

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.

b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB II

LANDASAN TEORI

2.1 Audio

Audio berasal dari kata *audible* yang artinya suara yang dapat didengarkan secara wajar oleh telinga manusia (Innayah, 2012). *Audio* adalah suara atau bunyi yang dapat di hasilkan dari getaran suatu benda yang agar dapat di tangkap oleh manusia haruslah kuat dan minimal 20 kali/detik. Suara merupakan getaran yang di hasilkan dari gesekan, pantulan dan yang lainnya antara benda-benda.

2.1.1 Format Audio

Pengertian dari Format *Audio* itu sendiri adalah kode informasi untuk penyimpanan file komputer. Sedangkan *Codec* adalah program komputer yang mengompres dan mendekomposisi data di file media. File *audio* dibagi menjadi 3 kategori utama yaitu :

1. *Uncompressed Audio Format*

Uncompressed Audio Format adalah sebuah format *audio* tanpa kompresi. Format ini dapat menyimpan *audio* dengan tetap mempertahankan keutuhan setiap bit dan *sample rate* dari sumber dihasilkan. File *uncompressed* yang dihasilkan dari proses rip sebuah *audio* cd juga mempunyai spesifikasi yang sama dengan source nya (2 channel PCM, 44 kHz *sample rate* dan 16 bit depth). Contohnya AIFF dan WAV (Taupik, 2010).

2. *Lossless Compressed Audio Format*

Lossless Compressed Audio Format adalah sebuah format *audio* yang dapat menyimpan data dengan ukuran yang lebih kecil dari *uncompressed* format, namun dengan kualitas yang identik. Secara teori *lossless audio* memiliki ukuran 55% - 65% dari ukuran *uncompressed*, namun angka ini bisa lebih rendah atau lebih tinggi tergantung dari banyak dan kompleksnya data yang di encode dari sumbernya. Contohnya FLAC (*Free Lossless Audio Codec*) dan *Lossless Windows Media Audio* (WMA) (Taupik, 2010).

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3. *Lossy Compressed Audio Format*

Lossy Compressed Audio Format adalah sebuah format *audio* yang dapat menyimpan data dengan ukuran yang lebih kecil lagi dari *lossless* format. Secara teori tujuan utama *lossy compression* adalah membuang berbagai informasi pada *audio* yang dipersepsikan sebagai suara yang tidak terlalu dapat terdengar oleh telinga manusia (metode *Psychoacoustics*) namun pada praktiknya akan ada *noticeable differences* antara *lossy*, *lossless* dan *uncompressed* terutama jika dikomparasikan pada *high-end sound system*. Contohnya MP3, *Vorbis* (OGG), AAC (*Advanced Audio Coding*), MID (MIDI) dan RA (*Real Audio*) (Taupik, 2010).

2.1.2 Audio WAV (Waveform Audio File Format)

File WAV (*Waveform Audio File Format*) merupakan format file digital *audio* yang disimpan dalam bentuk digital dengan ekstensi WAV. File tersebut menyimpan amplitudo dan frekuensi suara sehingga membutuhkan tempat penyimpanan yang cukup besar. File WAV itu dikembangkan oleh Microsoft dengan format PCM. File WAV merupakan salah satu file yang dikenali dalam Matlab *audio* (Wiedjaja, Pangbun, I.S, & Wahab, 2004).

File WAV dibagi menjadi tiga bagian dalam menyimpan informasi dan masing-masing bagian di sebut *chunk*. Tiga bagian tersebut diantaranya :

1. *Chunk* yang pertama adalah *chunk RIFF*. *Chunk* ini menyatakan bahwa file tersebut adalah file WAV.
2. *Chunk* yang kedua adalah *chunk FORMAT*. *Chunk* ini berisi parameter atau sifat file WAV, misalnya parameter *sample rate*.
3. *Chunk* yang ketiga adalah *chunk DATA*. *Chunk* ini berisi data dari sinyal sesungguhnya.

2.2 Suara

Suara adalah fenomena fisik yang dihasilkan oleh getaran benda getaran suatu benda yang berupa sinyal analog dengan amplitudo yang berubah secara kontinu terhadap waktu. Keragaman suara dapat terlihat dari persepsi fisik manusia



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

terhadap suara antara lain: *formant*, jenis suara, *pitch*, *timbre* dan *volume*. Persepsi fisik tersebut dapat terdengar secara jelas apabila diucapkan oleh seseorang berjenis kelamin laki-laki ataupun perempuan begitu juga suara anak-anak, remaja, dewasa dan orang yang sudah tua (Bhaskoro & Riedho, 2012)

1. Frekuensi

Frekuensi merupakan komponen dasar pembentuk suara. Frekuensi adalah kecepatan perubahan amplitudo terhadap waktu. Suara muncul karena adanya getaran pada udara. Jumlah getaran pembentuk suara dinyatakan dalam frekuensi atau jumlah getaran dalam satu satuan waktu. Frekuensi dinyatakan dalam satuan hertz (Hz). Misalkan suatu suara memiliki frekuensi sebesar 440Hz, maka berarti dalam satu detik terjadi 440 buah getaran.

