FPGA Implementation of Quranic Recitation for Pronunciation Correction: A Review

Noor Jamaliah Ibrahim, University of Malaya
Zulkifli Mohd Yusoff, University of Malaya
Zaidi Razak, University of Malaya
Mohd Yamani Idna Idris, University of Malaya
Noor Naemah Abdul Rahman, University of Malaya, et al.

Available at: https://works.bepress.com/noorjamaliah_ibrahim/8/
FPGA Implementation of Quranic Verse Recitation for Pronunciation Correction: A Review

Noor Jamaliah, I. (Corresponding author)
Centre of Quranic Research (CQR), Universiti Malaya
59990 Kuala Lumpur, Malaysia
Tel: +603-2246 3341 Email: jamaliah.um@gmail.com

M.Y. Zulkifli, M.Y.
Centre of Quranic Research (CQR)/ Academy of Islamic Studies
Universiti Malaya, 50603 Kuala Lumpur, Malaysia
Tel: +603-2246 3340 Email: zulkifliy@um.edu.my

Zaidi, R.
Centre of Quranic Research (CQR)/ Faculty of Computer Science & Information Technology,
Universiti Malaya, 50603 Kuala Lumpur, Malaysia
Tel: +603-7967 6412 Email: zaidi@um.edu.my

Mohd Yamani I dna, I.
Centre of Quranic Research (CQR)/ Faculty of Computer Science & Information Technology,
Universiti Malaya, 50603 Kuala Lumpur, Malaysia
Tel: +603-7967 6414 Email: yamani@um.edu.my

Noor Naemah, A. R.
Academy of Islamic Studies, Universiti Malaya
50603 Kuala Lumpur, Malaysia
Tel: +603-7967 6011 Email: naemah@um.edu.my

Emran, M. T.
Centre of Quranic Research (CQR)/ Faculty of Computer Science & Information Technology,
Universiti Malaya, 50603 Kuala Lumpur, Malaysia
Tel: +603-7967 6437 Email: emran@um.edu.my

Noorzaily, M. N.
Centre of Quranic Research (CQR)/ Faculty of Computer Science & Information Technology,
Universiti Malaya, 50603 Kuala Lumpur, Malaysia
Tel: +603-7967 6413 Email: zaily@um.edu.my
Abstract

This paper proposes recognition of Quranic verse recitation focus on Tajweed rules mainly for pronunciation correction using Hidden Markov Models with Field Programmable Gate Array (FPGA) board. It introduces the basic principle of speech recognition system by using two main algorithms which is Mel Frequency Cepstral Coefficients (MFCC) and Hidden Markov Model (HMM). In term of HMM module, two different algorithms will be applies, which is Baum-Welch Algorithm for HMM training and Viterbi Algorithm for HMM testing. The most important part on this project is the hardware implementation by using Xilinx Field Programmable Gate Array (FPGA) board.

Keywords: Quranic recitation, Speech Recognition, Hidden Markov Model (HMM), Mel-Frequency Cepstral Coefficients (MFCC), Field Programmable Gate Array (FPGA)

Introduction

Reciting Al-Quran in appropriate way is very important for all Muslims and indispensable in Islamic worship, such as prayers. Moreover, in learning Al-Quran as shown by our Prophets, different systems and methodologies are essential in putting the word of God in its rightful place. Since the revelation of the Holy Quran, Al-Quran learning process through talaqqi & musyafahah method is the only way to learn how to recite Al-Quran correctly until twentieth century. The development of Quranic lesson learns have been successfully produced a lot of Quranic scholars and at the same time promoting the Quranic standard into high priority level.

The development of the ICT also has change the world into many ways, both positive and negative aspects. It is also considered as one of the rapid channel of technology to spread information to all Muslim at all over the world. Therefore, each of Muslim must be able to identify the appropriate and practical ways of selecting the right type of information obtained from this new technology. Even though the world has changed drastically, the development in Quranic studies have never become outdated. World globalization era as well as high technology, also could not prevent the academia in Quranic studies from been influenced by the current trends of technology (N. Jamaliah, 2010).

