

EKSTRAKSI CIRI MEL FREQUENCY CEPSTRAL COEFFICIENT (MFCC) DAN RERATA COEFFICIENT UNTUK PENGECEKAN BACAAN AL-QUR'AN

Heriyanto⁽¹⁾, Sri Hartati⁽²⁾, Agfianto Eko Putra⁽³⁾

Fakultas Teknik Industri, Universitas Pembangunan Nasional "Veteran" (UPN) Yogyakarta⁽¹⁾,
Departemen Ilmu Komputer dan Elektronika, Fakultas Matematika dan Ilmu Pengetahuan Alam,
Universitas Gadjah Mada, Yogyakarta^(2,3)

heriyanto.yanto@mail.ugm.ac.id⁽¹⁾, shartati@ugm.ac.id⁽²⁾, agfi@ugm.ac.id⁽³⁾

Abstrak

Belajar membaca Al-Qur'an menggunakan alat bantu aplikasi sangat diperlukan dalam mempermudah dan memahami bacaan Al-Qur'an. Pengecekan bacaan Al-Qur'an salah satu metode dengan MFCC untuk pengenalan suara cukup baik dalam *speech recognition*. Metode tersebut telah lama diperkenalkan oleh Davis dan Mermelstein sekitar tahun 1980. MFCC merupakan metode ekstraksi ciri untuk mendapatkan *cepstral coefficient* dan *frame* sehingga dapat digunakan untuk pemrosesan pengenalan suara agar lebih baik dalam ketepatan. Tahapan MFCC mulai dari *pre-emphasis*, *frame blocking*, *windowing*, *Fast Fourier Transform* (FFT), *Mel Frequency Wrapping* (MFW), *Discrete Cosine Transform* (DCT) dan *cepstral liftreing*. Hasil pengecekan bacaan Al-Qur'an diujikan dalam sebelas surat mulai dari surat Al-Fatihah, Al-Baqarah, Al-Imran, Al-Hadid, Al-Ashr, Ar-rahman, Al-Alaq, Al-Kautsar, Al-Ikhlash, Al-Falaq dan An-Nas menghasilkan akurasi sebesar rata-rata 51,8%.

Kata Kunci : Suara, Bacaan, MFCC, Kesesuaian, Ekstraksi Ciri, Referensi, Bobot, Dominan.

1. PENDAHULUAN

Belajar membaca Al-Qur'an menggunakan alat bantu aplikasi sangat diperlukan dalam mempermudah dan memahami bacaan Al-Qur'an. Pengecekan bacaan Al-Qur'an salah satu metode dengan MFCC untuk pengenalan suara cukup baik dalam *speech recognition*. MFCC merupakan metode ekstraksi ciri untuk mendapatkan *cepstral coefficient* dan *frame* sehingga dapat digunakan untuk pemrosesan pengenalan suara agar lebih baik dalam ketepatan. Tahapan MFCC mulai dari *pre-emphasis*, *frame blocking*, *windowing*, *Fast Fourier Transform* (FFT), *Mel Frequency Wrapping* (MFW), *Discrete Cosine Transform* (DCT) dan *cepstral liftreing*. Hasil pengecekan bacaan Al-Qur'an diujikan dalam sebelas surat mulai dari surat Al-Fatihah, Al-Baqarah, Al-Imran, Al-Hadid, Al-Ashr, Ar-rahman, Al-Alaq, Al-Kautsar, Al-Ikhlash, Al-Falaq dan An-Nas.

2. TINJAUAN PUSTAKA

Metode MFCC diperkenalkan pertama kali oleh Davis dan Mermelstein sekitar tahun 1980. MFCC salah satu metode yang cukup baik dalam pengenalan suara dalam bidang *speech recognition* (Davis dan Mermelstein, 1980). MFCC merupakan ekstraksi ciri yang paling banyak digunakan dalam bidang *speaker recognition* dan *speech recognition*.

MFCC merupakan ekstraksi ciri yang menghasilkan fitur berupa parameter *cepstral coefficient* (Abriyono dan Harjoko, 2012). Ekstraksi ciri *Mel Frequency Cepstral Coefficient* (MFCC) mengubah gelombang suara menjadi beberapa tipe parameter seperti *cepstral coefficient* yang mempresentasikan *audio file* (Chamidy, 2016). Selain itu MFCC menghasilkan fitur vektor yang mengonversi sinyal suara menjadi beberapa vektor untuk pengenalan fitur suara (Putra dkk., 2011).

