

Evaluation of Suitability of Voice Reading of Al-Qur'an Verses Based on Tajwid Using Mel Frequency Cepstral Coefficients (MFCC) and Normalization of Dominant Weight (NDW)

¹Heriyanto, ²Sri Hartati, ³Agfianto Eko Putra

¹Faculty of Teknik Industry (FTI), University of Pembangunan Nasional "Veteran" University, (UPN) Tambak bayan Babarsari Yogyakarta Indonesia;

^{2,3}Departement of Computer Science and Electronics, Faculty of Mathematics and Natural Science (FMIPA), Gadjah Mada University (UGM), Sekip Bulaksumur, Yogyakarta 55281 Indonesia;
Heriyanto.yanto@mail.ugm.ac.id; shartati@ugm.ac.id; agfi@ugm.ac.id

ABSTRACT

The recitation of the Qur'an has its own uniqueness, among others having a special rule in reading and pronunciation, which is called tajwid science. At the time of the Qur'an is recited, there are often mistakes due to the limitations of knowledge of Tajwid. Therefore, the availability of tools to facilitate in checking the appropriateness of recitation is very much needed by those who recite the Qur'an and face limitations in understanding the science of tajwid. Checking the Qur'an reading is a problem that must be solved according to the rules. So far, voice identification studies have problems with feature extraction, compatibility or suitability testing, and accuracy. The issue of feature extraction, suitability, and impermanence testing have been improved in this study, which consists of two stages. The first stage is the extraction of the sound character of the Qur'an reading and the second stage is the testing of the conformity of the Qur'anic recitation and accuracy. In the first stage feature extraction is handled using MFCC and Normalization of Dominant Weight (NDW). Characteristics of reading the Qur'an as reference table is taken from one reader of Al-Qur'an who has competence in the field of science tajwid, for sampling 5-7 people as a source for testing. The process of the second stage of conformity testing of Qur'an reading is done starting from filtering, sequential multiplication of reference table and Conformity Uniformity Pattern (CUP). The sample of reading conformity test is taken from 11 Qur'anic letters containing 8 reading laws and 886 records. The test is performed on the dominant frame, the number of cepstral coefficient and the number of frames. The reading conformance test provides an average accuracy of 91.37% on the nine dominant frames. The test for the number of cepstral coefficients in the c-23 can be an average of 96.65%, while the number of frames on the F-10 is the best average of 96.65%.

Keywords: voice; reading; Al-Qur'an; MFCC; suitability; feature extraction; reference table

1 Introduction

The Qur'anic recitation sounds has its own uniqueness, among others having a special rule in reading and pronunciation by Zarkasyi [1], called tajwid science. Tajwid science is the ordinance in rightly reciting the Qur'an. At the time of the Qur'an is recited, there are often mistakes due to the limitations of knowledge of Tajwid. Therefore, the availability of tools to facilitate in checking the appropriateness of reading sounds is very much needed by those who read the Qur'an and they have limitations in understanding the science of tajwid. Correct reading of the Qur'an is a problem that requires its own

challenge, tailored to the rules. Research identification of the solution is done by testing matching or suitability and accuracy of reading. The issue of characteristic extraction, suitability testing, and accuracy were corrected in this study.

2 Literature Review

Research on voice identification is required for Voice To Text (VTT) applications. VTT research in Bahasa Indonesia by Suyanto and Hartati [2] identifies speech signals into vocabularies that include phonemes, syllables and accuracy segmentation of 98.93%. A similar study was conducted by Suyanto and Putra [3] using Mel Frequency Cepstrum Coefficient (MFCC) and Hidden Markov Model (HMM), which resulted in the introduction of phoneme segmentation in Indonesian. Another research conducted by Cahyarini et al.[4], which identifies the pause speech between phonemes that have an accuracy of 80%. Sound research in Nigerian by Yang et al. [5] identifies sayings 1 through 9, which results in 90% accuracy. Similar research involving words like "cash", "bought" and "man" was performed by Bodruzzaman et al. [6], with an accuracy of 90%. Sound research to control the robot as done by Chen et al. [5], in the form of the command words "front", "back", "right" and "left" with a level of accuracy of 80%. Robot greeting command study was also conducted by Tomasouw and Irawan [7] which resulted in a level of accuracy of 97.1%.

So far, feature extraction of the most widely used speech signals is Mel-Frequency Cepstrum Coefficients (MFCC) in both speaker recognition and speech recognition by Davis and Mermelstein [8]. Feature extraction with MFCC according to Aibinu et al. [9] has a level of accuracy of 58%. The recognition accuracy 75 % for MFCC by Hidayat [10].

To date, voice identification research has constraints ranging from feature extraction, testing and accuracy. This research is intended to improve the level of accuracy by developing the concept of extraction of Qur'anic recitations and the development of concepts for testing the appropriateness of reading the Qur'an.

3 Research Method

The evaluation of Al-Qur'an reading conformity study uses MFCC characteristic extraction and the Dominant Weighted Normalization model (NWD) in order to produce the features available in the reference reading conformity table. The reference table is derived from the source reading of its features and tested the reading of the source with the test reader. Records of recorded sources are from one expert who has competence in the field of science tajwid, while sample testers are taken from 5-7 people who are not experts. Source and tester samples were taken by recording with murotal reading. The murotal recording is a single word or two word verses of the Qur'an as an example of Al-Fatehah's letter reading one word in the first verse "Bismillah hirrohman nirrohman", the recording taken is "bismillah", "hirrohman" and "nirrohman". Taking a recording of two-word readings eg "ghoirill maghdhu bi alaihim", then the recording taken is "ghoiril magdhu". The Qur'anic letters taken are not all but samples of 11 letters containing 8 reading laws and 886 recording files. Furthermore, recording the recitation of the Qur'an is done extraction feature and testing the suitability of reading. Figure 1 shows a model for evaluating the conformity of the Qur'an reading.

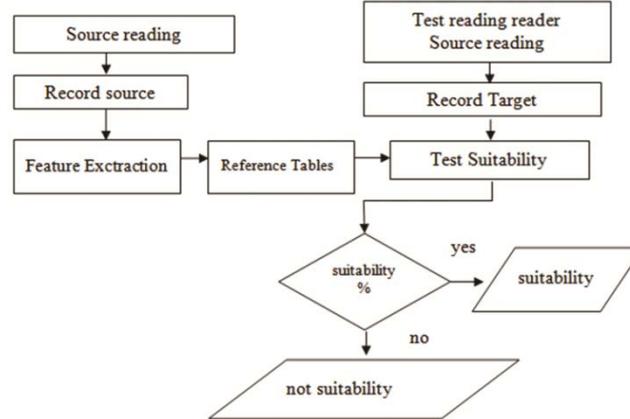


Figure 1: Model of conformity evaluation of Qu'anic reading

To conduct research the reading of Al-Qur'an readiness consists of two stages. The first stage feature extraction by using Mel Frequency Cepstrum Coefficient (MFCC) and developing the model of Normalization of Dominant Weight (NDW) to be able to produce a reference reading conformity table. In the second stage, a conformity test is performed by applying the Conformity Uniformity Pattern (CUP) to obtain a reading fit with a good degree of accuracy. Tests conducted are the dominant frame, the number of cepstral coefficient and the number of frames.

