

IMPLEMENTASI *HIDDEN MARKOV MODELS* (HMM) SEBAGAI FILTER UNTUK MEREDUKSINOISE PADA *ESOPHAGEAL SPEECH*

Aprilia Fitriatul Aisyah

Program Studi S1 Fisika, FMIPA, Universitas Negeri Surabaya

Email : apriliah.f.aisyah15@gmail.com

Drs. Andy Noortjahja, M.T.

Dosen Fisika, FMIPA, Universitas Negeri Surabaya

Email : anduro_k29@yahoo.com

Abstrak

Suara manusia yang digunakan sebagai alat komunikasi verbal terjadi melalui proses pengencangan laring. Sehingga akibat pengencangan laring tersebut, laring bergetar. Adanya tekanan udara dari paru-paru kemudian disempurnakan dengan getaran udara yang berasal dari laring tersebut melalui bagian-bagian yang terdapat pada rongga mulut seperti lidah, gigi, dan bibir. Namun terkadang ada suatu kasus dimana seseorang tidak bisa lagi berbicara (bersuara) normal. Hal tersebut terjadi karena adanya kerusakan pada pita suara. Banyak faktor yang menyebabkan seseorang mengalami kerusakan pita suara, namun yang paling parah disebabkan karena adanya kanker pada pita suara. Apabila seseorang telah terkena kanker pita suara, dan sel kanker telah 80% menutupi *glottis* maka tindakan terbaik pada kondisi ini yaitu operasi pengangkatan pita suara (*laryngectomy*). Untuk mengembalikan kemampuan bicara pasien *laryngectomy* salah satunya dengan *esophageal speech* (ES). Akan tetapi, terlalu banyak *noise* pada ES sehingga suara penderita tuna laring hampir tidak bisa dimengerti. Pada penelitian yang bertujuan untuk mereduksi *noise* ini, penulis menggunakan *Hidden Markov Models* (HMM) sebagai filter. Sebagai perbandingan, penelitian ini juga menggunakan suara normal. Sehingga penulis dapat membandingkan perbedaan suara normal dan ES baik dari segi sinyal suara, spectrogram, spektrum, dan hasil analisa HMMnya. Hasil akhir yang diperoleh berupa gelombang sinus hubungan antara gain yang bersatuan desibel (dB) dan frekuensi (Hz). Selain itu, juga diperoleh nilai MSE dan SNR. Pada ES diperoleh nilai MSE sebesar 8,2357 dan SNR sebesar 36,0305dB.

Kata Kunci : *esophageal speech*, HMM, MSE, dan SNR

Abstract

The human voice is used as a means of verbal communication occurs through a process of tightening of the larynx. So as a result of tightening of the larynx, the larynx vibrates. The air pressure from the lungs and then refined by air vibrations emanating from the larynx through the parts contained in the oral cavity such as the tongue, teeth and lips. But sometimes there is a case in which a person can no longer speak (voice) normal. This happens because the damage to the vocal cords. Many factors cause someone damaged vocal cords, but the most severe due to a cancer of the vocal cords. If a person has been exposed to vocal cord cancer, and cancer cells have 80% cover the glottis, the best course of action at this condition is surgical removal of the vocal cords (laryngectomy). To restore the ability to speak one laryngectomy patients with esophageal speech (ES). However, too much noise on the ES so that persons with hearing the sound of the larynx is almost incomprehensible. In a study that aims to reduce this noise, the authors use Hidden Markov Models (HMM) as a filter. As a comparison, this study also use normal voice. So I can compare the differences in a normal voice and ES both in terms of sound signals, spectrogram, spectrum, and the results of the analysis HMM. The final results obtained in the form of a sine wave that the relationship between dimensionless gain decibels (dB) and frequency (Hz). Moreover, it also obtained the value of MSE and SNR. At ES MSE values obtained by 8.2357 and SNR for 36.0305dB.

Keyword : *esophageal speech*, HMM, MSE, and SNR

PENDAHULUAN

Sinyal suara (*speech signal*) merupakan gelombang yang dihasilkan oleh tekanan udara yang berasal dari paru-paru, tekanan udara tersebut berproses melewati lintasan suara menuju mulut dan rongga hidung, sehingga frekuensinya berbeda-beda (Al-Akaidi, 2007). Sinyal suara juga dapat dianggap

sebagai kumpulan sinyal sinusoida kontinyu yang dibangkitkan oleh sistem suara manusia mulai dari paru-paru sampai dengan mulut dan hidung dengan kombinasi amplitudo, frekuensi dan fasa yang tak terhingga.

