ	رُ	رِ	را	رُو	رِي
س	Seen	سْ	سَ	سُ	سِ	سا	سُو	سِي
ش	Sheen	شْ	شَ	شُ	شِ	شا	شُو	شِي
ص	Saad	صْ	صَ	صُ	صِ	صا	صُو	صِي
ض	Daad	ضْ	ضَ	ضُ	ضِ	ضا	ضُو	ضِي
ط	Taa	طْ	طَ	طُ	طِ	طا	طُو	طِي
ظ	Zaa	ظْ	ظَ	ظُ	ظِ	ظا	ظُو	ظِي
ع	Ayn	عْ	عَ	عُ	عِ	عا	عُو	عِي
غ	Ghayn	غْ	غَ	غُ	غِ	غا	غُو	غِي
ك	Kaaf	كْ	كَ	كُ	كِ	كا	كُو	كِي
ق	Qaaf	قْ	قَ	قُ	قِ	قا	قُو	قِي
ف	Faa	فْ	فَ	فُ	فِ	فا	فُو	فِي
ل	Laam	لْ	لَ	لُ	لِ	لا	لُو	لِي
ن	Noon	نْ	نَ	نُ	نِ	نا	نُو	نِي
م	Meem	مْ	مَ	مُ	مِ	ما	مُو	مِي
ه	Haa	هْ	هَ	هُ	هِ	ها	هُو	هِي
و	Waw	وْ	وَ	وُ	وِ	وا	وُ	وِي
ي	Yaa	يْ	يَ	يُ	يِ	يا	يُو	يِ
ء	Hamza	ئْ	أ	أ	إ	آ	أُو	إِي
ا	Alif	n/a	n/a	n/a	n/a	n/a	n/a	n/a

Table 2. Symbols and Meanings in Quran Recitation [8].

Symbol	Name	Meaning	Example
ْ	Sokoon	Consonant letter	حبْس
◌َ	Fatha	Short vowel	كَذَبَ
◌ُ	Damma	Short vowel	كُل
◌ِ	Kasra	Short vowel	عِند
◌ّ	Shadda	Letter doubling	شدَّة
◌ً	Tanween el-fatha	Adds [an] to the letter	ايضاً
~	Madda	Turns the hamza into a long vowel	آدم

The aim of designing and developing Qiraat voice recognition applications is mainly to guide and assist the users who wish to practice reading or memorization of the Quran where they may not be a *hafiz* or as independent effort in addition to what is learnt from a *hafiz*. With the use of voice comparison analysis, mobile applications can be developed to provide assistance where there is no human experts or in addition to the efforts of human experts in reading the Quran. Generally, the application will have a collection of expert Quran reciter's voice features stored in the database which are subsequently used to compare with real-time recitations by learners, to provide responses based on whether they recited correctly or incorrectly. Then the comparison will acknowledge the users' level of accuracy. To achieve the research objectives, researchers often adopt different stages as described below:

(i) Samples of Quran recitation is collected from various famous reciters.
(ii) Those collected samples will go through spectrogram analysis to build a database of different verses of the Quran.
(iii) The collected spectrogram analysis will be tested with user recorded recitation in order to find accuracy of recognition.
(iv) Mobile application will be created to connect the process together in one single or integrated system after reasonable level of performance.
(v) The mobile application is evaluated with different reciters from the database to measure the performance and effectiveness.

Table 3. Summary of Speech Recognition Techniques.

Method	Description	Settings
Principal Component analysis (PCA)	Principal component analysis is a useful statistical technique that has applications in fields such as face recognition and image compression, and is a common technique for finding patterns in data of high dimension [9].	Non-linear feature extraction method, linear map, fast, eigenvector-based [10].
Independent Component Analysis (ICA)	Independent Component Analysis (ICA) is a statistical technique, perhaps the most widely used, for solving the blind source separation problem [11].	Extract voice feature vectors based on these higher order statistics from natural scenes and music sound. These features are localized in both time (space) and frequency. Nonlinear feature extraction method, linear map, iterative non- Gaussian [12].
Linear Predictive coding	Linear predictive coding (LPC) is defined as a digital method for encoding an analog signal in which a particular value is predicted by a linear function of the past values of the signal [13].	Uses mathematical approximation of the vocal tract represented by this tube of a varying diameter. At a particular time, t, the speech sample s(t) is represented as a linear sum of the p previous samples. Static feature extraction method, 10 to 16 lower order coefficient [13].
Cepstral Analysis	Cepstral analysis is used to separate the speech into its source and system components without any a prior knowledge about source and / or system [14].	Cepstral analysis is used for transforming the multiplied source and system components in the frequency domain to linear combination of the two components in the cepstral domain. Static feature extraction method and Power spectrum [10].