MFCC mempunyai tahapan yaitu *pre-emphasis*, *frame blocking*, *windowing*, *Fast Fourier Transform* (FFT), *Mel Frequency Wrapping* (MFW), *Discrete Cosinus Transform* (DCT) dan

cepstral liftering yang menghasilkan parameter sebagai fitur yaitu *frame* dan *cepstral coefficient* (Putra, 2011).

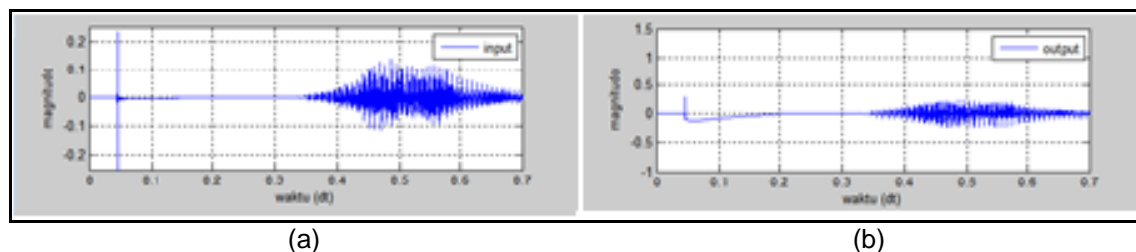
2.1 Pre-emphasis

Pre-emphasis menurut Chitode (2010) melakukan penekanan pada komponen frekuensi tinggi, menyelaraskan pada frekuensi rendah dan frekuensi tinggi. *Pre-emphasis* menurut Tokunbo Ogunfunmi (2015) mengurangi kebisingan untuk memperbaiki *Signal to Noise Rasio* (SNR) dan mengurangi suara yang tidak diinginkan. *Pre-emphasis* merupakan pemrosesan signal sederhana yang pada dasarnya adalah *filter linear* dan masih dalam ranah waktu (Holmes, 2003).

Pre-emphasis merupakan tahap awal dalam proses MFCC. Tahap ini, dilakukan karena sinyal sering sekali mengalami gangguan *noise*, sehingga diperlukan untuk mengurangi *noise* atau derau. Permasalahan *noise* dengan cara yang sangat sederhana dengan cara *filter* yaitu *pre-emphasis*. *Pre-emphasis* mempunyai tujuan agar level *base band* pada bagian frekuensi tinggi masih memiliki kualitas sinyal yang baik. Proses *pre-emphasis* menurut Proakis dan Manolakis (1996) dengan nilai α di antara 0 sampai 1 atau di antara $0,9 \leq \alpha \leq 1,0$ menggunakan Persamaan (2.1)

$$y(n) = s(n) - \alpha s(n - 1) \quad (2.1)$$

Dalam hal ini, $y(n)$ adalah signal hasil *pre-emphasis*, sedangkan $s(n)$ adalah signal sebelum *pre-emphasis*, simbol n merupakan nomor urut signal, α merupakan konstanta *filter pre-emphasis* di antara 0,9-1,0 dan s adalah signal. Proses *pre-emphasis* yang dilakukan seperti pada Gambar 1 menunjukkan bagian (a) merupakan suara masukan sebelum dilakukan *pre-emphasis*, sedangkan Gambar 1 bagian (b) adalah hasil *output* pengolahan signal setelah proses *pre-emphasis*.



Gambar 1. *Pre-emphasis* (a) sebelum dan (b) sesudah (Putra, 2011)

Pengambilan signal ke- n pada *pre-emphasis* dilakukan sepanjang bacaan satu kata atau dua kata dengan waktu satu sampai tiga detik.

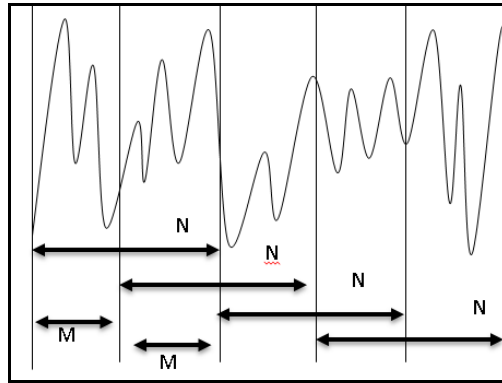
2.2 Frame blocking

Sinyal yang telah melalui *pre-emphasis* selanjutnya dilakukan proses *frame blocking* yang diblok dalam *frame* dengan N sampel dan digeser sebesar M sampel sehingga $N=2M$ dengan $M < N$. Gambar 3.2 menunjukkan ilustrasi *frame blocking* (Abriyono dan Harjoko, 2012). Lebar *frame* dinotasikan dengan N , sedangkan lebar pergeseran setiap *frame* sebagai M . Lebar *overlap* dihitung selisih $N-M$.