In this paper, we review the new implementation of FPGA (Field Programmable Gate Array) in recognizing the Tajweed rules in Quranic verse recitation. Before this, the same method also has been introduced by Zaidi, R. et al. (2011) in their research, but the input of languages used just focused on Arabic speech. In order to preserve and uphold the dignity of Al-Quran for the good sake of Muslim community, new ideas and research need to explore due to expand the broader field of techniques, focused in improving the Quranic learning process, through Quranic reading skill of talaqqi and musyafahah method. The development of the ICT is considered as one of the rapid channel of technology to spread information to all Muslim at all over the world. Thus, this channel can be used as a main medium, due to spread the words of Allah SWT into the appropriate manners (I. Noor Jamaliah, M.Y. Zulkifli M.Y.,Zaidi. R., 2011).This application of ICT invented is extremely important and beneficial due to strengthen the learning and teaching process of Al-Quran, especially for j-QAF. J-QAF educational program, which has been introduced by Malaysian Education Ministry were aims to build the character of the Malaysian pupils in the primary grades to outclass the teachings of the Noble Qur'an, the Islamic sciences, Arabic as well as other foreign
languages. It is mainly to inculcate moral and ethical values, as well as nurture responsible citizens among young Muslim generation in Malaysia.

This method is described as face to face learning process between students and teachers (Mudarris), where listening, correction of Al-Quran recitation and repetition of the correct Al-Quran recitation took place. This factor is important, so that students will know how the hijaiyah letters are pronouncing correctly. The process only can be done, if the Mudarris and students follow the art, rules and regulations while reading the Al-Quran, known as “Rules of Tajweed” (Tabbal, 2006; R. Zaidi, Noor Jamaliah, I., M. Yamani, I. I., Emran, M.T., Zulkifli, M.Y., and Noor Naemah, A. R., 2008).

**Literature Review**

According to this research, the basic of speech recognition technology was mainly focus here. However, the implementation is totally different, in which applied to the different type of application or languages, such as Quran. The different contents of input, would probably affect the percentage of accuracy during the recognition process. Thus, the reliability and effectiveness of the system was depending on the language and system design created. The algorithms used for the certain application using the basic principle of speech recognition technique has been explored in (Ehab, 2007; Tabbal, 2006; Xilinx, 2010) and Morgan (2000). Under the same field of speech recognition, the basic stages of speech processing need to be considered. Generally, there are 5 main stages of the speech recognition process, which includes; 1) Pre-processing 2) Feature Extraction 3) Training 4) Identification 5) Verification.

Based on the research conducted by Ahmad, A. M. et al. (2004) and Noor Jamaliah, I. et al. (2008), Mel Frequency Cepstral Coefficient (MFCC) algorithm has been used for feature extraction process. The survey provides recognition rates and description of test data for the approaches considered between LPCC and MFCC. From the results obtained from Ahmad (2004), LPCC is the best performance for recognizing the Arabic alphabets of Quran, with the percentage of 99.3% more efficient compared to MFCC. However, MFCC is still the most popular feature set with 98.6% efficient, in which computed on a warped frequency scale based on known human auditory perception.

Meanwhile for training, identification and verification stages the algorithm used is mainly stress on the Markov Chain Code algorithm for speech recognition. A Markov chain is characterization of a system that transits from one state to another, in a chainlike manner. It concerns any random process endowed with the Markov property. It is a useful tool for statistical modeling in almost all fields of modern applied mathematics. Markov chain is the basis for Hidden Markov models, which are an important tool in such diverse fields in speech recognition application of this proposed research. It involves with states, transitions and observations map into speech recognition task (Juang, 1991). This Markov Chain algorithm was used in developing the software part by using mathematical modeling of MATLAB program, due to be integrated with the hardware part. MATLAB program has become the preferred language of computing for the researchers. Moreover, in this research MATLAB language will be used as a development environment for the FPGA implementation that will be discussed in details later at implementation part.
Research Methodology

