

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kebutuhan akan teknologi dapat memenuhi tuntutan zaman semakin meningkat. Teknologi dikembangkan untuk lebih mempermudah setiap kegiatan manusia. Maka kecepatan dan keakuratan diutamakan untuk mengimbangi setiap aktivitas manusia.

Ucapan merupakan tindakan memproduksi suara melalui variasi tekanan udara yang dipancarkan oleh sistem articulator. Sementara *speech synthesis* adalah sistem yang diproduksi atau buatan manusia melalui mesin (komputer). Oleh sebab itu proses tersebut dapat digolongkan dalam tiga hal yang saling berkaitan, antara lain bagaimana mesin tersebut membaca text yang masuk, bagaimana mesin tersebut dapat mengeluarkan suara, serta unsur-unsur lain yang saling berkaitan agar dapat menjadi mesin tersebut bisa membaca dan berbicara. (Sudirman Melangi, 2018) .

Berbagai upaya pengembangan untuk menghasilkan kualitas suara yang lebih alami dalam *speech synthesis* dengan mengaitkan antar komputer dan bahasa komputasi. Komputer sebagai salah satu pengolahan data dapat menghasilkan *speech synthesis* untuk berbagai macam kegunaan praktis. Pengembangan kemampuan komputer mengubah teks menjadi suara tersebut selanjutnya dikenal sebagai *Text-to-Speech System*.(Sudirman Melangi, 2018). *Text-to-Speech System* (TTS) adalah sistem yang mengubah teks input menjadi sinyal suara yang disintesis. Sistem TTS

tradisional, yang disebut *unit-selection-based* atau *concatenation-based* TTS, menghasilkan sinyal suara, atau unit, dari database suara, berdasarkan teks input dan kemudian menggabungkan fragmen tersebut. (Dhiaulhaq et al., 2022)

Time Domain PSOLA merupakan suatu teknik untuk memodifikasi *pitch* sinyal suara yang dilakukan dengan mengubah durasi sinyal suaranya, memajang untuk mengurangi *pitch* dan memedekkan untuk meningkatkan *pitch*. Tapi dalam PSOLA gelombang suara dibagi terlebih dahulu menjadi beberapa segmen tumpang tindih kecil, kemudian segmen dipindahkan lebih dekat atau terpisah tergantung apakah akan menambah atau mengurangi *pitch*. *Time Domain PSOLA* berfungsi untuk mensinkronkan dekomposisi *pitch* sinyal suara dengan *periodenya* ketika proses *overlapping*, yang sebelumnya dilakukan *padding* terlebih dahulu.

Terdapat beberapa tantangan dalam penelitian di bidang sintesis ucapan yaitu kejelasan (*intelligibility*), kealamian (*naturalness*), efektivitas biaya (*cost-effectiveness*), dan ekspresifitas (*expressivity*). Di bidang industri tidak hanya mengharapkan kejelasan, kealamian, tetapi juga efektivitas biaya. Diharapkan sintesis ucapan hanya membutuhkan beberapa mega byte untuk menyimpan (*small footprint*), kapasitas kecil CPU (*small CPU*), mudah dikembangkan ke bahasa dan kemungkinan menciptakan suara baru secepat mungkin.

Dengan adanya konversi file teks ke audio mempermudah membantu berkomunikasi bagi individu dengan gangguan bicara atau kondisi lainnya yang mempengaruhi kemampuan bicara, mendapatkan informasi tanpa harus meluangkan waktu untuk membaca secara fisik, dan lebih mudah dipahami, dan lebih menarik.

Metode *synthesizer concatenation* dengan pendekatan dengan algoritma *time domain psola* untuk menghasilkan suara yang alami dan dapat dimengerti dengan jelas. Metode ini memecah suara rekaman menjadi *diphone* yang memiliki transisi antar dua bunyi yang berdekatan, lalu menggabungkan algoritma PSOLA. Algoritma ini memastikan perangkaian antar diphone menjadi halus, sehingga sintesis ucapan yang dihasilkan dapat lancar dan datar tanpa intonasi.

Sesuai dengan latar belakang diatas, saya akan membuat skripsi tentang aplikasi teks ke audio dengan judul “IMPLEMENTASI METODE *SYNTHESIZER CONCATENATION* ALGORITMA *TIME DOMAIN PSOLA* PADA APLIKASI KONVERSI FILE TEKS KE AUDIO”.