Frame blocking menurut Holmes (2003) adalah menganalisis signal ucapan ke dalam bentuk *frame*. Dalam hal ini, setiap *frame* diwakili oleh fitur vektor tunggal digambarkan dalam spektrum rata-rata untuk interval waktu dalam *frame* diambil antara 20-40 milidetik menurut Chamidy (2016). *Frame* diambil sepanjang mungkin untuk mendapatkan resolusi frekuensi yang baik, sedangkan waktu sependek mungkin dimaksudkan untuk mendapatkan ranah waktu yang terbaik. Perhitungan jumlah *frame blocking* menggunakan Persamaan (2.2)

$$f_l(n) = y(Ml + n). \quad (2.2)$$

Dalam hal ini, $f_l(n)$ merupakan hasil *frame blocking*, simbol n adalah $0, 1, \dots, N-1$. Simbol N adalah jumlah sampel, M adalah panjang *frame*, l adalah $0, 1, \dots, L-1$. Simbol L adalah seluruh signal dan y adalah hasil *pre-emphasis*.



Gambar 2. Ilustrasi *Frame blocking* (Abriyono dan Harjoko, 2012)

Gambar 2 menunjukkan M adalah *frame* pertama signal suara dalam formula disimbolkan f_l kemudian $M+M=N$.

2.3 Windowing

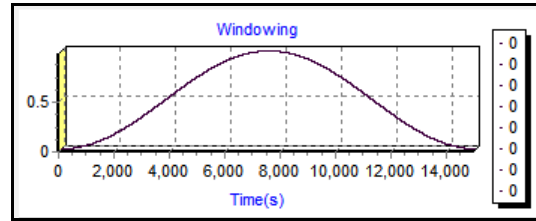
Windowing menurut Proakis dan Manolakis (1996) mempunyai fungsi efek menghaluskan pada spectrum setelah melalui proses *frame blocking*. *Windowing* mempunyai tujuan untuk mengurangi efek *diskontinue* pada ujung-ujung *frame* yang dihasilkan oleh proses *frame blocking*. *Windowing* yang digunakan yaitu *Rectangular Window*, *Hamming Window* dan *Hanning Window* (Chamidy, 2016). Dari ketiga fungsi *windowing* tersebut peneliti menggunakan *windowing Hanning* dikarenakan lebih halus dibandingkan dengan yang lain (Putra, 2008a). Representasi fungsi *windowing* menggunakan Persamaan (2.3)

$$X(n) = f_l(n)w(n). \quad (2.3)$$

Dalam hal ini, fungsi $X(n)$ adalah signal hasil *windowing*, dengan f_l adalah hasil *frame blocking*, dengan n adalah $0, 1, \dots, N-1$. Simbol N adalah jumlah sampel dalam masing-masing *frame* dan $w(n)$ adalah fungsi *window*. Sedangkan fungsi *windowing Hanning* menggunakan Persamaan (2.4)

$$w(n) = 0,5 \left(1 - \cos \left(\frac{2\pi n}{M-1} \right) \right). \quad (2.4)$$

Dalam hal ini, $w(n)$ adalah fungsi *window* menggunakan *hanning*, dengan n adalah $0, 1, \dots, M-1$, M adalah panjang *frame*. Gambar 3 menjelaskan tentang hasil proses *windowing* menggunakan *Hanning window*.



Gambar 3. Ilustrasi *windowing* (Proakis dan Manolakis, 1996)

3. METODE PENELITIAN

3.1 Fast Fourier Transform (FFT)

Fast Fourier Transform adalah pengembangan dari algoritma *Discrete Fourier Transform* (DFT) yang digunakan untuk mengubah sinyal digital pada ranah waktu ke ranah frekuensi (Abriyono dan Harjoko, 2012). Sinyal ini pada dasarnya merepresentasikan dekomposisi sinyal dalam hal komponen sinusoidal. Sinusoidal adalah sinusoid dari frekuensi yang sama tetapi amplitudo dan fase yang berbeda. FFT adalah algoritma yang dikembangkan oleh Cooley dan Turki merupakan proses signal dari ranah waktu menjadi ranah frekuensi.

Transformasi *fourier* menurut Proakis dan Manolakis (1996) salah satu dari beberapa alat matematika yang berguna dalam analisis dan desain sistem *Linier Time Invariant* (LTI) dan deret *fourier*.