2. Amplitudo

Amplitudo merupakan simpangan dari suatu getaran yang menghasilkan suara tersebut. Amplitudo menentukan keras lemahnya suatu suara. Semakin keras suatu suara semakin besar amplitudonya.

3. Pitch

Pitch adalah persepsi pendengaran manusia terhadap perbedaan frekuensi suatu suara. Biasanya *pitch* dikenal dalam bentuk nada, misalnya: C, D, E dan sebagainya.

4. Volume

Volume adalah persepsi manusia terhadap keras lemahnya suatu suara saat didengar. berbanding lurus dengan amplitudo. Semakin besar volume semakin keras suara yang di dengar berarti semakin besar amplitudonya

5. Timbre

Timbre adalah warna suara yang merupakan karakteristik suara dan merupakan pembeda antara satu suara dengan suara lainnya. Misalkan suara trompet dan suara piano, walaupun keduanya memainkan nada yang sama tetapi memiliki warna suara yang berbeda. *Timbre* biasanya digambarkan dalam *time* domain.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

6 Formant

Formant adalah frekuensi resonansi alami yang terjadi di dalam rongga bidang suara, tergantung pada bentuk dan ukuran bidang sua umumnya mempunyai tiga *formant* F1, F2 dan F3. Sedangkan *pitch* (F0) merupakan periode pengulangan pulsa yang disebabkan oleh membuka dan menutupnya pita suara. Satuan dari *formant* adalah (Hz).

2.3 Speech Recognition (Pengenalan Ucapan)

Pengenalan suara dalam istilah bahasa Inggris *automatic speech recognition* (ASR) adalah suatu pengembangan teknik dan sistem yang memungkinkan komputer untuk menerima masukan berupa kata yang diucapkan. Teknologi ini memungkinkan suatu perangkat untuk mengenali dan memahami kata-kata yang diucapkan dengan cara digitalisasi kata dan mencocokkan sinyal digital tersebut dengan suatu pola tertentu yang tersimpan dalam suatu perangkat. Kata-kata yang diucapkan diubah bentuknya menjadi sinyal digital dengan cara mengubah gelombang suara menjadi sekumpulan angka yang kemudian disesuaikan dengan kode-kode tertentu untuk mengidentifikasi kata-kata tersebut (Lestary, 2009).

Pengenalan ucapan (*speech recognition*) adalah semua kegiatan yang melibatkan identifikasi pembicara. Berdasarkan suaranya dan pengelompokan pembicara berdasarkan kesamaan dari suara mereka. Pengenalan ucapan didefinisikan sebagai proses pengubahan sinyal suara ke bahasa (*linguistic*) mesin dalam bentuk data digital (Charisma, 2013). Dengan kata lain, pengenalan suara menyatakan kemampuan untuk mencocokkan pola dari yang didapatkan atau yang pengenalan diperoleh perbendaharaan kata terhadap sinyal suara ke dalam bentuk yang tepat. Pengenalan ucapan adalah suatu proses di mana komputer dapat mengenal kata-kata yang diucapkan oleh manusia. Secara umum sistem pengenalan ucapan merupakan bagian dari pengenalan suara memiliki dua proses.

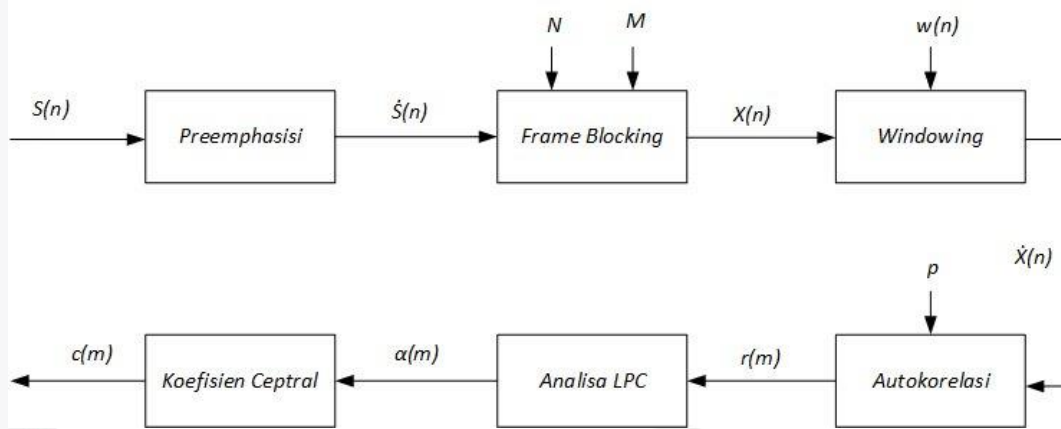
Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.4 Linear Predictive Coding (LPC)

Metode yang sebagian besar digunakan untuk pemrosesan sinyal *audio* dari sinyal digital pembicaraan dikompresi, *Linear Predictive Coding* (LPC) merupakan salah satu metode yang baik digunakan dalam teknik analisa suara dan salah satu metode yang paling berguna untuk *encoding* berkualitas pada *bit-rate* rendah (Sharma, 2012).