1.2 Rumusan Masalah

Berdasarkan uraian latar belakang diatas, rumusan masalah dalam penelitian ini adalah:

1. Bagaimana merancang dan membuat aplikasi konversi file teks ke audio?
2. Bagaimana menerapkan metode *synthesizer concatenation* dengan pendekatan algoritma *time domain psola* pada konversi file teks ke audio?

1.3 Batasan Masalah

Agar pembatasan lebih terarah dan sesuai dengan judul Tugas Skripsi yang telah ditentukan penulis hanya memahas pokok-pokok bahasan sebagai berikut.

1. Aplikasi yang dibangun merupakan aplikasi berbasis web.
2. Pembuatan aplikasi menggunakan bahasa pemrograman PHP.

3. Menggunakan metode *synthesizer concatenation* dengan pendekatan algoritma *time domain psola* untuk membuat aplikasi konversi file teks ke audio.
4. Aplikasi konversi file teks ke audio menggunakan bahasa indonesia.
5. Memuat maksimal 250 kata
6. Aplikasi ini hanya mengkonversi teks ke suara.

1.4 Tujuan Penelitian

Tujuan yang ingin dalam penelitian Tugas Skripsi ini adalah:

1. Untuk merancang dan membuat aplikasi konversi file teks ke audio menggunakan metode *synthesizer concatenation*.
2. Untuk menerapkan metode *synthesizer concatenation* pada file teks ke audio.

1.5 Manfaat Penelitian

Adapun manfaat penelitian ini adalah sebagai berikut:

1. Dapat menghasilkan aplikasi konversi file teks ke audio.
2. Dapat menambah referensi dalam bidang aplikasi konversi file teks ke audio, khususnya yang menggunakan metode *synthesizer concatenation*.
3. Dapat membantu berkomunikasi bagi individu dengan gangguan bicara atau kondisi lainnya yang mempengaruhi kemampuan bicara.

1.6 Metode Penelitian

Metodologi penelitian yang digunakan pada penelitian ini adalah:

1. Pada tahap ini dilakukan studi keputusan yaitu proses mengumpulkan informasi dengan melakukan pengumpulan, mempelajari, dan membaca berbagai referensi yang berkaitan dengan aplikasi konversi teks ke audio menggunakan metode *synthesizer concatenation*.

2. Analisis dan Perancangan

Pada tahap ini dilakukan analisis spesifikasi aplikasi dan melakukan perancangan aplikasi, seperti perancangan proses dan antarmuka yang meliputi desain, sketsa dan lain sebagainya.

3. Pengkodean

Pada tahap ini dilakukan terhadap aplikasi yang telah dibangun dan tingkat keakuratan dari sistem aplikasi yang telah dibuat.

4. Pengujian Aplikasi

Pada tahap ini dilakukan pengujian terhadap aplikasi yang telah dibangun, dan tingkat keakutaran dari sistem aplikasi yang telah dibuat.

5. Penyusunan Laporan

Pada tahap ini dilakukan penulisan dokumentasi dan laporan dari aplikasi yang dikembangkan

1.7 Sistematika Penulisan

Sistematika penulisan Tugas Skripsi ini dibagi atas beberapa bab, dimana masing-masing bab dibagi atas beberapa, sub agar mempermudah penjelasan mengenai penelitian yang dilakukan dan mempermudah pembaca dalam memahami

isi penelitian. Adapun sistematika penulisan Tugas Skripsi ini adalah sebagai berikut :

BAB 1 PENDAHULUAN

Pendahuluan berisi tentang Latar Belakang Masalah, Rumusan Masalah, Tujuan, Manfaat, Batasan Masalah, Metodologi Penelitian dan Sistematika Penulisan dalam pembuatan Tugas Skripsi.

BAB 2 TINJUAN PUSTAKA

Bab ini berisi teori-teori pengetahuan dasar yang di peroleh dari studi kepustakaan atau literature dan dokumentasi *internet* yang digunakan untuk memahami permasalahan yang disajikan antara lain tentang aplikasi system pendukung keputusan serta metode *synthesizer concatenation*.

BAB 3 METODE PENELITIAN

Bab ini menguraikan tahapan-tahapan sistematika yang digunakan untuk melakukan kajian penelitian. Tahapan-tahapan tersebut merupakan kerangka yang dijadikan pedoman penelitian untuk mencapai tujuan yang telah ditetapkan. Tahapan tersebut berisi waktu dan tempat penelitian untuk membuat aplikasi konversi file teks ke audio menggunakan metode *synthesizer concatenation*

BAB 4 HASIL DAN PEMBAHASAN

Bab ini berisi tentang hasil pembahasan dari aplikasi konversi file teks ke audio menggunakan metode *synthesizer concatenation*.