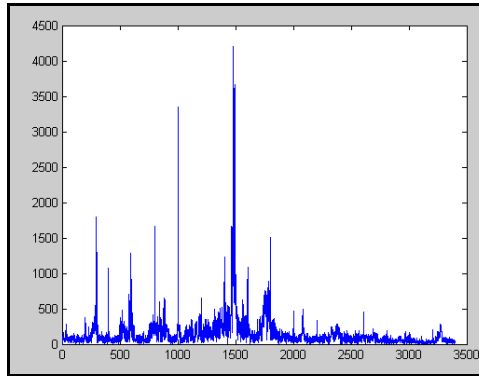
FFT adalah salah satu metode algoritma cepat untuk dapat mengimplementasikan *Discrete Fourier Transform* (DFT). DFT menurut Proakis dan Manolakis (1996) alat komputasi yang memainkan peran yang sangat penting dalam banyak aplikasi pemrosesan sinyal digital, seperti analisis frekuensi, estimasi spektrum daya, dan *filter linear*. Waktu komputasi DFT terlalu lama dan tidak efisien kemudian FFT dapat melakukan efisiensi perhitungan. Seperti yang dikatakan menurut Proakis dan Manolakis (1996) bahwa metode FFT digunakan sebagai cara yang efisien untuk dapat menghitung DFT. *Discrete Fourier Transform* (DFT) menggunakan Persamaan (2.5)

$$d[k] = \sum_{n=0}^{N-1} X(n) e^{-j \frac{2\pi}{N} nk} ; k = 0, 1, 2, \dots, N-1. \quad (2.5)$$

Dalam hal ini, $d[k]$ adalah hasil perhitungan DFT, simbol $X(n)$ adalah hasil *windowing*. Simbol N adalah bilangan natural, N adalah jumlah sampel yang akan diproses ($N \in \mathbb{N}$). Simbol k adalah variabel *frequency discrete* bernilai ($k = N/2, k \in \mathbb{N}$). *Fast Fourier Transform* mempunyai tujuan mendekomposisi sinyal menjadi sinyal sinusoidal yang berupa unit *real* dan *unit imajiner*. *Fast Fourier Transform* menggunakan Persamaan (2.6)

$$T(k) = \sum_{n=0}^{N-1} X(n) \cos\left(\frac{2\pi kn}{N}\right) - \sum_{n=0}^{N-1} X(n) \sin\left(\frac{2\pi kn}{N}\right) \quad (2.6)$$

Dalam hal ini, fungsi $T(k)$ adalah hasil perhitungan *Fast Fourier Transform* ke- k , simbol $X(n)$ adalah hasil perhitungan *windowing* ke- n . Simbol n adalah nomor urut sinyal. Simbol k adalah indeks dari frekuensi ($1, 2, \dots, N$). Gambar 4 menunjukkan hasil spektrum dengan FFT.



Gambar 4. FFT menghasilkan spektrum (Kumar, 2013)

3.2 Mel Frequency Wrapping (MFW)

Mel Frequency Wrapping (MFW) merupakan *filter* berupa *filterbank* untuk mengetahui ukuran energi dari *frequency band* tertentu dalam signal suara (Miftahuddin dan Hakim, 2017; Putra, 2011). MFW menurut Laha (2007) mengubah frekuensi ke dalam mel.

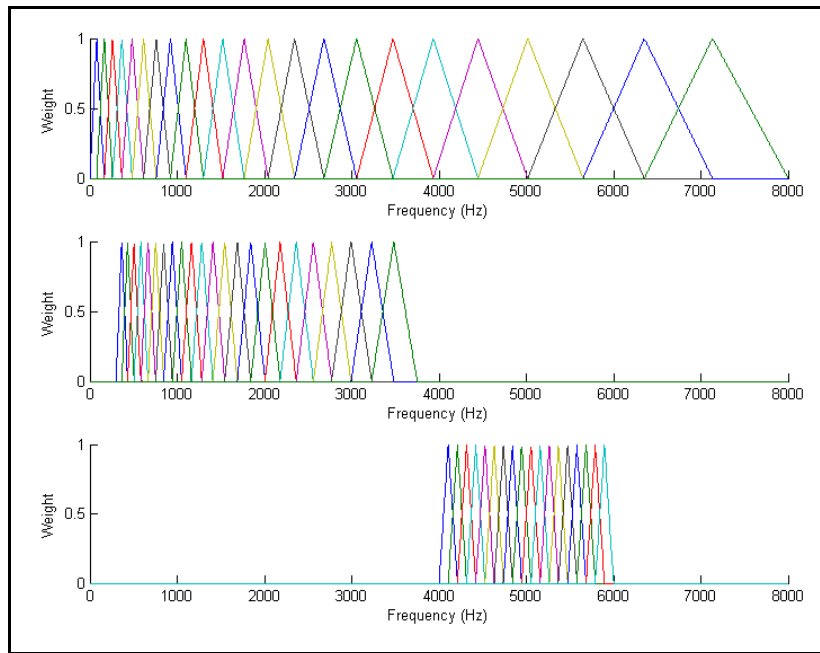
Sedangkan MFW menurut Tshilidzi Marwala (2012) berisi *filterbank*, yang diberi jarak pada skala mel. *Filterbank* memiliki respon frekuensi lewat jalur segitiga yang jarak dan besarnya ditentukan oleh interval frekuensi yang konstan. Proses *output* yang diperoleh dari *filter* dikenal dengan spektrum mel. MFW mempunyai tujuan menghasilkan spektrum mel menggunakan Persamaan (2.7)