Mel-frequency cepstrum (MFFCs)	MFCC have the ability to transform the frequency from a linear scale to a nonlinear. MFCC is based on human hearing perceptions which cannot perceive frequencies over 1Khz. MFCC is based on known variation of the human ear's critical bandwidth with frequency [15].	Have two types of filter which are spaced linearly at low frequency below 1000 Hz and logarithmic spacing above 1000Hz. A subjective pitch is present on Mel Frequency Scale to capture important characteristic of phonetic in speech. Power spectrum is computed by performing Fourier Analysis [16].
Hidden Markov Model	Hidden Markov Models (HMMs) provide a simple and effective framework for modelling time-varying spectral vector sequences. Each word is trained independently to get the best likelihood parameters [15].	Based on mathematical framework and implementation structure. In terms of the mathematical framework, method's consistent statistical methodology and the way it provides straightforward solutions to related problems. Implementation structure: the inherent flexibility the method provides in dealing with various sophisticated speech-recognition tasks and the ease of implementation [16].
Artificial Neural Network	Artificial Neural Network (ANN), also known as Neural Network (NN). ANN is also mainly used as feature matching or recognition for speech processing.	It is normally used to classify a set of features, which represent the spectral-domain content of the speech (regions of strong energy at particular frequencies). The features then will be converted into phonetic-based categories at each frame. Search is used to match the neural-network output scores to the target words (the words that are assumed to be in the input speech), in order to determine the word that was most likely uttered
Vector Quantization	VQ codebook consists of a small number of representative feature vectors, which used as an efficient means of characterizing speaker-specific features [17].	A speaker-specific codebook is generated by clustering the training feature vectors of each speaker. In the recognition stage, an input utterance is vector-quantized using the codebook of each reference speaker and the VQ distortion accumulated over the entire input utterance is used to make the recognition decision [17] [18].
Spectrogram	The interpretation of spectrum to spectrogram is to show the speech signal varying with time [3]. Most importantly, spectrogram is used widely to calculate and analyze speech signal. The horizontal axis represents time while vertical axis represents frequency in a spectrogram [4].	Spectrogram usually encodes amplitude of the signal using gray color instead of colors [4]. The peaks of the spectrum are black and the valley is white in gray color. Fast Fourier Transform (FFT) coding is utilized to extract noise of each data. The formant frequencies are the average value of the appropriate phoneme [3].

2.3 Review on speech recognition method with Quran

In previous sections, all different methods or approaches have been discussed. Those methods and approaches are used in general voice recognition or voice analysis. In this part, a look into the various studies and researches conducted related to Quran voice recognition is discussed. Most of the algorithms are developed using Matlab and Java development platform which is tested on a personal computer. In the last part of this review, a summary that gives an overview of previous conducted studies in this area is presented.

Studies in [19] and [20] developed a system using spectrographic extraction method. This spectrographic analysis is conducted on different frequency band intensity. The performance rate of accuracy is 93.33%.

Likewise, study in [21] developed a system called Arabic Text-to-Speech (TTS). The system was developed in Human Language

Technologies laboratory of IBM, Egypt. It uses IBM trainable concatenate speech synthesizer. The study result is presented in a mean opinion score for the synthesized speech.

In [22] voice recognition technology for recognizing Arabic speech using the neural network was developed. This technique is able to classify nonlinear problems based on research that was performed using neural networks. They proposed a fully-connected hidden layer, between the input and state nodes and the output. Furthermore, they investigated and showed that this hidden layer makes the learning of complex classification tasks more efficient. The difference between LPCC (linear predictive cepstrum coefficients) and MFCC (Mel-frequency cepstral coefficients) in the feature extraction process is also investigated. To test the effectiveness of the system they use six speakers (a mixture of male and female) in a quiet environment.

Another work, [23], introduced a new method where the verses is extracted from audio files using speech recognition technique. The technique implemented utilized the Sphinx IV framework to develop the system. The developed system uses MFCC features to extract those verses from audio files. Hidden Markov Model is used as recognizer to translate into Arabic word using hash map and breadth first search, combined with beam search for the dictionary database.

HMM Toolkit used for Arabic automatic speech recognition engine was developed by [24], which can recognize both continuous speech and isolated words. One of the main functions included was an Arabic dictionary composing words for its phones. The speech feature vectors are extracted using Mel Frequency Cepstral Coefficients (MFCC). Subsequently, the training of the engine based on tri-phones is employed to estimate the parameters for a HMM. The system was tested with thirteen Arabian native speakers.

The research work in [10] developed a system to conduct an automated *Tajweed* checking rules' engine for Quranic verse recitation. The feature extraction technique used is the Mel frequency Cepstral Coefficients (MFCC) which was utilized to extract the characteristics from Quranic verses' recitation. Similarly, Hidden Markov Model (HMM) was employed for classification and recognition purposes. The researcher noted that the developed system was able to achieve recognition rate that exceeded 91.95% and 86.41%.

