Automated Tajweed Checking Rules Engine for Quranic Learning

Noor Jamaliah Ibrahim
Centre of Quranic Research (CQR), University of Malaya, Kuala Lumpur, Malaysia

Mohd Yamani Idna Idris and Zaidi Razak
Department of Computer System & Technology, Faculty of Computer Science & Information Technology, University of Malaya, Kuala Lumpur, Malaysia

Noor Naemah Abdul Rahman
Academy of Islamic Studies, University of Malaya, Kuala Lumpur, Malaysia

Abstract

Purpose - The purpose of this paper is to provide a structural overview of speech recognition system for developing Quranic verse recitation recognition with Tajweed checking rules function. This function has been introduced, due to support the existing and manual method of *talaqqi & musyafahah* method in Quranic learning process, which described as face to face learning process between students and teachers. Here, the process of listening, correction and repetition of the correct Al-Quran recitation took place in real time condition. However, this method is believed become less effective and unattractive to be implemented, especially towards the young Muslim generation who are more attracted to the latest technology.

Design/methodology/approach – This paper focus on development of software prototype, mainly for developing an automated Tajweed checking rules engine, purposely for Quranic learning. It has been implemented and tested towards the j-QAF students at primary school in Malaysia.

Findings - The paper provides empirical insight about the viability and implementation of Mel-Frequency Cepstral Coefficients (MFCC) algorithm of feature extraction technique and Hidden Markov Model (HMM) classification for recognition part, with the results of recognition rate reached to 91.95% (ayates) and 86.41% (phonemes), after been tested on *Sourate Al-Fatihah*.

Originality/value - Based on the result, proved that the engine has a potential to be used as an educational tool, which helps the students read Al-Quran better, even without the presence of teachers or parents to monitor them. Automated system with tajweed checking rules capability functions could be another alternative due to support the existing method of manual skills of Quranic learning process, without denying the main role of teachers in teaching Al-Quran.

Keywords - Quranic learning, Mel-Frequency Cepstral Coefficient (MFCC), Hidden Markov Model (HMM), Tajweed rules, Talaqqi & Musyafahah, Log-Likelihood

Article Classification - Research Paper

1. Introduction

In this technological era, speech recognition is highly demanded and has many useful applications. This research stimulates speech recognition technology, which incorporates with the various components in artificial intelligent, natural sciences, speech processing technology and human computer interaction (Hewet *et al.,* 1996). The research presented in this paper is concerned on speech recognition engine, which has been implemented with Quranic verse recitation, related to "Tajweed Rules" based on recitation recognition process. This Tajweed checking rules engine is capable to educate the students and adults by using the interactive learning system with Tajweed checking rules (Al-Quran reading rules) correcting capability. This interesting tool was able to attract the students' interest and increased their motivation to learn Al-Quran with a better understanding of Tajweed rules, thus improved the educational system in Al-Quran learning process (Zaidi *et al.,* 2008). Moreover, the existing product/technology available, were only capable to show Al-Quran texts and/or play stored Al-Quran recitation.

Our system offers students to recite Al-Quran, due to be processed through the system, where the recitation will be revised and corrected in real time. This function is very beneficial to students and adults in learning Al-Quran, because this engine capable to help the students and assist them, even without the presence of teachers (Mudarris) to monitor them. Meaning that, this engine can be used as an educational tool or teaching aid for students as well as teachers (Mudarris), in order to support the existing method through manual skills of *talaqqi & musyafahah* method in learning and teaching Al-Quran (The New Strait Times Press, 2007). This method is described as face to face learning process between students and teachers, where listening, correction of Al-Quran recitation and repetition of the correct Al-Quran recitation took place (Hewet *et al.,* 1996). This factor is important, so that students will know how the *hijaiyah* letters are pronouncing correctly. The process only can be done, if the teachers and students follow the art, rules and regulations while reading the Al-Quran, known as "Rules of Tajweed" (Tabbal *et al.,* 2006).

2. Acoustic Model of Holy Qur'an Recitation

It is known as fact that, each person's voice is totally different. Thus, the recitations of the Holy Quran tend to differ a lot from one person to another. Although the same verse of Quranic recitations were particularly recite by the same recitor, but the way of the sentence in Al-Quran been recited or delivered may be different, due to the flexibility of the laws of the tajweed (Tabbal *et al.,* 2006; Noor Jamaliah *et al.,* 2008). There are many difficulties arise when dealing with Arabic language, especially in Al-Quran, regarding to the differences between written and recite the Holy Quran (Tabbal *et al.,* 2006; Maamouri *et al.,* 2006; Kirchhoff *et al.,* 2004; Laura *et al.,* 2003). The Quranic Arabic alphabets consist of 28 letters, known as *hijaiyah* letters (from alif (ا) until ya (ى)) (Kirchhoff *et al.,* 2004; Kirchhoff *et al.,* 2003; Vergyri *et al.,* 2004; Ahmad *et al.,* 2004). Those letters includes 25 letters, which represent consonant and 3 letters for vowels (/i: /, /a: /, /u :/) (Laura *et al.,* 2003) and the corresponding semivowels (/y/ and /w/), if applicable. A letter can have two to four different shapes: isolated, beginning of a (sub) word, middle of a (sub) word, and end of a (sub) word (Kirchhoff *et al.,* 2003). Letters are mostly connected and there is no capitalization.