4 Feature Extraction

In the first stage feature extraction is performed to obtain the reference table using Mel Frequency Cepstrum Coefficients (MFCC) and the Dominant Weighted Normalization model (NDW). MFCC consists of 7 processes, namely preemphasing, framing blocking, windowing, Fast Fourier Transform (FFT), Mel Frequency Wrapping (MFW), Discrete Cosine Transform (DCT) and Cepstral Liftering which produces cepstral coefficient feature. Then the cepstral coefficient is processed by using the Normalization of the Dominant Weight (NDW) to obtain the reference table for the suitability of the reading.

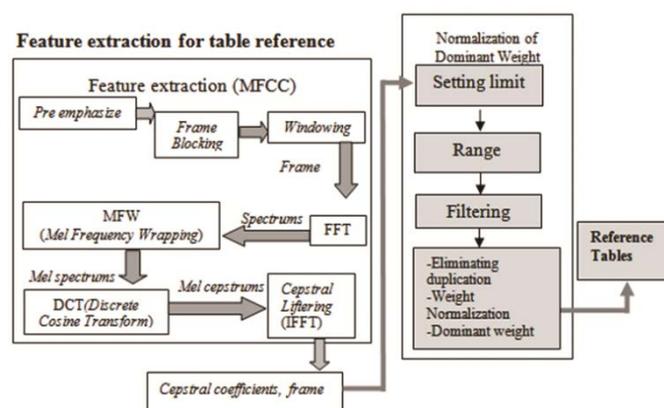


Figure 2: MFCC feature extraction section and NDW Model generate reference table

It is shown in Figure 2 that after going through the MFCC, then the NDW model is used in determining the limits, range, filtering, eliminating duplication, normalizing weights and dominant weights to generate reference tables. Here is an explanation of MFCC and NDW process stages.

4.1 Stages of MFCC

MFCC is a mel frequency and mel cepstral which produces the features of cepstral coefficients representing linear and non linear data in the form of cosine. Mel Frequency Cepstral Coefficients

(MFCCs) are a feature widely used in automatic speech and speaker recognition. They were introduced by Davis and Mermelstein in the 1980[11]. MFCC algorithm is used for voice recognition by Chakraborty [12]. Some of the advantages of this method according to Manunggal [13] is able to capture the voice characteristics that are very important for voice recognition, or in other words, it can capture important information contained in the voice signal. The MFCC stages are described as follows.

4.1.1 Preemphasis

The first stage in the MFCC process is preemphasis. Preemphasis is a noise signal filter that reduces noise, maintains a high signal. The preemphasize process is based on the equation $y(n) = s(n) - \alpha s(n - 1)$, with the value of α between 0 to 1. In general, the commonly used α values between 0.9 to 1.0 are based on following.

$$y(n) = s(n) - \alpha s(n - 1) \tag{1}$$

- Whereas,
- 1) $y(n)$: signal result of preemphasize filter,
 - 2) $s(n)$: signal before preemphasize filter,
 - 3) n : order of samples,
 - 4) α : coefficient of filter (0-1), s : signal.

4.1.2 Frame Blocking

Frame blocking is a process that divides the sound sample into multiple frames with a certain time and length. The time in the frame is taken between 10-30 milli seconds. Frame is taken as long as possible to get good frequency resolution, while shortest time may be intended to get the best time domain. The percentage of frame overlapping is between 30% - 50%. The frame blocking equation is as follows.

$$f(l; ns) = s^l(ns + ms(l - 1)) \tag{2}$$

- Whereas ,
- 1) ns : 0..N-1,
 - 2) l : 1..L,
 - 3) f : frame,
 - 4) s : signal, dan ns, ms : sampel.

The standard timing takes 25 milliseconds with a sample rate of 44.100 Hz thus the sample size is equal to 25 milliseconds to 0.025×44.100 for 1,102.5 samples. The calculation also considers 50% overlap of 1,102.5 ie 551.25 samples so the frame rate is $44.100 / (1,102,5-551,25) = 80$ frames per second

4.1.3 Windowing

Windowing is an interlocutor performed for each frame. Windowing aims to reduce the discontinuity effect at the ends of frames generated by the frame blocking process. Windowing consists of Triangular, Hamming and Hanning. The researchers apply windowing hanning because it is finer according to Agfianto [14]. The following is a representation of the window function of the input sound signal.

$$X(n) = Xi(n)w(n) \tag{3}$$

- Whereas,
- 1) n : 0,1,...,N-1,
 - 2) $X(n)$: sample signal value of windowing result,
 - 3) $Xi(n)$: the sample value of the frame signal to i ,
 - 4) $w(n)$: window function,

5) N : frame size, is a multiple of 2.

The Hanning window function is as follows :

$$w(n) = 0.5(1 - \cos \frac{2x\pi i}{M-1}) \quad (4)$$

- Whereas,
- 1) $n : 0, 1, \dots, M-1$,
 - 2) M : frame length,
 - 3) $w(n)$: function of window.

4.1.4 Fast Fourier Transform (FFT)

FFT is the process of signaling from time domain to frequency domain. Fourier is an enabling method for analysis of spectral properties that converts signals from time domain to frequency domain. FFT is a fast algorithm for implementing Discrete Fourier Transform (DFT). The DFT equation is as follows:

$$s[k] = \sum_{n=0}^{N-1} s(n) e^{-\frac{2\pi i}{N} nk} \quad 0 \leq k \leq N-1 \quad (5)$$

- Whereas,
- 1) N : number of samples to be processed,
 - 2) $s(n)$: the sample signal value,
 - 3) k : The discrete frequency variable that is valuable ($k=N/2, k \in N, n:0.1, \dots, M-1$).

4.1.5 MFW (Mel Frequency Wrapping)

MFW is a mel scale of filterbank with frequency domain. Mel frequency is linear for frequencies below 1000 Hz and is logarithmic for frequencies above 1000 Hz. The purpose of MFW is to produce a mel spectrum. The following is the equation used in the calculation of filterbanks and MFW.

$$Y[i] = \sum_{j=1}^{Nm} s[j][Hi[j]] \quad (6)$$

- Whereas,
- 1) Nm : number of magnitude spectrum ($N \in N$),
 - 2) $s[j]$: magnitude spectrum at frequency j ,
 - 3) $Hi[j]$: coefficient filterbank at frequency j ($1 \leq i \leq M$),
 - 4) M : the number of channels in the filterbank.

$$\text{Mel}(f) = 2595 \times \log_{10} (1 + f/700) \quad (7)$$

- Whereas,
- 1) Mel : ScaleMel,
 - 2) f : frequency.

4.1.6 Discrete Cosine Transform (DCT)

DCT is a correlate of mel spectrum. DCT comes from FFT which in inverse get the value of liftering. The value of the liftering is processed again The second FFT yields the value of cepstrum. DCT goals produce mel cepstrum to improve recognition quality. The equation used to calculate DCT.

$$C_n = \sum (\log S_k) \cos \left[n \left(k - \frac{1}{2} \right) \frac{\pi}{K} \right]; n = 1, 2, \dots, K \quad (8)$$

- Whereas,
- 1) C_n : Coefficient of MFCC,
 - 2) S_k : output from filterbank process on index k ,
 - 3) K : number of expected coefficient.

