

Pemodelan Harmonik untuk Pelafalan *Makhraj* Huruf Hijaiah

Muhammad Fadhlullah*¹, Catur Atmaji²

¹Program Studi Elektronika dan Instrumentasi, FMIPA UGM, Yogyakarta, Indonesia

²Departemen Ilmu Komputer dan Elektronika, FMIPA UGM, Yogyakarta, Indonesia

e-mail: *¹m.fadhlullah.khalilullah@mail.ugm.ac.id, ²catur_atmaji@ugm.ac.id

Abstrak

Proses belajar melafalkan huruf hijaiyah perlu dinilai secara objektif, sehingga dirasa perlu dibentuk audio digital hasil dari sintesis pemodelan Harmonic Plus Residual (HPR). Penelitian dilakukan dengan dua metode pelafalan, yaitu taskinul harf (taskin) dan tasydidul harf (tasydid). Metode penelitian yang digunakan merupakan metode eksperimen yang alurnya terdiri dari akuisisi data, pemotongan sinyal, framing dan windowing, deteksi frekuensi fundamental dan harmonik, sintesis pemodelan HPR, hingga menghasilkan sinyal sintetik. Hasil sinyal-sinyal sintetik kemudian dianalisis secara kualitatif dengan cara analisis spektrogram sinyal dan skoring.

Dari hasil eksperimen, dapat disimpulkan bahwa penelitian ini pada akhirnya tidak dapat menentukan parameter sintesis HPR model yang terbaik, namun menyimpulkan metode tasydid sebagai metode terbaik untuk belajar pelafalan dan untuk bahan sintesis model HPR. Hal ini karena metode pelafalan tasydidul harf dengan kombinasi parameter yang berbeda-beda namun kesemuanya dapat menghasilkan sinyal sintetik yang sama bagusnya, baik secara analisis perbandingan spektrogram sinyal yang mirip maupun dari hasil skoring dengan rata-rata nilai 10. Di lain sisi, metode pelafalan taskinul harf menunjukkan hasil yang tidak memuaskan, dimana suara hasil sintetik kebanyakan terdengar hanya derau, sehingga hasil skoring juga rendah dengan nilai rata-rata tidak mencapai 9, dan diperkuat dari hasil analisis perbandingan spektrogram antara sinyla asli dan sintetik dimana persebaran magnitudonya berbeda.

Kata kunci— *Harmonic Plus Residual, Taskin, Tasydid, parameter optimal*

Abstract

Learning to pronounce hijaiyah letters needs to be assessed objectively, so it is necessary to form digital audio resulting from the synthesis of Harmonic Plus Residual (HPR) modeling, which conducted with two pronunciation methods, taskin and tasydid. The experiment consists data acquisition, signal cutting, framing and windowing, detection of fundamental and harmonic frequencies, synthesis of HPR, to produce synthetic signals. The results of the synthetic signals then analyzed qualitatively by signal spectrogram analysis and scoring.

From the experimental results, it can be concluded that this study was ultimately unable to determine the best HPR parameters, but concluded that the tasydid method was the best method for learning pronunciation and for the HPR model synthesis. This is because the tasydid method with different parameters but all of them can produce good synthetic signal, both in terms of comparative analysis of similar signal spectrograms and from the results of scoring with an average value of 10. On the other hand, the taskin method harf shows unsatisfactory results, where the synthetic sound is mostly just noise, so the scoring results is under 9, and is reinforced by the results of the spectrogram comparison between the original and synthetic signals which visually different.

Keywords— *Harmonic Plus Residual, Taskin, Tasydid, optimal parameters*

1. PENDAHULUAN

Setiap muslim memiliki kewajiban membaca Alquran, karena membacanya terkait erat dengan ibadah ritual seperti salat, haji, dan sebagainya. Kesalahan dalam membaca Alquran akan sangat mempengaruhi makna dan pengertiannya [1]. Sehingga membaca Alquran dengan pelafalan *makhraj* dan tajwid yang benar menjadi kewajiban bagi setiap muslim [2]. Permasalahan dalam membaca Alquran yang seringkali ditemukan adalah terkait membunyikan simbol hijaiyah dengan tepat sesuai *makhraj*-nya [1]. *Makhraj* adalah tempat keluarnya huruf, dan setiap huruf hijaiyah memiliki *makhraj*-nya masing-masing. Perbedaan pengucapan antara bahasa Arab dengan bahasa Indonesia menjadi kendala bagi orang-orang yang terbiasa berbahasa Indonesia, dimana pengucapan bahasa Indonesia tidak perlu memperhatikan *makhraj*. Kendala selanjutnya adalah terdapat beberapa huruf hijaiyah yang suara pelafalannya terdengar mirip, misalnya huruf sin (س), syin (ش), sa (ث) yang biasanya sulit dibedakan [3]. Oleh karena itu, seseorang yang ingin belajar membaca Alquran dimulai dari belajar mengucapkan huruf secara tepat dari *makhraj*-nya, agar dapat membedakan antara satu huruf dengan huruf lainnya [4].

Pada umumnya, seseorang belajar Alquran secara langsung dengan seorang guru yang telah menguasai *makhraj* huruf. Guru berperan untuk mengecek bacaan Alquran dan mengoreksi jika ada kesalahan dalam pelafalannya [2]. Namun, dalam proses pembelajaran ini ditemukan beberapa kendala, antara lain frekuensi pertemuan dan jam belajar yang singkat, serta ketidakcocokan antara pembelajar dan guru dengan metode yang digunakannya dalam mengajar [5]. Kendala lain adalah bahwa setiap guru tentu memiliki penilaiannya masing-masing, sehingga penilaian guru bersifat subjektif. Hal tersebut karena setiap guru sangat mungkin berbeda pemahamannya tentang cara melafalkan *makhraj* huruf dengan benar. Akibatnya, belajar dengan dua guru yang berbeda dapat menyebabkan penilaian yang berbeda pula sehingga terkadang membuat pembelajar bingung. Selain itu, banyak faktor yang dapat mempengaruhi penilaian seorang guru, misal faktor usia dan kesehatan, atau faktor kebisingan lingkungan. Hal tersebut dapat menyebabkan penilaian pada hari ini, bisa jadi akan berbeda dengan penilaian pada hari esoknya.

Untuk mengetahui tempat asal atau *makhraj* setiap huruf hijaiyah dapat ditempuh dengan dua cara, yaitu taskinul harf dan tasydidul harf. Cara pertama yaitu dengan taskinul harf atau men-sukun-kan suatu huruf dan menambahkan hamzah washal sebelumnya. Cara kedua yaitu dengan tasydidul harf atau men-tasydid-kan suatu huruf dan menambahkan hamzah washal sebelumnya. Dua cara tersebut akan menunjukkan letak *makhraj* huruf ketika berhenti saat dilafalkan [4].