BAB 5 KESIMPULAN DAN SARAN

Bab ini berisi kesimpulan dari keseluruhan uraian bab-bab penulisan skripsi dan saran yang diajukan untuk pengembangan lebih lanjut.

BAB II

Tinjauan Pustaka

2.1 Aplikasi

2.1.1 Pengertian Aplikasi

Aplikasi merupakan teknologi yang berkembang pesat. Kemajuan teknologi membantu pengelolaan data atau informasi yang tersedia dapat berlangsung secara cepat, efisien dan akurat. Aplikasi dapat mempermudah proses pengerjaan atau penyelesaian suatu persoalan. Pengertian aplikasi secara umum adalah alat terapan yang difungsikan secara khusus dan terpadu sesuai kemampuannya yang dimilikinya, aplikasi merupakan suatu perangkat computer yang siap pakai bagi *user*. (Firmansah et al., 2020)

Program aplikasi adalah program siap pakai atau program yang direkam untuk melakukan suatu fungsi bagi pengguna atau aplikasi yang lain. Aplikasi juga diartikan sebagai penggunaan atau penerapan suatu konsep yang menjadi pokok pembahasan atau sebagai program computer yang dibuat untuk menolong manusia dalam melaksanakan tugas tertentu. Aplikasi *software* yang dirancang untuk penggunaan praktisi khusus, klasifikasi luas ini dapat dibagi menjadi 2 (dua) yaitu:

- a. Aplikasi *software* spesialis, program dengan dokumentasi tergabung yang dirancang untuk menjalankan tugas tertentu.
- b. Aplikasi paket, suatu program dengan dokumentasi tergabung yang dirancang untuk jenis masalah tertentu.

Dari kedua pengertian di atas dapat disimpulkan bahwa aplikasi adalah suatu sekumpulan perintah atau kode yang disusun sistematis untuk menjalankan suatu perintah yang diberikan oleh manusia melalui komponen atau *hardware* komputer yang digunakan oleh manusia dalam menjalankan program aplikasi dengan demikian membantu manusia untuk memberikan solusi dari apa yang diinginkan.

Pengertian aplikasi menurut para ahli adalah sebagai berikut:

- a. Menurut Ali Zaki dan Smitsey *community*, Aplikasi adalah komponen yang berguna melakukan pengolahan data maupun kegiatan-kegiatan seperti pembuatan dokumen atau pengolahan data.
- b. Menurut Jogiyanto, Aplikasi adalah penggunaan dalam suatu komputer instruksi(*instruction*) atau pernyataan (*statement*) yang disusun sedemikian rupa sehingga computer dapat memproses input menjadi output.
- c. Menurut Hengki W. Pramana, Aplikasi adalah suatu unit perangkat lunak yang dibuat untuk melayani kebutuhan beberapa aktivitas seperti sistem perniagaan, game pelayanan masyarakat, periklanan, atau semua yang hampir dilakukan manusia.

2.2 Website

2.2.1 Pengertian Website

Website atau situs web merupakan kumpulan halaman web yang saling terkait dan dapat diakses melalui internet. Setiap halaman web dalam sebuah website biasanya berisi teks, gambar, video, dan elemen interaktif lainnya. Situs web digunakan untuk

berbagai tujuan, seperti berbagai konten, atau membangun komunitas online. (Haryodi, 2024)

Secara umum situs web dapat dikategorikan kedalam 3 jenis, yaitu website statis, website dinamis, dan website interaktif. Berikut penjelasan dari masing-masing jenis website beserta contohnya.

1. Website Statis

Situs web statis (static website) merupakan jenis situs web yang terdiri dari halaman-halaman yang tidak berubah secara dinamis. Konten pada situs web statis biasanya ditampilkan dengan cara yang tetap dan tidak berinteraksi dengan pengguna.

Contohnya adalah situs web perusahaan yang berisi informasi tentang produk, layanan, kontak dan profil perusahaan. Halaman-halaman pada situs web statis dapat berupa HTML sederhana atau menggunakan bahasa markup lainnya seperti CSS dan JavaScript.