Wicara (*speech*) dihasilkan dari sebuah kerjasama antara paru-paru (*lungs*), *glottis* (dengan *vocal cords*) dan *articulation tract* (mulut dan rongga

hidung). Untuk menghasilkan sebuah suara ucapan (*voiced sounds*), paru-paru menekan udara melewati *epiglottis*, pita suara bergetar, udara dimasukkan melalui aliran udara dan tercipta sebuah gelombang dengan tekanan semu secara periodik (Fadlisyah, dkk, 2013).

Kehilangan pita suara sama saja dengan kehilangan kemampuan untuk memproduksi suara dan berbicara. Banyak faktor yang menyebabkan seseorang harus kehilangan pita suaranya, ada yang diakibatkan karena terdapat virus, seseorang yang mengalami kecelakaan hingga pita suaranya harus diangkat, atau karena terdapat sel kanker pada pita suara. Berbagai cara agar pasien yang kehilangan pita suara (*laryngectomy*) dapat mengembalikan kemampuan berbicaranya telah dilakukan. Terdapat tiga macam metode yaitu *Tracheoesophageal speech*, *Electrolarynx*, dan *Esophageal speech* (ES) (Nakamura 2010), (Nusdwiningtyas, 2007).

Dari ketiga metode tersebut, ES merupakan cara yang paling sederhana dan efisien karena tidak memerlukan alat bantu dan biaya mahal. ES merupakan terapi wicara dengan memanfaatkan pernapasan abdominal (pernapasan perut) seperti meniup, bersendawa, dan mengeluarkan suara dasar dengan menggunakan esofagus. Terapi wicara ini membutuhkan waktu 3-6 bulan, bahkan lebih. Terapi wicara menghasilkan kualitas suara frekuensi dasar yang intensitasnya rendah, getaran pada kerongkongan dan pernapasan tidak teratur dan banyaknya gangguan yang lain, sehingga suara yang dihasilkan terlalu banyak *noise*, kasar, serak, bahkan sulit dimengerti. Banyaknya *noise* pada ES menjadi permasalahan yang mendasar pada penelitian ini. Selanjutnya dengan menggunakan HMM diharapkan bisa mereduksi *noise* pada ES.

Speech processing dan *speech analysis* merupakan sebuah cara untuk menganalisis frekuensi suara. Melalui pengolahan sinyal *speech processing* ataupun *speech analysis*, dapat diketahui dari data pasien ES yang berupa data **wav*. Data **wav* dapat dianalisis untuk mengetahui karakteristik suara. Melalui karakteristik tersebut, dapat diketahui dan dianalisis bentuk spectrogram, *Linear Predictive Coding* (LPC), *format frequency*, *energi*, dan *pitch*. Dengan demikian,

pengenal karakteristik suara pada *esophageal speech* akan diketahui bentuk kerusakannya apabila dibandingkan dengan *normal speech*

Hidden Markov Models (HMM) merupakan model statistik dari sebuah sistem yang diumpamakan sebagai sebuah "Proses Markov" dengan parameter yang tidak terdeteksi, kita dapat menentukan parameter

yang tidak terdeteksi tersebut melalui parameter yang terdeteksi. HMM sudah banyak diterapkan di bidang *signal processing*, dan *speech processing*.

Berdasarkan uraian di atas, diperoleh rumusan masalah pada penelitian ini adalah bagaimanakah hasil reduksi *noise* pada ES apabila menggunakan HMM sebagai filter.

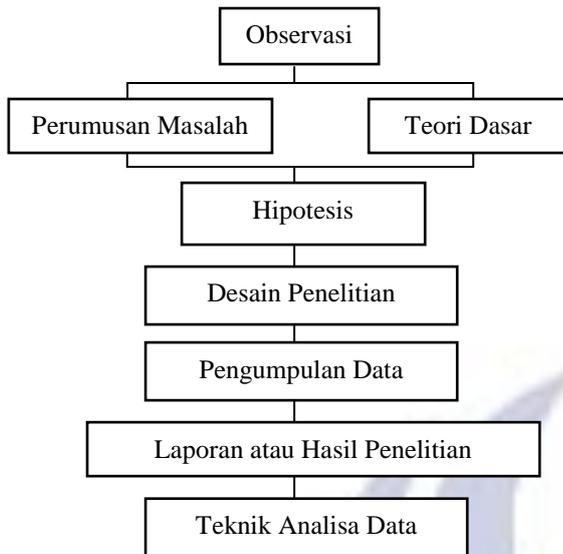
Tujuan penelitian ini adalah mengetahui hasil reduksi *noise* pada ES apabila menggunakan HMM sebagai filter.

Penelitian ini difokuskan pada permasalahan mereduksi *noise* pada ES. Selanjutnya mengaplikasikan HMM sebagai filter untuk mereduksi *noise* tersebut. Output dari pemodelan tersebut sebagai hasil yang akan menunjukkan adanya reduksi *noise* pada ES setelah mengaplikasikan HMM sebagai filter.