In Al-Quran, the phonemes of pronunciation are marked by diacritics, such as consonants doubling (phoneme in Arabic). It is indicated by the sign of *"shadda"* and *"tanween"* sign, word final adverbial markers, which add /n/ to the pronunciation (Maamouri *et al.,* 2006; Kirchhoff *et al.,* 2004; Ahmed, 1991) as shown below in table 1. These signs can reflect the differences of pronunciation and the diacritics sign are really important in setting the grammatical functions, which lead to the acceptable text understanding and correct reading or analysis (Maamouri *et al.,* 2006). The entire set of those diacritics is listed below in table 1 (Kirchhoff *et al.,* 2003; Vergyri and Kirchhoff, 2004).

Diacritics indicated the correct Arabic (Modern Standard Arabic (MSA), known as the vocalization or the voweling, while reciting Al-Quran (Laura *et al.,* 2003). Besides those diacritics as listed in table 1, there were also some additional character of Arabic letters listed, called as *Hamza*. Another non-basic character is *Taa-Marbuwta*, which is always at the end of the word (Vergyri and Kirchhoff, 2004). These characteristics are the most important consideration in our system design of Tajweed checking rules engine discussed later.

Table 1: Arabic diacritics (Vergyri and Kirchhoff, 2004)

Example	Symbol Name	Meaning
أَ	fatha	/a/
إِ	Kasra	/i/
أُ	Damma	/u/
رّ	Shadda	Consonant doubling
دَرْس	Sukuwn	Absence of vowel after consonant
اً	Tanwyn al-fatha	/an/
اٍ	Tanwyn al-kasra	/in/
اٌ	Tanwyn ad-Damm	/un/
ى	ʻalif maqsuwra	/a:/ sound historical
هٰذه	Dagger ʻalif	/a:/ sound historical
آ	Madda	Double alif
في ٱلبيت	Wasla	On ʻalif in *al*
لا	Laam ʻalif	Combination of laam and ʻalif
ة	Taa marbuwta	Morphophonemic marker

3. Methods and Experiments

The project of this research is mainly focus on the basic research of speech recognition technology using Al-Quran recitation. The differences of input that implemented in this engine, will affect the percentage of accuracy during recognition process. So, the reliability and effectiveness of the system is depended on languages and system architecture created for this engine, which will be discussed briefly in this section.

3.1 System architecture and Algorithms

According to Essa (1998) and Nelson & Kristina (1985), the Quranic Arabic recitation is best described as long, slow pace rhythmic, monotone utterance. The sound of Quranic recitation recognizably unique and reproducible according to a set of pronunciation rules of tajweed, which designed for clear and accurate presentation of the text. The input of the system is the speech signal and phonetic transcription of the speech utterance. Thus, this project need to have speaker (input speech sample), features extraction, features training and pattern classification/matching, which are components that important for Quranic verse recitation recognition formulation of architecture. Here, the main architecture involved is shown below in figure 1.