$$Y[i] = \sum_{j=1}^G T[j]H_i[j] \quad (2.7)$$

Dalam hal ini, $Y[i]$ adalah hasil perhitungan *mel frequency wrapping* ke- i dengan G adalah jumlah *magnitude spectrum* ($G \in \mathbb{N}$), simbol $T[j]$ adalah hasil FFT, $H_i[j]$ adalah *coefficient filterbank* pada frekuensi j ($1 \leq i \leq E$), dan E adalah jumlah *channel* dalam *filterbank*. Pendekatan yang digunakan dalam bentuk mel menggunakan Persamaan (2.8)

$$mel(f) = 2595 \log_{10} \left(1 + \frac{f}{700} \right). \quad (2.8)$$

Dalam hal ini, mel menggunakan frekuensi dengan skala mel, f sebagai *frequency*. MFW menghasilkan berupa spektrum mel. *Mel frequency scale* adalah *linear* frekuensi skala pada frekuensi di bawah 1.000 Hz dan merupakan *logarithmic* skala pada frekuensi di atas 1.000 Hz (Putra, 2011). Gambar 5 menunjukkan proses spektrum mel.



Gambar 5. Spektrum mel (Hidayat dkk., 2015; Hassan dkk., 2007)

Gambar 5 menunjukkan skala mel dibuat dari *filter bank* menggunakan jenis *filter triangular* dengan *low*, *high* dan *max* dalam warna.

3.3 Discrete Cosine Transform (DCT)

DCT menurut Smith (2000) merupakan kerabat dari transformasi *fourier* terdekomposisi sinyal ke gelombang *cosinus*. DCT juga banyak dipakai untuk mengolah *image processing* misalkan file JPEG. Pada dasarnya konsep dari DCT memiliki kesamaan dengan *inverse fourier transform*. Namun hasil dari DCT adalah mendekati *Principle Component Analysis* (PCA). PCA adalah metode *static* klasik yang digunakan secara luas dalam analisis data dan kompresi.

DCT diasumsikan menggantikan *inverse fourier transform* dalam proses ekstraksi ciri MFCC (Putra, 2011). *Discrete Cosine Transforms* (DCT) adalah anggota dari kelas transformasi unit sinusoidal (Britanak dkk., 2007). DCT mempunyai tujuan menghasilkan seprum mel untuk meningkatkan kualitas pengenalan. DCT menggunakan Persamaan (2.9)

$$C_m = \sum_{k=1}^K (\log_{10} Y[k] \cos \left[m(k - \frac{1}{2}) \frac{\pi}{K} \right]); m = 1, 2, \dots, K. \quad (2.9)$$

Dalam hal ini, C_m adalah *Coefficient*, di mana $Y[k]$ adalah keluaran dari proses *filterbank* pada indek, m adalah banyaknya *coefficient* dan K merupakan jumlah *coefficient* yang diharapkan. Proses DCT menghasilkan seprum mel.

3.4 Cepstral Liftering

Cepstral coefficient menurut Proakis dan Manolakis (1996) merupakan *the fourier series coefficients*. Septrum didefinisikan oleh Rabiner dan Schafer (2007) adalah kebalikannya transformasi *fourier*. *Cepstral liftering* meningkatkan akurasi digunakan untuk *pattern matching*, baik *speaker recognition* maupun *speech recognition* (Putra, 2011). *Cepstral coefficient* menggunakan Persamaan (2.10)

$$w(b) = 1 + \frac{C}{2} \sin \left(\frac{b\pi}{C} \right); b = 1, 2, \dots, C \quad (2.10)$$

Dalam hal ini, $w(b)$ adalah fungsi *window* terhadap *cepstral features*, C adalah *cepstral coefficients*, simbol b adalah index dari *cepstral coefficients*. Pengolahan *cepstral liftering* mempunyai hasil berupa *frame* dan *cepstral coefficients* selanjutnya diproses ke pemilihan fitur. Pemilihan fitur dijelaskan pada selanjutnya.