4.1.7 Cepstral Liftering

Cepstral liftering is an implementation of the window function against cepstral coefficient features. The formula is as follows.

$$w(ni) = 1 + \frac{L}{2} \sin\left(\frac{nix\pi}{L}\right); ni = 1,2..L \tag{9}$$

- Whereas,
- 1) $w[ni]$: window to cepstral features,
 - 2) L : number of cepstral coefficients,
 - 3) ni : index of cepstral coefficients.

Generally, the feature of cepstral coefficients for one word taken as many as 11-12 cepstral coefficients Suyanto and Hartati [2]. Al-Qur'an reading study consisting of one or two words so that $2 \times 12 = 24$ features cepstral coefficients.

Table 1. MFCC frame results and cepstral coefficients

Name	Reading	Frame	Coefficient to	Cepstral Coefficients
Barkoni-lqra01.wav	lqra	0	0	27.98
Barkoni-lqra01.wav	lqra	0	23	41.67
Barkoni-lqra01.wav	lqra	1	0	49.38
Barkoni-lqra01.wav	lqra	1	23	0.84
Barkoni-lqra01.wav	lqra	10	0	37.87
Barkoni-lqra01.wav	lqra	10	23	34.70

Table 1 is a feature extraction with MFCC in the form of cepstral coefficient and frame features. Cepstral coefficient is processed to the next stage by using the model of Normalization of Dominant Weight (NDW).

4.2 Stages of Normalization of Dominant Weight (NDW)

The Normalization Model of the Dominant Weight, hereinafter referred to as the NDW consisting of 6 stages, namely determining the limits, ranger, filtering, eliminating duplication, normalization of weight and dominant weight. The whole process is called the Normalization of Dominant Weight (NDW). The NDW model can be explained as follows.

4.2.1 Setting Limit

Limit is a line that becomes the boundary of a field, separator between two fields, one part with another part. Boundary selection is also based on the selection algorithm. The election algorithm is an algorithm to find the smallest number of k (the largest number of k) in a list. This algorithm is also called the order of statistics. The statistical order is used to determine the limits based on the features of cepstral coefficient starting from the smallest value, the largest value, the smallest value plus the largest value divided by two (median) and the mean value. The formulation is as follows.

$$\begin{aligned} Limit1 &= \min(\text{initial limit}), \\ Limit2 &= (\min + (\min + \max)/2)/2, \\ Limit3 &= \text{average}, \end{aligned}$$

$$\begin{aligned} \text{Limit4} &= (\min + \max)/2, \\ \text{Limit5} &= \left(\left(\frac{\min+\max}{2}\right) + \max\right)/2, \\ \text{Limit6} &= \max (\text{final limit}). \end{aligned} \tag{10}$$

The fetching limit becomes 6, as it uses the quarter and average system. Limit is divided into 5 parts consisting of 4 parts with quarter system and one part with average seen in Figure 3.

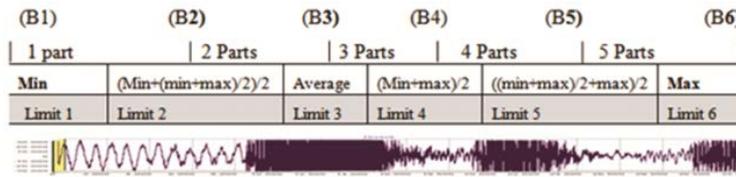


Figure 3: Illustration of limit 1 to limit 6

Limits are to be used as one of the features that exist in the reference table suitability reading. Figure 4 shows that the 1st limit is B1 up to the 6th or B6 boundary consisting of readings and frames.

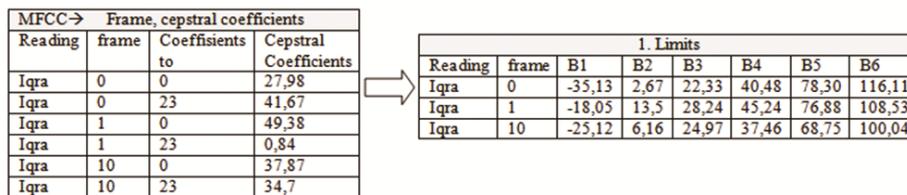


Figure 4: Setting a limit of 1 to 6

4.2.2 Setting Range

Range is also called the span in the data variant. Range is the difference between the largest value data with the smallest value. Range is taken based on previously created limits consisting of 6 limits. The determination of the range is taken from the 1st or B1 limit up to the 6th or B6 limit resulting in several conditions formed with the following equation.

1. Condition of R1 is if $((\text{limit1}) \text{ minimum value} = cc)$ then $P1\text{Weight} = 1$,
2. Condition of R2 is if $((\text{limit1}) \geq cc)$ and $(cc < \text{limit2})$ then $P2\text{Weight} = 1$,
3. Condition of R3 is if $((\text{limit2}) \geq cc)$ and $(cc < \text{limit3})$ then $P3\text{Weight} = 1$,
4. Condition of R4 is if $((\text{limit3}) \geq cc)$ and $(cc < \text{limit4})$ then $P4\text{Weight} = 1$,
5. Condition of R5 is if $((\text{limit4}) \geq cc)$ and $(cc < \text{limit5})$ then $P5\text{Weight} = 1$,
6. Condition of R6 is if $((\text{limit5}) \geq cc)$ and $(cc < \text{limit6})$ then $P6\text{Weight} = 1$,
7. Condition of R7 is if $((\text{limit6}) \text{ maximum value} = cc)$ then $P7\text{Weight} = 1$.

In this case, cc : features of cepstral coefficients.

Based on the above equation then formed into 7 conditions based on range, which then used to the filtering process.

4.2.3 Filtering

Filtering is the process of filtering. Filtering is a feature of source data based on range. Filtering itself aims to separate the features of the data in order to collect on each part. The feature collection of the data in each section is named P1, P2, P3, P4, P5, P6 and P7. Filtering process that is done if it meets the

conditions according to range, then given weight 1 on each of P1 until P7. The subsequent process is added to each of P1 to P7 and the sum is made to the whole P1 to P7 called K.P. The equation is as follows.

$$K.P = \sum_{PJ1}^{PJ7}$$

$$PJ1 = \sum_{f0}^{fm} P1$$
(12)

The equaton is also applicable to PJ2, PJ3, PJ4, PJ5, PJ6, PJ7

- Whereas,
- 1) K.P: total number of PJ1 to PJ7,
 - 2) f0: the 0th frame,
 - 3) fm: frame to-m,
 - 4) PJ1: the number of P1,
 - 5) PJ2: the number of P2,
 - 6) PJ3: the number of P3,
 - 7) PJ4: the number of P4,
 - 8) PJ5: the number of P5,
 - 9) PJ6: the number of P6,
 - 10) PJ7: the number of P7,
 - 11) P1, P2, P3, P4, P5, P6, P7: filtering results in the form of weight.

Figure 5 shows that filtering P1 through P7 with the 0th to-m frame taking turns there is duplication of one reading to another.