Sinyal wicara (*speech signal*) adalah sinyal yang heterogen sebagai fungsi waktu, dan waktunya tergolong lambat, ketika diteliti pada periode yang sangat singkat (yaitu periode 5-100ms) karakteristik sinyal yang dihasilkan masih statis (tetap) (Prativi, dkk, 2012). Namun, apabila diteliti pada periode yang lebih panjang yaitu pada periode $>0,2$ detik, karakteristik sinyal berubah untuk menggambarkan suara yang dimunculkan oleh pembicara.

Transformasi yang umum digunakan untuk merubah sinyal suara dari domain waktu ke domain frekuensi adalah transformasi Fourier, dimana sinyal dapat diuraikan menjadi komponen-komponen sinusoida. Dengan penguraian tersebut, maka sinyal dapat ditunjukkan dalam *Domain Frekuensi*. Pada sinyal yang teratur, penguraian menjadi komponen sinusoida disebut *Deret Fourier*. Dengan melakukan analisis frekuensi, dapat dilihat representasi matematis dari komponen-komponen frekuensi yang terdapat dalam suatu sinyal. Rangkaian frekuensi tersebut disebut dengan *spectrum*. Cara yang paling dikenal untuk mempresentasikan sinyal wicara adalah *spectrogram* yang merupakan suatu bentuk *gray scale image* yang mempresentasikan nilai frekuensi sinyal pada waktu tertentu.

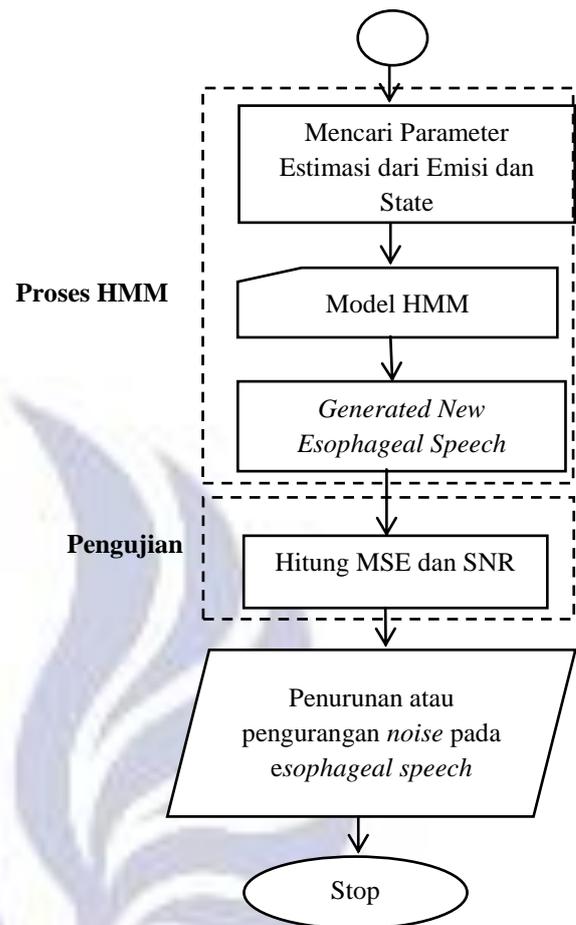
METODE



Gambar 1. Alur Rancangan Penelitian

Penelitian ini dilakukan dengan cara merekam suara penderita tuna laring yang berada di bawah naungan Perkumpulan Wicara Esophagus (PWE) yang berada di Surabaya.

Pengambilan data terhadap subyek yaitu pasien penderita tuna laring. Pengambilan suara tersebut berbentuk “data.wav” yang bebas dari *noise* lingkungan. Kemudian suara tersebut dianalisa menggunakan matlab dengan menerapkan HMM.



Gambar 2. Flowchart Penelitian

Parameter pada penelitian ini adalah sebagai berikut:

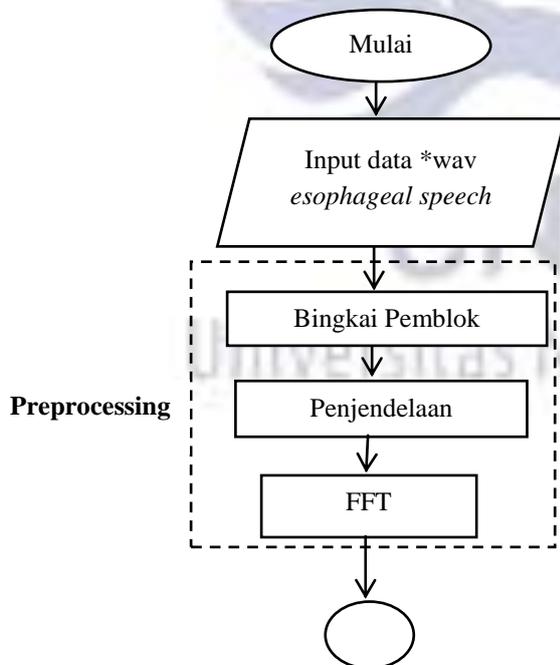
- a) Variabel kontrol : Pemodelan HMM ($A, \pi, b, \log P$).
- b) Variabel manipulasi : Suara pasien tuna laring (ES signal).
- c) Variabel respon : Pengurangan *noise* pada ES.

Terdapat beberapa langkah pengolahan data pada penelitian ini. Langkah-langkah tersebut meliputi:

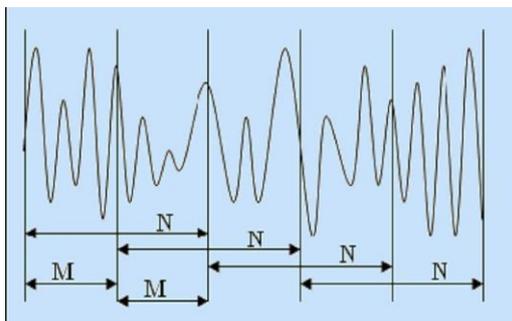
a) Frame Blocking (Bingkai Pemblok)

Frame blocking merupakan pembagian sinyal suara menjadi beberapa *frame* dan satu *frame* terdiri atas beberapa sampel. Pengambilan sampel tersebut tergantung dari berapa detik suara yang akan dijadikan sampel dan berapa besar frekuensi sampelnya. Representasi fungsi *frame blocking* sebagai berikut :

$$x(n) = y(M+n) \dots \dots \dots (1)$$



dimana, $x(n)$ = sinyal sesudah di *frame blocking*
 y = sinyal hasil *preemphasis*
 M = *overlapping frame*, dan $n = 1,2,3$,



Gambar 3. Bentuk sinyal yang di *frame blocking*

(Sumber : Maulidia, N, 2009)

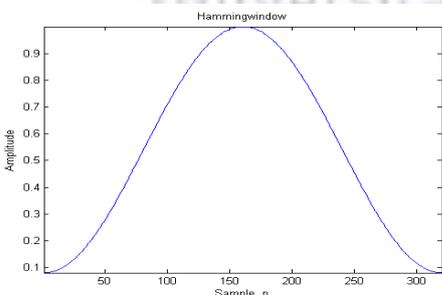
Proses *framing* ini umumnya dilakukan secara *overlapping* untuk setiap *frame*-nya. Panjang daerah *overlapping* yang umum digunakan sekitar 30% sampai 50% dari panjang *frame*.

b) Windowing (Penjendelaan)

Sinyal suara yang dipotong-potong menjadi beberapa *frame* akan menyebabkan kesalahan data pada proses *fast fourier transform*. Windowing dibutuhkan untuk mereduksi efek diskontinuitas dari sinyal yang dipotong. Jenis *windowing* ada beberapa macam yaitu, Hanning, Bartlet, Hamming, Blackman, dan Rectanguler. *Windowing* yang dipakai pada proses ini adalah *windowing hamming*. Persamaan *windowing hamming* adalah sebagai berikut:

$$w(k) = 0,54 - 0,46 \cos\left(\frac{2\pi k}{K-1}\right) \dots (2)$$

dimana, $k = 1,2,3, \dots$; $\pi = 3.14$; $K =$ panjang *frame*.



Gambar 4. Bentuk sinyal *windowing hamming*
 (Sumber: Nick Bardici et al, 2006)

c) Fast Fourier Transform (FFT)

Fast fourier transform ini dilakukan untuk mentransformasikan sinyal dari domain waktu ke domain frekuensi. Hal ini bertujuan agar sinyal dapat diproses dalam spectral subtraksi.

FFT adalah bentuk khusus dari persamaan integral *fourier* :

$$H = \int h(t) e^{-j\omega t} dt \dots \dots \dots (3)$$

Dengan mengubah variabel-variabel, waktu (t), frekuensi (ω) kedalam bentuk diskrit diperoleh transformasi fourier diskrit (DFT) yang persamaannya adalah:

$$H(k\omega_0) = \sum_{n=0}^{N-1} (nT) a^{-jk\omega_0 nT} \dots \dots \dots (4)$$

Disederhanakan dengan $T =$ sampel waktu, $N =$ sampel frekuensi k sehingga menjadi :

$$H(k) = \sum h(n) e^{-j2\pi kn/N} \dots \dots \dots (5)$$

Dimana $k = 0,1,2, \dots, N-1$

d) Parameter Pereduksi Noise

Parameteryang digunakan untuk mengetahui kemampuan pereduksinoise yang dibuat adalah *Signal to Noise Ratio* (SNR) dan *Mean Square Error* (MSE).