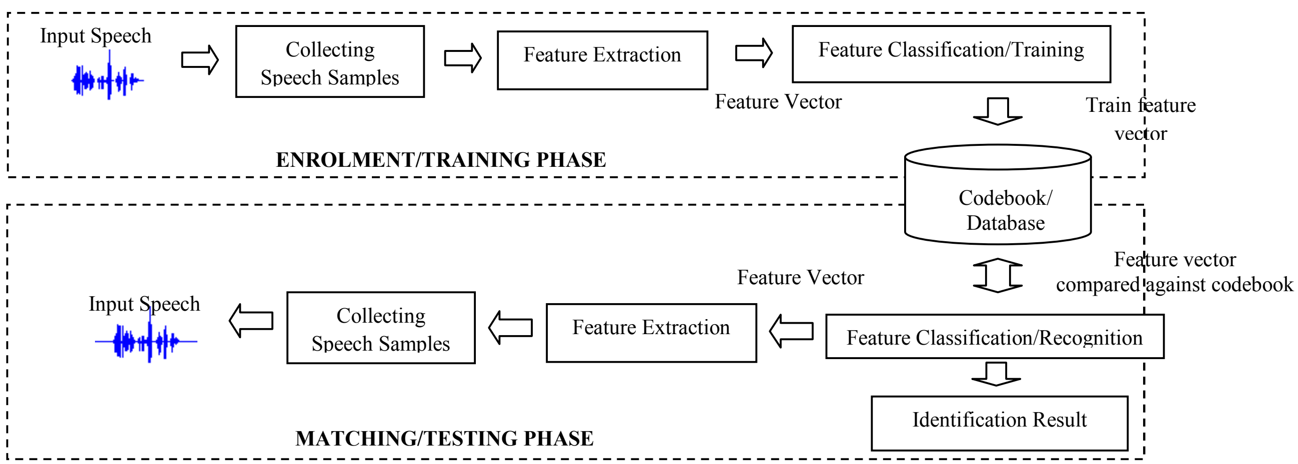

Figure 1: Tajweed checking rules engine architecture

Figure 1 depicts an automated Tajweed Checking rules engine architecture, which illustrated a speaker recognition system for Quranic verse recitation. In this architecture, 2 distinguished phases have been represented, which are *enrolment/training phase* and *matching/testing phase*. The execution process for both phases are really important, due to extract, store and analyze Al-Quran recitation parameters. The Mel-Frequency Cepstral Coefficient (MFCC) and Hidden Markov Model (HMM) based algorithm is selected for feature extraction and classification method. Speech recording process (Speech samples collection), feature extraction, feature training and pattern recognition were used to formulate the Quranic verse recitation recognition methodology. Both 2 phases of *enrolment/training phase* and *matching/testing phase* will be discussed in details in the next sections of 3.1.1 and 3.1.2.

3.1.1 Enrolment/ Training phase

In this phase, each recitor needs to provide the samples of Quranic recitation for the training process of the engine. Then, the features obtained from this process will be stored as a reference model into database using Hidden Markov Model (HMM), input of speech signal obtained is well characterized as a parametric random process and the parameters of stochastic process can be determined in precise and well-defined manner. The HMM model parameter need to be updated regularly, due to make the system able to fit a sequence for particular application. Thus, the training process of HMM model is so important, to represent the utterances of words. This model will be used later on in the testing and recognition part, by calculating the probability values (sequence of vectors) of the model. A complete specification of HMM is required for observation symbols of model parameters, **N** and **P**, as well as 3 sets of probabilities measure **A, B** and initial state distribution, **pi0**. The complete parameter set of an HMM model is denoted as below (Hemantha *et al.,* 2006):

$$\lambda = (A, B, pi0) \text{ .... (1)}$$

In equation (1), the parameter set need to be adjusted, due to estimate the parameters of the model λ, which maximizes P (O/λ). This model parameter λ will represent the likelihood of the training set observation vectors as shown in figure 2. The probability for observation sequence O = {O_1 O_2 O_3… O_T} of this model can be carried out, by finding which model that most likely has produced the observation sequence through the Forward and Backward Algorithm.

i) *Forward Algorithm*

Every possible sequence of states of length T can be evaluated, through the equation below:

$$P(O \mid \lambda) = \sum_{q_1, q_2, \ldots, q_T} \pi_{q_1} \prod_{t=2}^{T} a_{q_{t-1} q_t} b_{q_t}(o_t) \text{ .... (2)}$$

The Forward Algorithm is based on the forward variable $\alpha_t(i)$ defined by:

$$\alpha_t(i) = P(o_1 \ o_2 \ \ldots o_t \ , q_t = i \mid \lambda) \text{ .... (3)}$$

From the above equation (3), $\alpha_t(i)$ is the probability at time *t* and in state *i*, given by the model, which generated the partial observation sequence from the first observation until observation number *t*, $o_1 o_2 \ldots o_t$.

ii) *Backward Algorithm*

If the recursion process is described as to calculate the forward variable in reverse way, then the $\beta_t(i)$ will be the backward variable. This variable is described with the following equation:

$$\beta_t(i) = P(o_{t+1} \, o_{t+2} \ldots o_T \mid q_t = i, \lambda) \text{ .... (4)}$$

From the equation (4), $\beta_t(i)$ is the probability at time *t* and state *i*, given by the model in which having generated the partial observation sequence from *t + 1* observation until observation number *T*, $o_{t+1}o_{t+2}...o_T$.
Vocabulary of the words and phonemes to be recognized is modelled by a distinct HMM for developing pattern for the Tajweed Checking Rules database during the training phase. Each of word and phoneme in the vocabulary has a training set of k utterances by different speakers (Rabiner and Juang, 1993).

3.1.2 Matching/Testing Phase

In the recognition stage, both process for matching and testing will be executed. Here, features vector will be generated from the input speech sample, through the same extraction procedure in the training stage, discussed earlier. Those samples of input speech will be matched with stored reference model, and thus a decision/evaluation can be made (recognition). The output data obtained from Hidden Markov Model (HMM) need to be responded and compared, by referring to the database created during the training process. Meanwhile, at the same time the system itself will act upon to any feedback result, and give the answer, either the output data produced can match the stored data in database or not. If the output is slightly different from the stored data in database, the system will assume those output data (Al-Quran recitation) as false/wrong.

This invented engine of automated Tajweed checking rules will act upon any Quranic recitation, whenever it receives the input speech signal. Here, any input speech that passing through the system will give an output score and cause the engine to make decision/evaluation. Thus, the score value measuring the confidence of a recognized word needs to find out. The output score is obtained, after the word is selected using the Viterbi algorithm, which the model of likelihood is maximum as given in figure 2.