Saat ini telah banyak dikembangkan alternatif belajar melafalkan huruf hijaiyah dengan sistem komputer pemrosesan sinyal dengan metode DTW, LPC dan MFCC [6] [7] [8]. Dalam beberapa penelitian lain, sintesis dilakukan dengan pemodelan harmonik namun yang diteliti bukan lafal huruf hijaiyah [9] [10] [11]. Berdasarkan berbagai faktor yang menjadi kendala dan dengan tidak mengubah fungsi guru untuk mengoreksi pelafalan dalam pembelajaran tatap muka, maka diusulkanlah penelitian ini, agar supaya belajar *makhraj* huruf hijaiyah dapat terkomputerisasi sehingga siapapun dapat belajar, latihan dan mengoreksi pelafalan *makhraj*-nya secara mandiri, dengan penilaian atau pengoreksian yang stabil dan konsisten menggunakan pemodelan harmonik yang umum digunakan dalam sintesis sinyal. Dalam penelitian ini, pemodelan harmonik yang digunakan ialah *Harmonic Plus Residual* (HPR), dimana komponen harmonik dijumlahkan dengan komponen residu sinyal sebagai unsur lingkungan agar suara sintetik yang dihasilkan lebih alami.

2. METODE PENELITIAN

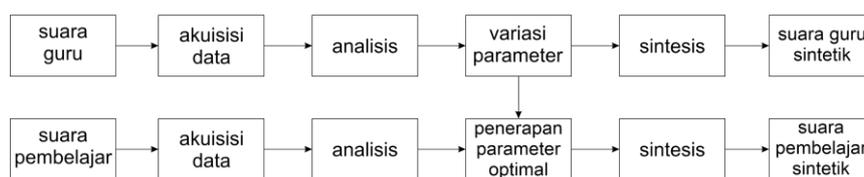
2.1 Analisis dan Rancangan Sistem

Dalam penelitian ini, dilakukan *analysis by synthesis* pemodelan *harmonic plus residual* (HPR), yaitu menganalisis sinyal suara pelafalan dengan cara mensintesis suara pelafalannya. Pelafalan huruf hijaiyah akan dilakukan oleh dua subjek, yaitu guru dan murid/pembelajar. Guru adalah laki-laki dewasa yang menguasai pelafalan *makhraj* huruf hijaiyah dengan baik, dan terbiasa

mengajar membaca Alquran. Murid/pembelajar adalah laki-laki dewasa yang belum menguasai pelafalan *makhraj* huruf hijaiyah, namun pada pelafalannya berusaha untuk melafalkan semirip mungkin dengan pelafalan yang dicontohkan guru. Secara umum, penelitian ini terdiri dari dua alur, alur pertama yaitu pemrosesan sinyal guru, dan alur kedua yaitu pemrosesan sinyal murid/pembelajar.

Tahap pemrosesan sinyal guru bertujuan untuk mencari parameter yang paling optimal dari sinyal suara guru dalam melafalkan *makhraj* huruf hijaiyah. Oleh karenanya, pada tahap ini terdapat proses variasi jenis dan nilai parameter-parameter yang dibutuhkan dalam merekonstruksi sinyal guru. Kemudian tahap pemrosesan sinyal pembelajar adalah tahap dimana parameter yang paling optimal yang diperoleh dari alur pemrosesan sinyal guru diterapkan kepada sinyal pembelajar untuk menilai apakah hasil sintetik pembelajar bisa lebih baik daripada sinyal aslinya dari segi ketepatan *makhraj* setiap huruf hijaiyah.

Pada Gambar 1 dapat diperhatikan bahwa sistem secara umum terdiri dari dua alur. Alur pertama adalah tahap pemrosesan sinyal guru hingga menghasilkan suara guru sintetik. Alur kedua adalah tahap pemrosesan sinyal pembelajar hingga menghasilkan suara pembelajar sintetik. Pada dasarnya kedua alur tersebut terdiri dari tahap-tahap yang sama, yaitu akuisisi data, analisis, variasi parameter dan sintesis. Dalam perancangan sistem, pemrograman atau coding tidak sepenuhnya dirancang sendiri, melainkan diambil dari salah satu *course* di Coursera berjudul *Audio Signal Processing for Music Applications* [12].



Gambar 1 Diagram Blok Sistem

2.2 Akuisisi Data

Proses akuisisi data untuk alur pertama maupun alur kedua dilakukan di awal, yaitu dengan merekam suara 1 orang guru dan 7 orang pembelajar. Data yang akan diambil adalah pelafalan beberapa huruf hijaiyah, yaitu hamzah (ء), ba (ب) mim (م), dad (ض), sad (ص), sin (س), ta (ت), ta (ط), zal (ذ), za (ظ), ha (ح), ha (ه). Perekaman menggunakan *microphone* dengan frekuensi cuplik sebesar 44.100 Hz. Setiap proses perekaman terdiri dari dua sesi. Satu sesi menggunakan metode pelafalan *taskinul harf* dan sesi yang lain menggunakan metode *tasydidul harf* untuk dapat mengetahui letak *makhraj* huruf hijaiyah yang dibaca. Tujuan menggunakan dua metode tersebut karena dianggap akan menghasilkan luaran yang berbeda, sehingga dapat diukur dan dibandingkan metode mana yang menghasilkan luaran yang lebih efektif dalam melafalkan huruf hijaiyah.

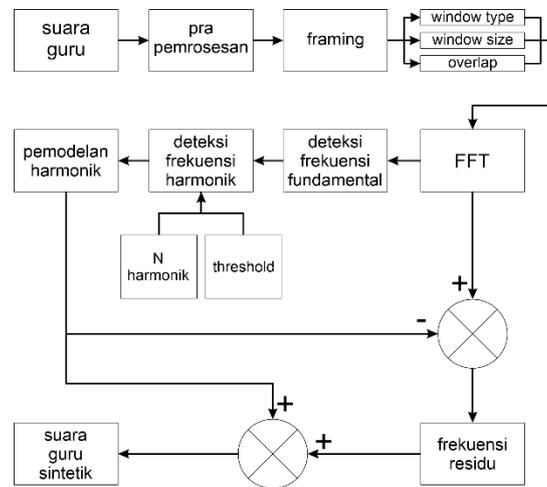
2.3 Pra-Pemrosesan Sinyal

Sebelum diproses, data sinyal lebih dulu dipotong-potong per huruf sehingga setiap sampel data sinyal menjadi 12 data sinyal mewakili setiap huruf hijaiyah yang telah direkam. Durasi pemotongan ialah 1 detik untuk setiap huruf pada sampel *taskin*, dan 2 detik untuk setiap huruf pada sampel *tasydid*. Hal tersebut karena setiap pelafalan huruf dengan metode *tasydidul harf* lebih panjang daripada *taskinul harf*.

