

Text to Speech untuk Bahasa Arab Menggunakan Perangkaian Diphone

(Text to Speech for Arabic Using Diphone Concatenation)

Achmad Fauzan¹, Sri Hartati²

¹Teknik Informatika, Universitas Muhammadiyah Purwokerto

²Pascasarjana Program Studi Ilmu Komputer UGM, Yogyakarta

¹mr.achmadfauzan@gmail.com

²shartati@ugm.ac.id

Abstrak – Bahasa Arab di Indonesia yang mayoritas penduduknya beragama Islam, lebih banyak digunakan untuk mempelajari Al-Qur'an maupun hadist. Namun berdasarkan data BPS tahun 2015, sebanyak 54 persen muslim di Indonesia masih buta aksara Arab.

Penelitian ini bertujuan merancang dan menyusun sistem *text to speech* yang dapat membangkitkan suara pengucapan untuk teks masukan berbahasa Arab. Rangkaian kode fonem dari teks masukan diperoleh berdasarkan ketentuan (aturan) pengucapan dasar bahasa tersebut. Aturan-aturan yang diterapkan adalah hukum bacaan nun sukun dan tanwin, vokal rangkap, bacaan panjang (madd), tanwin, tasydid, ta marbutah, dan sukun. Rangkaian kode fonem selanjutnya dibuat berpasang-pasangan (diphone). File-file suara diphone dirangkai menjadi satu pada proses pembangkitan suara akan menghasilkan sebuah file suara utuh.

Pengujian dilakukan dengan melibatkan 13 responden yang memahami ilmu tajwid. Pada pengujian menggunakan teks masukan hukum bacaan idghaam bilaagunnah, idghaam bighunnah, iqlab, ikhfa' adna, ta marbutah, dan sukun, 100% responden menilai suara yang dihasilkan sudah tepat. Sedangkan dengan teks masukan hukum bacaan idzhaar, responden yang menilai suara dihasilkan sudah tepat mencapai 92%. Persentase lebih rendah terdapat pada hukum bacaan vokal rangkap (diftong), ikhfa' ausath, tasydid, ikhfa' aqrab, dan bacaan panjang (madd) yang masing-masing memperoleh 81%, 77%, 73%, 69%, dan 65%.

Kata kunci – *text to speech*, bahasa Arab, perangkaian diphone, finite state automata.

Abstract – Arabic in Indonesia, which is predominantly Muslim, is more widely used to study the Qur'an and hadith. However, based on data from BPS in 2015, as many as 54 percent of Muslims in Indonesia are still illiterate Arabic.

This study aims to design and compose a text to speech system that can generate sound pronunciation for Arabic input text. The phoneme code sequence of the input text is obtained under the language basic pronunciation rules. The rules applied are the law of reading nun sukun and tanwin, double vowels, long reading (madd), tanwin, tasydid, ta marbutah, and breadfruit. The sequence of phoneme codes is then made in pairs (diphone). The sound files diphone assembled into one in the sound generation process will produce a sound file intact.

The test was conducted by involving 13 respondents who understand the science of tajwid. In the test using the text of the legal input reading idghaam when amar, idghaam bighunnah, iqlab, ikhfa' adna, ta marbutah, and breadfruit, 100% of respondents assess the sound produced is correct. While the text input legal reading idzhaar, respondents who assess the resulting sound has reached exactly 92%. The lower percentage is found in the law of diphthong (diphthong), ikhfa' ausath, tasydid, ikhfa' aqrab, and long reading (madd), 81%, 77%, 73%, 69%, and 65% respectively.

Keywords – *text to speech*, Arabic Language, diphone concatenation, finite state automata.

I. PENDAHULUAN

Indonesia merupakan negara dengan penduduk muslim terbesar di dunia, sehingga bahasa Arab tidaklah terasa asing, karena digunakan dalam Al-Qur'an sebagai kitab suci agama Islam. Namun hasil survey BPS pada tahun 2015 menyebutkan 54 persen penduduk muslim di Indonesia belum dapat membaca tulisan Arab seperti dalam Al Qur'an.

