Classification of the Correct Quranic Letters Pronunciation of Male and Female Reciters

IOP Conf. Series: Materials Science and Engineering **260** (2017) 012004 doi:10.1088/1757-899X/260/1/012004

Classification of the Correct Quranic Letters Pronunciation of Male and Female Reciters

Safiah Khairuddin, Salmiah Ahmad, Abdul Halim Embong, Nik Nur Wahidah Nik Hashim, Tareq M.K. Altamas, Syarifah Nuratikah Syd Badaruddin, Surul Shahbudin Hassan*

Department of Mechatronics Engineering, International Islamic University Malaysia, Kuala Lumpur, Malaysia

* Centre for Languages and Pre-University Academic Development, International Islamic University Malaysia, Kuala Lumpur, Malaysia

Abstract – Recitation of the Holy Quran with the correct Tajweed is essential for every Muslim. Islam has encouraged Quranic education since early age as the recitation of the Quran correctly will represent the correct meaning of the words of Allah. It is important to recite the Quranic verses according to its characteristics (*sifaat*) and from its point of articulations (*makhraj*). This paper presents the identification and classification analysis of Quranic letters pronunciation for both male and female reciters, to obtain the unique representation of each letter by male as compared to female expert reciters. Linear Discriminant Analysis (LDA) was used as the classifier to classify the data with Formants and Power Spectral Density (PSD) as the acoustic features. The result shows that linear classifier of PSD with band 1 and band 2 power spectral combinations gives a high percentage of classification accuracy for most of the Quranic letters. It is also shown that the pronunciation by male reciters gives better result in the classification of the Quranic letters.

Keywords *– Quranic Letters, Formants, Power Spectral Density, Linear Discriminant Analysis*

1. Introduction

Nowadays, speech processing applications has been widely used in various sectors such as wireless communication, education, medicine and many more. Speech is used as a communication platform for human to connect with each other. Arabic language is one of the languages that is widely used in speech all over the world, especially for Muslims, as it is the language of the Quran, which is the holy book of Islam. Pronunciation of Quranic utterances is considered as a part of general speech production process. It is compulsory for all Muslims to read the Quran regularly throughout their life. Therefore, it is important to start the education process in learning Quranic recitation from early age. People usually learn to recite the Quran through a teacher that will teach them the correct ways of

pronunciation and recitation of Quranic verses. The student will also read the Quran in front of the teacher and the teacher will listen and corrects the wrong pronunciation or recitation [1].

This conventional way of learning Quran has been practiced for hundreds of years, ever since the Quran was sent down to Prophet Muhammad s.a.w.. The process of learning will usually take a couple of years before the student becomes expert and can recite the Quran by themselves [2]. Technically, the Quran should be recited using the correct Tajweed, or the rules of pronunciation during recitation of Quran. Tajweed literally means to make better or to improve. It can also be defined as articulating every letter from its articulation point (Makhraj) and giving them its right and dues of characteristics (Sifaat) [3]. Point of articulation of a specific letter is the place that is used to emit the letter. It is necessary to use the right articulation points while producing the utterances. The study on makhraj using speech processing has been done in [4] [5] and [6].

To date, there are no published study on identifying Makhraj in the pronunciation of all 28 Quranic letters that compares male and female pronunciation. Thus, the motivation of this paper is to identify the possible features that can represent each of the Quranic letters uniquely and to see whether there are significant differences between male and female Quranic letters pronunciation.

2. Database collection and formulation

2.1 Data Collection

Our sample consisted of 22 sets of audio data which were obtained via recording session with 11 males and 11 females who were selected based on their expertise in Quranic recitation. All data were verified by a certified expert which specialized in Quranic recitation, Ust. Dr. Surur Shahbudin. The audio data was recorded using a portable high-quality field recorder, TASCAM DR-05, with frequency response of 40Hz to 20 kHz. The recordings were done in a normal room environment such as classrooms and labs in engineering faculty of the university, which will later may represent a normal learning environment. During the data recording, the participant will need to recite 28 combinations of Quranic letters, which are based on *Rasm 'Uthmani* narrated by Hafs bin Sulaiman for *Qira'at Asim* [7]. Table 1 shows the *sokoon* (°) combination of *Quranic* letters and its pronunciation that is recited for the recording. The participants are required to recite each *sukoon* letters only once without repeating them. This *sokoon* (°) combination of *Quranic* letters dataset can be used to describe the *Makhraj* (point of articulations) and *Sifaat* (characteristics) of each letter.

Table 1. The *sokoon* (°) combination of Quranic letters and its pronunciation.