2.4 Pemrosesan Sinyal Guru

Setelah akuisisi data, data sinyal guru kemudian memasuki tahap analisis dan sintesis. Tahap analisis adalah tahap dimana parameter tertentu ditentukan jenis atau nilainya yang dibutuhkan untuk merekonstruksi sinyal asli dalam tahap sintesis. Parameter yang dimaksud ialah *window type*, *window size*, maksimum harmonik dan *threshold*. Tahap ini mencakup proses *framing* dan *windowing*, FFT (*fast fourier transform*), deteksi frekuensi fundamental dan frekuensi harmonik, hingga perhitungan frekuensi residu. Tahap sintesis adalah tahap dimana

sinyal direkonstruksi berdasarkan bahan-bahan yang diperoleh dari tahap analisis sebelumnya, yaitu frekuensi fundamental, frekuensi harmonik dan frekuensi residu, sehingga menghasilkan sinyal suara sintetik. Alur pemrosesan sinyal guru selengkapnya dapat diperhatikan pada Gambar 2.



Gambar 2 Alur Pemrosesan Sinyal Guru

2.4.1 Framing dan Windowing

Proses *framing* berfungsi untuk memotong sinyal suara dengan durasi yang panjang menjadi durasi yang lebih pendek, agar didapatkan karakteristik sinyal suara yang lebih stabil [13]. Sinyal dibaca dari frame ke frame dengan nilai *overlap* tertentu lalu dilakukan windowing pada setiap frame [14]. Setiap *frame* sinyal suara dikalikan dengan fungsi *windowing* yang berukuran sama dengan ukuran *frame*. Representasi fungsi *window* terhadap sinyal suara yang diolah, ditunjukkan pada persamaan (1). Dimana $x(n)$ adalah nilai sampel sinyal, x_i adalah nilai sampe dari *frame* sinyal ke- i , $w(n)$ adalah fungsi dari *window* yang akan digunakan, dan N adalah ukuran atau panjang *frame* yang merupakan kelipatan 2 [15]. Untuk tingkat *overlap* direpresentasikan dengan *hop size* yang merupakan pengaturan jarak waktu proses *frame* awal dan *frame* berikutnya pada teknik pembagian sinyal agar *frame-frame* tersebut saling tumpang tindih (*overlap*) satu sama lain dengan ukuran $h < N$ [16].

$$x(n) = x_i(n)w(n) \quad (1)$$

$$n = 0, 1, \dots, N - 1$$

Proses *windowing* berfungsi untuk menghilangkan diskontinuitas yang diakibatkan proses *framing* [17] serta untuk meredam *noise* yang muncul di kedua ujung *frame* [18]. Adapun *overlap* digunakan untuk menghaluskan tahap sintesis sinyal suara. Sehingga dalam proses ini, parameter yang akan divariasikan dan dicari kualitas optimalnya ialah jenis dan lebar *window* yang akan diterapkan pada setiap *frame*.

2.4.2 Deteksi Frekuensi Fundamental dan Frekuensi Harmonik

Sinyal suara yang diproses dengan *Fast Fourier Transform* (FFT) akan menghasilkan sinyal suara domain frekuensi. Lalu dari frekuensi tersebut akan dianalisis frekuensi fundamental dan frekuensi harmoniknya menggunakan fungsi pemodelan harmonik sebagaimana pada persamaan (2), (3) dan (4). Frekuensi fundamental dan frekuensi harmonik yang diperoleh selanjutnya disebut komponen harmonik. Komponen harmonik inilah yang digunakan untuk sintesis sinyal digital.

$$h_1[n] = \sum_{k=-L(n_a^i)}^{L(n_a^i)} a_k(n_a^i) e^{j2\pi k f_0(n_a^i)(n-n_a^i)} \quad (2)$$

$$h_2[n] = \Re \left\{ \sum_{k=1}^{L(n_h^i)} A_k(n) \exp^{j2\pi k f_0(n_h^i)(n-n_h^i)} \right\} \quad (3)$$

$$h_3[n] = \sum_{k=0}^{L(n_h^i)} a_k(n) \cos(\varphi_k(n)) \quad (4)$$

Nilai magnitudo dari frekuensi harmonik semakin menghilang seiring dengan bertambahnya nilai waktu, sehingga setiap *frame* akan mengandung frekuensi harmonik yang berbeda. Nilai magnitudo disimpan dengan komponen frekuensi. Banyaknya frekuensi harmonik berpengaruh terhadap suara yang dihasilkan. Oleh karena itu, banyaknya frekuensi harmonik juga menjadi parameter dalam penelitian ini.

2.4.3 Analisis Frekuensi Residu

Frekuensi residu berfungsi untuk dijadikan model frekuensi *noise* dalam proses sintesis suara, yang diperoleh dari sinyal asli yang disubtraksi dengan komponen harmonik. Frekuensi residu merupakan komponen yang diasumsikan sebagai unsur lingkungan agar hasil sintesis sinyal menjadi lebih alami.

2.4.4 Sintesis Sinyal

Sintesis sinyal dilakukan dengan model *harmonic plus residual* (HPR) yang pada dasarnya melibatkan model sintesis sinusoidal terhadap komponen harmonik, dimana yang disintesis adalah *main lobe* dari *window blackman* dengan *Inverse Discrete Fourier Transform* (IDFT) sebagaimana pada persamaan (5). Komponen harmonik yang telah disintesis selanjutnya dijumlahkan dengan frekuensi residu menjadi satu sinyal, sehingga menghasilkan luaran satu data sinyal sintetik.

$$y(n) = IDFT \left(\sum_{r=0}^R A_r \text{mWl}[k - k_r] e^{j(pWl[k-k_r] + \varphi_r)} \right) \quad (5)$$

2.5 Pengujian Signal to Noise Ratio

Pada tahap ini, sinyal sintetik guru diukur kemiripannya terhadap sinyal asli guru menggunakan *Signal to Noise Ratio* (SNR). Tujuannya adalah menyeleksi variasi parameter yang dapat menghasilkan sinyal sintetik paling mirip dengan sinyal aslinya. SNR adalah rasio antara daya sinyal yang diinginkan dengan daya derau (*noise*). Estimasi SNR dilakukan berdasarkan statistik sinyal dengan menentukan koefisien dari spektrum dan energi sinyal. Nilai koefisien korelasi sinyal diperoleh dari mengoperasikan fungsi korelasi sinyal yang berderau dengan waktu tunda tertentu. Nilai koefisien korelasi sinyal menunjukkan persentase perbedaan sinyal. Persamaan SNR menggunakan metode korelasi ditunjukkan pada persamaan (6). Persamaan (9) menunjukkan nilai SNR, dengan ρ adalah koefisien korelasi antar sinyal. Koefisien korelasi bernilai antara $-1 < \rho < 1$. Persamaan (10) menunjukkan perbandingan daya antar sinyal dalam satuan dB.