1) Signal-to-Noise Ratio (SNR)

SNR adalah perbandingan antara daya sinyal asli terhadap daya sinyal *noise*. Nilai SNR yang semakin tinggi menunjukkan bahwa kualitas sistem pereduksinoise semakin baik. Besarnya SNR dapat dihitung dengan persamaan :

$$SNR(\text{dB}) = 10 \log \left[\frac{P_1}{|P_2 - P_1|} \right] \left[\frac{P_1}{|P_2 - P_1|} \right] \dots \dots \dots (6)$$

P_1 adalah kekuatan sinyal awal, $|P_2 - P_1|$ adalah kekuatannoise yang tersisa, dan P_2 adalah kekuatan sinyal hasil perhitungan.

2) Mean Square Error (MSE)

Perbedaan antara sinyal perhitungan dan sinyal awal menyebabkan adanya *error*. Kualitas pereduksi *noise* dapat dikatakan baik, jika nilai MSE yang dihasilkan semakin mendekati nol. Nilai MSE dapat dihitung dengan persamaan :

$$MSE = \frac{1}{N} \sum_{n=0}^{N-1} [u(n) - e(n)]^2 \dots \dots \dots (7)$$

dengan $u(n)$ adalah sinyal awal, $e(n)$ adalah sinyal perhitungan, dan N adalah panjang sampel sinyal.

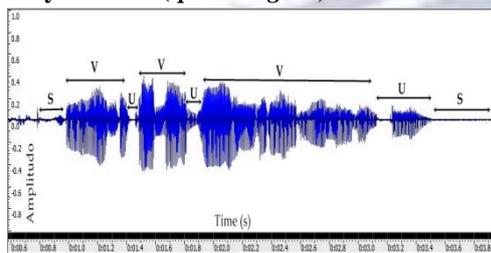
Analisa data pada penelitian ini menggunakan Uji-t. Uji-t adalah jenis pengujian statistika untuk mengetahui apakah ada perbedaan dari nilai yang diperkirakan dengan nilai hasil perhitungan statistika. Uji-t pada penelitian ini menggunakan microsoft excel. Terdapat beberapa pilihan untuk melakukan uji-t, sesuai yang diperoleh:

- a) **t-Test: Two-Sample Assuming Equal Variances.** Sarana ini berguna untuk menguji kecocokan nilai rata-rata dua populasi dengan uji-t, asumsi varians kedua populasi sama dan dikenal dengan nama **uji-t homoskedastik**.
- b) **t-Test: Two-Sample Assuming Unequal Variances.** Sarana ini berfungsi untuk menguji kecocokan nilai rata-rata dua populasi dengan uji-t, asumsi varians kedua populasi tidak sama dan dikenal dengan **uji-t heteroskedastik**. Uji ini digunakan apabila kedua sampel berasal dari dua populasi yang berbeda.
- c) **t-Test: Paired Two Sample For Means.** Sarana ini digunakan untuk menguji perbedaan nilai rata-rata sepasang data dengan menggunakan uji-t. Uji ini tidak mengasumsikan kecocokan varians kedua populasi. Berfungsi untuk menguji perbedaan nilai rata-rata dua set data yang berpasangan, misalnya hasil pengamatan sebelum dan sesudah adanya metode terhadap sekelompok populasi yang sama.

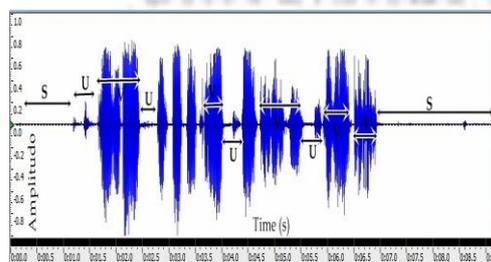
HASIL DAN PEMBAHASAN

Hasil penelitian ini berupa sinyal suara, spectrogram, spektrum, analisa uji-t, dan hasil implementasi HMM.