2. Website Dinamis

Website dinamis adalah jenis situs web yang menghasilkan konten yang dapat berubah atau disesuaikan secara dinamis sesuai dengan permintaan pengguna. Konten di dalam situs web dinamis diperbaharui secara otomatis melalui basis data atau sumber lainnya.

Contohnya adalah situs berita yang menampilkan berita terkini yang dapat diperbaharui secara langsung.

3. Website Interaktif

Web interaktif adalah jenis situs web yang memungkinkan pengguna berinteraksi secara langsung dengan konten atau fitur yang disediakan. Situs web interaktif sering kali melibatkan pengguna dalam kegiatan, seperti mengisi formulir, mengirim komentar, mengunduh file, dan berpartisipasi dalam forum diskusi.

Salah satu contoh web interaktif adalah Facebook, pengguna dapat berinteraksi dengan teman. Mengposting status, mengirim pesan, dan berbagi konten.

2.3 HTML (*Hypertext Markup Language*)

HTML (*Hypertext Markup Language*) adalah bahasa pemrograman yang digunakan untuk membuat dan mengatur struktur konten sebuah halaman web. HTML menggunakan elemen-elemen markup untuk memberi instruksi kepada situs web tentang bagaimana tampilan dan reaksi dari elemen-elemen tersebut. HTML dibuat oleh Tim Berners-Lee ketika masih bekerja dengan CERN dan dipopulerkan pertama kali oleh *browser Mosaic*.

HTML terdiri dari dua komponen utama, yaitu elemen dan atribut. Elemen adalah bagian terkecil dalam HTML yang diberi tanda dengan tag-tag. Setiap elemen dimulai dengan tag pembuka dan diakhiri dengan tag penutup. Dengan menggunakan tanda (tag) HTML, pengembangan web dapat menentukan elemen-elemen seperti teks, gambar, tautan, tabel, dan formulir pada halaman web. (Haryodi, 2024)

2.4 CSS (*Casding Style Sheets*)

CSS (*Casding Style Sheets*) adalah bahasa pemrograman yang digunakan untuk mengatur tampilan dan gaya suatu halaman web. CSS berperan penting dalam pengembangan web karena memisahkan tampilan dan struktur konten dari halaman web. Dengan CSS dapat mengubah tampilan elemen-elemen HTML, seperti warna, ukuran font, layout, dan banyak lagi. (Haryodi, 2024)

CSS memungkinkan pengembangan web untuk menciptakan halaman web yang menarik secara visual, konsisten dan mudah diubah.

2.5 PHP (*Hypertext Preprocessor*)

PHP (*Hypertext Preprocessor*) adalah bahasa pemrograman yang umum digunakan untuk pengembangan web. PHP dikembangkan khusus untuk membangun aplikasi web dinamis, dimana konten web dapat berubah secara server, atau data yang diterima. Dalam pengembangan web, PHP berperan sebagai bahasa server-side, yang berarti kode PHP dijalankan di sisi server sebelum hasilnya dikirim ke peramban pengguna. (Haryodi, 2024)

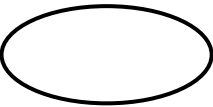
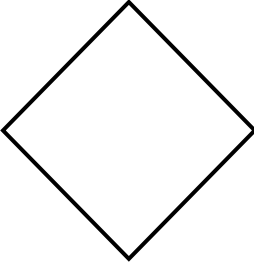
2.6 *Flowchart*

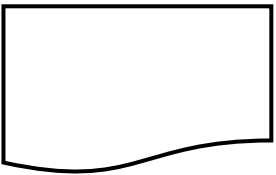
Flowchart adalah sebuah jenis diagram yang mewakili algoritma. Alir kerja atau proses, yang menampilkan langkah-langkah dalam bentuk simbol-simbol grafis, dan urutannya dihubungkan dengan panah. Bagan ini mewakili ilustrasi atau penggambaran penyelesaian masalah. Bagan alir digunakan untuk menganalisis, mendesain, mendokumentasi atau manajemen sebuah proses atau program diberbagai bidang, (Wikipedia, 2024)

Bagan alir digunakan untuk mendesain dan mendokumenatsi proses atau program sederhana. Seperti jenis diagram lainnya, diagram ini terjadi dan dengan demikian membantu mengerti sebuah proses. Dan mungkin saja menentukan kekurangan fitur, atau bagian yang kurang jelas didalam sebuah proses.