$$r_{12} = \sum_{n=0}^{N-1} x_1(n)x_2(n) \quad (6)$$

$$r_{12} = \frac{1}{N} \sum_{n=0}^{N-1} x_1(n)x_2(n) \quad (7)$$

$$\rho = \frac{\frac{1}{N} \sum (x - \bar{x}) \cdot (y - \bar{y})}{\sqrt{\frac{1}{N} \sum (x - \bar{x})^2} \cdot \sqrt{\frac{1}{N} \sum (y - \bar{y})^2}} \quad (8)$$

$$SNR = \frac{\text{ sinyal }}{\text{ derau }} = \frac{\rho}{(1 - \rho)} \quad (9)$$

$$SNR = 10 \log_{10} \left(\frac{\rho}{1 - \rho} \right) \text{ dB} \quad (10)$$

2.6 Pemrosesan Sinyal Pembelajar

Tahapan ini dilakukan setelah alur pertama telah selesai diulang-ulang berdasarkan jumlah variasi parameternya. Hal itu karena pada alur ini bergantung dari parameter sinyal guru. Secara alur, tahapan-tahapan pemrosesan sinyal pembelajar ini sama dengan tahapan pada alur pemrosesan sinyal guru. Hanya saja, pada alur ini tinggal menerapkan parameter paling optimal dari sinyal guru saja.

2.7 Penilaian Kualitas Sinyal

Setiap data sinyal asli dan sinyal sintetis dari parameter optimal dinilai kualitasnya berdasarkan kesesuaian lafal huruf terhadap *makhraj*-nya oleh guru. Penilaian (skoring) dilakukan dengan cara data sinyal didengarkan dan diberi nilai dalam rentang 1 hingga 10, dimana semakin tinggi skor maka semakin sesuai dengan *makhraj* huruf yang bersangkutan.

3. HASIL DAN PEMBAHASAN

Dari proses akuisisi data dari 8 orang pelafal, 2 metode pelafalan yakni taskin dan tasydid, dan 3 kali pengulangan skema perekaman suara sehingga menjadi sampel 1, sampel 2 dan sampel 3 untuk masing-masing metode taskin dan metode tasydid. Dari proses akuisisi data diperoleh 48 data rekaman dengan frekuensi cuplik 44100 Hz dimana setiap data rekaman mengandung pelafalan 12 huruf hijaiyah. Oleh karena satu data masih mengandung 12 huruf hijaiyah, maka pada tahap ini, setiap data rekaman dipotong-potong per huruf setiap 1 detik untuk metode taskin dan per huruf setiap 2 detik untuk metode tasydid. Sehingga diperoleh total 576 data dari 48 data rekaman dikalikan 12 huruf. Dimana pemotongan dilakukan tidak ada aturan khusus mengenai posisi sinyal, sehingga proses pemotongan memposisikan sinyal dalam rentang waktu 1 atau 2 detik.

Tabel 1 Hasil SNR Terbaik

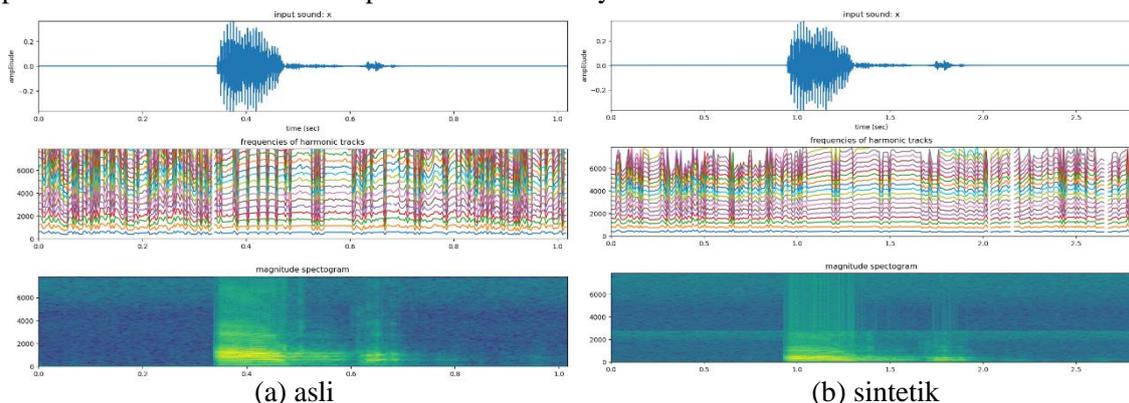
Nilai SNR		
Taskin	Sampel 1	-0.000758
	Sampel 2	-0.000723
	Sampel 3	-0.000711
Tasydid	Sampel 1	-0.000635
	Sampel 2	-0.000639
	Sampel 3	-0.000876

Proses pengolahan sinyal guru berfokus pada pencarian parameter HPR terbaik dari hasil sinyal suara pelafalan sintetis dengan pertimbangan kesesuaian *makhraj* huruf dan kemiripan dengan sinyal asli. Kemiripan sinyal sintetis dengan sinyal asli ditentukan dengan *Signal to Noise Ratio* (SNR) dimana yang memiliki nilai SNR paling besar dianggap menjadi sinyal sintetis yang paling mirip dengan sinyal asli. Dari proses variasi 4 jenis parameter, diperoleh 375 variasi kombinasi parameter. Jumlah variasi tersebut diperoleh dari 3 *window type*, 5 nilai *window size*, 5 nilai *threshold* dan 5 nilai *harmonic max*. *Window type* yang digunakan adalah hanning, hamming dan blackman. Nilai *window size* yang digunakan yaitu 600, 700, 800, 900 dan 1000. Nilai *threshold* yang digunakan yaitu -100, -110, -120, -130 dan -140. Nilai *harmonic max* yang digunakan yaitu 10, 20, 30, 40 dan 50. Jumlah data sebanyak 576 dikalikan 375 kombinasi parameter, maka diperoleh data sinyal sintetis sejumlah 216.000 dengan rincian 36.000 untuk setiap sampel 1, 2 dan 3 masing-masing metode taskin dan tasydid.