4. HASIL DAN PEMBAHASAN

Ekstraksi ciri MFCC mempunyai tujuan untuk mendapatkan fitur berupa parameter-parameter. MFCC mempunyai tujuh tahap. Tahap pertama adalah *pre-emphasis*, kedua *frame blocking*, ketiga *windowing*, keempat *Fast Fourier Transform* (FFT), kelima *Mel Frequency Wrapping* (MFW), keenam *Discrete Cosine Transform* (DCT), dan ketujuh *cepstral liftering*.

4.1. Pre-emphasis

Pre-emphasis dilakukan untuk mengurangi *noise* atau derau. *Pre-emphasis* berfungsi sebagai *filter* awal. Perhitungan *pre-emphasis* dilakukan pada sinyal-sinyal digital dalam ranah waktu dan menggunakan Persamaan (2.1). *Pre-emphasis* yang digunakan dalam penelitian ini mempunyai nilai α sebesar 0,94 yang menunjukkan bahwa *filtering* yang digunakan sebesar 94%. *Pre-emphasis* juga dilakukan dengan α sebesar 0,97 menunjukkan sinyal yang diekstraksi merupakan 97% sinyal aslinya (Chamidy, 2016).

4.2. Frame Blocking

Sinyal dalam ranah waktu kemudian dilakukan segmentasi. Sinyal ucapan yang telah *pre-emphasis* dibagi menjadi beberapa *frame* (bingkai) dengan masing-masing *frame* memuat N sampel sinyal dan *frame* yang saling berdekatan dipisahkan sejauh M sampel. Panjang *frame* yang membagi setiap sampel menjadi beberapa *frame* berdasarkan waktu di antara 20 hingga 40 milidetik (Chamidy, 2016). *Frame blocking* mempunyai Persamaan (2.2). Penelitian ini menggunakan waktu 31 milidetik di antara 20 hingga 40 milidetik.

4.3. Windowing

Windowing dilakukan, karena pengaruh efek dari *frame blocking* sinyal suara menjadi diskontinu. Salah satu cara untuk menghindari *diskontinu* pada ujung *window* adalah dengan meruncingkan sinyal menjadi nol atau mendekati nol, sehingga mengurangi kesalahan. Penelitian ini memilih menggunakan *Hanning Window* karena lebih halus menurut Putra (2008a). *Windowing* mempunyai Persamaan (2.3), sedangkan proses *Hanning Windowing* mempunyai Persamaan (2.4). *Windowing* mempunyai hasil berupa *frame*.

4.4. Fast Fourier Transform (FFT)

Setiap hasil *frame* dari fungsi *window* dilakukan Transformasi Fourier agar proses komputasi menjadi lebih efisien. FFT menghasilkan spektrum, persamaan ditunjukkan pada Persamaan (2.6).

4.5. Mel Frequency Wrapping (MFW)

FFT melakukan *wrapping* terhadap spektrum yang dihasilkan sehingga menjadi skala mel. Jangkauan frekuensi dalam sangatlah luas dan sinyal suara tidak mengikuti skala *linear*, sehingga spektrum terkomputasi data dipetakan dalam skala mel menggunakan *filter* segitiga yang saling tumpang tindih. Perhitungan MFW mempunyai Persamaan (2.7) dan (2.8). Sinyal yang dihasilkan adalah spektrum mel.

4.6. Discrete Cosine Transform (DCT)

DCT merupakan proses mendekorelasikan seprum mel sehingga menghasilkan penyajian yang baik dari *property spectral local* (Putra, 2011). DCT mempunyai tujuan menghasilkan seprum mel untuk meningkatkan kualitas pengenalan. Perhitungan DCT menggunakan Persamaan (2.9) dan menghasilkan seprum mel.

4.7. Cepstral Liftering

Cepstral coefficient menjadi salah satu standar yang diterapkan untuk meminimalisasi sensitifitas kesalahan meningkatkan kualitas pengenalan. *Cepstal liftering* mempunyai Persamaan (2.10). *Cepstral liftering* menghasilkan *cepstral coefficient*. MFCC menghasilkan fitur berupa *frame* dan *cepstral coefficient*.