Bagian terpenting dari pengenalan suara adalah ekstraksi ciri (*feature extraction*). Kegunaan dari ekstraksi ciri adalah mengurangi informasi-informasi dari sensor dan mengonversi informasi-informasi yang penting dari sinyal untuk pengenalan pola agar menghasilkan format yang lebih sederhana dengan kelas-kelas yang jelas. LPC secara khusus merupakan metode yang cocok dalam pengolahan sinyal suara. Metode ini dapat juga diterapkan dalam pengenalan kata. Tujuan dari digunakannya metode ini adalah untuk mencari nilai koefisien LPC dari suatu sinyal.



Gambar 2.1 Diagram Blok Proses Analisa LPC

Prinsip dasar dari ekstraksi ciri sinyal dengan menggunakan LPC adalah bahwa contoh sinyal suara $s(n)$ pada waktu ke- n dapat diperkirakan sebagai kombinasi linear p sampel sinyal suara sebelumnya yaitu :

$$s(n) \approx a_1s(n - 1) + q_2s(n - 2) + \dots + a_p s(n - p) \dots\dots\dots(2.1)$$

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dengan koefisien a_1, a_2, \dots, a_p diasumsikan bernilai konstan selama *frame* analisis ucapan. Gambar 2.1 menunjukkan blok diagram analisis LPC (Z, Hidayatno, & Saksono, 2008).

2.4.1 Pre-Emphasis

Sinyal suara digital akan melewati tahap *pre-emphasis* atau tahap *normalize* sebelum diproses lebih lanjut. *Pre-emphasis* merupakan sebuah tahap untuk mendeteksi adanya aktivitas suara atau tidak. Dengan mendeteksi aktivitas suara pada sinyal maka akan dilakukan penghapusan terhadap waktu yang tidak memiliki aktivitas suara. Pada tahap *pre-emphasis* akan dilakukan proses menghilangkan *noise* suara dari data suara masukan, *noise* ini merupakan sinyal yang memiliki perbedaan yang signifikan di antar nilai di sekitar (Dewi, Zulkarnain, & Lestari, 2018).

$$\tilde{s}(n) = s(n) - \tilde{\alpha}s(n - 1) \quad 0,9 \leq \tilde{\alpha} \leq 1,0 \dots \dots \dots (2.2)$$

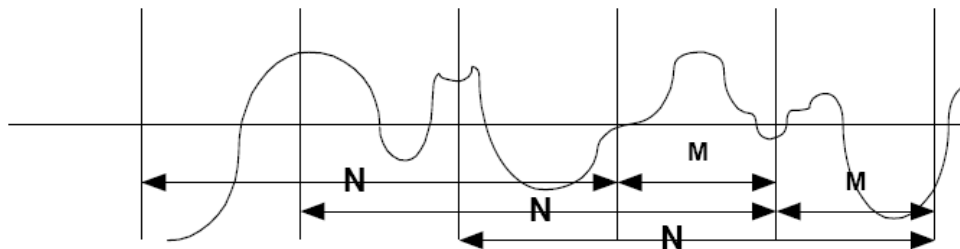
$\tilde{s}(n)$ = Sinyal suara yang telah di-*emphasis*

$s(n)$ = Sinyal suara awal

$\tilde{\alpha}$ = Koefisien filter yang umumnya bernilai 0.9375

2.4.2 Frame Blocking

Pada tahap ini sinyal yang telah dilakukan *pre-emphasis* di kelompokkan ke dalam bingkai-bingkai dengan ukuran masing-masing bingkai sebesar N sampel, kemudian dibagi menjadi sejumlah *frame*. Tiap *frame* berisi N *sample* dengan jarak setiap *frame* dipisahkan sebanyak M *sample*. Biasanya $M = 1/3 N$. Proses *frame blocking* bisa dilihat pada gambar 2.2.



Gambar 2.2 Frame Blocking, Blok pemilihan bingkai dari sinyal suara

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Dapat diperhatikan pada gambar, *frame* pertama mempunyai panjang N *sample*. *Frame* kedua dimulai M *sample* setelah *frame* pertama dan terjadi *overlap* sebanyak $N - M$ *sample*, begitu juga dengan *frame* ketiga yang dimulai dari $2M$ *sample* dan terjadi *overlap* sebanyak $N - 2M$ *sample*. Proses ini akan berlanjut sampai semua *frame* dikunjungi. Parameter yang harus diperhatikan untuk mendapatkan *frame* yang *overlap* adalah ukuran $M \leq N$ dan hasil estimasi koefisien LPC akan berkorelasi antar *frame*.

Pada tahap ini sinyal suara yang telah ter *emphasis* selanjutnya dilakukan proses *frame blocking*. Data suara akan dibagi menjadi beberapa *frame* yang dilakukan secara manual dengan menggunakan *software*. Suara di potong menjadi *frame-frame* berdasarkan banyaknya data yang sudah ditentukan. Proses *frame blocking* yaitu untuk memisah kan masing-masing kelas target.