Tabel 2 Parameter Terbaik Berdasarkan Hasil Skoring Sinyal Guru

Sampel	Parameter	Skor Rata-rata
Taskin 1 (parameter 1)	Window : hanning Window size : 700 Threshold : -140 Harm. max : 40	8,3
Taskin 2 (parameter 1)	Window : hanning Window size : 600 Threshold : -130 Harm. Max : 30	5,8
Taskin 3 (parameter 1)	Window : hamming Window size : 700 Threshold : -120 Harm. Max : 30	1,0
Tasydid 1 (parameter 4)	Window : blackman Window size : 600 Threshold : -100 Harm. Max : 50	10,0
Tasydid 2 (parameter 2)	Window : hanning Window size : 700 Threshold : -100 Harm. Max : 40	10,0
Tasydid 3 (parameter 10)	Window : hamming Window size : 800 Threshold : -140 Harm. max : 30	10,0

Seluruh data sintetik kemudian diukur kemiripannya terhadap sinyal aslinya dengan *Signal to Noise Ratio* (SNR), dimana nilai hasil uji SNR semakin besar maka menunjukkan sinyal sintetik semakin mirip dengan sinyal asli secara hitungan matematis. Hasilnya diperoleh SNR terbesar yang berbeda-beda untuk setiap sampel taskin dan tasydid, sebagaimana ditunjukkan pada Tabel 1. Hasilnya, terdapat beberapa variasi parameter dengan nilai SNR yang sama besarnya. Pada metode taskin; sampel 1 diperoleh 3 variasi parameter, sampel 2 diperoleh 9 variasi parameter, dan sampel 3 dengan 2 variasi parameter. Sedangkan pada metode tasydid; sampel 1 diperoleh 8 variasi parameter, sampel 2 diperoleh 3 variasi parameter, dan sampel 3 terdapat 12 variasi parameter. Sehingga total parameter dengan nilai SNR terbesar diperoleh 14 parameter untuk taskin dan 23 parameter untuk tasydid.

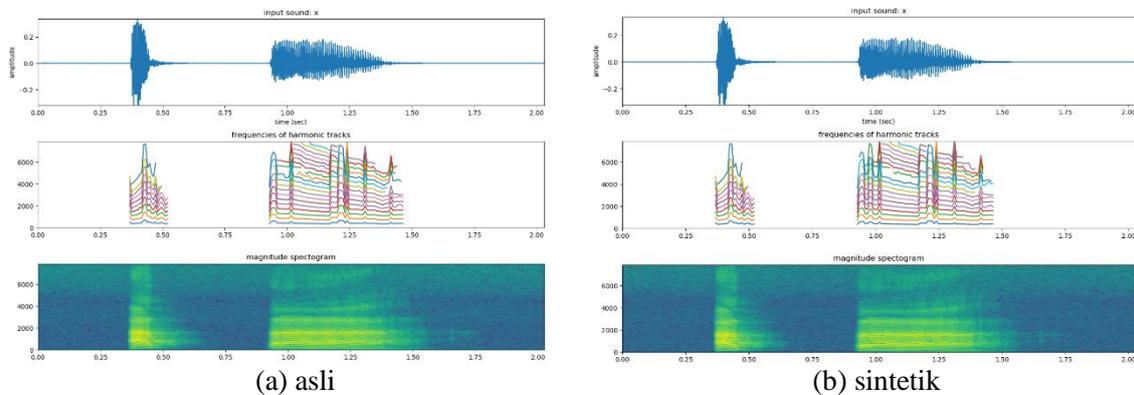


Gambar 3 Spektrogram Sinyal Guru (huruf Ta, taskin sampel 2-parameter 6)

Sejumlah parameter terbaik menurut sistem tersebut kemudian diuji lagi dengan mendengarkan hasil suara sintetik sekaligus diberi penilaian skor secara langsung oleh guru. Dari tahap seleksi ini didapatkan hasil akhir parameter terbaik sebagaimana ditunjukkan pada Tabel 2. Dari rangkuman data hasil eksperimen Tabel 2 dapat disimpulkan bahwa metode taskin dinilai

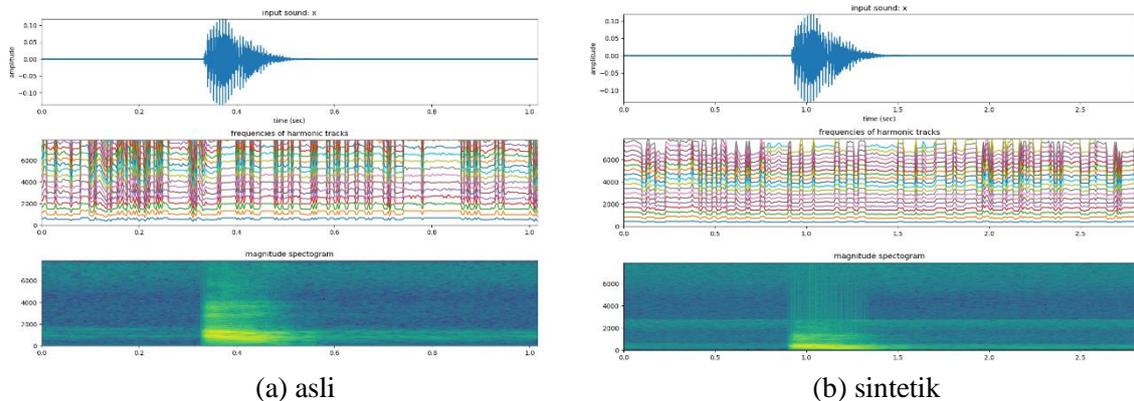
sangat kurang dalam proses sintesis sinyal suara pelafalan huruf hijaiyah. Terutama taskin sampel 3 yang semua hasil sintetiknya sangat buruk dengan skor 1. Sebaliknya, metode tasydid dinilai sangat baik dengan skor 10.

Untuk metode taskin, sampel data yang diambil untuk dibahas adalah huruf ط (Ṭa) dari taskin sampel 2-parameter 6, parameternya yaitu *window* hamming, lebar *window* 800, *threshold* -130 dan *harmonic max* 20. Selain hasil skoring rendah, dari hasil analisis spektrogram juga menunjukkan magnitudo frekuensi yang tidak sepadat sinyal asli sebagaimana ditunjukkan pada Gambar 3. Namun ada keanehan pada spektrogram sinyal sintetik, dimana panjang sinyal yang aslinya 1 detik menjadi 3 detik pada sinyal sintetik. Hasil *tracking* frekuensi harmonik juga terdeteksi sepanjang sinyal pada Gambar 3 yang menunjukkan sinyal asli dan sinyal sintetik. Pada Gambar 3a, sinyal terdeteksi dalam rentang detik ke 0,3 sampai 0,7 namun frekuensi harmonik terbaca dari detik 0 hingga detik 1. Pada Gambar 3b menunjukkan sinyal sintetik berbeda dengan sinyal aslinya, mulai dari panjang sinyal yang terbaca pada detik 0,8—1,8, persebaran frekuensi dalam spektrogram yang tidak sepadat dan sekuat frekuensi sinyal asli, maupun pola frekuensi harmonik juga berbeda. Hal ini menjelaskan mengapa sinyal tersebut memperoleh skor rendah dengan nilai 4.