Combination of Quranic letters	Pronunciation
أء	aa
أب	ab
أت	at
أث	ath
أج	aj
أح	ah
أخ	akh
أد	ad
أذ	az
أر	ar
أز	azz
أس	as
أش	ash
أص	asd
أض	adh
أط	athd
أظ	azd
أع	a’
أغ	agh
أف	af
أق	aq
أك	ak
أل	al
أم	am
أن	an
أه	ahh
أو	aww
أي	aii

2.2 Pre-processing

The audios were digitized with a sampling rate of 99.6 kHz and edited using *Audacity,* open source software for audio signal. The processing of data was started with de-identification of personal information as well as removal of interviewer’s voice and undesirable sound. Unwanted background noise was also removed using *Audacity*. For the sake of simplicity, all stereo data have been converted to mono channel prior to the noise reduction step. The audio signal has been normalized with the amplitude in between -1 to 1.

3. Feature extraction

Features extraction was done in speech recognition system to identify the components of the audio signal that are good for representing the acoustic content.

3.1 Formant

A formant is a concentration of acoustic energy around a frequency in the speech wave [8]. Each formant corresponds to a resonance in the vocal tract. They represent the frequencies that pass most of the acoustic energy from the source to the output. The formant frequency representation shows an efficient and compact representation of time-varying characteristics of speech [9]. Usually the first two formants are enough to distinguish between the vowels. Formant frequencies of the audio signal were first obtained using spectrogram features in *PRAAT* software.

The first and second formant frequency (F1 and F2) are extracted from 22 samples, which contains 11 male data and 11 female data. The average formant frequency was calculated. Table 2 shows the average F1 and F2 value of 11 male experts and 11 female experts for all Quranic letters *sukoon* combinations.

Table 2. Average F1 and F2 value (Hz) for 11 males and 11 female samples.

Quranic	Male average		Female average	
letters	F1 (Hz)	F2 (Hz)	F1 (Hz)	F2 (Hz)
أءْ	770.99	1718.26	796.48	1692.11
أبْ	674.04	1752.69	658.21	1706.89
أثْ	773.82	1885.74	775.96	1888.05
أثْ	851.46	1973.26	807.26	1918.27
أجْ	708.54	2002.06	632.34	1908.15
أحْ	990.37	1965.40	946.95	1985.06
أخْ	898.11	1992.48	839.76	1652.43
أدْ	632.93	1825.73	657.17	1892.83
أذْ	702.64	1878.27	679.53	1821.29
أرْ	742.87	1669.10	723.77	1463.68
أزْ	734.17	1953.30	668.46	1850.18
أسْ	986.86	2171.90	845.37	2002.96
أشْ	1263.66	2442.06	995.29	2277.63
أصْ	947.03	2138.36	840.56	1924.64
أضْ	676.95	1638.52	671.06	1483.90
أطْ	755.98	1664.30	714.73	1427.16
أظْ	754.98	1789.21	607.38	1496.56

أغْ	928.86	1741.28	795.37	1607.24
أغْ	754.89	1749.50	662.78	1517.99
أفْ	987.61	2065.16	855.68	1911.81
أقْ	775.82	1626.50	733.76	1450.35
أكْ	799.22	1967.71	794.13	1754.01
ألْ	620.81	1860.35	616.13	1855.05
أمْ	661.91	1820.03	624.39	1677.02
أنْ	667.69	1906.23	538.91	1692.22
أهْ	850.59	1901.06	837.53	1788.08
أوْ	716.52	1679.28	675.86	1527.66
أيْ	656.27	2063.66	624.45	1846.38

3.2 Power Spectral Density (PSD)

Power Spectral Density is the distribution of power over a frequency in signals [10]. Three spectral energy bands were extracted using Matlab from all 28 letters of *sukoon* combinations of 22 sets of male and female data. The energy bands named PSD_1, PSD_2, and PSD_3 will be paired by two and undergoes data classification with the different band combinations.

4. Classification based on selected features

Linear Discrimant Analysis (LDA) or linear classifier used in this study as classifier. The analysis is done based on pairwise analysis classification between male and female sample. Resampling method was applied because of the small number of data sets. Jackknife (Leave-One-Out) was used as the resampling method in this project as it can give an optimum result [11]. The Jackknife method leaves one sample from the data set for testing and constructs the classification function with the remaining N-1 samples as the training data [10].

5. Results and discussion

All the features that have been extracted are classified using LDA with Jackknife method as the resampling technique. In this study, we evaluated the features that have been extracted and see what is the feature combinations that will give the best result. All data are classified according to male and female group, and the data is further being classified by each quranic letters *sukoon* combination to group them into male or female class. The percentage indicates the accuracy of the classifier, which represent the correct classification of the data into its class. There would be 28 categories of Quranic letters after the classification process which is represented by male and female recitation.