a) **Sinyal Suara (speech signal)**



Gambar 5. Sinyal suara normal

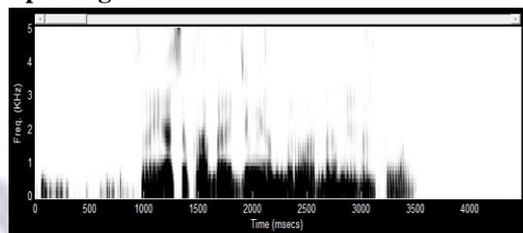


Gambar 6. Sinyal esophageal speech

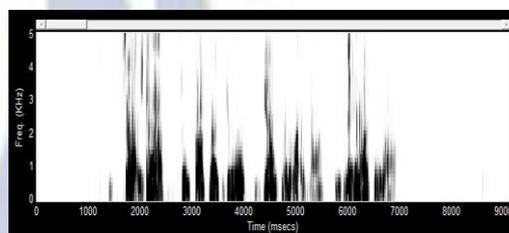
Berdasarkan stimulasi yang terciptaselama proses produksi suara, sinyal suara dibagi menjadi tiga bagian yaitu *silence*(S), *unvoiced*(U), dan *voiced*(V).

Silence merupakan sinyal pada saat tidak terjadi proses produksi suara ucapan, sinyal *unvoiced* terbentuk ketika pita suara tidak bergetar, dimana sinyal stimulasi berupa sinyal acak, dan sinyal *voiced* terbentuk ketika pita suara mulai bergetar, yaitu ketika sinyal stimulasi berbentuk sinyal quasi-periodik. Selama terbentuk sinyal *voiced* ini pita suara bergetar pada frekuensi dasar

b) **Spectrogram**



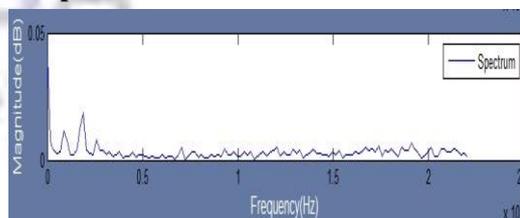
Gambar 7. Spectrogram suara normal



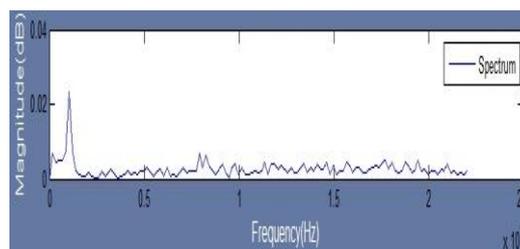
Gambar 8. Spectrogram esophageal speech

Pada gambar 7 dan gambar 8 diatas, kita dapat melihat spectrogram dari suara normal dan suara *esophageal*. Dari spectrogram tersebut, kita dapat menganalisa tingkat energi dari suara normal maupun suara *esophageal*. Spectrogram mempresentasikan frekuensi sinyal yang berubah terhadap waktu. Waktu ditunjukkan sepanjang sumbu x, frekuensi pada sumbu y, dan tingkat energi sinyal pada waktu tertentu ditunjukkan dengan tingkat warna.

c) **Analisis Fast Fourier Transform(FFT) dan Spektrum**



Gambar 9. Spektrum suara normal



Gambar 10. Spektrum esophageal speech

Sinyal dapat diuraikan menjadi komponen-komponen sinusoida. Dengan penguraian komponen sinusoida tersebut, sinyal dikatakan direpresentasikan dalam domain frekuensi. Domain frekuensi akan terlihat ketika sumbu waktu tidak lagi terlihat dan hanya nilai-nilai positif dari magnitude tiap-tiap sinus yang diplot, disebut magnitude spektrum.

d) Uji-t

Tabel 1. t-Test ES sebelum dan sesudah HMM

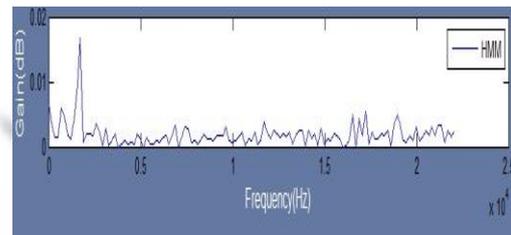
	Variable 1	Variable 2
Mean	0,002068116	0,003231581
Variance	9,10E-06	1,35718E-05
Observations	69	69
Pearson Correlation	0,377381429	
Hypothesized Mean Difference	0	
Df	68	
t Stat	2,556912558	
P(T<=t) one-tail	0,006398593	
t Critical one-tail	1,667572281	
P(T<=t) two-tail	0,012797187	
t Critical two-tail	1,995468931	

Tabel 2. t-Test: Suara Normal dan suara esophageal setelah HMM

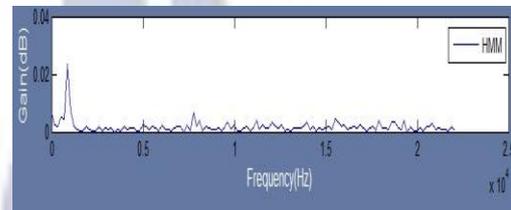
	Variable 1	Variable 2
Mean	0,002068116	0,003016864
Variance	9,10456E-06	4,90191E-06
Observations	69	69
Hypothesized Mean Difference	0	
df	125	
t Stat	-2,105770385	