3. Filtering									
Reading	frame	K.P	P1	P2	P3	P4	P5	P6	P7
Iqra	0	903	1	306	202	306	76	12	0
Iqra	1	825	1	303	140	306	63	0	0
Iqra	10	889	0	304	174	306	84	0	0

2. range	
If (B1) =	then P1_weight = 1
If (B2) >= cc	<(B3) then P2_weight = 1
If (B3) >= cc	<(B4) then P3_weight = 1

Figure 5: Result of filtering at P1 through P7

The next process is processing by eliminating duplication, normalization of weights and dominant weights so as to produce a reference table of reading conformity. Explanation of the process as follows.

4.2.4 Eleminating Duplication

The process of eliminating duplication by means of feature selection of duplicated data in each frame and readings ranging from P1 to P7 to be omitted or to 0. Figure 6 shows that there is a lot of duplication in P1 almost entirely. P2 is not entirely duplicated. Duplication process starts from P1 to P7 in each frame and reads in a way compared to one overall.

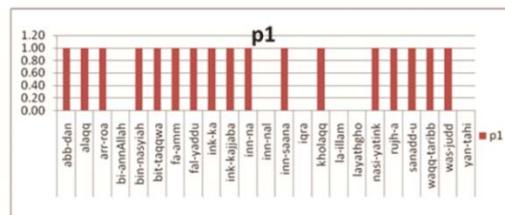


Figure 6: The overall P1 is duplicated

In Figure 6 that the P1 in frame 0 looks the same spread even if it is entirely the same starting from "abb / and" and "yan / tahi" so that the deletion is done to 0. Figure 7 shows that there is a duplication of P2 so that it is eliminated to 0. Based on the selection it eliminates the duplication of features obtained from data P2, P3, P5 and P6, whereas P1, P4 and P7 are removed because of duplication.

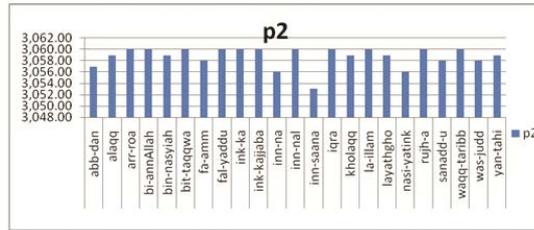


Figure 7:P2 is partly duplicated

Figure 8 shows that there is duplication so that it is eliminated and P2 becomes PF2, P3 becomes PF3, P5 becomes PF5 and P6 becomes PF6.

3. Filtering									
Reading	frame	K.P	P1	P2	P3	P4	P5	P6	P7
Iqra	0	903	1	306	202	306	76	12	0
Iqra	1	825	1	303	140	306	63	0	0
Iqra	10	889	0	304	174	306	84	0	0

4. Eliminating duplication									
Reading	frame	K.P	pf1	pf2	pf3	pf4	pf5	pf6	pf7
Iqra	0	903	0	0	0	0	76	12	0
Iqra	1	825	0	0	140	0	63	0	0
Iqra	10	889	0	0	174	0	84	0	0
amount		9348	0	0	1.268	0	636	80	0

Figure 8:Eliminates duplicate features from the data

Eliminating the duplication can not be a reference table of reading suitability because there is still a possibility of similarity in the process of calculation. The next process is then carried out the process of normalization.

4.2.5 Weight Normalization

The notion of normalization comes from E. F.Codd [15], one of the pioneers of database technology about relational data structure. Normalization is a technique for organizing data into tables to meet the needs of users within an organization, eliminate duplicate data by Jogiyanto [15]. The goal of normalization is to eliminate duplicate data, reduce complexity and facilitate data modification. The normalization process is normalization of weights. Normalization of weights is done by making variable amounts in advance of the variable then the number becomes a divisor. Variable sum is the sum of from each pf can be seen in Figure 9.

4. Eliminating duplication									
Reading	frame	K.P	pf1	pf2	pf3	pf4	pf5	pf6	pf7
Iqra	0	903	0	0	0	0	76	12	0
Iqra	1	825	0	0	140	0	63	0	0
Iqra	10	889	0	0	174	0	84	0	0
amount		9348	0	0	1.268	0	636	80	0

Variable amount					
Reading	frame	K.P	PB2	PB3	PB6
Iqra	0	903,00	0	1610361	405769
Iqra	1	825,00	0	1610361	405769
Iqra	10	889,00	0	1610361	405769

$PB2^2 = \sum_{pf=0}^6 pf^2$

Figure 9:Variable amount

Variable amount (PB) is done after completion through the process of eliminating duplication. The equation is as follows.

$$PB2^2 = \sum_{f0}^{fp} pf2 \tag{13}$$

The equation is also applicable to $PB3^2, PB5^2, PB6^2$

- Whereas,
- 1) $PB2$: variable amount $pf2$,
 - 2) $PB3$: variable amount $pf3$,
 - 3) $PB5$: variable amount $pf5$,
 - 4) $PB6$: variable amount $pf6$,
 - 5) $f0$: frame 0, fp : frame p .

If the variable amount (PB) has been obtained, then the process continues to normalization. The normalization used is the normalization of weights because the features of the filtered P2, P3, P5 and P6 data meet a given range given the weights so that it becomes $pf2, pf3, pf5$ and $pf6$. Next, we calculate the normalization of weights by calculating each $pf2$ in the frame divided by the number of $PB2$. The results of $pf2, pf3, pf5$ and $pf6$ have normalized weights to $npf2, npf3, npf5$ and $npf6$. The normalization equation weights as follows.

$$npf2 = \sum_{f0}^{fk} pf2 / PB2 \tag{14}$$

The equation is also applicable to $npf3, npf5, npf6$

- Whereas,
- 1) $npf2$: number of normalized weights $pf2, pf2$: duplicated filter results P2,
 - 2) $npf3$: number of normalized weights $pf3, pf3$: duplicated filter results P3,
 - 3) $npf5$: number of normalized weights $pf5, pf5$: duplicated filter results P5,
 - 4) $npf6$: number of normalized weights $pf6, pf6$: duplicated filter results P6,
 - 5) $f0$: frame 0, fk : frame k .

The following normalized weights can be seen in Table 2 on each number of features of the pre-pf comparison data and features of the npf data after normalization of weights.

Table 2:Result of normalization of weights

5.Normalization of weights										
Reading	frame	K.P	Before Normalization				After Normalization			
			pf2	pf3	pf5	pf6	npf2	npf3	npf5	npf6
lqra	0	903	0	0	76	12	0	0	0.11	0.14
lqra	1	825	0	140	63	0	0	0.11	0.09	0
lqra	2	895	0	190	0	0	0	0.14	0	0
lqra	3	848	0	158	68	12	0	0.12	0.10	0.14
lqra	4	821	0	140	59	12	0	0.11	0.09	0.14
lqra	5	826	0	149	58	9	0	0.11	0.09	0.11
lqra	6	811	0	145	50	6	0	0.11	0.07	0.07
lqra	7	858	0	172	64	11	0	0.13	0.10	0.13
lqra	8	802	0	0	50	10	0	0	0.07	0.12
lqra	9	870	0	0	64	8	0	0	0.10	0.09
lqra	10	889	0	174	84	0	0	0.13	0.13	0
amount		9348	0	1269	637	81	0	0.99	1.00	0.99

Normalization of weights also can not be a reference table of reading suitability so it needs to be done next process which is the selection of dominant weight. Selection of dominant weights aims to

determine which areas are the most dominant or superior to be used as reference table reading suitability. Explanation of the dominant weight process as follows.