Perkembangan ilmu pengetahuan di bidang komputer cerdas memungkinkan dikembangkannya alat untuk membantu mengetahui bacaan teks berbahasa Arab dengan cara mengkonversi teks tersebut menjadi

suara pengucapannya. Sistem semacam ini dikenal dengan sistem *text to speech* (TTS). Berbeda dengan *talking machine* biasa, TTS memiliki kelebihan dalam keotomatisannya mengucapkan kata-kata baru.

Terdapat tiga teknik perangkaian suara dalam aplikasi TTS. Teknik pertama adalah perangkaian bunyi suku kata yang bekerja dengan cara memenggal tiap suku kata dari teks masukan dan merangkai tiap suara dari masing-masing penggalan suku kata tersebut. Sebuah sistem *text to speech* dalam bahasa Arab yang pernah dikembangkan menggunakan bunyi suku kata dan modul *Microsoft Speech API*, menghasilkan 80% kata dan 70% kalimat yang berhasil diucapkan dengan baik [1]. Teknik kedua adalah perangkaian bunyi fonem yang bekerja dengan cara mengubah teks masukan ke dalam bentuk fonem, kemudian menyatukan tiap bagian suara sesuai fonemnya menjadi satu suara ucapan utuh. Pemilihan fonem sebagai unit ucapan dapat menghemat penggunaan memori secara signifikan. Namun pada pengujian yang dilakukan Hanane [2] dan Tritoasmoro [3], diketahui bahwa teknik ini belum dapat menghasilkan kualitas ucapan yang baik untuk masukan teks yang relatif panjang. Sedangkan teknik ketiga adalah perangkaian bunyi dari dua buah fonem (*diphone concatenation*) yang bekerja dengan cara mengkonversi teks masukan ke dalam bentuk pasangan fonem dan merangkai suara dari masing-masing pasangan fonem tersebut [4]. Teknik ini dapat menghasilkan suara yang lebih baik dibandingkan teknik-teknik sebelumnya [5]. Kualitas suara pengucapan hasil proses pembangkitan suara dapat ditingkatkan dengan menambahkan model intonasi dengan mempertimbangkan struktur kalimat, pola intonasi dari contoh kalimat, dan kaidah umum pengucapan dalam bahasa yang digunakan [6]. Alat bantu untuk membangkitkan suara dengan metode ini diantaranya yaitu Mbrola, yang juga sudah menyediakan basis data suara untuk bahasa Arab, namun terbatas penggunaannya karena tidak dapat dimanfaatkan selain dengan *tool* tersebut [7].

Penerapan sistem *text to speech* akan sangat tergantung pada jenis bahasa yang digunakan. Pengucapan suatu teks dalam bahasa Indonesia akan berbeda dengan bahasa lainnya, seperti bahasa Inggris atau bahasa Arab [2]. Pada bahasa Arab, aturan-aturan dasar pengucapan suatu teks/tulisan diistilahkan dengan tajwid. Penerapan tajwid sangat penting di dalam membangkitkan suara pengucapan, karena perbedaan pengucapan yang semestinya dapat menyebabkan perubahan makna [8]. Tanda baca yang terdapat pada suatu teks juga akan mempengaruhi cara pengucapan bacaannya [9]. Namun masih sangat sedikit penelitian-

penelitian tentang *text to speech* dalam bahasa Arab yang membahas dan menerapkan aturan tajwid. Oleh karena itu penelitian ini diharapkan dapat membantu dalam memahami bacaan teks berbahasa Arab, terutama bagi pengguna yang belum bisa membaca tulisan Arab. Selain itu juga dapat menjadi salah satu alat pendamping dalam proses pembelajaran bahasa Arab.