Under this part, the recognition task can be divided either identification or verification, which described as below:

a. Identification: Each unknown words to be recognized is measured from the observation sequence through the feature analysis of the speech, corresponding to the word. Here, the word is selected using Viterbi algorithm (Likelihood values is maximum). Symbol X is represented the identification result for the particular recitation of Al-Quran.

b. Verification: The input features are compared with the registered pattern that already stored in database. Any features that giving the highest score is identified as the selected/target speaker (recitor). Then, this input features are compared with the claimed speaker (recitor) and decision is made either to accept or reject the claimed. Symbol Y is represented the verification result, based on the threshold value set earlier.

As shown in figure 2 below, the parameter λ defines as probability measure for the observation sequence O, i.e., P (O/λ). This observation sequence O, need to be compared with a model λ = (*pi0, A, mu, sigma*), in order to find the optimal sequence of states q**,** to a given observation sequence and model. Due to maximize P (q|O,λ), the suitable algorithm to be used must be Viterbi algorithm (Rabiner, 1989). The Viterbi algorithm is used to find the best single state sequence for the given observation sequence (Rabiner and Juang, 1993).

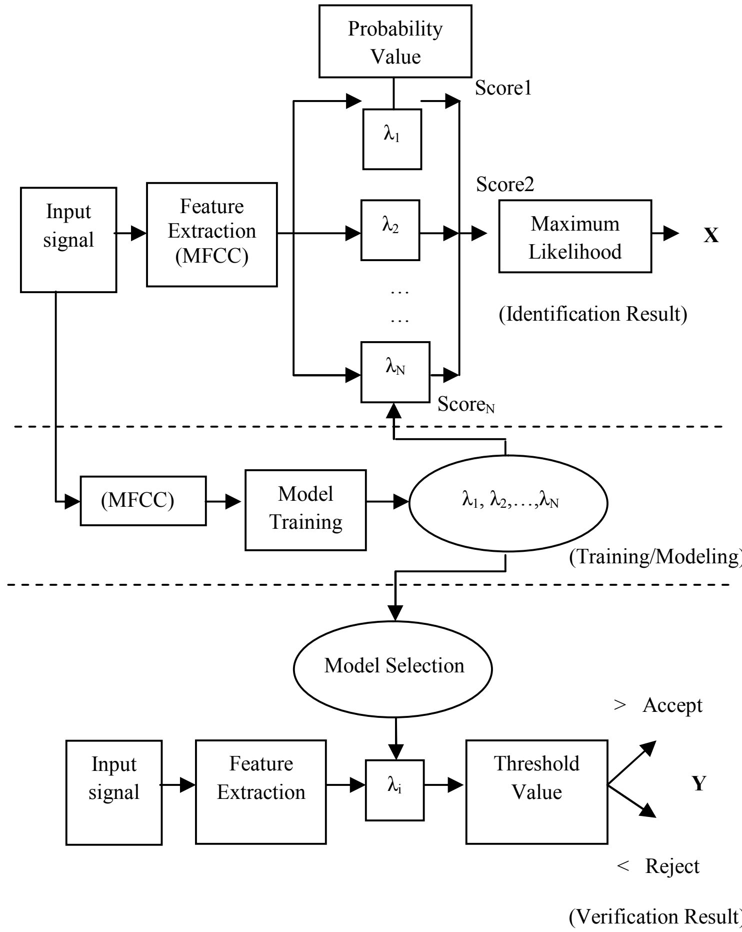

Figure 2: Automated Tajweed Checking Rules engine structure

In this research, the ayates and phonemes were classified under 2 different probabilities of HMMs (λ_1 and λ_2), either **'In Vocabulary (IV)'** or **'Out of Vocabulary (OOV)'**, due to ensure that the engine compatible in checking the tajweed rules. The basic idea for separating the IV and OOV phonemes/words are the likelihood difference between the best and 2nd best result of IV input are smaller than those of OOV input, because of unmatched model of OOV input. Here, a reasonable measure of the similarity using the logarithmic distance concept can be determined, by defining the distance measure of D (λ_1, λ_2) between two Markov models, denoted as below (Hemantha *et al.,* 2006):

$$D(\lambda_1, \lambda_2) = 1/N\left[\log_{10} P(O_2 / \lambda_1) - \log_{10} P(O_2 / \lambda_2)\right] \qquad (5)$$

The standard Log-Likelihood Ratio (LLR) and augment LLR are used, using the above equation (5). N, is the length of input utterance, $\log_{10} P(O_2 / \lambda_1)$ is the largest LLR and $\log_{10} P(O_2 / \lambda_2)$ is the second largest LLR.

The Log-Likelihood Ratio (LLR) calculation is performed in these 2 tests of experiments, in order to evaluate the system performance of this engine. Each of experiment executed, both the training and word templates uttered are from the same speaker, purposely to improve the accuracy and recognition rate. The results obtained from both experiments will be discussed in details in the next section.