4.2.6 Dominant Weight

The dominant weight is the weight that becomes dominant or superior or prominent. The selection of features of the data in the form of dominant weight is taken by taking the highest in each frame in npf2, npf3, npf5 and npf6 being npf2d, npf3d, npf5d and npf6d as the dominant weights. The dominant taking is taken on each frame and readings. Table 3 comparison before and after taken dominant weight.

Table 3:Results of dominant weight

Surat Al-Alaq													
Frame	Reading	Before dominant weight				After dominant weight							
		npf2	npf3	npf5	npf6	2-dn	npf2d	3-dn	npf3d	5-dn	npf5d	6-dn	npf6d
0	lqra	0	0	0.11	0.14	5	0	10	0	2	0.11	1	0.14
1	lqra	0	0.11	0.09	0	4	0	7	0.11	6	0.09	11	0
2	lqra	0	0.14	0	0	2	0	1	0.14	11	0	9	0
3	lqra	0	0.12	0.10	0.14	11	0	4	0.12	3	0.10	3	0.14
4	lqra	0	0.11	0.09	0.14	10	0	8	0.11	7	0.09	2	0.14
5	lqra	0	0.11	0.09	0.11	9	0	5	0.11	8	0.09	6	0.11
6	lqra	0	0.11	0.07	0.07	8	0	6	0.11	10	0.07	8	0.07
7	lqra	0	0.13	0.10	0.13	7	0	3	0.13	5	0.10	4	0.13
8	lqra	0	0	0.07	0.12	6	0	11	0	9	0.07	5	0.12
9	lqra	0	0	0.10	0.09	1	0	9	0	4	0.10	7	0.09
10	lqra	0	0.13	0.13	0	3	0	2	0.13	1	0.13	10	0

One example of feature retrieval from the dominant weight data in the npf6d column is 0.14. The columns contain frame 0, frame 3 and frame 4 and are ranked dominantly with the names of 6-dn ranks 1, 2 and 3. The difference lies in before using the dominant weights on npf2, until npf6 is retrieved all there is no dominant rank on each frame, while after using the feature of the dominant weight data there is a dominant rank on each frame. The dominant frame is searched by testing up to how many dominant frames to be the reference tables reading suitability. The reference table of conformity can be seen in Tables 4, 5 and 6.

Table 4:Result of limit reference table

Reference tables of Limit									
Reading	Frame	Reading Laws	K.P	B1	B2	B3	B4	B5	B6
lqra	0	Qalqalah	903	-35.13	2.67	22.33	40.48	78.30	116.11
lqra	1	Qalqalah	825	-18.05	13.5	28.24	45.24	76.88	108.53
lqra	10	Qalqalah	889	-25.12	6.16	24.97	37.46	68.75	100.04
		CUP	0.4/0.5						

Table 5:Result of variable amount reference table

Reference tables of <i>variable</i> amount						
Reading	frame	K.P	PB2	PB3	PB5	PB6
lqra	0	903.00	0	1610361	405769	6561
lqra	1	825.00	0	1610361	405769	6561
lqra	10	889.00	0	1610361	405769	6561

Table 6:Results of dominant weight table reference

Reference tables of dominant weight									
Frame	Reading	2-dn	npf2d	3-dn	npf3d	5-dn	npf5d	6-dn	npf6d
0	lqra	5	0	10	0	2	0.11	1	0.14
1	lqra	4	0	7	0.11	6	0.09	11	0
10	lqra	3	0	2	0.13	1	0.13	10	0

The reference table consists of limits B1 to B6, CUP, overall P (K.P), variable amount PB2 through PB6, and Normalization of Dominant Weight (NDW) on each frame and reading.

After yielding reference table readiness appropriateness obtained then processed to next step that is test to input reader test and also input reading source.

5 Testing Stage

In the process of testing the conformity stages of reading the Qur'an compare the existing features in the reference table with the features available on the testers reading. Testing is not only on the testers' reading but also on the source reading. Testing stages starting from the recording of test readers then performed the MFCC process to get the features cepstral coefficient. The cepstral coefficient feature is processed by first filtering into the reference table in the form of a 1st to 6th feature limit. After going through the filtering process with the range coming from the boundary feature, then the next process performs sequential multiplication count. The calculation of the reference table in the form of feature variable number and feature of dominant weight. After the sequential multiplication counting process is obtained, then the final count calculation is the Conformity Uniformity Pattern (CUP) formula. Tests for the readiness of the Qur'an recitation are sampled against 11 Qur'anic letters, 8 reading codes and 886 recording files. Stages of testing steps are carried out as follows.

5.1 Input Testers' Readings

The reading test is done by taking the test reader input of 5-7 people on 11 letters containing 8 reading laws.

5.2 Records of Testers

Taking sound recordings of testers as many as 886 recording files. The recording of the reading is extracted with MFCC which produces the cepstral coefficient feature, which is matched to the corresponding reading to the reference table.

5.3 MFCC Readers Test

The subsequent reading sound of the testers using MFCC produces the frame and cepstral coefficient features which are then performed to the conformity testing stage.

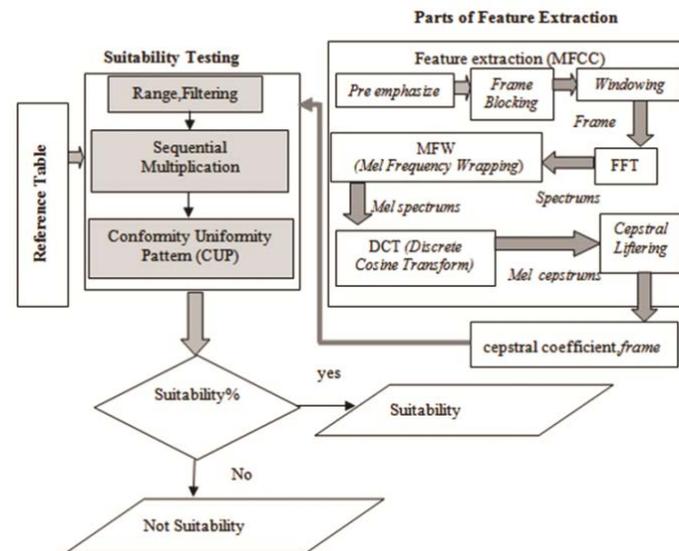


Figure 10:Section of extraction and conformity testing

The reading suitability test as shown in Figure 10 begins with range, filtering, sequential multiplication and the Conformity Uniformity Pattern (CUP). The testing process is carried out on the features in the reference reference reading table. Explanation of process range and filtering as follows.

5.4 Range and Filtering Testers

Filtering is the process of filtering testers with a certain range. Screening of test data readings starting from MFCC which produces cepstral coefficient feature then tested against the existing features in the reference table in the form of a boundary that forms the range with if it meets the conditions then given the weight equal to one. The equation is as follows.