II. METODE PENELITIAN

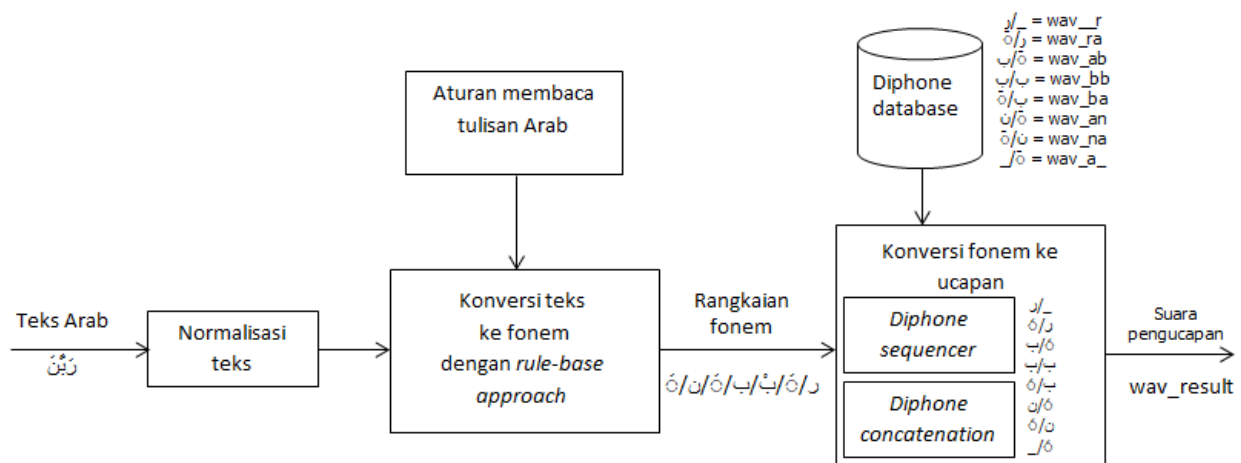
A. Gambaran Umum

Penerapan sistem *text to speech* dalam bahasa Arab menggunakan *input* berupa teks bahasa Arab dengan *output* berupa suara yang diucapkan sesuai aturan pengucapannya. *Input* akan melalui proses normalisasi untuk memastikan teks masukan hanya berisi huruf dan karakter yang diizinkan. Teks *input* ter-normalisasi selanjutnya akan dikonversi ke dalam fonem dengan pendekatan berbasis aturan (*rule-based approach*) yaitu dengan menyesuaikan pengucapan kata berdasarkan ejaannya sehingga memenuhi aturan atau tata cara pembacaan tulisan bahasa Arab.

Rangkaian kode fonem selanjutnya akan dibagi kedalam bentuk pasangan dua fonem (*diphone*) yaitu antara fonem ke *i* dan fonem ke *i+1* hingga terbentuk himpunan kode *diphone*. Kode *diphone* tersebut akan digunakan sebagai acuan dalam pengambilan file suaranya. Suara ucapan dari teks masukan akan dihasilkan dari penggabungan segmen-segmen suara *diphone* yang diambil (*diphone concatenation*). Proses sistem *text to speech* dalam bahasa Arab ditunjukkan pada Gambar 1.

B. Normalisasi Teks

Teks masukan yang diterima sistem dimungkinkan memiliki sejumlah karakter yang mempengaruhi pengucapan, ataupun memiliki pengucapannya sendiri. Sehingga setiap karakter tersebut akan diproses dengan cara mengubahnya ke dalam bentuk pengucapannya. Masukan sistem (*input*) yang digunakan berupa karakter Arab dengan mengacu pada *Arabic Unicode Character Table* yaitu tabel karakter unicode untuk bahasa Arab dengan kode antara 0600-06FF dalam *hexadecimal*. Teks masukan terlebih dahulu akan dikonversi ke dalam bentuk kode *unicode*. Proses normalisasi teks bertujuan untuk menghilangkan karakter yang tidak termasuk dalam daftar karakter yang diizinkan oleh sistem (karakter ilegal) dan mengubah karakter angka menjadi penulisan ejaannya.