Spectrogram voice analysis uses Fourier analysis technique. The sound waveform is transformed into spectrum which is the frequency representation of the signal. Then spectrogram is used to determine the formant frequency. The formant frequency of each collected signal is determined through observation on the spectrogram.

Table 2.4: Summary of Speech Recognition Technique used in Quran experiment.

Extraction Method	Recognition Technique	Performance	References
Spectrographic Analysis	Spectrographic Analysis based on different frequency band of intensity	93.33%	[19][20]
MFCC	Hidden Markov Model (HMM)	90.2%	[21]
MFCC **LPCC**	Recurrent Neural Network (RNN)	MFCC 95.9%– 8.6% LPCC 94.5%- 99.3%	[22]
MFCC	Hidden Markov Model (HMM)	85% - 92%	[23]
MFCC	Hidden Markov Model (HMM)	90.62%, 98.01 % and 97.99% for sentence correction, word correction and word accuracy respectively.	[24]
MFCC	Hidden Markov Model (HMM)	86.41% to 91.95%	[10]
FFT	Spectrogram	Over 90%	[4]
MFCC	Not Available	90%-92%	[25]
MFCC	VQ	82.1-95%	[18]

Another work by [4], conducted a research on spectrogram voice analysis where it was shown that second and third formant frequency (F2 and F3) increased as the articulation is made through the mouth. Again, [25] developed and tested a novel system called E-hafiz. The

system was built based on Mel-Frequency Cepstral Coefficient (MFCC) technique to extract voice features from Quranic verses recitation. Those extracted voice features are mapped with the data collected during the training phase, thereby able to identify the mismatch or mistakes using comparison analysis.

3 CONCLUSION & FUTURE WORK

Based on previous studies, it is understood that every letter in Arabic have different point formant. This finding could be used to study the implementation flow that could be crucial for the mobile application. Considerable samples should be examined to study the different formant point. These studies should not be limited to a type of participant (e.g. male/female) rather from different sources and genders. The focus of this treatise is to describe the background of the approaches to differential Qiraat processing based on voice recognition techniques. It provides the basic approaches and stages involved as well as issues and challenges related to voice recognition based Qiraat processing. Highlights of achievements so far were provided as well as significance and contribution of the study, specifically with respect to performance issues and implementation on portable mobile devices which are often utilized by Quranic readers. This paper reviews the various researches of spectrogram in mobile platform. Spectrogram uses FFT to extract data from audio files. FFT is believed to have a faster transform and requires less computational power to process a given amount of information. This therefore has appealing advantages of energy consumption in mobile multimedia devices such as smart phones [26]. Implementation on mobile platforms are seldom reported in this area which calls for a review and further research on implementation, system processes and interface design. These should be the future direction in this area.

REFERENCES

[1] Waibel, A., Hanazawa, T., Hinton, G., Shikano, K., and Lang, K. J., 1989, 'Phoneme recognition using time-delay neural networks,' IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 37, pp. 328-339.

[2] Hassan, A., 1984, Linguistik Am Untuk Guru Bahasa Malaysia, 5thed. Selangor, Malaysia: Fajar Bakti, 1984.

[3] Dzati, A. R., Salina, A. S., and Aini, H., 2007, 'Preprocessing Techniques for Voice-Print Analysis for Speaker Recognition', The 5th Student Conference on Research and Development – SCOReD 2007 11-12 December 2007, Malaysia.

[4] Abdul-Kadir, N.A., Sudirman, R. and Safri, N. M., 2010, 'Modelling of the Arabic Plosive Consonants Characteristics based on Spectrogram' 2010 Fourth Asia International Conference on Mathematical/Analytical Modelling and Computer Simulation. ISBN: 978-0-7695-4062-7/10, 2010 IEEE.

[5] Harrag, A. and Mohamadi, T., 2010, 'QSDAS: New Quranic Speech Database for Arabic Speaker Recognition', The Arabian Journal for Science and Engineering, Volume 35, Number 2C.

[6] Shafiqah, S., 2011, 'Independent Learning of Quran (ILoQ)-Alphabet', Degree Thesis, Faculty of Computer Systems & Software Engineering, University Malaysia Pahang, Pahang, Malaysia.

[7] Oberteuffer, J.A., 1995, 'Commercial applications of speech interface technology: An industry at the threshold. In Proceedings of the National Academy of Science'. Vol. 92, pp. 10007-10010

[8] Ramzi A.H., Omar E.A., 2007. 'CASRA+: A Colloquial Arabic Speech Recognition Application". American Journal of Applied Sciences 4(1):23-32, 2007 Science Publication.