Penelitian juga dilakukan dengan pengujian MFCC tanpa NBD. Perbandingan tersebut dilakukan untuk mengetahui perbedaan dampak menggunakan model usulan NBD dengan tanpa NBD terhadap surat Al-Fatihah, Al-Baqarah, Al-Imran, Ar-rahman, Al-Hadid, Al-Alaq, Al-Ashr, Al-Kautsar, Al-Ikhlash, Al-Falaq dan An-nas.

Tabel 1. Surat Al-Fatihah menggunakan MFCC

| No | Bacaan Surat Al-Fatihah | Pembaca | MFCC |
|----|-------------------------|---------|------|
| 1 | alai-him | 7 | 57,1 |
| 2 | a-lamin | 7 | 42,8 |
| 3 | al-hamdulillah | 7 | 85,7 |
| 4 | an-am-ta | 7 | 42,8 |
| 5 | arr-roh-man | 7 | 42,8 |
| 6 | bi-alai-him | 7 | 85,7 |
| 7 | Bismillah | 7 | 57,1 |
| 8 | ghoi-rill-magh-dhu | 7 | 42,8 |
| 9 | hi-robb-bil | 7 | 71,4 |
| 10 | hirr-rohman | 7 | 71,4 |
| 11 | ih-dinasy | 7 | 71,4 |
| 12 | iyy-ya | 7 | 42,8 |
| 13 | kanak-bu-du | 7 | 71,4 |
| 14 | kanas-tai-in | 7 | 57,1 |
| 15 | la-zii-naa | 7 | 57,1 |
| 16 | ma-li-ki | 7 | 71,4 |
| 17 | mus-ta-qim | 7 | 42,8 |
| 18 | nirr-rohim | 7 | 85,7 |
| 19 | syi-roo-toll | 7 | 100 |
| 20 | wa-iyy-ya | 7 | 71,4 |
| 21 | walad-dhool-lin | 7 | 28,5 |
| 22 | yau-midd-din | 7 | 57,1 |
| | RATA-RATA | | 61,7 |

Fitur dari ekstraksi ciri MFCC berupa fitur *frame* dan membuat rata-rata *cepstral coefficient* di setiap *frame* fitur langsung pengecekan dengan pembaca dengan referensi. Keakurasian sejauh ini menggunakan ekstraksi ciri MFCC yang mempunyai tingkat keakurasian sebesar 58% sampai dengan 75% (Aibinu, dkk., 2011).

Tabel 1 menunjukkan surat Al-Fatihah hasil rata-rata MFCC sebesar 61,7% sedangkan berdampak pada kesesuaian bacaan dan akurasinya dikarenakan terdapat banyak kesalahan. Kesalahan tersebut dalam referensi diakibatkan masih ditemukannya fitur *frame* dan *cepstral coefficient* yang belum unik, sehingga berpengaruh pada ketepatan dengan fitur yang ada pada suara pembaca.

Pengujian dilakukan terhadap bacaan dan jenis surat yang sama, seperti yang telah dilakukan pada pengujian-pengujian sebelumnya terhadap sebanyak sebelas surat dengan orang yang sama.

Tabel 2. Perbandingan seluruh kesebelas surat MFCC tanpa NBD

| No | Surat | MFCC |
|----|------------|------|
| 1 | Al-Fatihah | 61,7 |
| 2 | Al-Baqarah | 58,0 |
| 3 | Al-Imran | 59,7 |
| 4 | Ar-rohman | 51,3 |
| 5 | Al-Hadid | 59,1 |

| | | |
|----|------------|------|
| 6 | Al-Alaq | 60,1 |
| 7 | Al-Ashr | 56,4 |
| 8 | Al-Kautsar | 46,3 |
| 9 | Al-Ikhlas | 38,3 |
| 10 | Al-Falaq | 38,7 |
| 11 | An-nas | 40,0 |
| | Rata-rata | 51,8 |

Tabel 2 menunjukkan bahwa perbandingan MFCC tanpa NBD dengan hasil rata-rata sebesar 51,8%, Hal ini terjadi, karena MFCC tanpa NBD fitur *frame* dan *cepstral coefficient* belumlah sesuatu yang unik dan beda sehingga berdampak pada akurasi.

5. KESIMPULAN

Pengenalan suara bacaan Al-Quran menggunakan Metode MFCC. MFCC mulai dari *pre-emphasis*, *frame blocking*, *windowing*, *Fast Fourier Transform (FFT)*, *Discrete Cosine Transform (DCT)* dan *cepstral liftreing*. Hasil pengecekan bacaan Al-Qur'an diujikan dalam sebelas surat mulai dari surat Al-Fatihah, Al-Baqarah, Al-Imran, Al-Hadid, Al-Ashr, Ar-rahman, Al-Alaq, Al-Kautsar, Al-Ikhlas, Al-Falaq dan An-Nas menghasilkan akurasi sebesar rata-rata 51,8%.