2.4.3 Windowing

Tahap selanjutnya dalam proses LPC adalah *windowing*. *Windowing* adalah proses yang dilakukan untuk meminimalisir diskontinuitas antar *frame* yang dapat menyebabkan kehilangan informasi yang terdapat pada suatu sinyal. Konsepnya adalah meruncingkan sinyal ke angka nol pada permulaan dan akhir setiap *frame*. Bila *window* didefinisikan sebagai $w(n)$, $0 \leq n \leq N-1$, dengan N adalah jumlah sampel dalam tiap *frame*, maka hasil dari proses ini adalah sinyal :

$$\tilde{x}(n) = \tilde{s}(n) \times w(n), 0 \leq n \leq N - 1 \dots\dots\dots (2.3)$$

Keterangan :

- $\tilde{x}(n)$ = Sinyal hasil *widowing sample* ke-n
- $x(n)$ = Nilai *sample* ke-n
- $w(n)$ = Nilai *window* ke-n

Window memiliki beberapa bentuk, dalam kasus ini digunakan *hamming window* yang mempunyai persamaan sebagai berikut :

$$w(n) = 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right), 0 \leq n \leq N - 1 \dots\dots\dots (2.4)$$

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Keterangan :

- $w(n)$ = Hamming window ke-n
- n = Frame yang sedang diproses (0, 1, 2, 3,...,N-1)
- N = Jumlah *sample frame*

2.4.4 Analisa Autokorelasi

Setiap *frame* yang telah dilakukan proses *windowing* kemudian dilakukan *autocorrelation analysis* untuk mendapatkan nilai *autocorrelation analysis*. Adapun persamaannya bisa dilihat pada rumus di bawah ini :

$$r(m) = \sum_{n=0}^{N-m} \tilde{x}(n) * \tilde{x}(n + m) \dots\dots\dots(2.5)$$

Keterangan :

- $r(m)$ = Nilai autokorelasi
- $\tilde{x}(n)$ = Sinyal hasil *widowing* sample ke-n
- $\tilde{x}(n + m)$ = Nilai *windowing* orde ke m
- m = Nilai orde 0, 1, 2, ..., p

Nilai tertinggi dari autokorelasi tersebut p adalah orde dari analisa LPC yang akan dilakukan. Nilai yang umum untuk orde analisa LPC adalah antara 8 sampai 16. Keuntungan dari penggunaan metode autokorelasi adalah bahwa nilai ke-nol, $r(0)$ adalah energi dari sinyal yang dibuat autokorelasinya (Thiang & Saputra, 2005).

Sebelum mendapatkan nilai autokorelasi $r(m)$, terlebih dahulu harus mendapatkan nilai dari $\tilde{x}(n)$ dan nilai $\tilde{x}(n + m)$. Nilai dari $\tilde{x}(n)$ adalah nilai *windowing* maka nilai $\tilde{x}(n)$ tersebut telah di dapat sebelumnya pada proses *windowing*. Selanjutnya menentukan nilai dari $\tilde{x}(n + m)$

$$\tilde{x}(n + m) = \tilde{s}(n + m) * w(n + m) \dots\dots\dots(2.6)$$

Keterangan :

- $\tilde{x}(n + m)$ = Nilai *windowing* orde ke m
- $\tilde{s}(n + m)$ = Suara telah di-*emphasis* ditambah orde ke m
- $w(n + m)$ = Hamming window ke-n ditambah orde ke m
- m = Nilai orde 0, 1, 2, ..., p

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Proses perhitungan di atas berhenti ketika $\tilde{x}(n+m) = n$. Sebelum mendapatkan nilai $r(m)$ terlebih dahulu mencari nilai $r(n+m)$ dengan mengalikan nilai *windowing* $\tilde{x}(n)$ dengan $\tilde{x}(n+m)$ yang telah didapat sebelumnya, sehingga dengan merata-ratakan nilai $r(n+m)$ ini maka akan didapatlah nilai $r(m)$.

$$r(n+m) = \tilde{x}(n) * \tilde{x}(n+m) \dots \dots \dots (2.7)$$

2.4.5 Ekstraksi Ciri LPC

Proses selanjutnya adalah ekstraksi ciri dengan menggunakan LPC *Analysis*, yaitu merubah setiap *frame* dari *autocorrelation* menjadi himpunan LPC parameter, yang berupa koefisien LPC dan koefisien *error* yang kemudian akan diiterasikan sebanyak jumlah *sample* pada data suara. Metode formal untuk melakukan konversi LPC parameter menjadi koefisien LPC dikenal dengan metode *Durbin*, dan dapat dijabarkan dengan (Bagus Prasetyo & Novia Wisesty, 2016):

$$E^{(0)} = r(0) \dots \dots \dots (2.8)$$

$$k_m = \frac{\{r(m) - \sum_{j=1}^{m-1} \alpha_j^{(m-1)} \cdot r(|m-j|)\}}{E^{(m-1)}} \dots \dots \dots (2.9)$$

$$\alpha_m^{(m)} = k_m \dots \dots \dots (2.10)$$

$$\alpha_j^{(m)} = \alpha_j^{(m-1)} - k_m \cdot \alpha_{m-j}^{(m-1)} \dots \dots \dots (2.11)$$

$$E^{(m)} = (1 - k_m^2) E^{(m-1)} \dots \dots \dots (2.12)$$

$$G(m) = \frac{(\alpha^{(m)} + \alpha^{(m-1)})}{Q} \dots \dots \dots (2.13)$$

Keterangan :