[9] Nidaa, A. A., 2009, 'Speech Scrambling Based on Principal Component Analysis', MASAUM Journal of Computing, October 2009, Vol.1(3).

[10] Jamaliah, N., 2010, 'Automated Tajweed Checking Rules Engine for Quran verse Recitation', Master Thesis, Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia.

[11] Ganesh, R. N. and Dinesh, K K., 2011, 'An Overview of Independent Component Analysis and Its Applications', School of Electrical and Computer Engineering, RMIT University, Australia

[12] J.-H. Lee, H.-J. Jung, T.-W. Lee, and S.-Y. Lee, 1998, 'Speech feature extraction using independent component analysis', Proc. ICASSP, vol. 2, pp.1249 -1252 1998.

[13] Bradbury, J., 2000, 'Linear Predictive Coding', [Online] Available at: http://my.fit.edu/~vKepuska/ece5525/lpc_paper.pdf retrieved on 25 December 2013.

[14] Sakshat Virtual Labs, n.d., 'Speech signal processing laboratory: Cepstral analysis of speech', [Online] Available at: http://iitg.vlab.co.in/?sub=59&brch=164&sim=615&cnt=6 retrieved on 25 December 2013.

[15] Gales, M. and Young, S., 2007, 'The Application of Hidden Markov Models in Speech Recognition', Cambridge University Engineering Department, Trumpington Street, Cambridge, Foundations and

Trends in Signal Processing, Vol. 1, No. 3 (2007) 195–304.

[16] B. H. Juang and L. R. Rabiner, 1991, 'Hidden Markov Models for Speech Recognition', Speech Research Department AT&T Bell Laboratories. American Statstical Association and the American Society for Quality Control, 1991, VOL. 33, NO. 3.

[17] Kanade, T., Jain, A., Ratha, N. K, 2005, 'Audio and Video Based Biometric Person Authentication', 5th International Conference, AVBPA 2005, Hilton Rye Town, NY, USA, July 2005 Proceedings

[18] Ismail, A., Idris, M. Y. I., Noor, M. N., Razak, Z., Yusoff, Z. M., 2014 'MFCC-VQ Approach for Qalqalah Tajweed Rule Checking', American Journal of Applied Sciences 27(4):275-293, 2014.

[19] Bashir, M.S., Rasheed, S.F., Awais, M.M., Masud, S. and Shamail, S., 2003. Simulation of Arabic phoneme identification through spectrographic analysis. Department of Computer Science LUMS, Lahore Pakistan.

[20] Anwar, M.J., Awais, M.M., Masud, S. & Shamail, S.," Automatic Arabic Speech Segmentation System." Department of Computer Science, Lahore University of Management Sciences, Lahore, Pakistan.

[21] Youssef, A. & Emam, O., 2004, 'An Arabic TTS based on the IBM Trainable Speech Sythesizer', Department of Electronics & Communication Engineering, Cairo University, Giza, Egypt.

[22] Ahmad, A.M., Ismail, S., Samaon, D.F., 2004, 'Recurrent Neural Network with Backpropagation through Time for Speech Recognition,' IEEE International Symposium on Communications & Information Technology, 2004. ISCIT '04.Volume 1, pp. 98 – 102.

[23] Tabbal, H., El-Falou, W. and Monla, B., 2006, 'Analysis and Implementation of a "Quranic" verses delimitation system in audio files using speech recognition techniques',In: Proceeding of the IEEE Conference of 2nd Information and Communication Technologies, 2006. ICTTA '06.Volume 2, pp. 2979 – 2984.

[24] Bassam, A. Q. Al-Qatab, Ainon, R. N., 2010, 'Arabic Speech Recognition Using Hidden Markov Model Toolkit (HTK)', Software Engineering Department, Faculty of Computer Science and Information Technology, University Of Malaya, Malaysia.

[25] Waqar, M.M., Rizwan, M., Aslam, M., Martinez, E. A. M.,2010, 'Voice Content Matching System for Quran Readers', Ninth Mexican International Conference on Artificial Intelligence, 2010 IEEE Hasan, M.R., Jamil, M., Rabbani, M.G. & Rahman, M.S., 2004, 'Speaker Identification Using Mel Frequency Cepstral Coefficients', 3rd International Conference on Electrical & Computer Engineering ICECE 2004, 28-30 December 2004, Dhaka, Bangladesh ISBN 984-32-1804-4 565.

[26] Andersson, M., 2012, 'A Faster Fourier Transform: A mathematical upgrade promises a speedier digital world', [Online] Available at: http://www2.technologyreview.com/ article/427676/a-faster-fourier-transform/ retrieved on 25 December 2013.