5.1 Formant Frequency

F1 and F2 features extracted from the samples undergo classification process to classify them into male or female experts group. Table 3 shows the percentage of correct classification of all data according to its class.

Table 3. Male and Female experts' data classification with all data (Feature: F1 and F2).

Classifier	**Percentage (%)**		
LDA (resampled with Jackknife method)	All	Male	Female
	62.0	60.7	63.3

Using LDA, classification of all data using the first two formants gives 62.0% accuracy as it gives the result of 60.7% of correct classification for male group while female group exhibit 63.3% of the correct classification. The percentage of correct classification was also calculated using the same method to each of the Quranic letters *sukoon* combination. The results were tabulated in Table 4.

Table 4. Percentage of correct classification for each letter's *sukoon* combination using LDA (Feature: F1 and F2).

Letter Combination	أغْ	أبْ	أثْ	أشْ	أجْ	أحْ	أخْ	أذْ	أذْ
All (%)	50.0	31.8	0.0	40.9	59.1	50.0	72.7	59.1	45.5
Male (%)	36.4	36.4	0.0	45.5	54.5	36.4	63.6	54.5	45.5
Female (%)	63.6	27.3	0.0	36.4	63.6	63.6	81.8	63.6	45.5
Letter Combination	أرْ	أزْ	أسْ	أشْ	أصْ	أضْ	أطْ	أظْ	أغْ
All (%)	63.6	59.1	72.7	72.7	63.6	50.0	63.6	68.2	77.3
Male (%)	63.6	63.6	63.6	72.7	63.6	45.5	63.6	63.6	90.9
Female (%)	63.6	54.5	81.8	72.7	63.6	54.5	63.6	72.7	63.6
Letter Combination	أغْ	أفْ	أقْ	أكْ	ألْ	أمْ	أنْ	أهْ	أوْ	أيْ
All (%)	81.8	72.7	54.5	59.1	0.0	59.1	72.7	54.5	68.2	68.2
Male (%)	81.8	72.7	54.5	63.6	0.0	63.6	81.8	54.5	63.6	72.7
Female (%)	81.8	72.7	54.5	54.5	0.0	54.5	63.6	54.5	72.7	63.6

The result shows that the combination of F1 and F2 give the highest percentage of correct classification for the letter combination 'أغْ' for male experts, which is 90.9%. For female experts group, the letter combination of 'أخْ' 'أسْ' and 'أغْ' can be represented correctly using F1 and F2 features with 81.8% of accuracy respectively. Meanwhile, the result indicates that the combination of F1 and F2 features cannot represent the letter 'أتْ' and 'ألْ' correctly at all, as the percentage of classification accuracy is at zero percent. In average, this feature combination is not a good representation to specifically classify the Quranic letters into groups of male and female as the result shows low percentage accuracy for most of the letters. There are also not many differences in male and female percentage result using formants as acoustic features. Therefore, it can be concluded that formants are not a good feature representation in differentiating the gender in Quranic letters pronunciation.

5.2 Power Spectral Density (PSD)

Three PSD energy bands that have been extracted (PSD_1, PSD_2 and PSD_3) undergoes data classification process with three different combination of bands as the features. The combination of energy bands is in Table 5.

Table 5. Combination of PSD energy band.

Combination	PSD band
1	PSD_1 and PSD_2
2	PSD_1 and PSD_3
3	PSD_2 and PSD_3

The result of data classification that was done using LDA with Jackknife as resampling method was recorded in Table 6.

Table 6. Percentage of correct classification of PSD bands combination using LDA.

Letters		Combination	All (%)	Male (%)	Female (%)
1	أءْ	1	77.3	72.7	81.8
		2	77.3	72.7	81.8
		3	77.3	63.6	90.9
2	أبْ	1	90.9	81.8	100
		2	90.9	81.8	100
		3	81.8	72.7	90.9
3	أتْ	1	77.3	81.8	72.7
		2	77.3	81.8	72.7
		3	77.3	81.8	72.7
4	أثْ	1	90.9	81.8	100
		2	90.9	81.8	100
		3	95.5	90.9	100
5	أجْ	1	86.4	90.9	81.8
		2	77.3	90.9	63.6
		3	81.8	90.9	72.7
6	أحْ	1	63.6	54.5	72.7
		2	54.5	54.5	54.5
		3	68.2	63.6	72.7
7	أخْ	1	77.3	72.7	81.8
		2	68.2	81.8	54.5
		3	59.1	63.6	54.5
8	أدْ	1	40.9	45.5	36.4
		2	45.5	45.5	45.5
		3	50.0	36.4	63.6
9	أذْ	1	50.0	54.5	45.5
		2	54.5	63.6	45.5
		3	54.5	63.6	45.5
10	أرْ	1	59.1	45.5	72.7
		2	54.5	36.4	72.7
		3	54.5	36.4	72.7
11	أزْ	1	68.2	72.7	63.6
		2	50.0	27.3	72.7
		3	45.5	27.3	63.6
12	أسْ	1	95.5	100	90.9
		2	72.7	81.8	63.6
		3	86.4	90.9	81.8