4. Result and Discussion

4.1 Speech Samples Collection

The collecting of speech samples from Quranic recitation for 5 different speakers (recitors) were carried out, through recording process. Each of distinct word (ayates in *sourate Al-Fatihah*) will be recorded, and those speech samples were saved for further processing. The numbers of speech samples collected were 52 words (ayates) and 82 probable samples of phonemes for those ayates in different samples of Quranic recitation. Speech samples were recorded in a constraint environment, where 5 selected speakers (recitors) were choose and highly trained in Quranic recitation based on the 'Tajweed rules'. Here, the first chapter of Al-Quran (*Sourate Al-Fatihah*) were recited, with approximately recited in 4 seconds (time length) each, in '.wav' of file format. Table 2 shows the summary of the collected speech samples of *sourate Al-Fatihah*.

Table 2: Except from the dictionary of *Sourate Al-Fatihah*

The word in the dictionary (Wave file assigned)	**The utterances (Phonemes)**	**The ayates in Quran**
Bismillahirrah manirrahim *(Bismillah.wav)*	*Bismi Llahii Rraohimani Rraohiiim*	بِسْمِ ٱللَّهِ ٱلرَّحْمَٰنِ ٱلرَّحِيمِ
Alhamdu lillahi rabbi alAAalameen *(fatihah1.wav)*	*Allhamdu Lillahhirabbil A'alamiinna*	ٱلْحَمْدُ لِلَّهِ رَبِّ ٱلْعَٰلَمِينَ
Arrahmaanirrahiim *(fatihah2.wav)*	*Alrrahmani Alrraheemi*	ٱلرَّحْمَٰنِ ٱلرَّحِيمِ
Maalikiyawmiddiini *(fatihah3.wav)*	*Maaliki Yawmi Alddeeni*	مَٰلِكِ يَوْمِ ٱلدِّينِ
Iyyakana'Abudu waiyyaka nastaeen *(fatihah4.wav)*	*Iyyaka naA'Abudu waiyyaka nastaAAeen*	إِيَّاكَ نَعْبُدُ وَإِيَّاكَ نَسْتَعِينُ
Ihdinaassiratholmustakiim *(fatihah5.wav)*	*Ihdina Alssiratho Almustaqeema*	ٱهْدِنَا ٱلصِّرَٰطَ ٱلْمُسْتَقِيمَ

SiraathollazinaAn'amta'Alaihim ghayrillmaghdoobi'Alaihim waladdholeen *(fatihah6.wav & fatihah7.wav)*	*Siratho Allatheena An'Aamta 'AAalayhim Ghayri Almaghdoobi 'AAalayhim Wala Alddhalleen*	صِرَٰطَ ٱلَّذِينَ أَنۡعَمۡتَ عَلَيۡهِمۡ غَيۡرِ ٱلۡمَغۡضُوبِ عَلَيۡهِمۡ وَلَا ٱلضَّآلِّينَ

4.2 Result of Training and Enrolment

Two references model were developed during this training process, which are ***Word based Model*** and ***Phoneme based Model.*** Both models coming from 2 separate template of HMM models in training corpus, which are word (ayates) template and phoneme-like templates, respectively.

The training corpus used 2 tests to compose the samples of Quranic recitation. From the corpus, 82 samples of Quranic recitation phonemes like templates are produced and converted into phoneme strings using the Quranic pronunciation rules. Those templates were particularly taken from 8 words (ayates) of *Sourate Al-Fatihah*, which manually arranged into 7 files of model and stored into the database as HMM model (.mat). These models not just recognize the phonemes but also checks for the tajweed rules that govern the recitation of the Holy Quran (Bashir, M.S. et al., 2003). Each of experiment executed, both the training and word templates uttered are from the same speaker (recitor).

4.3 Result of Matching/Testing

As mentioned earlier in training/enrolment part, data templates were used as reference model (template matching), purposely for recognition task. In this task, any input that passing through this engine will be compared with the stored template, and any template pattern will be identified as recognized word (ayates) or phrase/phoneme (Identification Result). In other hand, the result of verification process can be obtained after a specific threshold value has been computed from the training process, in order to obtain the correct data as to make it as a reference for the upcoming process for recognition. The algorithm developed for recognition phase as discussed earlier, is simulated using MATLAB version 7. As mentioned before, there are 2 tests performed which are:

4.3.1 Testing – Word (Ayates) Like-Template

The testing process has been carried out, due to evaluate the system performance and those results were tabulated as shown in table 3 and 4. All the values shown in table 4 have been determined based on the maximum Log-Likelihood Ratio (LLR) of the score values, which had been discussed earlier in section of 3.1.2. In this part, the LLR threshold value is -1100 with the different ratio value of 0.2. If the value of LLR >-1100, it is considered as IV word, meanwhile, if LLR < -1100, it will considered as OOV word. Moreover, result obtained after the implementation from the equation (5), gives the different ratio values of IV almost bigger than

0.2, while most of OOV input give the values less than 0.2. Based on the result shown in table 3, those values represent under the $\mathbf{x_1}$ column indicated as IV words (above the ratio value of 0.2).