Tabel berikut adalah simbol-simbol bagan alir yang umum digunakan:

Tabel 2. 1 Simbol-simbol dalam *flowchart*

Gambar	Nama	Keterangan
	Terminal	Menunjukkan awal atau akhir sebuah proses
	Garis Alir	Menunjukkan arah aliran algoritma, dari satu proses ke proses berikutnya
	Proses atau Langkah	Menyatakan kegiatan yang akan terjadi dalam diagram alir
	Titik Keputusan	Proses atau langkah dimana perlu adanya keputusan atau adanya kondisi tertentu. Dimana titik ini selalu ada dua keluaran untuk melanjutkan aliran kondisi

		yang berbeda
	Masukan atau Keluaran	Digunakan untuk mewakili data masuk, atau data keluar. Hanya bisa dimulai dari menuju keluaran, bukan sebaliknya
	Simbol Predefine Proses	Digunakan untuk menunjukkan suatu proses yang kompleks, sehingga tidak bisa dijelaskan di diagram alir ini dan merujuk diagram alir yang terpisah
	Simbol Manual Operator	Menunjukkan pengolahan yang tidak dilakukan oleh komputer
	Simbol Dokumen	Simbol yang menyatakan input berasal dari kartu atau output dicetak ke kertas

2.7 XAMPP

XAMPP adalah singkatan i (X-platform, Apache, MySQL, PHP, Pel). Perangkat lunak berbasis *website server* yang bersifat *open source* (bebas), serta mendukung di berbagai sistem operasi, baik Windows, Linux, atau Mac OS.

XAMPP digunakan sebagai standal *one server* (berdiri sendiri) atau bisa disebut dengan localhost. Hal tersebut memudahkan dalam proses dalam pengeditan, desain, dan pengembangan aplikasi. Penggunaan dari XAMPP sangat dibutuhkan untuk dapat mengembangkan *software* atau tamoilan *website* dengan lebih mudah, cepat, dan terstruktur. Terdapat tiga komponen penyusun uatama tools ini yaitu htdocs, Control Panel, dan PhpMyAdmin. Gunakan aplikasi *website web server* ini sebagai tools bantu untuk mulai belajar tahapan pengembangan perangkat lunak sesuai dengan kebutuhan proyek bisnis. (Rina Noviana, 2022)

2.8 Algoritma

Menurut Saputra, dkk menjelaskan bahwa algoritma adalah deretan instruksi yang jelas untuk memecakan masalah, yaitu untuk memperoleh keluaran yang diinginkan dari suatu masukan. Ada 3 definisi tentang algoritma yang dijelaskan oleh Suarga diantaranya:

1. Teknik penyusunan langkah-langkah penyelesaian masalah dalam bentuk kalimat dengan jumlah kata terbatas tetapi tersusun secara logis dan sistematis.
2. Suatu prosedur yang jelas untuk menyelesaikan suatu persoalan dengan menggunakan langkah-langkah tertentu dan terbatas jumlahnya.

3. Susunan langkah yang pasti, yang bila diikuti maka akan mentransformasi data input menjadi output yang berupa informasi. Algoritma merupakan suatu prosedur untuk menyelesaikan suatu masalah yang tersusun secara logis dan sistematis serta akan memperoleh data masukan menjadi keluaran yang diinginkan berupa informasi.

2.9 Sytem Text to Speech Synthesizer

Pengertian secara umum sistem TTS sintesis merupakan sebuah aplikasi dengan bertujuan mengolah input berupa teks kemudian menghasilkan output audio yang menyerupai ucapan manusia. Speech signal merupakan gelombang kontinue (gelombang akustik). Sebuah sinyal ucapan dapat dianalisis dengan berbagi metode pendekatan karena mengandung nilai akustik, fonetik, fonologi, morfologi, sintatik, dan semantic. Adanya intonasi sebagai aspek akustik sinyal suara sangat membantu di dalam mengidentifikasi setiap segmen. (Sudirman Melangi, 2018)

Menurut metode format dimana setiap fonem yang dihasilkan terutama oleh sistem vokal selama artikulasi mempengaruhi dinamika spektrum spektral suara. Variasi intonasi secara subtansial mempengaruhi idetitas pengucapan suatu kata sehingga fonem dapat menjadi panjang atau pendek, keras atau lemah, dan memiliki pola pitch (nada) yang bervariasi pula. Secara umum sintesis ditentukan oleh 3 karakteristik yaitu:

1. Tingkat Pemahaman (intellegibility): Apakah teks yang dihasilkan komputer tersebut dapat dimengerti oleh pendengar (sejauh mana pendengar dapat memahami suara yang dihasilkan oleh komputer secara tepat).