1. Condition of R1 is if (limit1) minimum value = cc then $P1WeightTest = 1$,
2. Condition of R2 is if $((limit1) \geq cc) \text{ and } (cc < limit2)$ then $P2WeightTest = 1$,
3. Condition of R3 is if $((limit2) \geq cc) \text{ and } (cc < limit3)$ then $P3WeightTest = 1$,
4. Condition of R4 is if $((limit3) \geq cc) \text{ and } (cc < limit4)$ then $P4WeightTest = 1$,
5. Condition of R5 is if $((limit4) \geq cc) \text{ and } (cc < limit5)$ then $P5WeightTest = 1$,
6. Condition of R6 is if $((limit5) \geq cc) \text{ and } (cc < limit6)$ then $P6WeightTest = 1$,
7. Condition of R7 is if (limit6) maximum value = cc then $P7WeightTest = 1$. (15)

In this case, cc: features of cepstral coefficients.

Checking filtering is done based on existing features in the reference table is the limit and range features.

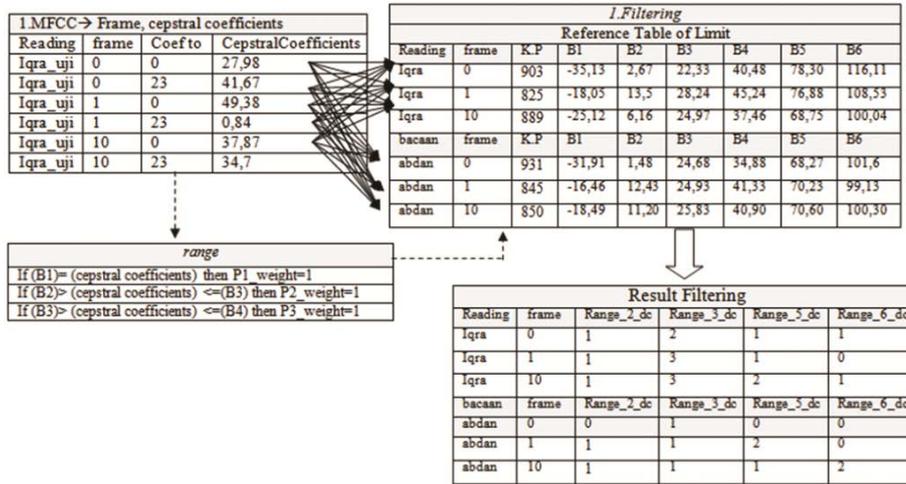


Figure 11:Range and filtering

The reference table used is the limit from B1 to B6. Figure 11 shows that the features of the frame and cepstral coefficient are skewed with the filtering features present in the limit reference table. The matching processes that are based on a range that, when fulfilled, are weighted and accumulated are stored in the range_2_dc, range_3_dc, range_5_dc and range_6_dc. The sequential multiplication process is described as follows.

5.5 Sequential Multiplication

The sequential multiplication is the process of multiplying the reading of the testers by the existing reading of the features in the reference table of the suitability of the sequential reading. The equation is as follows.

$$P2uji = \sum_{f_0}^{fr} (range_2_dc \times npf2d \times PB2^2) \tag{16}$$

This equation is also applicable to P3uji, P5uji, P6uji

- Whereas,
- 1) P2Uji: the number of P2 test,,
 - 2) P3Uji: the number of P3 test,
 - 3) P5Uji: the number of P5 test,
 - 4) P6Uji: the number of P6 test,
 - 5) range_2_dc: filter range P2,
 - 6) range_3_dc: filter range P3,
 - 7) range_5_dc: filter range P5,
 - 8) range_6_dc: filter range P6,
 - 9) npf2d: normalization of dominant weight P2,
 - 10) npf3d: normalization of dominant weight P3,
 - 11) npf5d: normalization of dominant weight P5,
 - 12) npf6d: normalization of dominant weight P6,
 - 13) PB2: variable amount of P2,
 - 14) PB3: variable amount of P3,
 - 15) PB5: variable amount of P5,

16) PB6: variable amount of P6.

On the equation above, it is done one by one process of filtering result sequentially on the feature of variable amount and also to feature of reference table that is feature of dominant weight as shown in Figure 12.

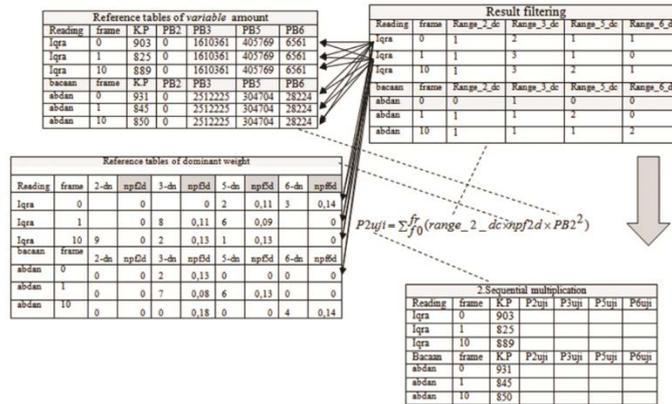


Figure 12: The sequential multiplication process of the test's reader

The sequential multiplication calculation is then processed into the Conformity Uniformity Pattern (CUP). The calculation phase of the CUP is described as follows.

5.6 Conformity Uniformity Pattern(CUP)

Conformity Uniformity Pattern (CUP) is the multiplication of the average number of P2Uji, P3Uji, P5Uji, P6Uji with the average P total (K.P) and divides by the average variable of the number raised. The reading suitability count with Conformity Uniformity Pattern(CUP) as follows.

$$CUPP2 = \frac{\overline{(P2uji)} \times \overline{(K.P)}}{\overline{(PB2^2)}} \quad (17)$$

The equation is also applicable to $CUPP2, CUPP3, CUPP5, CUPP6$

$$CUPP2P3P5P6 = \frac{\overline{(P2uji)} + \overline{(P3uji)} + \overline{(P5uji)} + \overline{(P6uji)} \times \overline{(K.P)}}{\overline{(PB2^2)} + \overline{(PB3^2)} + \overline{(PB5^2)} + \overline{(PB6^2)}}$$

- Whereas,
- 1) CUPP2: Conformity Uniformity Pattern P2,
 - 2) CUPP3: Conformity Uniformity Pattern P3,
 - 3) CUPP5: Conformity Uniformity Pattern P5,
 - 4) CUPP6: Conformity Uniformity Pattern P6,
 - 5) P2Uji: P2 test,
 - 6) P3Uji: P3 test,
 - 7) P5Uji: P5 test,
 - 8) P6Uji: P6 test,
 - 9) PB2: variable amount of P2,
 - 10) PB3: variable amount of P3,
 - 11) PB5: variable amount of P5,

12) PB6: variable amount of P6,

13) K.P: overall of P (P2, P3, P5, P6).

The CUP calculations on the features contained in the reference table are performed to measure how well the resource reads the testers' reading. The test is also conducted by measuring the level of suitability level of reading that the CUP in can from the existing features in the reference table approaching between 0.4 to 0.5 as in Table 7.