Meanwhile, other LLR values under column $\mathbf{x_2}$ until $\mathbf{x_8}$, represent the OOV words (LLR values less than 0.2). Meaning that, all 8 ayates of *sourate Al-Fatihah* shown in table 3, were categorized under the IV words. In relation with the application of this engine, whenever any of input claimed as OOV word/ayates, there is notification of the incorrect recitation of *sourate Al-Fatihah*, as well as the notifications of Tajweed Rules references, made for evaluation purposes. Whenever an IV input identified as an IV, there is a correct IV notification detection, altogether with the ayates of *sourate Al- Fatihah* identified with the correct recitation will be heard all along.

Table 3: Result of Likelihood Ratio (LLR) for 8 recitations of speech samples (1.0 x 10^3)

Ayates (.wav)	**Sequence**	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
Bismillah	**logP(X│Θ_1)**	**0.2112**	-3.9878	-4.4179	-4.6103	-5.1018	-5.4842	-5.6575	-5.7628
	MLM	1	3	6	7	8	5	2	4
Fatihah1	**logP(X│Θ_2)**	**0.4394**	-4.7675	-4.8948	-4.9438	-5.1501	-5.2021	-5.5128	-5.8265
	MLM	2	7	8	5	6	1	4	3
Fatihah2	**logP(X│Θ_3)**	**0.2472**	-3.6302	-3.9481	-4.5353	-4.8883	-5.0468	-5.1351	-5.1712
	MLM	3	1	6	7	2	8	4	5
Fatihah3	**logP(X│Θ_4)**	**0.7253**	-3.9347	-4.1251	-4.2471	-4.4244	-4.5630	-4.6807	-4.8629
	MLM	4	6	1	5	7	3	8	2
Fatihah4	**logP(X│Θ_5)**	**0.2659**	-4.8868	-5.7913	-5.9782	-6.6163	-7.6434	-7.6572	-8.4972
	MLM	5	6	7	1	8	3	4	2
Fatihah5	**logP(X│Θ_6)**	**0.2667**	-4.4097	-4.8590	-4.8904	-4.9843	-5.3690	-5.8303	-7.7457
	MLM	6	7	5	8	1	3	4	2
Fatihah6	**logP(X│Θ_7)**	**0.6612**	-4.6829	-5.1914	-5.3106	-5.4521	-6.4570	-6.9848	-7.8626
	MLM	7	6	8	1	5	3	4	2
Fatihah7	**logP(X│Θ_8)**	**0.8678**	-4.3930	-4.6584	-4.8508	-5.1978	-5.8682	-6.4164	-7.1213
	MLM	8	1	7	5	6	4	3	2

MLM = Most Likely Model

Table 4: Test result for 8 recitations of speech samples (ayates of *sourate Al-Fatihah*)

Ayates/Articulation	**# of utterances**	**Correct**	**Wrong**	**% Accuracy**	**% Word Error Rate**
بِسْمِ ٱللَّهِ ٱلرَّحْمَـٰنِ ٱلرَّحِيمِ	5	5	0	100	0
ٱلْحَمْدُ لِلَّهِ رَبِّ ٱلْعَـٰلَمِينَ	5	5	0	100	0
ٱلرَّحْمَـٰنِ ٱلرَّحِيمِ	7	7	0	100	0
مَـٰلِكِ يَوْمِ ٱلدِّينِ	6	6	0	100	0
إِيَّاكَ نَعْبُدُ وَإِيَّاكَ نَسْتَعِينُ	9	8	1	88.89	11.1
ٱهْدِنَا ٱلصِّرَٰطَ ٱلْمُسْتَقِيمَ	9	9	0	100	0
صِرَٰطَ ٱلَّذِينَ أَنْعَمْتَ عَلَيْهِمْ غَيْرِ ٱلْمَغْضُوبِ	6	4	2	66.67	33.33

ٱلضَّآلِّينَ عَلَيْهِمْ وَلَا	5	4	1	80	20
Total	**52**	**48**	**4**	**91.95**	**8.05**

The final result for the first test as in table 4 shown that the extracted features of 8 ayates of *Sourate Al-Fatihah* (compared to the Word based Templates) is perfectly reached to 91.95%, with 8.05% of Word Error Rate (WER). Compared with other research likely, the result obtained in table 4 is better than the result of previous researchers, carried out by Ehab *et al.* (2007) and Anwar *et al.* (2006) with the accuracy rate of recognition are 85% and 89% respectively.

4.3.2 Testing - Phonemes-Like Template

In this second test for phonemes-like templates, the Tajweed rules for the particular ayates of Quranic rules recitation have been carried out. Note that, threshold value is -500 with the different ratio of 0.01. Here, the threshold setting is totally different from the previous testing process, where if LLR<-500 it is considered as IV phoneme, meanwhile, if LLR>-500, it will be considered as OOV phonemes. If the particular utterance has been detected as OOV phoneme, the identification and verification process of pronunciation rules error (Tajweed rules) will assume those pronunciations as false/incorrect. For a better understanding related to Tajweed checking rules process, author has represented a comparison between correct and incorrect recitation as shown in table 5.