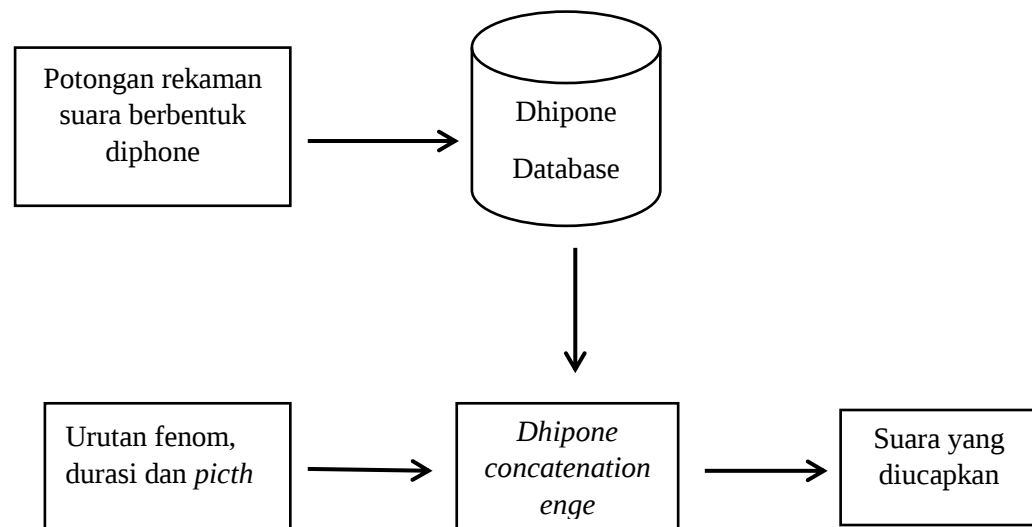
2. Kelancaran Pengucapan (fluidity): Apakah teks yang dilafalkan oleh komputer tersebut lancar (transisi antara suku kata-kata lancar penyebutannya).

3. Kealamian Pengucapan (Naturalness): Apakah teks yang dilafalkan oleh komputer tersebut terdengar sesuai pengucapan manusia pada umumnya (Sudirman Melangi, 2018)

2.10 Concatenation Synthesizer

Concatenation Synthesizer merupakan synthesizer yang mampu memproduksi sinyal ucapan secara otomatis melalui transkripsi grafem-ke-fenom untuk kalimat yang diucapkan. *Concatenation* yaitu, penggabungan-gabungan segmen-segmen bunyi yang telah direkam sebelumnya. Setiap segmen berupa *Dhipone* (gabungan dua buah fenom). Pada perekaman suara yang dilakukan beberapa kali agar mendapatkan hasil yang akurat.

Umumnya metode ini memang menghasilkan output sintesa ucapan yang natural (mendekati ucapan manusia). Namun dari segi intelligibility (sejauh mana output ucapan dapat dimengerti oleh manusia) metode ini masih memiliki kekurangan. Proses metode ini dilakukan dengan cara menggabungkan (concatenation) unit-unit ucapan yang telah direkam sebelumnya. Rekaman dari segmen-segmen ucapan tersebut disimpan di dalam *database* dalam bentuk gelombang (uncoded) atau dalam bentuk encoded disesuaikan dengan pengkodean ucapan (speed coding method) yang dipakai. (Sudirman Melangi, 2018)



Gambar 2. 1 Skema Proses Concatenation TTS

Gambar diatas menampilkan cara kerja sintesis concatenation berdasarkan pada rentetan bagian ucapan yang telah direkam. Dasar perangkaian sintesis adalah dengan menggabungkan segmen-segmen dari gelombang ucapan alami yang disimpan dalam *database*. Segmen-segmen tersebut dapat berupa kata-kata (words), unit sub-kata (subword unit) seperti fonem (phonemes), *diphones* dan suku kata (syllables). Sintesis ucapan perangkaian banyak digunakan luas, bekerja dengan prinsip pada pemilihan unit (unit selection). Sistem sintesis perangkaian pemilihan unit yang populer adalah *unisyn*, *clunits* dan *multisyn* (Sudirman Melangi, 2018)