Gambar 1. Sistem text to speech dalam bahasa Arab

C. Konversi Teks ke Fonem

Tahapan ini bertujuan untuk mendapatkan ketentuan pengucapan dasar dari setiap kata teks yang telah dinormalisasi. Konversi huruf menjadi fonem biasanya dilakukan dengan dua pendekatan dasar yaitu pendekatan berdasarkan kamus, dan pendekatan menurut aturan (*rule-based approach*). Pada sistem ini, aturan-aturan yang digunakan didasarkan pada aturan/cara membaca tulisan berbahasa Arab (tajwid) [8]. Aturan-aturan tersebut diantaranya yaitu:

1) *Hukum Nun Sukun dan Tanwin*. Terdapat beberapa hukum di dalamnya, yaitu idzhaar, idghaam bilaaghunnah, idghaam bighunnah, iqlab, ikhfa' aqrab, ikhfa' ausath, dan ikhfa' adna. Idzhar berarti apabila terdapat nun sukun (نْ) atau tanwin (نً) bertemu dengan salah satu huruf ا, ح, خ, ع, غ, atau و maka harus dibaca dengan terang dan jelas, tidak berdengung. Idghaam bilaaghunnah berarti apabila terdapat nun sukun (نْ) atau tanwin (نً) bertemu dengan salah satu huruf ل atau ر maka pengucapan bacaannya dimasukkan pada huruf berikutnya tanpa disertai dengung. Idghaam bighunnah berarti apabila terdapat nun sukun (نْ) atau tanwin (نً) bertemu dengan salah satu huruf و, ن, م, or ي maka pengucapan bacaannya dimasukkan pada huruf berikutnya disertai dengung. Iqlab berarti apabila nun mati (نْ) atau tanwin (نً) bertemu dengan huruf ب maka dibaca dengan cara menukar huruf nun mati atau tanwin tersebut menjadi bunyi huruf mim dengan merapatkan dua bibir serta mendengung. Ikhfa' aqrab berarti apabila nun mati (نْ) atau tanwin (نً) bertemu dengan salah satu huruf ط, د, or ث maka ketika menyuarakan nun mati, ujung lidah hampir menyentuh pangkal dua buah gigi atas sesuai posisi makhrojnya. Ikhfa' ausath berarti apabila nun sukun (نْ) atau tanwin (نً) bertemu dengan salah satu huruf ض, ص, ش, س, ز, ج, ذ, or ث, atau ف maka ketika mengucapkan nun mati, sikap lidah/bibir dipersiapkan menempati makhroj huruf yang dihadapi. Ikhfa' adna berarti apabila nun mati (نْ) atau tanwin (نً) bertemu dengan salah satu huruf ك or ق maka cara pengucapannya menjadi seperti “ng”. Dilanjutkan dengan penerapan aturan vokal rangkap (diftong). Vokal rangkap merupakan bunyi pengucapan dari dua buah vokal berdekatan. Ada dua vokal rangkap pada bahasa Arab yaitu vokal يَ yang berbunyi ai dan وُ yang berbunyi au. Bunyi vokal ai akan teridentifikasi apabila ditemukan adanya karakter *fathah* (َ) yang diikuti dengan karakter *ya* (ي) dan *sukun* (ْ). Sedangkan bunyi vokal au akan teridentifikasi apabila ditemukan adanya karakter *fathah* (َ) yang diikuti dengan karakter *wau* (و) dan *sukun* (ْ).

2) *Mad (bacaan panjang)*. Bacaan panjang disebut juga dengan Mad. Aturan bacaan panjang yang akan diterapkan adalah mad thobi'i. Aturan ini membuat suatu bacaan harus dibaca panjang hingga dua harakat (dua ketukan). Aturan ini diterapkan pada tiga buah kondisi yaitu karakter *fathah* (َ) yang diikuti alif (ا), karakter *dhomah* (ُ) yang diikuti *wau* (و) dan *sukun* (ْ), dan karakter *kasroh* (ِ) yang diikuti *ya* (ي) dan *sukun* (ْ).

3) *Tanwin*. Harakat tanwin menyatakan bahwa huruf dengan harakat tersebut diucapkan layaknya bertemu dengan huruf nun mati. Tanwin terdiri dari tiga buah harakat, yaitu fathah tain (وَّ dibaca وَّ), kasrah tain (ِ dibacaِ), dan dlommah tain (ُ dibacaُ).

4) *Tasydid (konsonan ganda)*. Harakat tasydid (ّ) melambangkan penekanan pada suatu konsonan dan dibaca sebagai konsonan ganda dimana konsonan pertama dibaca sebagai huruf mati dan kedua dibaca harakat aslinya.