Table 7: Results of Al-Alaq's letter of CUP

frame	Reading	K.P	P1	P2	P3	P4	P5	P6
0	abbdan	931.00	1.00	306.00	210.00	306.00	83.00	24.00
1	abbdan	845.00	0.00	304.00	139.00	306.00	72.00	24.00
2	abbdan	847.00	0.00	303.00	137.00	306.00	73.00	28.00
3	abbdan	877.00	0.00	304.00	169.00	306.00	77.00	21.00
4	abbdan	887.00	0.00	304.00	169.00	306.00	80.00	28.00
5	abbdan	873.00	0.00	304.00	154.00	306.00	78.00	31.00
6	abbdan	878.00	0.00	304.00	157.00	306.00	80.00	31.00
7	abbdan	832.00	0.00	303.00	134.00	306.00	68.00	21.00
8	abbdan	906.00	0.00	304.00	179.00	306.00	83.00	34.00
9	abbdan	1,041.00	0.00	306.00	290.00	306.00	103.00	36.00
10	abbdan	850.00	0.00	304.00	144.00	306.00	72.00	24.00
	amount	9,767.00	1.00	3,346.00	1,882.00	3,366.00	869.00	302.00
	Average	887.91		304.18	171.09	306.00	79.00	27.45
	square			11,195,716.00	3,541,924.00		755,161.00	91,204.00
	CUP	0.43/0.5		0.02	0.04		0.09	0.27

CUP between 0.4 up to 0.5 is used as reference table can be done base to do level of reading suitability testing. Tests of reading confidence with CUP can be seen in Figure 13 after the sequential multiplication in P2uji, P3uji, P5uji and P6uji are included into the Conformity Uniformity Pattern (CUP). The results of the CUP seen highest between 0.5 and 0.4 are dominant and superior, so it looks appropriate or not according to the existing reading on the reference table. CUP test with the result 0.4 looks more in accordance with the existing count on the features contained in the reference table so that the reading tested by reference is the sound of "iqra" while in reading "abbdan" the CUP result is not close to 0.4 or 0.5 so it is not according to the reading.

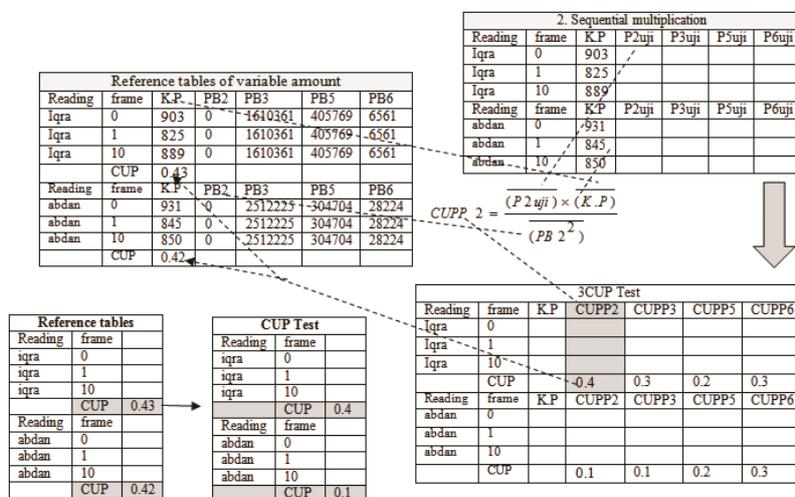


Figure 13: Counting the CUP process

CUP test taken highest 0.4 closer to the reference table of "iqra" sound, while the reading of "abbdan" at CUP Test is taken the highest 0.3 is also not included between 0.4 and 0.5

6 Test Result

The research was conducted on 11 letters, 8 reading law, and 886 recording files. The first test conducted is the test of the dominant frame consisting of 11 frames. The second test is performed on the number of cepstral coefficient, and the third test is carried out on the number of frames.

6.1 Result of Frame Dominant Testing

This test is done to find the most dominant frame area between frame 0 to 10th frame. The dominant frame test was obtained with an accuracy rate of 91.37% on the dominant 9 frames, while the dominant frame 5 (5-dm) up to 1 (1-dm) frame decreased as shown in Table 8.

Table 8: Results of Quranic reading test

No	Letter	1-dm	2-dm	3-dm	4-dm	5-dm	6-dm	7-dm	8-dm	9-dm	10-dm	11-dm
1	Al-Fatehah	21.43	37.01	67.53	77.92	87.01	89.61	90.91	90.91	90.91	90.91	90.26
2	Al-Baqarah	43.48	72.46	75.36	82.61	82.61	86.96	86.96	86.96	86.96	86.96	85.51
3	Al-Imran	41.67	65.28	81.94	81.94	83.33	84.72	84.72	84.72	87.50	87.50	87.50
4	Ar-rohman	53.85	71.79	73.08	82.05	84.62	83.33	82.05	83.33	84.62	84.62	84.62
5	Al-Hadid	51.52	65.15	68,18	78.79	84.85	84.85	78.79	86.36	86.36	86.36	86.36
6	Al-Alaq	42.75	62.32	69.57	84.78	93.48	95.65	95.65	96.38	95.65	96.38	96.38
7	Al-Ashr	56.36	65.45	69.09	70.91	74.55	72.73	92.73	94.55	100.00	100.00	100.00
8	Al-Kautsar	51.85	83.33	81.48	83.33	83.33	83.33	100.00	98.15	100.00	98.15	100.00
9	Al-Ikhlas	58.33	53.33	66.67	83.33	81.67	83.33	86.67	91.67	93.33	91.67	91.67
10	Al-Falaq	58.67	74.67	80.00	81.33	90.67	92.00	93.33	93.33	92.00	92.00	93.33
11	An-nas	63.08	73.85	84.62	84.62	83.08	87.69	84.62	87.69	87.69	86.15	86.15
	Average	49.36	65.88	74.32	81.06	84.47	85.84	88.77	90.37	91.37	90.97	91.07

Description dm = dominant

Testing the accuracy value against the dominant frame of 11 frames indicates that there are 9 dominant frames (9-dm) with good accuracy value this is seen as in Table 9.

Table 9: Results of 9-dm dominant reading tests

NO	Letter	Number	L	P	Juz	Person	Record	Average (9-dm)	amount Record
	Al-Qur'an	Letter							
1	Al-Fatehah	1	3	4	1	7	22	90.91	154
2	Al-Baqarah	2	3	0	1	3	23	86.96	69
3	Al-Imran	3	3	0	1	3	24	87.50	72
4	Ar-rohman	55	3	0	27	3	26	84.62	78
5	Al-Hadid	57	3	0	27	3	22	86.36	66
6	Al-Alaq	96	3	3	30	6	23	95.65	138
7	Al-Ashr	103	3	2	30	5	11	100.00	55
8	Al-Kautsar	108	3	3	30	6	9	100.00	54
9	Al-Ikhlas	112	2	3	30	5	12	93.33	60
10	Al-Falaq	113	2	3	30	5	15	92.00	75
11	An-nas	114	3	2	30	5	13	87.69	65
	amount					51	200	91.37%	886

Testing of 11 letters indicates that the dominant 5-dm down is getting smaller or its accuracy decreases.

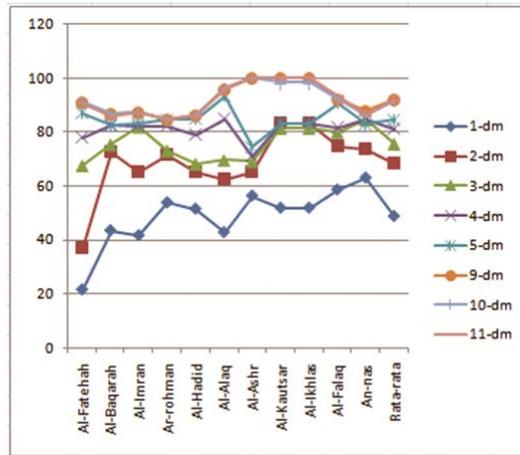


Figure 14: Comparison of the best dominant frames of 11 letters

Figure 14 shows that the dominant taking is preferably above 5-dm as it has an increased accuracy rate.