Beberapa penelitian terdahulu sudah dikaji oleh peneliti, membahas tentang *teks to speech* menggunakan metode *synthesizer concatenation*: Sudirman Pelangi menghasilkan aplikasi teks ke suara dengan hasil secara umum dapat mengucapkan kata-kata dalam Bahasa Indonesia dengan cukup lancar dan dapat dimengerti oleh

pendengar. Namun ucapan sintesis yang dihasilkan belum memiliki pola intonasi yang baik. Sintesis

2.11 Algoritma *Time Domain Psola*

Time Domain PSOLA merupakan suatu teknik untuk memodifikasi *pitch* sinyal suara yang dilakukan dengan mengubah durasi sinyal suaranya, memajang untuk mengurangi *pitch* dan memedekkan untuk meningkatkan *pitch*. Tapi dalam PSOLA gelombang suara dibagi terlebih dahulu menjadi beberapa segmen tumpang tindih kecil, kemudian segmen dipindahkan lebih dekat atau terpisah tergantung apakah akan menambah atau mengurangi *pitch*.

Pada algoritma *Time Domain PSOLA* setiap periode *pitch* akan dianalisis satu frame. Syarat yang dibutuhkan untuk ini adalah kemampuan mengidentifikasi permulaan periode (*epochs*) pada *signal speech*. Gelombang ucapan dibagi menjadi frame-frame dengan menggunakan *hamming window* yang dikenakan pada periode *pitch* sebelum dan sesudah *epoch*, bingkai hasil *window* ini direkomendasikan dengan meletakkan *epoch* frame tersebut pada posisi yang sama seperti sebelumnya, kemudian menambahkan bagian yang bertumpang tindih (oleh sebab itu dinamakan *overlap add*). Setelah proses ini dilakukan, akan dihasilkan gelombang ucapan yang hampir mirip dengan gelombang ucapan yang asli.

Gelombang ucapan berupa frame-frame kecil dimana setiap frame tersebut dianggap sebagai *time invariant system*. Sebuah frame ucapan $x[n]$ didapat dari gelombang penuh signal ucapan $s[n]$ yang dikalikan dengan fungsi “jendela” (*window*) $w[n]$ dalam domain waktu :

$$x[n] = w[n]s[n] \dots\dots\dots(2.1)$$

Proses windowing berfungsi untuk mengurangi efek diskontinuitas pada ujung-ujung frame yang dihasilkan oleh proses framing. Fungsi window yang digunakan adalah *hamming window* yang memiliki persamaan sebagai berikut:

$$\text{hamming } w(n) = 0.54 - 0.46 \cos \left(\frac{2\pi n}{n-1} \right), 0 \leq n \leq (n-1) \dots\dots\dots(2.2)$$

ada banyak jenis PSOLA, pada tugas akhir ini menggunakan *time domain* psola karena memungkinkan untuk mengubah pitch sinyal suara tanpa mengubah durasinya atau sebaliknya. *Time Domain PSOLA* berfungsi untuk mensinkronkan dekomposisi *pitch* sinyal suara dengan *peridenya* ketika proses *overlapping*, yang sebelumnya dilakukan *famming* terlebih dahulu. Hal-hal yang penting dilakukan pada proses TD-PSOLA adalah: *pitch detection*, *framing*, dan *windowing*.

2.12 Suara

Suara adalah urutan gelombang tekanan yang merambat melalui media kompredibel (udara dan air). Gelombang tekanan tersebut dapat dipantulkan, dibisakan ataupun dilemahkan oleh media. Suara juga dapat didegisikan sebagai bunyi yang dikeluarkan mulut manusia/ suara dapat terjadi ketika berbicara, bernyanyi, tertawa ataupun menangis.

2.13 Sistem Produksi Suara Manusia

Proses produksi suara pada manusia dapat dibagi menjadi tiga buah proses fisiologis, yaitu : pembentukan aliran udara dari paru-paru , perubahan aliran udara dari paru-paru menjadi suara, baik voiced maupun unvoiced yang dikenal dengan

istilah phonation, dan artikulasi yaitu proses modulasi atau pengaturan suara menjadi bunyi yang spesifik. (Sudirman Melangi, 2018)