I. Software Development

![System Block Diagram](image1)

A. Speech Pre-Processing

Under the pre-processing stage, the input speech signal (analog) will be converted through the conversion process using Analog to Digital Converter (ADC). This process will be carried out after the input speech has been successfully recorded. The preliminary segmentation also executed in this stage, where start and endpoint detection, as well as noise reduction will be identified. Process of segmentation during this stage is totally important, due to differentiate between the speech segments and silence segments. The silence segments need to removed, because it is considered as unnecessary segments, thus the time frame can be narrow down and overall processing time able to be reduced.

B. Speech Analysis

i. Mel-Frequency Cepstral Coefficients (MFCC)

According to (Bateman, 1992; Ehab, 2007; I. Noor Jamaliah, Zaidi, R., Zulkifli, M.Y., M. Yamani I., Emran, M.T., 2008), MFCC is perhaps the popular feature extraction method used recently. This algorithm is based on the known variation of the human ear’s critical bandwidth with frequency. MFCC consists of 8 main computation steps, which include pre-emphasis, framing, windowing (Hamming Window), Discrete Fourier Transform (DFT), Mel Scale Filterbank, Logarithm and lastly, the inverse DFT (DCT).

![MFCC Block Diagram](image2)

The output and results obtained from the final stage of feature extraction were known as
Mel cepstrum signal. At this stage, the inverse Fourier Transform (IDFT) of the algorithm of the magnitude spectrum was computed, in order to obtain the Mel frequency Cepstrum Coefficients (MFCC). Thus, MFCC are ready to be formed in a vector format, known as features vector.

ii. Hidden Markov Model (HMM)

Hidden Markov Model (HMM) act as a statistical method of characterizing and matching the spectral pattern of input speech. In HMM statistical approach, the input speech of Quranic recitation is represented accordingly with some probability distributions. If the observation is a probabilistic function of state, it is known as Hidden Markov Models, because it consists of doubly embedded stochastic process with underlying that is not directly observable (hidden). However, it can be observed through another set of stochastic process only, which may produce the sequence of observations (Rabiner, 1993). As mentioned before, HMM has been implemented at the training, identification and verification stages.

![Figure 3: Example of Hidden Markov Model (E.M Essa, 2008)](image)

**HMM Training**

In this part, **Forward-Backward Recursions** process will be executed. It also known as **Baum-Welch algorithm**, in which used to determine the reference pattern, as the best matches between input features vectors by comparing stochastic possibility scores. HMM model (template) will be the results of this process, after the re-estimation process successfully been executed.

**HMM Testing/Recognition**

Acoustic feature sequence through decoding or aligning requires the prior specification of parameter from HMM. As mentioned earlier, the HMM model has a role of stochastic templates, for comparing the observations. These templates consist of several sentences, which represent different phonemes of Quranic recitations. Each of templates can be determined and identified through the estimation of HMMs parameter, specified by a certain database which contained the observations sequences, either supervised or unsupervised way of learning method. Here, the ayates and phonemes of the Quranic recitations are recognized after comparing the testing model with the help of **Viterbi algorithms**. This algorithm will test the HMM set, due to find the optimal state path for an observation sequence.

iii. MATLAB – The technical computing language

MATLAB is a high level of language, which equipped with an interactive environment that perform computationally intensive tasks faster than traditional language, such as C, C++ and Fortran (E.M Essa, 2008).
II. Hardware Development

Field Programmable Gate Array (FPGA) is an integrated circuit (semiconductor device), which designed to be configured by a customer or manufacturer. This device can be programmed using a hardware description language (HDL), which allow us to program product features and functions, adapt to new standards, and reconfigure hardware for specific applications even after the product has been installed to any hardware function.