5) *Ta Marbutah*. Ta Marbutah adalah varian dari huruf ta (ت) yang melambangkan fonem /ت/ atau /ة/. Terdapat dua kondisi untuk pengucapan ta marbutah, yaitu diucapkan seperti huruf ت bila diikuti dengan vokal hidup (fathah, kasroh, atau dhomah) dan diucapkan seperti huruf ة bila diikuti sukun.

6) *Sukun*. Karakter sukun melambangkan suatu huruf yang dibaca mati. Apabila ditemukan huruf yang diikuti karakter sukun di belakangnya, maka untuk mematikan huruf tersebut cukup dengan menghilangkan karakter sukun (fonem sukun).

D. Konversi Fonem ke Ucapan

Konversi fonem ke ucapan dilakukan dalam beberapa tahap. Teks masukan yang telah diubah menjadi rangkaian fonem, akan terlebih dahulu dibuat berpasangan. Pasangan fonem yang disebut dengan diphone akan digunakan untuk mencari file suaranya. File-file suara yang ditemukan akan dirangkai (*concatenation*) menjadi sebuah file suara. File suara hasil perangkaian tersebut merupakan suara pengucapan dari teks masukan.

1) *Diphone sequencer*. Masukan yang akan diproses oleh modul *diphone sequencer* adalah rangkaian kode fonem hasil proses konversi teks ke fonem. Rangkaian kode fonem akan dibagi menjadi bentuk pasangan dua fonem (*diphone*) yaitu antara fonem ke *i* dan fonem ke *i+1* hingga terbentuk himpunan kode *diphone*. Masing-masing kode *diphone* tersebut akan digunakan sebagai acuan dalam pengambilan file suaranya.

2) *Diphone Database*. Basis data *diphone* merupakan kumpulan data suara pengucapan dua buah fonem. File suara diperoleh melalui proses perekaman dari bunyi *diphone* tersebut. File suara diletakkan di sebuah direktori tertentu dan diberi nama berupa gabungan kode desimal *unicode* dari kedua fonem tersebut. Jenis file suara yang digunakan yaitu file berekstensi *.wav*.

3) *Diphone Concatenation*. Rangkaian kode *diphone* akan digunakan untuk memanggil file suara *diphone* yang telah dipersiapkan sebelumnya. Pemanggilan file dilakukan berdasarkan pengalamatan (*path*) direktori disertai dengan nama masing-masing file yang juga merupakan kode *diphone* yang digunakan.

File-file suara akan dipanggil secara berurutan. Format suara berupa *.wav* ditentukan dari format suara file yang pertama kali dipanggil. File suara pertama akan dirangkai dengan file suara kedua. Hasil perangkaian keduanya kemudian akan dirangkai lagi dengan file ketiga. Proses tersebut dilakukan hingga semua file suara dirangkai dan menjadi satu suara utuh.

III. HASIL DAN PEMBAHASAN

A. Implementasi Sistem

Sistem *text to speech* dibangun menggunakan teknologi berbasis web dengan bahasa pemrograman JSP, Java, HTML, dan javascript. Masukan sistem (*input*) yang digunakan berupa karakter Arab dengan mengacu pada *Arabic Unicode Character Table* yaitu tabel karakter unicode untuk bahasa Arab dengan kode antara 0600-06FF dalam hexadecimal serta menyesuaikan dengan jenis karakter yang tersedia pada tombol *keyboard* sistem operasi *windows* dengan pengaturan *keyboard layout Arabic (101)* dan *Central Kurdish*. Batasan karakter tersebut diterapkan pada daftar karakter yang diterima oleh sistem.