Table 5: Comparison between correct and incorrect Tajweed rules for ayates "Bismillahir ***<rahmaanir>*** rahimi"

	Correct Recitation	**Incorrect Recitation**
Ayates	بِسْمِ ٱللَّهِ ٱلرَّحْمَٰنِ ٱلرَّحِيمِ	بِسْمِ ٱللَّهِ ٱلرَّحْمَٰنِ ٱلرَّحِيمِ
The utterances (Articulation)	Bismillahir **RAHMAANIR** rahimi	Bismillahir **RAHMUUNIR** rahimi
Log-Likelihood (LLR)	**LLR:** 1.0e+003 * Columns 1 through 6 **-0.5685** -0.6033 -0.6398 - 0.6541 -0.6604 -0.7845 Columns 7 through 9 - 0.7968 -0.8684 -0.8995 - 1.0446 -1.0463 -1.1091 Columns 13 through 17 - 1.1624 -1.3030 -1.5521 - 2.1808 -2.2018	**LLR:** 1.0e+003 * Columns 1 through 6 **0.0544** -0.4929 -0.5621 - 0.7123 -0.7422 -0.7777 Columns 7 through 12 - 0.8362 -0.8738 -0.9155 - 0.9279 -0.9294 -0.9670 Columns 13 through 17 - 0.9958 -1.0265 -1.2703 - 1.4974 -2.2708
Tajweed Rules	-	Mad Asli Mutlak

In this case, a phoneme from the ayates "Bismillahir ***<rahmaanir>*** rahimi" has been successfully tested. Here, the result of LLR value highlighted is **-0.5685**, which are below the LLR threshold value (LLR< -500) and been classified under the IV phonemes (Correct recitation with LLR value less than 0.01). Meanwhile, another LLR value highlighted is **0.0544**, which represent as OOV phoneme, with the value above 0.01. This value were categorized under OOV phonemes (Incorrect recitation), since this value were specified above the LLR threshold value (LLR > -500). For the first phoneme, the incorrect recitation of tajweed pronunciation error is claimed as 'Mad Asli Mutlak', where the phoneme tested need to pronounce as ***'rahmaanir'*** and not ***'rahmuunir'*** with 2 harakattes of recitation.

In other hand, based on the results shown in table 6, the features vector of 28 samples from Al-Quran input phonemes was perfectly match the phoneme based template with the percent accuracy reached to 86.41%, with 14.34% of error rate only. Although the percent accuracy obtained is quite smaller compared to previous result in table 4, but the result is still under the expectation. It is because this experiment involved with a large amount of samples, particularly for testing purposes.

Table 6: Test result for 28 recitations of speech samples (Phoneme)

Ayates	Phonemes	# of utterance	Correct	Wrong	% Accuracy	% WER
Bismillah.wav	*Bismi / Llahii / Rraohimani / Rraohiiim*	17	16	1	94.12	5.88
fatihah1.wav	*Allhamdu / Lillahhirabbil / A'alamiinna*	9	8	1	88.89	11.1
fatihah2.wav	*Alrrahmani / Alrraheemi*	6	6	0	100	0
fatihah3.wav	*Maaliki / Yawmi / Alddeeni*	8	8	0	100	0
fatihah4.wav	*Iyyaka / naA'Abudu / waiyyaka / nastaAAeenu*	11	8	3	72.72	33.3
fatihah5.wav	*Ihdina / Alssiratho / Almustaqeema*	9	8	1	88.89	11.11
fatihah6.wav	*Siratho / Allatheena / An'Aamta / 'AAalayhim / Ghayri /Almaghdoobi*	12	8	4	66.67	33.33
fatihah7.wav	*'AAalayhim / Wala / Alddhalleena*	10	8	2	80	20
Total	**28**	**82**	**70**	**12**	**86.41**	**14.34**

The rationale behind this result, due to the complexity in pronouncing this ayates, as well as the difficulties in matching and recognizing the exact utterance properly. Here, the current method used is much simpler than LLR, which only need to calculate the perturbed value with easier calculation. Besides, the accuracy of recognition rate also depends upon the clarity of pronunciation of the speakers (recitors).

5. Conclusion

The work presented in this paper is an attempt towards automation of the process of speech recognition, specifically for Quranic verse recitation. The developed algorithms are tested on varieties of samples drawn from different speakers (recitors). For the earlier stage of research, the engine developed showed the promising results, although it was only tested against the small Quranic chapter (*Sourate Al-Fatihah*).There are some variations existed between both results in table 4 (ayates) and table 6 (phonemes). The rationale behind this result, probably because of the complexity in pronouncing the certain ayates, as well as difficulties in matching and recognizing the exact utterance properly. The implementation of Al-Quran in speech recognition system, especially in checking the Tajweed Rules always be a new developments in this technology, in which allow more researchers and creativity to get involved. The proper actions need to be put into consideration, due to improve the development of this engine and make it more compatible and useful for the end users, as it is designated for that purposes.

6. Acknowledgement

Special thanks are posed to University of Malaya for its funding in carrying out this research project. The authors are grateful to the project leader Professor Dato' Dr M.Y. Zulkifli Mohd Yusoff for valuable guidance and ideas, and to Emran Mohd. Tamil and Noorzaily Mohamed for helping in software development, as well as cultivating ideas in several aspects of this work. We gratefully acknowledge their support for this research.