- $E^{(m)}$ = nilai *Error* pada LPC ke-i
- k_m = nilai koefisien pantulan
- $\alpha_i^{(m)}$ = nilai koefisien LPC ke-i
- $G(m)$ = nilai vektor ciri suara ke-m
- Q = nilai koefisien *cepstral*

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Kemudian koefisien LPC yang dihasilkan adalah sebagai berikut:

$$a_m = \alpha_m^{(p)} \quad 1 \leq m \leq p \dots\dots\dots(2.14)$$

Pengubahan parameter LPC menjadi koefisien *cepstral* Parameter LPC yang sangat penting yang bisa diturunkan dari koefisien LPC adalah koefisien *cepstral* LPC $c(m)$. Persamaan yang digunakan untuk menghitungnya adalah:

$$c_m = a_m + \sum_{k=1}^{m-1} \left(\frac{k}{m}\right) \cdot c_k \cdot a_{m-k} \quad 1 \leq m \leq p \dots\dots\dots(2.15)$$

$$c_m = \sum_{k=1}^{m-1} \left(\frac{k}{m}\right) \cdot c_k \cdot a_{m-k} \quad m > p \dots\dots\dots(2.16)$$

Koefisien *cepstral* ini adalah koefisien dari representasi transformasi *Fourier* pada spektrum logaritmik (Thiang & Saputra, 2005).

2.5 Klasifikasi Algoritma K-NN (*K-Nearest Neighbor*)

Klasifikasi merupakan proses untuk menemukan model atau fungsi yang menjelaskan atau membedakan konsep atau kelas data, dengan tujuan untuk dapat memperkirakan kelas dari suatu objek yang labelnya tidak diketahui.

Tujuan dari algoritma K-NN adalah untuk mengklasifikasi objek baru berdasarkan atribut dan training sampel. Dimana hasil dari sampel uji yang baru diklasifikasikan berdasarkan mayoritas dari kategori pada K-NN. Pada proses pengklasifikasian, algoritma ini tidak menggunakan model apapun untuk dicocokkan dan hanya berdasarkan pada memori. Algoritma K-NN menggunakan klasifikasi ketetanggaan sebagai nilai prediksi dari sampel uji yang baru. Jarak yang digunakan adalah jarak *Euclidean Distance*. Jarak *Euclidean* adalah jarak yang paling umum digunakan pada data numerik (Nobertus Krisandi, 2013).

2.6 *Euclidean Distance*

Euclidean distance adalah sebuah metode yang digunakan untuk mengukur jarak (*distance*). *Euclidean distance* sebenarnya merupakan generalisasi dari *teorema pythagoras*.

Berikut ini adalah contoh perhitungan dengan menggunakan *Euclidean distance*. Jika terdapat dua buah titik pada sebuah bidang dua dimensi (R^2), $u = (x_1,$

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

y_1) dan $v = (x_2, y_2)$ maka untuk mengukur jarak dari kedua buah titik tersebut dapat digunakan persamaan *pythagoras*.

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \dots\dots\dots(2.17)$$

Keterangan :

x_1, x_2 = koordinat sumbu x dari sebuah titik

y_1, y_2 = koordinat sumbu y dari sebuah titik

Jarak tersebut menyebabkan sebuah metrik pada R^2 yang disebut sebagai *Euclidean* metrik pada R^2 ((Hustinawaty, Jati, & Humaira, 2014).

Bila terdapat dua buah vektor dengan n dimensi , $a = (a_1, a_2, \dots, a_n)$ dan $b = (b_1, b_2, \dots, b_n)$ maka formula *pythagoras* dapat digeneralisasikan ke dalam n dimensi (R^2) menjadi :

$$d = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_n - b_n)^2} \dots\dots\dots(2.18)$$

$$d = (a_n - b_n)^2 \dots\dots\dots(2.19)$$

$$d = \sqrt{(a_n - b_n)^2} \dots\dots\dots(2.20)$$

Keterangan :

a = hasil ekstraksi ciri data uji

b = hasil ekstraksi ciri data latih

n = nilai ciri ke-n

2.7 Tajwid Hukum Bacaan Dalam Surat Al-Fatihah

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِیْمِ
 الْحَمْدُ لِلّٰهِ رَبِّ الْعٰلَمِیْنَ ۝ الرَّحْمٰنِ الرَّحِیْمِ ۝
 مٰلِكِ یَوْمِ الدِّیْنِ ۝ اِیَّاكَ نَعْبُدُ وَاِیَّاكَ نَسْتَعِیْنُ ۝
 اِهْدِنَا الصِّرَاطَ الْمُسْتَقِیْمَ ۝ صِرَاطَ الَّذِیْنَ اَنْعَمْتَ
 عَلَیْهِمْ ۝ غَیْرِ الْمَغْضُوْبِ عَلَیْهِمْ وَلَا الضَّالِّیْنَ ۝

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Ilmu tajwid yaitu ilmu yang mempelajari tentang teknik mengeluarkan huruf sesuai dengan makhrajnya dan memberikan hak dan karakteristiknya dengan tujuan menghindari kesalahan lisan dalam mengucapkan huruf (Baharuddin, 2012).