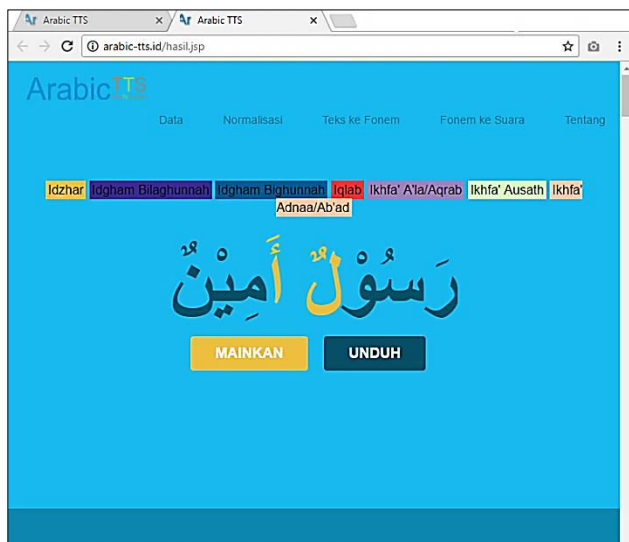
Sistem *text to speech* untuk Bahasa Arab dibangun dengan teknologi web sehingga dapat diakses menggunakan mesin perambah internet (*browser*). Halaman utama yang ditampilkan ditunjukkan pada Gambar 2.

Teks *input* akan diproses oleh server dan hasil pemrosesan akan ditampilkan pada halaman hasil identifikasi, seperti ditunjukkan pada Gambar 3.

Halaman hasil pemrosesan menampilkan teks Arab dengan keterangan warna yang menunjukkan letak dan jenis hukum bacaannya. Selain itu juga terdapat tombol "MAINKAN" untuk memutar file suara pengucapan teks masukan, dan tombol "UNDUH" untuk menyimpannya ke dalam komputer. Pemrosesan teks masukan menjadi suara ditunjukkan dalam beberapa bagian yang meliputi data teks masukan, normalisasi teks, konversi teks ke fonem, dan konversi fonem ke suara.



Gambar 2. Halaman utama sistem berisi kolom input



Gambar 3. Halaman hasil pemrosesan

B. Pengujian Ketepatan Hasil Pengucapan

Ketepatan pengucapan sesuai identifikasi hukum bacaan (tajwid) diuji dengan melakukan survey yang melibatkan responden dengan kemampuan membaca dan mengerti hukum bacaan (tajwid) pada tulisan bahasa Arab. Terdapat 13 jenis pertanyaan yang masing-masing pertanyaan akan mewakili sebuah pengujian terhadap hukum bacaan (tajwid) tertentu. Responden memilih jawaban “Ya”, “Tidak”, atau “Ragu-ragu” yang masing-masing pilihan jawaban tersebut memiliki bobot 1, 0, dan 0,5. Nilai akhir (NA) dapat diketahui dengan menerapkan perhitungan seperti pada persamaan 1.

$$NA_p = \frac{\text{jumlah bobot pertanyaan}_p}{\text{jumlah responden}} \times 100\% \quad (1)$$

NA_p merupakan nilai akhir suatu pertanyaan (dalam bentuk persen), dan *jumlah bobot pertanyaan_p* merupakan jumlah total bobot nilai suatu pertanyaan. Hasil perhitungan nilai akhir (NA) keseluruhan ditunjukkan pada TABEL I.

TABEL I
PENGUJIAN KETEPATAN HASIL PENGUCAPAN
SESUAI TAJWID

No	Jenis hukum bacaan (tajwid)	Persentase ketepatan pengucapan
1	Idzhaar	92%
2	Idghaam bilaaghunnah	100%
3	Idghaam bighunnah	100%
4	Iqlab	100%
5	Ikhfa' aqrab	69%
6	Ikhfa' ausath	77%
7	Ikhfa' adna	100%
8	Vokal rangkap (diftong)	81%
9	Bacaan panjang (madd)	65%
10	Tanwin	81%
11	Konsonan rangkap (tasydid)	73%
12	Ta marbutah	100%
13	Sukun	100%

Hasil pengujian pada TABEL I menunjukkan teks masukan dengan hukum bacaan idghaam bilaagunnah, idghaam bighunnah, iqlab, ikhfa' adna, ta marbutah, dan sukun, suara yang dihasilkan dinilai sudah tepat oleh 100% responden. Sedangkan dengan teks masukan hukum bacaan idzhaar, responden yang menilai suara dihasilkan sudah tepat mencapai 92%. Persentase lebih rendah terdapat pada hukum bacaan vokal rangkap (diftong), ikhfa' ausath, tasydid, ikhfa' aqrab, dan bacaan panjang (madd) yang masing-masing memperoleh 81%, 77%, 73%, 69%, dan 65%. Hal ini bisa disebabkan proses perekaman suara *diphone* yang digunakan belum tepat.