Gambar 4 Spektrogram sinyal guru (huruf Ta, tasydid sampel 1-parameter 4)

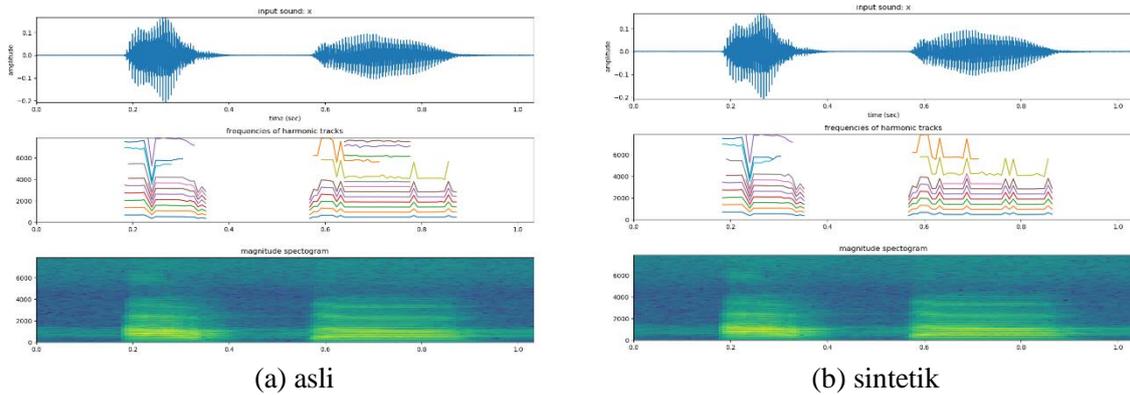
Untuk metode tasydid, sampel data yang dibahas ialah huruf ت (Ta) pada tasydid sampel 1-parameter 4, parameternya yaitu *window* blackman, lebar *window* 600, *threshold* -100 dan *harmonic max* 50. Hasil spektrogram dan pola frekuensi harmonik sinyal tersebut dapat diperhatikan pada Gambar 4 yang secara visual sangat mirip antara sinyal asli dan sinyal sintetik, sehingga spektrogram tersebut menegaskan hasil skoring yang sangat bagus dengan skor 10 untuk sintesis metode tasydid.



Gambar 5 Spektrogram sinyal murid 1 (huruf Ḍad, taskin sampel 2-parameter 6)

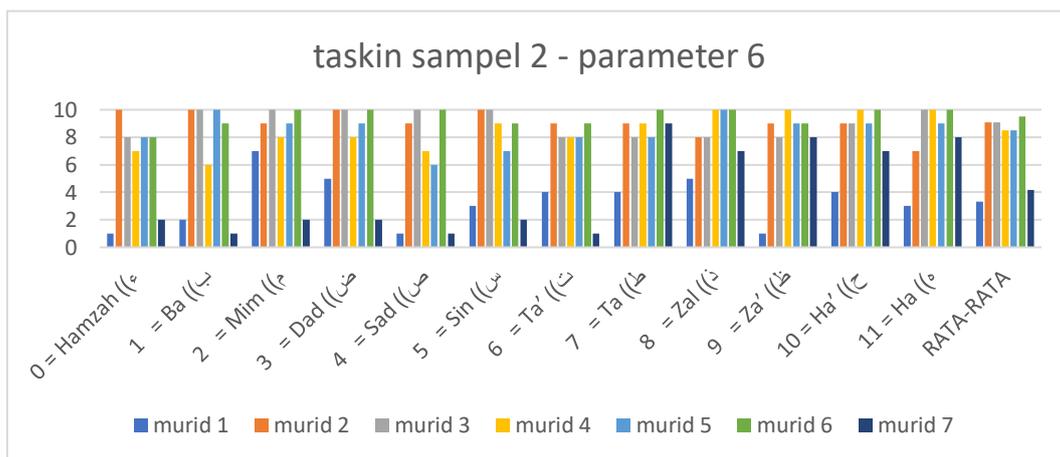
Parameter terbaik yang diperoleh dari tahap pemrosesan sinyal guru kemudian diuji diterapkan pada sinyal pembelajar. Tujuannya ialah untuk menguji kemampuan parameter *Harmonic Plus Residual* (HPR) dalam mensintesis sinyal suara pembelajar yang belum ahli dalam

melafalkan huruf hijaiyah sesuai *makhraj*-nya. Gambar 5 menunjukkan spektrogram sinyal dari data sintetis terbaik metode taskin yang memperoleh skor rata-rata tertinggi di antara ketiga sampel taskin dengan skor 7,4, yaitu taskin sampel 2 parameter 6. Sampel yang ditampilkan adalah data huruf ض (Ḍad) dari murid 1 yang memperoleh skor 5 (skor selengkapnya dapat diperhatikan pada Gambar 7). Hasilnya tidak jauh berbeda dengan hasil pengolahan sinyal guru metode taskin yang sama tidak bagusnya, dari skoring yang rendah (rata-rata tertinggi di antara ketiga sampel namun skornya hanya 7,5), juga dari hasil spektrogram sinyal sintetis yang berbeda dari spektrogram sinyal sintetis (Gambar 4), magnitudo frekuensi yang tidak sepadat dan sekuat frekuensi sinyal asli, dan pola frekuensi harmonik yang berbeda dengan sinyal asli.