2.14 Bahasa Indonesia

Bahasa Indonesia merupakan bahasa resmi sekaligus bahasa nasional di Indonesia. Bahasa Indonesia merupakan varietas yang dibakukan dari bahasa Melayu, sebuah bahasa rumpun Austronesia yang digolongkan ke dalam rumpun Melatik yang sendirinya merupakan cabang turunan dari cabang Melayu-Polinesia. Bahasa Indonesia telah sejak lama digunakan sebagai basantara di wilayah kepulauan Indonesia yang rata-rata memiliki kemajemukan linguistika. Dengan jumlah penuturan bahasa yang lumayan besar ditambah dengan populasi diaspora yang tinggal diluar negeri. Bahasa Indonesia masuk sebagai salah satu bahasa yang paling banyak digunakan atau dituturkan diseluruh dunia. (Wikipedia, 2024)

2.15 Peneliti Terdahulu

Penelitian terdahulu sudah dikaji oleh peneliti, membahas tentang *text-to-speech* dengan metode *synthesizer concatenation* yaitu Pembuatan Text- To- Speech *Synthesizer System* Untuk Penuturan Berbahasa Indonesia yang dimana *Synthesizer concatenation* terdiri dari 3 bagian, yaitu *text pre-processing*, pembangkitan *prosody* dan *concatenation*. Kelebihan dari *concatenation synthesizer* yaitu :

- 1) Natural : *Concatenation synthesizer* dapat menghasilkan suara yang alami dan mirip dengan suara manusia.
- 2) Kejelasan: sistem ini dapat menghasilkan suara yang jelas dan mudah dipahami oleh manusia.

- 3) Flexibility: dapat digunakan untuk berbagai bahasa aplikasi serta dapat disesuaikan dengan kebutuhan spesifik.

Sedangkan kekurangan menggunakan *concatenation synthesizer*:

- 1) Kompleksitas: memerlukan perangkat lunak yang kompleks dan teknologi yang canggih untuk menghasilkan suara yang alami.
- 2) Keterbatasan: memiliki keterbatasan dalam menghasilkan suara yang alami dan mirip dengan suara manusia, terutama dalam hal intonasi dan prosodi.
- 3) Biaya: pembuatan system ini dapat memerlukan biaya yang tinggi, terutama jika menggunakan teknologi yang canggih dan perangkat lunak yang kompleks.

Pada penggunaan algoritma *time domain psola* peneliti terdahulu telah dikaji yaitu membahas tentang implementasi konversi suara menggunakan algoritma *pitch shifting* dengan *time domain pitch synchronous overlap add* (TD-PSOLA), hal yang penting dilakukan dalam proses *time domain psola* yaitu, deteksi nada pada sinyal suara yang diproses hanya sinyal suara *voiced* saja, karna suara *voiced* yang memiliki *pitch* yang bisa dihitung dan dimodifikasi. Pengbingkaian (*framing*) yang digunakan sangat mempengaruhi keberhasilan dalam analisis spectral. Dan yang terakhir jendela (*windowing*) bertujuan untuk mnegurangi terjadinya kebocoran *spectral* atau *alisang* yang mana merupakan suatu efek dari timbulnya sinyal baru yang memiliki frekuensi yang berbeda dengan sinyal aslinya.

Kelebihan dari *time domain psola* yaitu:

- 1) Natural: dapat menghasilkan suara yang mirip dengan suara asli.
- 2) Flexibility: dapat digunakan diberbagai aplikasi dan dapat disesuaikan dengan kebutuhan spesifik.
- 3) Ekspresifitas: dapat mengurangi efek diskontinuitas pada ujung-ujung *frame* yang dihasilkan oleh proses *framing*

Kekurangan dari *time domain psola* yaitu:

- 1) Kompleksitas: memerlukan perangkat lunak yang kompleks dan teknologi yang canggih untuk menghasilkan suara yang alami.
- 2) Keterbatasan: masih memiliki keterbatasan suara yang sangat alami dan mirip dengan suara manusia, terutama dalam hal intonasi.
- 3) Biaya: memerlukan biaya yang tinggi, terutama jika menggunakan teknologi yang canggih dan perangkat lunak yang kompleks.

2.16 Pengujian

Terakhir, tahap pengujian merupakan salah satu hal yang penting untuk dilakukan dengan tujuan untuk memeriksa kesalahan-kesalahan terjadi didalam sistem yang dibangun dan meminimalisir terjadinya kerugian dari kesalahan-kesalahan yang mungkin terjadi di sistem. Pengujian dalam perangkat lunak menggunakan satu metode pengujian yaitu pengujian *black box testing*. Pengujian *black box* juga disebut sebagai pengujian fungsional berdasarkan spesifikasi dari klien dan penguji sistem tidak memiliki akses untuk ke kode program dari sistem tersebut. (Praniffa et al., 2023)