Here, a Field Programmable Gate Array (FPGA) board will be used to act as a co-processor, due to its capability to perform recognition at a much higher rate than software implementation. Besides, FPGA also lies between general processors and ASICs (application-specific ICs) on the spectrum of processing elements, which are highly flexible and high performance. Generally, FPGAs are programmed using a hardware description language such as VHDL (Very High Speed Integrated Circuit Hardware Description Language) and Verilog (Thompson, 2001). Several types of FPGA have been produced by manufacturers, and the current market leaders identified are Xilinx and Altera, both the long-time industry rivals (Stephen J. Melnikoff, 2002). The differences between both two technologies are shown in table 1 below:

<table>
<thead>
<tr>
<th>Xilinx</th>
<th>Altera</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Flexible I/Os</td>
<td>• Simple tools</td>
</tr>
<tr>
<td>• More complex tools</td>
<td>• Quick compile</td>
</tr>
<tr>
<td>• Good at arithmetic functions</td>
<td>• Less capable arithmetic</td>
</tr>
<tr>
<td>• More capable tools for the power user</td>
<td>• No small blocks of embedded RAM</td>
</tr>
<tr>
<td>• Smaller and cheaper devices</td>
<td>• RAM has one read and one write port, not proper dual port.</td>
</tr>
<tr>
<td>• Longer compile times</td>
<td>• Less flexible tools for the power user.</td>
</tr>
<tr>
<td>• Both small and large blocks of embedded RAM</td>
<td>• Flex and Apex make it tricky to make fast bi-directional I/O.</td>
</tr>
<tr>
<td>• Proper dual port RAM</td>
<td>•</td>
</tr>
</tbody>
</table>

Implementation of FPGA in this research required us to use AccelDSP synthesis tool to transform a MATLAB floating-point design into FPGA. This transformation is mainly used to perform stimulus creation, evaluation of algorithm and post-processing results. AccelDSP synthesis tool is a high-level MATLAB® language based tools for designing the DSP blocks of Xilinx FPGAs (McCain, 2005).

As mentioned before, mathematical modeling of MATLAB language will be used in developing the software part with appropriate algorithms regardless to the targeted application. This software part will be integrated with the hardware part through AccelDSP synthesis tool. Meaning that, the MFCC and HMM algorithm developed at software development part will be integrated with the hardware system on the Xilinx FPGA board.

Expected Result & Discussion

The main focus in this research paper is towards the implementation of FPGA, as the core part at the hardware development section in Quranic verse recitation recognition. The process
begin with the software development, using the MATLAB language for both MFCC (input feature extraction) and HMM (training & recognition) algorithms.

Meanwhile, AccelDSP synthesis tool will be used to transform a MATLAB floating point design into block modules for Xilinx FPGA, in order to perform stimulus creation, algorithm evaluation and the results for post-processing. In other hand, Xilinx ISE Design Suite is used to combine all block modules into one system, while the Xilinx ISE Simulator is used to test the system functionality. After that, the system will be uploaded onto FPGA board, and lastly, the final testing need to executed, due to ensure the functionality of the desired system on the FPGA board. Figure 4 below, shows the overall flow of the proposed method.

![Figure 4: Flow chart of research activity](image)

In this research paper, the proposed method discussed earlier highlights the potential device purposely to create speaker independent device for Quranic verse recitation system, which can be used for checking the Quranic verse pronunciation (revised and corrected in real time), either it is true or false. The expected outcome of this application approximately reaches to 95% accuracy.

**Conclusion**

The main idea of this proposed research, were mainly focus towards the improve design of Quranic verse recitation recognition with capability function of tajweed checking rules. The
integration process between software and hardware system is really beneficial for future commercialized purposes and for sake of the Noble Quran for Muslim community. The advantages of hardware implementation (such as Xilinx FPGA, DSP) over software is, its capability of real-time processing and response, thus improving speed and energy efficiency.

References