13	أشْ	1	81.8	81.8	81.8
		2	68.2	63.6	72.7
		3	59.1	63.6	54.5
14	أصْ	1	63.6	72.7	54.5
		2	72.7	63.6	81.8
		3	63.6	63.6	63.6
15	أضْ	1	72.7	90.9	54.5
		2	72.7	90.9	54.5
		3	72.7	81.8	63.6
16	أطْ	1	68.2	72.7	63.6
		2	54.5	54.5	54.5
		3	50.0	45.5	54.5
17	أظْ	1	68.2	72.7	63.6
		2	63.6	81.8	45.5
		3	68.2	72.7	63.6
18	أغْ	1	72.7	90.9	54.5
		2	72.7	90.9	54.5
		3	72.7	90.9	54.5
19	أعْ	1	45.5	45.5	45.5
		2	50.0	54.5	45.5
		3	50.0	54.5	45.5
20	أفْ	1	59.1	63.6	54.5
		2	59.1	63.6	54.5
		3	59.1	63.6	54.5
21	أقْ	1	54.5	54.5	54.5
		2	54.5	54.5	54.5
		3	50.0	72.7	27.3
22	أكْ	1	100	100	100
		2	100	100	100
		3	100	100	100
23	ألْ	1	77.3	72.7	81.8
		2	81.8	81.8	81.8
		3	77.3	72.7	81.8
24	أمْ	1	77.3	81.8	72.7
		2	50.0	45.5	54.5
		3	72.7	72.7	72.7
25	أنْ	1	68.2	81.8	54.5
		2	68.2	72.7	63.6
		3	54.5	45.5	63.6

26	أهْ	1	54.5	45.5	63.6
		2	36.4	45.5	27.3
		3	22.7	36.4	9.1
27	أوْ	1	100	100	100
		2	59.1	54.5	63.6
		3	95.5	100	90.9
28	أيْ	1	100	100	100
		2	81.8	72.7	90.9
		3	95.5	90.9	100

Based on the observation of the percentage result, it was found that combination 1 which is the combination of band 1:2 (PSD_1 & PSD_2) gave highest classification accuracy for all, as well as male and female data for each of the Quranic letters *sukoon* pronunciation using LDA methods. This method also shows to be better in classifying male speaker in the letters pronunciation as it demonstrates a higher percentage in correctly classifying the letters into male group. With LDA, the combination of band 1:2 indicates the highest percentage accuracy in classifying the letter أبْ, and أثْ in female group, the letter أسْ in male group, and أكْ , أوْ , and أيْ in all of the groups, which shows 100% accuracy.

However, this combination may not be suitable in classifying the letters أذْ , أذْ , أغْ , أقْ , and أهْ as they displays a low percentage classification accuracy. This may be because the features are not enough to distinguish the unique characteristic of the letters to be classified into its own group. On the other hand, the other letters *sukoon* combination shows a good result in classification percentage with band 1:2 combination.

5.3 Power Spectral Density and Formant

We've combined the features of PSD and Formant to see if the result would give a better percentage accuracy. Table 7 shows the percentage accuracy of the classification of each letter into groups of all, male and female.

Table 7. Percentage classification of combined features of each letter using LDA.

Letter Combination	أغْ	أبْ	أثْ	أشْ	أجْ	أحْ	أخْ	أذْ	أذّ	
All (%)	77.3	90.9	77.3	90.9	86.4	63.6	77.3	40.9	50.0	
Male (%)	72.7	81.8	81.8	81.8	90.9	54.5	72.7	45.5	54.5	
Female (%)	81.8	100	72.7	100	81.8	72.7	81.8	36.4	45.5	
Letter Combination	أرْ	أزْ	أسْ	أشْ	أصْ	أضْ	أطْ	أظْ	أعْ	
All (%)	59.1	68.2	95.5	81.8	63.6	72.7	68.2	68.2	72.7	
Male (%)	45.5	72.7	100	81.8	72.7	90.9	72.7	72.7	90.9	
Female (%)	72.7	63.6	90.9	81.8	54.5	54.5	63.6	63.6	54.5	
Letter Combination	أغْ	أفْ	أقْ	أكْ	ألْ	أمْ	أنْ	أهْ	أوْ	أيْ
All (%)	45.5	59.1	54.5	100	77.3	77.3	68.2	54.5	100	100
Male (%)	45.5	63.6	54.5	100	72.7	81.8	81.8	45.5	100	100
Female (%)	45.5	54.5	54.5	100	81.8	72.7	54.5	63.6	100	100