DAFTAR PUSTAKA

- Abriyono dan Harjoko, A., 2012. Pengenalan Ucapan Suku Kata Bahasa Lisan Menggunakan Ciri LPC, MFCC, dan JST. *Indonesian Journal of Computing and Cybernetics Systems*, 6(2), hal.23–34.
- Aibinu, A.M., Salami, M.J.E., Najeeb, A.R., Azeez, J.F. dan Rajin, S.M.A.K., 2011a. Evaluating the effect of voice activity detection in isolated Yoruba word recognition system. *2011 4th International Conference on Mechatronics: Integrated Engineering for Industrial and Societal Development, ICOM'11 - Conference Proceedings*, (May), hal.17–19.
- Barkoni, 2017. Tajwid. *Pakar Tajwid*.
- Bodruzzaman, M., Kuah, K., Jamil, T., Wang, C. dan Li, X., 1993. System Using Artificial Neural Network. , hal.1–3.
- Bustami, Fadlisyah, I. mauliza, 2017. Sistem Pendeteksi Kesalahan Dalam Membaca Al-Qur'an Ayat 1-5 Menggunakan Metode Viterbi Bustami,. *TECHSI: Jurnal Penelitian Teknik Informatika Vol 9 NO. 1 hal 1-15*, Vol 9 No.1, hal.1–15. Available at: <http://ojs.unimal.ac.id/index.php/techsi/article/view/205>.
- Chamidy, T., 2016. Metode Mel Frequency Cepstral Coeffisients (MFCC) Pada klasifikasi Hidden Markov Model (HMM) Untuk Kata Arabic pada Penutur Indonesia. *Matics*, 8(1), hal.36–39. Available at: <http://ejournal.uin-malang.ac.id/index.php/saintek/article/view/3482>.
- Davis, S.B. dan Mermelstein, P., 1980. Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 28(4), hal.357–366.
- Holmes, J.H. and W., 2003. *Speech Synthesis and Recognition, Second Edition*,
- Laha, D., 2007. *Handbook of Computational Intelligence in Manufacturing and Production Manajemen*,
- Miftahuddin, Y. dan Hakim, M.R., 2017. Coefficient Dan Dynamic Time Warping Untuk Pengenalan Nada Pada Alat Musik Bellyra. , hal.120–127.

- Proakis, J.G. dan Manolakis, D.G., 1996. *Digital Signal Processing: Principles, algorithms, and applications*, Available at:
https://engineering.purdue.edu/~ee538/DSP_Text_3rdEdition.pdf.
- Putra, A.E., 2008. Frekuensi Cuplik pada FFT. *Tan Li, Processing, Digital Signal*, 1.
- Putra, D. dan Adi, R., 2011. Verifikasi Biometrika Suara Menggunakan Metode MFCC dan DTW. *Biometrika, Universitas Udayana*, 2(1), hal.8–21.
- Suyanto, S. dan Putra, A.E., 2014a. Automatic Segmentation of Indonesian Speech into Syllables using Fuzzy Smoothed Energy Contour with Local Normalization, Splitting, and Assimilation. *Journal of ICT Research and Applications*, 8(2), hal.97–112. Available at: http://journal.itb.ac.id/index.php?li=article_detail&id=1804.
- Suyanto, S. dan Putra, A.E., 2014b. Automatic Segmentation of Indonesian Speech into Syllables using Fuzzy Smoothed Energy Contour with Local Normalization, Splitting, and Assimilation. *Journal of ICT Research and Applications*, 8(2), hal.97–112. Available at: http://journal.itb.ac.id/index.php?li=article_detail&id=1804.
- Suyanto dan Hartati, S., 2013. Design of Indonesian LVCSR using Combined Phoneme The Approaches of LVCSR. *Icts*, hal.191–196.
- Tokunbo Ogunfunmi, R.T.M. (Sim) narasimha, 2015. *Speech and Audio Processing and Recognition*, springer.
- Tshildzi Marwala, 2012. *Condition Monitoring Using Computational Intelligence Methods*,
- Vladimir Britanak, Patrick C.Yip, K.R.R., 2007. *Discrete Cosine and Sine Transform*,
- Zarkasyi I, 1995. Pelajaran Tajwid. *Gontor Ponorogo*, Trimurti P, hal.Hal 1-3.
-