Gambar 6 Spektrogram sinyal murid 6 (huruf Ḍad, tasydid sampel 1-parameter 4)

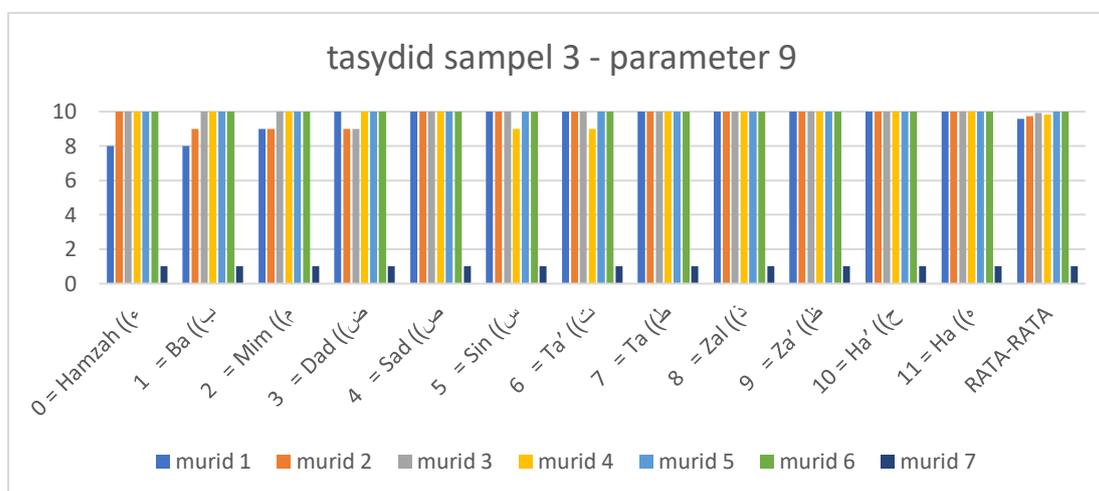
Gambar 6 menunjukkan spektrogram sinyal dari data sintetis terbaik metode tasydid yang memperoleh skor rata-rata tertinggi di antara ketiga sampel tasydid dengan skor 8,5, yaitu tasydid sampel 1 parameter 4. Sampel yang ditampilkan disamakan dengan sampel taskin agar dapat dibandingkan, yaitu data huruf ض (Ḍad) dari murid 1 yang memperoleh skor 9 (skor selengkapnya dapat diperhatikan pada Gambar 8). Hasil spektrogram dan pola frekuensi harmonik sinyal secara visual cukup mirip antara sinyal asli dan sinyal sintetis, sehingga spektrogram tersebut menegaskan hasil skoring yang sangat bagus dengan skor 10 untuk sintesis metode tasydid.



Gambar 7 Grafik Hasil Skoring Sinyal Murid (Parameter Terbaik Metode Taskin)

Untuk perbandingan hasil skor, dikumpulkan hasil skoring tertinggi masing-masing dari metode taskin (Gambar 7) dan tasydid (Gambar 8) yang dianggap sebagai parameter terbaik untuk dianalisis dan dibandingkan. Pada Gambar 7, hasil skoring setiap murid secara rata-rata belum dapat dikatakan sebagai hasil yang bagus. Misalnya murid 1 yang memiliki skoring cukup variatif, dengan skor terendah dengan nilai 1 pada pelafalan huruf Hamzah, namun pada pelafalan huruf

Mim dapat mencapai skor 7, dengan rata-rata skor murid 1 ialah 3. Hasil yang diperoleh dari pengolahan sinyal metode taskin, cukup jauh dibandingkan hasil sintetik sinyal tasydid. Gambar 8 menunjukkan hasil yang sangat bagus dengan rata-rata skor setiap murid ialah 9 atau bahkan 10. Kecuali pada murid 7 yang memang bermasalah pada sinyal aslinya yang saat proses perekaman suara murid 7 terlalu kecil.



Gambar 8 Grafik Hasil Skoring Sinyal Murid (Parameter Terbaik Metode Tasydid)

Dari pengolahan sinyal guru disimpulkan bahwa parameter paling optimal tidak dapat ditentukan, namun metode terbaik untuk sintesis dan untuk belajar melafalkan huruf hijaiyah adalah metode tasydid (*tasydidul harf*). Begitupun dengan hasil pengolahan sinyal murid/pembelajar, sampai pada tahap ini dapat disimpulkan bahwa parameter HPR dapat mensintesis sinyal murid yang belum ahli sekaligus menunjukkan kemampuan setiap murid dalam melafalkan huruf hijaiyah dengan standar suara guru yang telah disintesis, dan metode yang paling baik adalah metode tasydid. Dari Tabel 3 dan Tabel 4, dapat diperhatikan skor sinyal asli metode tasydid lebih baik daripada skor sinyal asli metode taskin, hal ini dianggap sebagai kemudahan metode tasydid dibanding metode taskin sebagai metode belajar pelafalan oleh murid yang belum ahli. Dari Tabel 3 dan Tabel 4 juga dapat diperhatikan skor sinyal sintetik metode tasydid juga lebih baik daripada metode taskin, yang menunjukkan pemodelan HPR lebih baik digunakan dalam mensintesis sinyal dengan metode tasydid daripada metode taskin.

Tabel 3 Perbandingan skor sinyal asli dan sinyal sintetik (taskin sampel 2–parameter 6)

Data	Skor rata-rata	
	Sinyal asli	Sinyal sintetik
Murid 1	9,3	3,3
Murid 2	6,3	9,1
Murid 3	9,9	9,1
Murid 4	9,9	8,5
Murid 5	10,0	8,5
Murid 6	8,9	9,5
Murid 7	6,7	4,2
Rata-rata	8,7	7,5

Dilihat secara keseluruhan dari Tabel 3 dan Tabel 4, dapat disimpulkan bahwa para murid/pembelajar terbilang mudah melafalkan dengan baik dan benar dengan metode tasydid, tapi tidak sebegitu dalam pelafalan huruf dengan metode taskin, khususnya murid 2 dan murid 7. Sedangkan Tabel 5 menunjukkan perbandingan kemampuan pelafalan setiap murid dibandingkan dengan guru.

Tabel 4 Perbandingan skor sinyal asli dan sinyal sintetik (tasydid sampel 1-parameter 4)

Data	Skor rata-rata	
	Sinyal asli	Sinyal sintetik
Murid 1	9,7	9,3
Murid 2	7,4	9,6
Murid 3	9,8	10,0
Murid 4	9,5	9,8
Murid 5	10,0	10,0
Murid 6	9,4	10,0
Murid 7	9,6	1,0
Rata-rata	9,3	8,5

Tabel 5 Nilai rata-rata skoring setiap sampel per subjek/penutur (suara asli)

Taskin								
Sampel	Guru	Murid1	Murid2	Murid3	Murid4	Murid5	Murid6	Murid7
1	10,0	10,0	7,0	10,0	9,9	10,0	10,0	7,8
2	10,0	9,3	6,3	9,9	9,9	10,0	8,9	6,7
3	10,0	9,5	7,7	10,0	9,4	10,0	8,5	6,3
Tasydid								
Sampel	Guru	Murid1	Murid2	Murid3	Murid4	Murid5	Murid6	Murid7
1	10,0	9,7	7,4	9,8	9,5	10,0	9,4	9,6
2	10,0	9,9	9,7	10,0	10	10,0	9,9	8,8
3	10,0	10,0	9,0	10,0	9,9	10,0	9,9	9,3

4. KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, dapat disimpulkan bahwa pemodelan harmonik, dalam penelitian ini adalah Harmonic Plus Residual (HPR), dapat digunakan dalam mensintesis pelafalan huruf hijaiyah sesuai *makhraj*-nya. Namun terdapat beberapa catatan penting, seperti metode taskin yang dinilai kurang baik untuk dijadikan bahan untuk membangkitkan sinyal sintetik. Hal ini dapat dilihat dari hasil skoring secara umum yang menunjukkan kualitas sinyal tidak konsisten, terlepas dari kemampuan murid berbeda-beda dalam pelafalan. Sedangkan metode tasydid terbukti lebih konsisten dengan hasil skoring secara umum sempurna dengan hanya beberapa huruf saja oleh murid yang diberi nilai 9. Selain itu, metode tasydid dianggap lebih mudah untuk diukur ketepatan pelafalannya oleh guru serta lebih mudah ditiru dan dipraktikkan dengan baik oleh murid. Dilihat dari hasil sintetis metode tasydid yang konsisten, tidak dapat ditentukan kombinasi parameter mana yang terbaik, karena setiap jenis jendela, besar jendela, nilai threshold hingga nilai maksimum harmonik berbeda-beda namun semua hasil sintetiknya dapat dikatakan sama bagusnya. Oleh karena parameter optimal tidak dapat ditetapkan mana yang paling optimal, maka disimpulkan bahwa metode belajar pelafalan huruf hijaiyah yang baik secara objektif ialah metode *tasydidul harf*.

DAFTAR PUSTAKA

- [1] Yulyawati, "Implementasi Metode At-Tibyan dalam Pembelajaran Membaca Al-Qur'an untuk Anak Usia Dini," Universitas Pendidikan Indonesia, Bandung, 2016.
- [2] A. N. Wahidah, M. Suriazalmi, M. Niza, H. Rosyati, N. Faradila, A. Hasan, A. Rohana and Z. Farizan, "Makhraj Recognition Using Speech Processing," 2011.
- [3] E. S. Wahyuni, "Arabic Speech Recognition using MFCC Feature Extraction and ANN Classification," *2017 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE)*, pp. 22-25, 2017.

- [4] H. A. Rahim, *Tajwid Praktis As-Shafa*, Makassar: Lembaga Penerbitan Universitas Hasanuddin, 2018.
- [5] D. Pratiwi, "Analisis Kesulitan Belajar Membaca Al-Qur'an pada Siswa Kelas VIII SMP Muhammadiyah 1 Surakarta Tahun Pelajaran 2016/2017," Universitas Muhammadiyah Surakarta, Surakarta, 2017.
- [6] M. Subali, M. Andriansyah and C. Sinambela, "Analisis Frekuensi Dasar dan Frekuensi Formant dari Fonem Huruf Hijaiyah untuk Pengucapan Makhraj dengan Metode DTW," *Prosiding PESAT (Psikologi, Ekonomi, Sastra, Arsitektur & Teknik Sipil)*, vol. 6, pp. 60-73, 2015.
- [7] M. Subali, M. Andriansyah and C. Sinambela, "Pengucapan Makhraj dari Unit Bunyi Terkecil Huruf Hijaiyah berdasarkan Frekuensi Dasar dan Frekuensi Formant untuk Media Pembelajaran Alquran," *ALQALAM*, vol. 32, no. 2, pp. 284-308, 2015.
- [8] T. AlTalmas, W. Sediono, N. Hashim, S. Ahmad and S. Khairuddin, "Analysis of Two Adjacent Articulation Quranic Letters Based On MFCC and DTW," *7th International Conference on Computer and Communication Engineering (ICCCE)*, pp. 187-191, 2018.
- [9] F. B. Ali and D. S. Larbi, "A Long Term Harmonic plus Noise Model for Narrow-Band Speech Coding at Very Low Bit-Rates," *International Conference on Telecommunications and Signal Processing (TSP)*, vol. 40th, pp. 372-376, 2017.
- [10] R. Rajan M, "Singing Voice Synthesis System for Carnatic Music," *International Conference on Signal Processing and Integrated Networks (SPIN)*, vol. 5th, pp. 831-835, 2018.
- [11] A. Ferreira, J. Silva, F. Brito and D. Sinha, "Impact of A Shift-Invariant Harmonic Phase Model in Fully Parametric Harmonic Voice Representation and Time/Frequency Synthesis," *International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 701-705, 2020.
- [12] Serra, Xavier, "Coursera: Audio Signal Processing for Music Applications," 2019. [Online]. Available: <https://www.coursera.org/learn/audio-signal-processing/home/welcome>. [Accessed 1 January 2021].
- [13] D. K. Pramesthi, "Cerita Nadytha Zildjian," 2011. [Online]. Available: <http://nadythazildjian.blogspot.com/2011/07/sekilas-tentang-mel-frequency-cepstral.html>. [Diakses 14 Juni 2020].
- [14] P. K. Sari, K. Priandana and A. Buono, "Perbandingan Sistem Perhitungan Suara Tepuk Tangan dengan Metode Berbasis Frekuensi dan Metode Berbasis Amplitudo," *Jurnal Ilmu Komputer Agri-Informatika*, vol. 2, no. 1, pp. 29-37, 2013.
- [15] T. Nasution, "Metoda Mel Frequency Cepstrum Coefficients (MFCC) untuk mengenali Ucapan pada Bahasa Indonesia," *Jurnal Sains dan Teknologi Informasi*, vol. 1, no. 1, pp. 22-31, 2012.
- [16] Y. C. H. Siki, "Perbandingan Berbagai Waktu - Frekuensi Musik Gong Timur Menggunakan Short Time Fourier Transform dan Continous Wavelet Transform," Institut Teknologi Sepuluh Nopember, Surabaya, 2015.
- [17] Heriyanto, S. Hartati and A. E. Putra, "Ekstraksi Ciri Mel Frequency Cepstral Coefficient (MFCC) dan Rerata Coefficient untuk Pengecekan Bacaan Al-Qur'an," *Telematika*, vol. 15, no. 02, pp. 99-108, 2018.
- [18] D. K. Putra, I. I. Triasmoro and R. D. Atmaja, "Simulasi dan Analisis Speaker Recognition Menggunakan Metode Mel Frequency Cepstrum Coefficient (MFCC) dan Gaussian Mixture Model (GMM)," *e-Proceeding of Engineering*, vol. 4, no. 2, pp. 1766-1772, 2017.