6.2 Test Results Number of Cepstral Coefficient

Tests were also performed on the amount of cepstral coefficient. Taking the number of cepstral coefficient with c-23 shows a good accuracy level with an average of 96.65%.

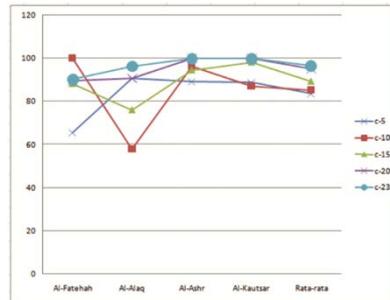


Figure 15: Comparison of the best amount of cepstral coefficient

Figure 15 shows that the samples of the five letters of Al-Fatehah, Al-Alaq, Al-Ash, Al-Kautsar show the highest on cepstral coefficient 23 or c-23 with an average of 96.65%.

6.3 Test Results Number Frame

Tests on the number of frames also performed starting from the frames F-5, F-7, F-9 and F-10. The test was conducted to find the number of frames that have the best accuracy on F-10 with an average of 96.65%. Figure 16 shows that the F-10 frame has the highest average compared to F-5, F-7 and F-9.

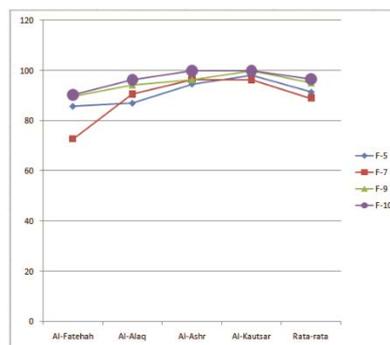


Figure 16: Comparison of the best frame counts

Based on the CUP testing, the dominant frame rate is found in the dominant 9 frames, the test for the number of cepstral coefficients in the can be c-23 with an average of 96.65%, while the best frame number in F-10 with an average of 96.65 %. Tests with the CUP formula can also determine the suitability of reading uniformity, which is contained in the reference table with CUP 0.4 - 0.5. Taking the source reading before the reference can be known feasibility if less than 90% of the source reading test should be re-recorded and re-reference.

7 Conclusion

Correct reading of the Qur'an is a problem that requires its own challenge, tailored to the rules. Research on reading conformity is done with the features in the reference table of reading fit, test and readability accuracy. Tests with the CUP calculations performed on dominant frames resulted in an average accuracy rate of 91.37% at 9 frames (9-dm). The test was also conducted on the number of cepstral coefficient to produce the best cepstral coefficient c-23 with an average of 96.65%, while testing the number of frames yielded the best F-10 with an average of 96.65%. This research yields some contribution that is:

1. The first contribution that the model extraction concept extraction uses MFCC and NDW model can be used to generate the reference table of Al-Qur'an reading conformity with the accuracy level of 91.37%. The retrieval of the reference table on the dominant frames suggested above 5 (5-dm).
2. The second contribution to the test model based on the level or level of reading conformity with the Conformity Uniformity Pattern (CUP) of 0.4 with the high level to 0.5 with the low level can know the ability of reading.

The third contribution in the process of testing the source reading before the reference table can be known the feasibility of reading the source is declared in accordance uniformity to be a reference table.

REFERENCES

- [1] Zarkasyi I, "Pelajaran Tajwid," *Gontor Ponorogo*, vol. Trimurti P, p. Hal 1-3, 1995.
- [2] Suyanto and S. Hartati, "Design of Indonesian LVCSR using Combined Phoneme The Approaches of LVCSR," *Icts*, pp. 191-196, 2013.
- [3] S. Suyanto and A. E. Putra, "Automatic Segmentation of Indonesian Speech into Syllables using Fuzzy Smoothed Energy Contour with Local Normalization, Splitting, and Assimilation," *J. ICT Res. Appl.*, vol. 8, no. 2, pp. 97-112, 2014.
- [4] R. Cahyarini, U. L. Yuhana, and A. Munif, "Rancang Bangun Modul Pengenalan Suara Menggunakan Teknologi Kinect," *J. Tek. Pomits*, vol. 2, no. 1, pp. 1-5, 2013.
- [5] M. L. Chen, S. K. Changchien, X. M. Zhang, and H. C. Yang, "The design of voice recognition controller via grey relational analysis," *Proc. 2011 Int. Conf. Syst. Sci. Eng. ICSSSE 2011*, no. June, pp. 477-481, 2011.
- [6] M. Bodruzzaman, K. Kuah, T. Jamil, C. Wang, and X. Li, "System Using Artificial Neural Network," pp. 1-3, 1993.
- [7] B. P. Tomasouw and M. I. Irawan, "Multiclass Twin Bounded Support Vector Machine Untuk Pengenalan Ucapan," *Pros. Semin. Nas. Penelitian, Pendidik. dan Penerapan MIPA, Fak. MIPA, Univ. Negeri Yogyakarta*, vol. 2, no. 2004, pp. 1-10, 2012.
- [8] M. A. Pathak, "Privacy-Preserving Machine Learning for Speech Processing," *PhD*, pp. 1-140, 2012.

- [9] A. M. Aibinu, M. J. E. Salami, A. R. Najeeb, J. F. Azeez, and S. M. A. K. Rajin, "Evaluating the effect of voice activity detection in isolated Yoruba word recognition system," *2011 4th Int. Conf. Mechatronics Integr. Eng. Ind. Soc. Dev. ICOM'11 - Conf. Proc.*, no. May, pp. 17–19, 2011.
- [10] S. Hidayat, R. Hidayat, and T. B. Adji, "Sistem Pengenal Tutur Bahasa Indonesia Berbasis Suku Kata Menggunakan MFCC, Wavelet Dan HMM," *Conf. Inf. Technol. Electr. Eng.*, no. September, pp. 246–251, 2015.
- [11] S. B. Davis and P. Mermelstein, "Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences," *IEEE Trans. Acoust.*, vol. 28, no. 4, pp. 357–366, 1980.
- [12] K. Chakraborty Asmita Talele Savitha Upadhya, "Voice Recognition Using MFCC Algorithm," *Int. J. Innov. Res. Adv. Eng.*, vol. 1, no. 10, pp. 2349–2163, 2014.
- [13] H. S. Manunggal, "Perancangan dan Pembuatan Perangkat Lunak Pengenalan Suara Pembicara Dengan Menggunakan Analisa MFCC Feature Extraction.," *Tugas Akhir Sarj. pada Jur. Tek. Inform. Fak. Teknol. Ind. Univ. Kristen Petra Surabaya*, 2005.
- [14] A. E. Putra, "Frekuensi Cuplik pada FFT," *Tan Li, Process. Digit. Signal*, vol. 1, 2008.
- [15] J. E.F Codd, P. Ritonga, and L. A. Reply, "Pengertian Normalisasi Database Dan Bentuk-," pp. 3–5, 2015.