P(T<=t) one-tail	0,018612524
t Critical one-tail	1,657135178
P(T<=t) two-tail	0,037225048
t Critical two-tail	1,979124109

e) Hasil implementasi Hidden Markov Models (HMM)



Gambar 11. HMM suara normal



Gambar 12. HMM esophageal speech

Implementasi HMM pada penelitian ini menghasilkan gain (penguatan) gelombang sinus. Gain pada penelitian ini memiliki satuan desibel (dB). Selain menganalisa besarnya penguatan, penelitian ini juga mendapatkan besarnya nilai *mean square error* (MSE) dan *signal to noise ratio* (SNR).

Untuk suara normal, diperoleh nilai MSE sebesar 1,5980 dan SNR sebesar 21,1942dB. Sedangkan untuk suara *esophageal* diperoleh MSE dan SNR masing-masing sebesar 8,2357 dan 36,030dB.

Pada suara *esophageal*, sinyal suara lebih banyak terpotong oleh sinyal *unvoiced*. Hal tersebut menyebabkan kurang jelasnya artikulasi kata yang diucapkan. Dari segi waktu pun, suara *esophageal* memiliki waktu yang relatif lebih panjang jika dibandingkan dengan suara normal (dengan pengucapan kalimat yang sama).

Semakin gelap (hitam) bidang spectrogram mengindikasikan adanya energi yang disebabkan adanya lipatan vokal, harmonik, atau getaran formant dalam sinyal suara. Dan sebaliknya, semakin terang (mendekati putih) bidang spectrogram mengindikasikan hanya terdapat sedikit energi. Apabila kita membandingkan spectrogram suara normal dan suara *esophageal*, maka akan terlihat bahwa suara *esophageal* memiliki tingkat energi yang lebih rendah daripada suara normal.

Komponen spektrum sebenarnya berupa garis yang berfungsi untuk menunjukkan posisi cosinus. Sebuah spektrum dapat digambarkan hanya dengan menunjukkan titik akhir setiap garis yang kemudian saling terhubung. Untuk spektrum yang memiliki banyak titik, celah antara garis hampir tidak terlihat dan spektrum akan mencul secara kontinu (berkelanjutan), dengan mengabaikan komponen spektrum yang disebut garis.

Magnitude spektrum tidak cukup untuk menentukan sinyal secara mutlak. Fase spektrum tiap komponen frekuensi harus diketahui dengan jelas untuk menggambarkan hal itu. Dengan begitu, jelas bahwa komponen spektrum pada frekuensi tertentu harus dijelaskan dengan dua parameter. Magnitude dan fase (*phase*), keduanya adalah salah satu contoh dari dua parameter.

Analisis domain frekuensi atau estimasi kerapatan spektral adalah proses teknis dekomposisi sinyal yang kompleks menjadi bagian-bagian yang lebih sederhana. Dengan analisis frekuensi, kita dapat mengetahui representasi matematis komponen-komponen frekuensi yang terdapat dalam sinyal. Urutan frekuensi yang terdapat dalam sinyal tersebut disebut spektrum. Banyak proses fisik yang digambarkan sebagai jumlah dari banyaknya komponen frekuensi individu. Setiap proses yang mengkuantifikasi jumlah (misalnya amplitudo, kekuatan, intensitas, atau fase) versus frekuensi, bisa disebut analisis spektrum. Analisis spektrum dapat dilakukan pada seluruh sinyal. Atau dapat dipecah menjadi segmen pendek (*frame*).

Melalui pengujian uji-t, diperoleh nilai t hitung lebih kecil apabila dibandingkan dengan t tabel. Pada penelitian ini, peneliti menguji perbedaan rerata nilai magnitude suara *esophageal* sebelum dan sesudah menerapkan metode HMM. Sesuai dengan syarat untuk menguji *t-Test: Paired Two Sample For Means*, maka peneliti menggunakan subjek yang sama, yaitu suara *esophageal* dengan kalimat "Saya suka baju yang berwarna merah tua". Pada pengujian ini diperoleh nilai t hitung sebesar $-2,5569$ dan t tabel dengan pengujian dua arah dengan taraf signifikan 2,5% sebesar $1,99547$. Hal itu berarti ($t_{hitung} = -2,5569$) < ($t_{0,025(68)} = 1,99547$). Artinya, nilai rerata dua populasi tidak sama dan itu menunjukkan tidak adanya korelasi dan pengaruh yang signifikan antara dua populasi yang diuji.