7. References

Ahmad, A.M., Ismail, S. and Samaon, D.F. (2004), "Recurrent Neural Network with Backpropagation through Time for Speech Recognition", in proceeding of IEEE International Symposium on Communication & Information Technology (ISCIT '04), Sapporo, Japan, Vol. 1, pp.98-102.

Ahmed, M.E. (1991), "Towards an Arabic Text-To Speech system", The Arabian Journal for Science and Engineering (AJSE), Vol. 16, No. 4B, pp.565-583.

Anwar, M.J., Awais, M.M., Masud, S. and Shamail, S. (2006), "Automatic Arabic Speech Segmentation System", International Journal of Information Technology, Vol.12, No. 6, pp.102-111.

Bashir, M.S, Rasheed, F.S., Awais, M.M., Masud, S., & Shamail, S. (2003),"Simulation of Arabic Phoneme Identification through Spectrographic Analysis", working paper, Department of Computer Science, University of Engineering & Technology, Lahore Pakistan.

Essa, O. (1998), "Using Prosody in Automatic Segmentation of Speech", in proceeding of the 36th ACM Southeast Regional Conference, NY, USA, pp. 44 – 49.

Ehab, M., Ahmad, S. and Mousa, A. (2007), "Speaker Independent Quranic Recognizer Based on Maximum Likelihood Linear Regression", in proceeding of World Academy of Science, Engineering and Technology, Vol. 20, pp. 376-382.

Hemantha, G.K., Ravishankar, M., Nagabushan, P. and Basavaraj, S. A. (2006), “Hidden Markov model-based approach for generation of Pitman shorthand language symbols for consonants and vowels from spoken English”, Sadhana Journal, Academy Proceedings in Engineering Sciences, Vol.31, pp.277-290.

Hewet, T.T., Baecker, R., Card, S., Carey, T., Gasen, J. Mantei,M., Perlman, G., Strong, G., and Verplank, W. (1996), *ACM SIGCHI Curricula for Human-Computer Interaction,* ACM Special Interest Group (SIG) on Computer-Human Interaction Curriculum Development Group, The Association for Computing Machinery, New York, USA.

The New Strait Times Press (2007), “Institute for Research in Islamic education”, 26 September, pp. 002.

Kirchhoff, K., Vergyri, D., Bilmes, J., Duh, K. and Stolcke, A. (2004), “Morphology-Based Language Modeling for conversational Arabic speech recognition”, in proceeding of Eighth International Conference on Spoken Language Processing (INTERSPEECH 2004 – ICSLP), Korea, pp. 2245-2248.

Kirchhoff, K., Bilmes, J., Das, S., Duta, N., Egan, M., Ji, G., He, F., Henderson, J., Liu, D., Noamany, P., Schone, R., Schwartz, D. and Vergyri (2003), “Novel approaches to Arabic speech recognition: report from the 2002 Johns-Hopkins Summer Workshop”, in proceeding of IEEE International Conference on Acoustics, Speech & Signal Processing (ICASSP’03), Hong Kong , Vol.1, pp. I-344-I-347.

Laura, M.T., Alan, W.B. and Kevin, A.L. (2003), “Arabic in my Hand: Small-footprint Synthesis of Egyptian Arabic”, in proceeding of the EuroSpeech ’03, Geneva, pp.2049-2052.

Maamouri, M., Bies, A. and Kulick, S. (2006), “Diacritization to Arabic Treebank Annotation and Parsing”, in proceeding of the Conference of the Machine Translation SIG, London, pp.35-47.

Kristina Nelson (2002), “The art of Reciting the Quran”, Cairo & New York: The American University in Cairo Press, pp. xxviii +246.

Noor Jamaliah, I., Zaidi, R., Zulkifli, M.Y., M. Yamani I., and Emran, M.T. (2008), "Quranic verse Recitation feature extraction using Mel- Frequency Cepstral Coefficients (MFCC)", in proceeding of The 4th International Colloquium on Signal Processing and its Application (CSPA), Kuala Lumpur, Malaysia, pp. 13-18.

Rabiner, L. and Juang, B.H. (1993), *Fundamental of Speech Recognition*, Prentice Hall, New Jersey, USA.

Rabiner, L.R. (1989), “A Tutorial on Hidden Markov Model and Selected Applications in Speech Recognition”, in proceeding of the IEEE, Vol. 7, Issue 2, pp. 257-286.

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 & Communication Technologies, Vol. 2, pp. 2979–2984.

Vergyri, D. and Kirchhoff, K. (2004), ”Automatic Diacritization of Arabic for Acoustic Modelling in Speech Recognition”, COLING Workshop on Arabic-script Based Languages, Geneva.

Zaidi, R., Noor Jamaliah, I., M. Yamani, I., Emran, M.T., Zulkifli, M.Y. and Noor Naemah, A.R. (2008), “Quranic Verse Recitation Recognition Module for Support in j-QAF Learning: A Review”, International Journal of Computer Science and Network Security, Vol. 8, No.8, pp.207-216.