Tajwid hukum bacaan yang ada dalam surat Al-Fatihah diantaranya :

1. *Lamjalalah Tarqiq*, yaitu hukum bacaan yang terdapat pada lafal Allah pada bacaan *bismillah*. Lafal Allah di dahului oleh huruf yang berharakat *kasrah* maka lafal Allah dibaca tipis.
2. *Alif Lam Syamsiah*, yaitu hukum bacaan *alif lam* (ال) apabila bertemu salah satu huruf *syamsiyah* (ت ث د ذ ر ز س ش ص ض ط ظ ن) hukum bacaannya disebut *Idgham syamsiyah* atau *lam syamsiyah*. Cara membacanya harus dimasukan atau di *idgham* kan kepada huruf *syamsiyah* atau dengan kata lain tulisannya tetap ada tetapi tidak dibaca *alif lam* nya.
3. *Alif Lam Qomariyah*, yaitu hukum bacaan *alif lam* (ال) apabila bertemu salah satu huruf *Qomariyah* (ا ب ج ح خ ع غ ف ق ك م و غ ه ء ي) *Alif lam Qomariyah* juga disebut dengan istilah *Idzhar Qomariyah* karena sesuai dengan hukum bacaan *Idzhar*, *Alim lam Qomariyah* terbaca dengan jelas atau dengan kata lain huruf *alim lam* dibaca sesuai dengan lafalnya yakni 'Al'.
4. *Mad 'Aridlissukun*, yaitu memanjangkan suatu bacaan ketika berhenti (*wakaf*) baik berhenti di akhir ayat ataupun juga di tengah ayat Al Qur'an. Panjang bacaan dari *Mad 'Aridlissukun* adalah boleh 1 *alif* (2 harakat), 2 *alif* (4 harakat), atau pun 3 *alif* (6 harakat).
5. *Idzhar Safawi*, ketika huruf hijaiiah *Mim Sukun* (مْ) ketemu dengan seluruh huruf hijaiiah, selain huruf hijaiiah *Mim* dan huruf hijaiiah *Ba*.
6. *Mad Thobi'i*, ketika ada huruf *alif* (ا) letaknya sesudah huruf *fathah* atau *Ya' sukun* (يْ) sesudah *kasrah* (— ◌) atau *Wau* (و) sesudah *Dammah* (— ◌) maka dihukum *mad thabi'i*. *Mad* artinya panjang, membacanya harus panjang 1 *alif* (dua harakat).
7. *Mad Lin* atau *Layyin*, ketika ada huruf *Wau sukun* (وْ) atau *Ya' sukun* (يْ) sedangkan huruf sebelumnya yaitu berharakat *fathah*, maka cara membacanya sekedar lunak dan lemas.

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

8. *Idzhar*, apabila terdapat *Nun sukun* (نْ) ataupun juga tanwin َ, ُ, ِ bertemu dengan huruf *Idzhar* Alif (ا), ‘Ain (ع), Ghain (غ), Ha (ح), Kha (خ), Ha’ (ه) dan Hamzah (ء) hurufnya bersambung atau dalam satu kata maka hukum adalah hukum *Idzhar Halqi* dan cara membacanya harus jelas atau terang dan tidak disertai dengung.
9. *Mad Lazim Mutsaqqal kilmi* atau *Muthawwal*, ketika huruf dari *Mad Thobi’i* (اِي وَ يِ) bertemu dengan huruf hijaiyah yang ber *tasydid* panjang bacaan pada *Mad Lazim Mutsaqqal kilmi* ini yaitu harus 3 alif atau 6 harakat

Tabel 2.1 Hukum Bacaan Surah Al-Fatihah Pada Ayat 1

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ		
Karena lafadz “allah” الله didahului oleh tanda (<i>kasroh</i>) maka bunyi vokalnya ‘a’.	<i>Lamjalat Tarqiq</i>	بِسْمِ اللَّهِ
Kerna ada “ <i>alif-lam</i> ” (ال) dan (<i>tasydid</i>) bertemu huruf <i>Syamsiyah</i> yaitu “ra” (ر).	<i>Alif-Lam Syamsiyah</i>	هِ الرَّ
Karena ada “ <i>fatah</i> ” berdiri pada huruf “ <i>mim</i> ” (م).	<i>Mad Ashli/Mad Thobi’i</i>	رَحْمَنِ
Kerna ada “ <i>alif-lam</i> ” (ال) dan (<i>tasydid</i>) bertemu huruf <i>Syamsiyah</i> yaitu “ra” (ر).	<i>Alif-Lam Syamsiyah</i>	نِ الرَّ
Karena ada huruf <i>Mad Ashli</i> “ <i>ya</i> ” (<i>sukun</i>) (يْ) sesudah (<i>kasrah</i>) menghadapi huruf hidup “ <i>mim</i> ” (م) yang di wakafkan.	<i>Mad ‘Aridlissukun</i>	رَحِيمِ
		5



Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Tabel 2.4 Hukum Bacaan Surah Al-Fatihah Pada Ayat 4

مَلِكِ يَوْمِ الدِّينِ		
Karena ada (<i>fatah</i>) berdiri pada huruf “ <i>mim</i> ” (م).	<i>Mad Ashli/Thabi'i</i>	مَلِكِ
Karena ada “ <i>wau</i> ” (<i>sukun</i>) (و) sesudah (<i>fatah</i>) pada huruf “ <i>ya</i> ” (ي)	<i>Mad Lin</i>	يَوْمِ
Kerna ada “ <i>alif-lam</i> ” (ال) dan (<i>tasydid</i>) bertemu huruf <i>Syamsiyah</i> yaitu “ <i>dal</i> ” (د).	<i>Alif-Lam Syamsiyah</i>	مِ الدِّ
Karena ada huruf <i>Mad Ashli</i> “ <i>ya</i> ” (<i>sukun</i>) (يْ) sesudah (<i>kasrah</i>) menghadapi huruf hidup “ <i>nun</i> ” (ن) yang di wakafkan.	<i>Mad ‘Aridlissukun</i>	دِينَ
		4

Tabel 2.5 Hukum Bacaan Surah Al-Fatihah Pada Ayat 5

إِيَّاكَ نَعْبُدُ وَإِيَّاكَ نَسْتَعِينُ		
Karena ada “ <i>alif</i> ” (ا) sesudah (<i>fatah</i>) pada “ <i>ya</i> ” (ي).	<i>Mad Ashli/Thabi'i</i>	إِيَّاكَ نَعْبُدُ
Karena ada “ <i>alif</i> ” (ا) sesudah (<i>fatah</i>) pada “ <i>ya</i> ” (ي).	<i>Mad Ashli/Thabi'i</i>	وَإِيَّاكَ
Karena ada huruf <i>Mad Ashli</i> “ <i>ya</i> ” (<i>sukun</i>) (يْ) sesudah (<i>kasrah</i>) menghadapi huruf hidup “ <i>nun</i> ” (ن) yang di wakafkan.	<i>Mad ‘Arid Lissukun</i>	نَسْتَعِينُ
		3



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

algoritma pengukuran jarak *Euclidean Distance*. Berikut penelitian-penelitian terdahulu yang juga sudah membahas metode LPC dan *Euclidean Distance* dengan ruang lingkup dan studi kasus yang berbeda :

2.9.1 *Speech Recognition Dengan Ekstraksi Fitur Linear Predictive Coding dan Jst Cerebellar Model Articulation Controller*

Pada penelitian ini menjelaskan bahwa semakin besar nilai Orde LPC bukan berarti semakin bagus hasil ekstraksi ciri nya, orde yang dihasilkan terletak pada nilai orde $p=10$. *Learning Rate* akan mempengaruhi kecepatan pembaharuan bobot jaringan. Pada penelitian tersebut diambil nilai rata-rata *learning rate* nya yang kemudian menghasilkan akurasi maksimal terletak pada *learning rate* yang besar yaitu $0.5 - 0.9$, dengan menggunakan data *training* dan testing dapat dilakukan secara selektif dengan demikian hasil pengujian menghasilkan performansi sistem yang cukup tinggi.

Pada penelitian tersebut tingkat akurasi mencapai 80% dengan menggunakan metode LPC dan JST *Articulation Controller* dengan menentukan nilai orde sebagai *learning rate*. Pada penelitian ini, penulis berharap bisa mendapatkan tingkat akurasi lebih tinggi penelitian sebelumnya dengan metode fitur pencocokan berbeda, karena pada penelitian kali ini penulis menggunakan *Euclidean Distance* (Fahmi, Suyanto, & Purnama, 2012).

2.9.2 *Perbandingan Metode Ekstraksi Ciri Suara Mfcc, Zcpa, Dan Lpc*

Pada kasus ini, sistem pengenalan suara mencoba untuk mensimulasikan sistem pendengaran manusia untuk mendapatkan hasil yang optimal dengan membandingkan beberapa metode yang biasa digunakan dalam proses ekstraksi suara. Langkah paling penting dalam proses pengakuan untuk mengekstraksi ciri suara, jadi ada banyak teknik ekstraksi ciri yang digunakan. Bila menggunakan CHMM sebagai klasifikasi peneliti melihat bahwa MFCC memberikan hasil yang baik dalam waktu yang terkecil dibanding dengan LPC dan PS-PS-ZCPA.