The result exhibit a slight increase in percentage accuracy of classification of letters when all the features are combined. However, the changes are not significant as we could see that the percentage results are almost the same with classification of PSD band 1:2. Only the letter أكْ , أوْ , and أيْ can be represented using the features correctly with 100% accuracy. The other letter shows the same behaviour as in classification using PSD band 1:2 only. This shows that the combination of features has no effect in increasing the accuracy of the classifier to classify each of the letter.

6. Conclusion

We have done classification of Quranic letters to Male and Female groups using Linear Discriminant Analysis with Formant and Power Spectral Density as the features. Among the two features, power spectral density with combination of band 1:2 exhibit a better result to classify the letters into male and female. Overall, the classification percentage was higher for male reciters identification. For future study, it is recommended to explore other acoustical features and classifiers that will give better results in classifying the Quranic letters into its group accurately.

Acknowledgement

The author would like to express sincere appreciations to Ustaz Dr. Surul Shahbudin Hassan from the Centre For Languages And Pre-University Academic Development, International Islamic University Malaysia, Kuala Lumpur, Malaysia for his efforts in making the experiment successfully conducted by providing the experts, also thanks to all Ustaz and Ustazah who have kindly agreed and spent time in recording the pronunciation and recitation. Thank you also to the undergraduate students under my supervision who were committed in performing their final year project successfully.

References

[1] Ibrahim, N. J., Yusoff, Z. M., & Razak, Z. (2008). Quranic verse Recitation Recognition Module for Educational Programme.
[2] Rashid, N. R. R., Venkat, I., Damanhoori, F., Mustaffa, N., Husain, W., & Khader, A. T. (2013, December). Towards Automating the Evaluation of Holy Quran Recitations: A Pattern Recognition Perspective. In *Advances in Information Technology for the Holy Quran and Its Sciences (32519), 2013 Taibah University International Conference on* (pp. 424-428). IEEE.
[3] Wardhani, A., & Nurhayati, T. Makhraj Al-Jauf Dalam Kajian Ilmu Tajwid.
[4] Wahidah, A., Suriazalmi, M., Niza, M., Rosyati, H., Faradila, N., Hasan, A., Farizan, Z. (2012). Makhraj recognition using speech processing. Paper presented at the Computing and Convergence Technology (ICCCT), 2012 7th International Conference.
[5] Altalmas, T., Ahmad, S., Sediono, W., & Hassan, S. S. (2015). Quranic Letter Pronunciation Analysis based on Spectrogram technique: Case Study on Qalqalah Letters.
[6] N.W.Arshad, S.N.Abdul Aziz, R. Hamid, R. Abdul Karim, F. Naim, N. F. Zakaria (2011), Speech Processing for Makhraj Recognition: The Design of Adaptive Filter for Noise Removal, International Conference on Electrical, Control and Computer Engineering, Pahang, Malaysia, June 21-22, 2011
[7] Arshad, N. W., Sukri, S. M., Muhammad, L. N., Ahmad, H., Hamid, R., Naim, R., & Naharuddin, N. Z. (2013). Makhraj Recognition for Al-Quran Recitation using MFCC. *International Journal of Intelligent Information Processing (IJIIP)*, 45-53.
[8] Rabiner, L., & Juang, B. H. (1993). Fundamentals of speech recognition. (pp. 26-28).
[9] Jorgensen, C., Lee, D. D., & Agabont, S. (2003, July). Sub auditory speech recognition based on EMG signals. In Neural Networks, 2003. Proceedings of the International Joint Conference on (Vol. 4, pp. 3128-3133). IEEE.
[10] Azam, H., Hashim, N. N. W. N., Sediono, W., Mukhtar, F., Ibrahim, N., Mokhtar, S. S. S., & Aziz, S. A. (2016, December). Classifications of clinical depression detection using acoustic measures in Malay speakers. In *Biomedical Engineering and Sciences (IECBES), 2016 IEEE EMBS Conference on* (pp. 606-610). IEEE.
[11] Hashim, N. N. W. N. (2011). Analysis of power spectrum density of male speech as indicators for high risk and depressed decision (Doctoral dissertation, Vanderbilt University).