IV. PENUTUP

A. Kesimpulan

Berdasarkan hasil penelitian diperoleh hasil kesimpulan:

1) Hasil pengujian menggunakan teks masukan hukum bacaan idghaam bilaagunnah, idghaam bighunnah, iqlab, ikhfa' adna, ta marbutah, dan sukun, menunjukkan 100% responden menilai suara yang

dihasilkan sudah tepat. Sedangkan dengan teks masukan hukum bacaan idzhaar hanya 92% responden. Persentase lebih rendah terdapat pada hukum bacaan vokal rangkap (diftong), ikhfa' ausath, tasydid, ikhfa' aqrab, dan bacaan panjang (madd) yang masing-masing memperoleh 81%, 77%, 73%, 69%, dan 65%.

2) Persentase tinggi menunjukkan suara pengucapan yang dihasilkan oleh sistem telah sesuai dengan hukum bacaan (tajwid)-nya.

3) Hasil penilaian (persentase) yang rendah disebabkan terjadinya distorsi akibat ketidakberlanjutan pada titik perangkaian, yang dapat dikurangi dengan memperbaiki suara *diphone* atau menerapkan suatu metoda untuk memperhalus sinyal ucapan.

B. Saran

Saran untuk pengembangan pada penelitian selanjutnya:

1) Perlu dilakukan percobaan dengan beragam teknik perekaman suara dan pengolahan sinyal suara sehingga menghasilkan kualitas suara *diphone* yang lebih baik.

2) Mencoba beragam teknik perangkaian suara *diphone* seperti PSOLA (*Pitch Synchronous Overlap Add*) dan pengembangannya, atau WSOLA (*Waveform Similarity based Overlap Add*).

3) Aplikasi yang dibangun dapat dikembangkan lebih lanjut pada perangkat *mobile* dengan kemampuan pengenalan tulisan Arab dari sebuah gambar yang diambil melalui kamera (*capture*).

UCAPAN TERIMA KASIH

Ucapan terima kasih yang sebesar-besarnya penulis sampaikan kepada segenap pimpinan Universitas Muhammadiyah Purwokerto atas beasiswa studi lanjut S2 dan bantuan penelitian tesis yang telah diberikan.

DAFTAR PUSTAKA

- [1] Jayousi, S.Q.M., 2007, *Arabic Text to Speech Synthesizer*, Tesis, Kuala Lumpur: University of Malaya.
- [2] Hanane, T., Maamar, M., dan Hamid, A., 2014, *TTS-SA (A Text-to-Speech System based on Standard Arabic)*, IEEE, ISBN: 978-1-4799-3724-0.
- [3] Tritasmoro, I.I., 2006, *Text to Speech Bahasa Indonesia Menggunakan Concatenation Synthesizer Berbasis Fonem*, *Seminar Nasional Sistem dan Informatika*, Bali.
- [4] Sutarman, 2006, *Sistem Text to Speech Bahasa Indonesia dengan Menggunakan Diphone Synthesis*, Tesis, Yogyakarta: UGM.
- [5] Hamad, M. dan Hussain, M., 2011, *Arabic Text-To-Speech Synthesizer*, *IEEE Student Conference on Research and Development*, Sudan.
- [6] Suyanto, Y., 2014, *Penentuan Intonasi Secara Otomatis dalam Sintesis Ucapan Bahasa Indonesia*, *Disertasi*, Yogyakarta: UGM.
- [7] Chabchoub, A. & Cherif, A., 2011, *An Automatic Mbrola Tool for High Quality Arabic Speech Synthesis*, *International Journal of Computer Application*, 36, 1, 0975-8887.
- [8] Humam, A., 2005, *Cara Cepat Belajar Tajwid Praktis*, Team Tadarus Angkatan Muda Masjid & Mushola, Yogyakarta.
- [9] Rebai, I. dan Benayed, Y., 2015, *Text-to-Speech Synthesis System with Arabic Diacritic Recognition System*, *Elsevier*, 43-60, 0885-2308.