Berdasarkan penelitian tersebut metode yang cukup efektif digunakan dalam proses ekstraksi ciri adalah MFCC dibandingkan LPC karena MFCC dapat mengekstrak hasil yang lebih bagus pada sistem pendengaran manusia. Pada penelitian kali ini metode yang digunakan dalam proses ekstraksi ciri adalah LPC

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dan *Euclidean Distance* yang nanti akan dibandingkan tingkat akurasi dengan MFCC dengan lingkup dan kasus berbeda (Riyanto & Sutejo, 2014).

2.9.3 Desain dan Implementasi *Automatic Video Captioning* Dengan *Speech Recognition* Menggunakan *Hidden Markov Model*

Dilihat dari hasil penelitian tersebut akurasi maksimum yang didapatkan untuk sistem *Automatic Video Captioning* dengan *speech recognition* menggunakan LPC untuk *feature extraction* dan *Hidden Markov Model* untuk *feature matching* adalah sebesar 75,50% yang didapat pada sistem dengan jumlah *cluster* 256, jumlah *state* 6 dan jumlah data latih 90 untuk setiap suku kata.

Pada penelitian ini penulis juga akan menggunakan metode LPC sebagai ekstraksi ciri dan di sini penulis tidak menggunakan *Hidden Markov Model* sebagai proses perhitungan tetapi penulis menggunakan *Euclidean Distance*. Dengan metode perhitungan dan kasus berbeda diharapkan nantinya memberikan nilai akurasi yang lebih (Dimasatria, Virgono, & R. Rumani M., n.d.).

2.9.4 *A Novel Approach To Stuttered Speech Correction*

Tujuan dari penelitian ini adalah penggunaan pengenalan suara untuk mendeteksi kebenaran pengucapan seseorang dalam berpidato, dalam penelitian tersebut juga digunakan metode LPC sebagai ekstraksi ciri dan ASR sebagai koreksi terhadap ucapan pidato yang baik. Pada kasus ini sampel kebenaran pengucapan saat berpidato adalah 3 sampel sehingga nantinya akan memberikan rekonstruksi pidato yang benar-benar baik

Berdasarkan uraian di atas penelitian tersebut menggunakan 3 sampel pidato yang benar yang nantinya akan di bandingkan dengan data *training* sehingga didapatlah persentase kecocokan antara data sampel dan data *training*. Pada penelitian ini penulis juga menggunakan data sampel dan *training* sebagai patokan untuk persentase kecocokan (Ajibola, Khair, Sediono, & Nur, 2016).

2.9.5 Pengenalan Suara Manusia Menggunakan Metode LPC Dengan Jst-BP

Pada penelitian ini dijelaskan bahwa sinyal suara menggunakan LPC memberikan hasil yang cukup baik karena mampu menunjukkan kemiripan fitur dari sinyal suara *training* terhadap sampel. Pelatihan jaringan dengan mengubah-

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
- b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

ubah jumlah *hidden neuron* pada *hidden layer* dapat mempengaruhi kinerja jaringan. Dari percobaan tersebut didapat, semakin besar jumlah *hidden neuron* maka semakin sedikit iterasi yang dilakukan untuk mencapai target error yang diinginkan tetapi waktu untuk mencapai target error menjadi lebih lama, dengan mengubah-ubah nilai *learning rate* dapat mempengaruhi kinerja jaringan dari percobaan didapat, semakin besar nilai *learning rate* maka semakin sedikit iterasi yang dilakukan untuk mencapai target error yang diinginkan.

Berdasarkan hasil pengujian dari penelitian ini, metode JST-BP memiliki akurasi pengujian 80% dengan melakukan pengenalan terhadap *blind-set database*. Pada penelitian kali ini, penulis berharap bisa mendapatkan tingkat akurasi lebih tinggi atau sama dengan penelitian sebelumnya dengan menggunakan metode ekstraksi fitur yang sama yakni LPC dan pengukuran jarak dengan *Euclidean Distance* (Chandra, Tritoasmoro, & Rizal, 2008).

2.9.6 Modified Perceptual Linear Prediction Liftered Cepstrum (MPLPLC) Model For Pop Cover Song Recognition

Pada penelitian kali ini cukup menarik untuk penulis cermati sebagai bahan pelajaran di penelitian yang akan penulis lakukan, model MPLPLC, merupakan model baru atau model percabangan ekstraksi fitur yang di adopsi dari LPC yang pada penelitian tersebut digunakan untuk ekstraksi fitur dari sinyal suara musik pop. Gagasan utama dari penelitian tersebut adalah memodifikasi model PLP yang merupakan suatu teknik dalam ASR dengan mengambil perbedaan antara pidato dan sinyal suara musik. Selanjutnya, pemrosesan dengan LPC yang digabungkan dalam model varian lain yang hasilnya eksperimen itu menunjukkan bahwa MPLPLC mencapai identifikasi akurasi yang tinggi.

Penulis berharap pada penelitian ini bisa mendapatkan nilai akurasi yang tinggi juga dan mengembangkan metode yang penulis gunakan sehingga kelak di penelitian yang akan datang bisa menciptakan model baru dengan nilai akurasi yang lebih baik dari penelitian sebelumnya (Downie, Xiao, Zhu, Zhu, & Chen, 2015)