Untuk kasus ke-2, pengujian antara suara normal dengan suara *esophageal* sesudah HMM menggunakan *t-Test: Two-Sample Assuming Unequal Variances*. Pengujian ini bertujuan untuk menguji kesamaan dua buah populasi dengan uji-t, dengan

asumsi varians kedua populasi tidak sama. Uji ini digunakan karena sampel berasal dari dua subjek yang berlainan. Pengujian ini memperoleh nilai t_{hitung} sebesar $-2,10577$ dan t_{tabel} sebesar $1,9791$ dengan pengujian dua arah, taraf signifikan 2,5% dan df 125, atau dapat ditulis ($t_{hitung} = -2,10577$) < ($t_{0,025(125)} = 1,9791$). Hal itu berarti adanya metode HMM tidak berpengaruh signifikan terhadap hasil suara *esophageal* apabila dibandingkan dengan suara normal. Subjek yang digunakan berbeda (suara normal dan suara *esophageal* setelah diterapkan HMM) namun kalimat yang diucapkan sama yaitu "Saya suka baju yang berwarna merah tua".

Hasil kedua uji-t tersebut, berpengaruh pada hasil MSE dan SNR. Dimana hasil MSE dan SNR pada penelitian ini masih jauh dari keinginan peneliti. Nilai MSE untuk *esophageal speech* pada penelitian ini diperoleh sebesar $8,2357$ dan SNR sebesar $36,030$ dB. Sesuai dengan teori yang ada, nilai MSE dan SNR berbanding terbalik. Hasil penelitian ini dikatakan baik apabila nilai MSE semakin mendekati nol dan nilai SNR semakin mendekati 100.

PENUTUP

Simpulan

Setelah dilakukan penelitian tentang implementasi HMM sebagai filter untuk mereduksi noise pada *esophageal speech* dapat diambil kesimpulan bahwa hasil penelitian ini belum sesuai dengan keinginan peneliti. Pada penelitian ini diperoleh nilai MSE pada *esophageal speech* sebesar $8,2357$ dan SNR hanya sebesar $36,030$. Berdasarkan teori yang ada, penelitian dikatakan baik (berhasil) apabila nilai MSE semakin mendekati nol dan SNR semakin mendekati 100. Dan hasil akhir masih berupa gelombang sinus, belum berupa suara.

Saran

- Perlu dilakukan penelitian lanjutan untuk pengembangan implementasi HMM sebagai filter untuk mereduksi noise pada *esophageal speech* agar diperoleh hasil akhir berupa suara.
- Perlu dilakukan penelitian lanjutan untuk pengembangan implementasi HMM sebagai filter pada aplikasi-aplikasi dibidang pengolahan suara.

DAFTAR PUSTAKA

- Fadlisyah, dkk. 2013. Pengolahan Suara. Yogyakarta : Graha Ilmu.
- Tompunu, A.N. 2012. Implementasi Algoritma *Least Mean Square* untuk Peningkatan Kualitas Suara Penderita Tuna Laring Berbasis Processor

- TMS320C6713, Vol.2, No.2. Palembang : Politeknik Negeri Sriwijaya Palembang.
- Noortjahja Andy, Sardjono TA, Purnaomo HM. *Comparative Study of Wavelet and ANFIS to Reduce Gaussian Noise in Esophageal Speech* (seminar internasional *BME DAYS* di ITS), 2010.
- Estrada, Richie. 2008. *Pengolahan Sinyal Suara Menggunakan Matlab*. Jakarta : Fakultas Teknik, Universitas Kristen Krida Wacana.
- B.H. Juang, L.Rabiner. 1993. *Fundamentals of Speech Recognition*. United States of America: PTR Prentice-Hall, Inc. Englewood Cliffs, New Jersey 07632.
- Prasetyo, M.E.B. 2010. *Teori Dasar Hidden Markov Model*. Makalah II2092 Probabilitas dan Statistik – Seminar I Tahun 2101/2011.
- Abdulla,W.H, Kasabov,N.K. 1999. *The Concept of Hidden Markov Model in Speech Recognition*. New Zealand : Knowledge Engineering Lab, Department of Information Science, University of Otago.
- Maulidia, N. 2009. *Pembuatan Program Aplikasi untuk Menampilkan Ciri Sinyal Wicara dengan Matlab*. Surabaya: Jurusan Teknik Telekomunikasi, Politeknik Elektronika negeri Surabaya, ITS.
- Nick Bardici, Bjorn Skarin. 2006. *Speech Recognition using Hidden Markov Model*. Department of Telecommunications and Signal Processing, Blekinge Institute of Technology.

