

SIGNAL-BASED FEATURE EXTRACTION FOR MAKHRAJ EMISSION POINT CLASSIFICATION

Nurul Wahidah Arshad^{1*}, Mohd Zamri Ibrahim¹, Rohana Abdul Karim¹,
Yasmin Abdul Wahab¹, Nor Farizan Zakaria¹, Tuan Sidek Tuan Muda²

¹Faculty of Electrical and Electronics Technology Engineering, Universiti Malaysia Pahang, Pekan, Pahang, Malaysia

²Centre for Human Sciences, Universiti Malaysia Pahang, Pekan, Pahang, Malaysia

*wahidah@ump.edu.my

Keywords: SPEECH PROCESSING, MAKHRAJ RECOGNITION, MAKHRAJ ARTICULATION, MFCC

Abstract

Due to the similar sound of one letter to the others, mistakes might happen when pronouncing a hijaiyah letter. The reciter will not read the Quran correctly if they do not understand the relationship between the hijaiyah letter sound and its point of articulation. This study addresses the issue to recognize the nine points of articulation (throat, uvular, molar, palatal, alveolar, dental, alveolar dental, lip, and interdental) from makhraj recitation using speech processing technique. As much as 181 non-distributive audio samples recorded in control environment. The input speech is a *sukun* combination of the *Hijaiyah* letter from an expert reciter. The research uses 5 type of signal-based feature extraction methods (MFCC, chroma, Mel spectrogram, spectral contract, and Tonnetz) and three type of classification methods (ANN, kNN, and SVM). The result shows the proposed method obtained a fair accuracy with the highest accuracy is 56% using ANN.

1 Introduction

Research in *makhraj* pronunciation is expanding, especially from engineering perspectives. *Makhraj* means having the correct position of the vocal organs to produce a *hijaiyah* alphabet sound that can be differentiated from others [1]. Correct *makhraj* is accomplished by understanding where the sound originates. *Hijaiyah*, on the other hand, is a collection of 28 alphabets for Quran recitation, each with its unique sound. To pronounce the *hijaiyah* alphabet correctly, knowing the *makhraj* is a must.

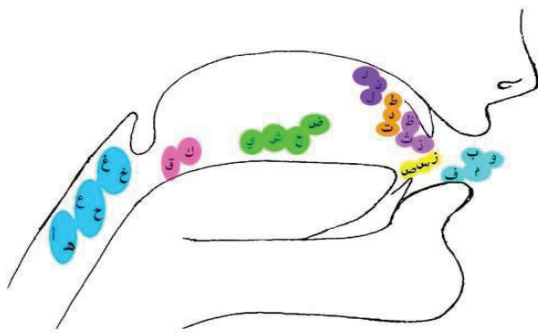


Fig. 1 The places of articulation for the *Hijaiyah* letter [2].

When pronouncing a hijaiyah letter, mistakes may occur; this can be due to a similar sound of one letter with the others, and it is difficult to pronounce correctly without some practice [3]. The letter *dal*, for example, is pronounced “*ad*” in *sukun*, while the letter *dhod* is pronounced “*adh*” in *sukun*. These two letters are pronounced differently on different *makhraj*, with *dal* originating from the dental part and *dhod* from the molar part. However, if one does not study the *makhraj* properly, confusion will arise due to their nearly identical sound. Therefore, Muslims are obligated to recite the Quran in the

correct *makhraj*. Without knowing the relationship between the *hijaiyah* letter sound with its point of articulation, the reciter will not read the Quran correctly. Fig 1 shows the point of articulation for each *hijaiyah* letter [2].

The consonantal inventory of the *hijaiyah* letter was discussed in [4], and the point of articulation was classified into eleven places. However, as shown in Table 1, nine points of articulation were investigated in this study: throat, uvular, molar, palatal, alveolar, dental, alveolar dental, lip, and interdental.

Table 1 *Hijaiyah* letter and corresponding point of articulation

Point of articulation (<i>makhraj</i>)	Letters emerging from the <i>makhraj</i>
Throat	خ, د, ح, ه, و
Uvular	ك, ق
Molar	ض
Palatal	ج, ش, ي
Alveolar	ن, ر, ل
Dental	ط, ظ, ث
Alveolar dental	ظ, ث, ذ
Interdental	ز, س, ص
Lip-bilabial	ف, و, م, ب

In this study, a system for determining the correct emission point from *makhraj* recitation using speech processing technique is developed. The input speech is a *sukun* combination of the *Hijaiyah* letter from an expert reciter. As stated in [5], this *sukun* (◌ْ) combination dataset is best to describe the *makhraj* (point of articulations) and *sifaat* (characteristics) of each letter. Mel-frequency Cepstrum Coefficient (MFCC), chroma, mel, contrast, and tonnetz were extracted from the input audio as signal-based features. In addition, the data was trained and tested using three different