

DEVELOPING QURANIC VERSE RECITATION LEARNING SOFTWARE BASED ON SPEECH RECOGNITION TECHNIQUES

B. Putra, B.T. Atmaja, D. Prananto

Department of Engineering Physics, Faculty of Industrial Technology

Sepuluh Nopember Institute of Technology

Sukolilo Campus, Surabaya 60111

Indonesia

email : giganggu@ep.its.ac.id, bagusta@scientist.com, d_prananto@ep.its.ac.id

Abstract

Indonesia is well known as the biggest Moslem citizen country in the world. The Muslim holy book, Al-Quran (Koran), need to be pronounced correctly. However, the limitation of teachers and time to study Quran verse recitation together in a class could be an obstacle in Qur'an recitation learning. Hence, a learning tool needed to be developed to overcome the obstacle. By implementing speech recognition techniques based on cepstral feature coefficient and Gaussian Mixture Model (GMM) modeling, we designed and developed Quranic verse recitation learning software, namely Virtual Quran Lecturer (VQL). VQL is interactive multimedia software which is developed to help and ease people to learn how to recite Qur'an correctly. This software has the ability to perform correction in Qur'an recitation and guide user to recite Qur'an in correct and proper manner with the help of tutorial provided.

Keywords: Quran verse recitation, learning software, speech recognition.

I. Introduction

As the biggest Muslim country in the world, Indonesia have mission to educate people of Indonesia in all sectors, no exception for Holy Quran recitation learning. The conventional learning method for Quran verse recitation learning is face to face method where student and teacher have to from teacher and student. The limited time to study together in a class and also the limited number of teachers result in not optimal learning process.

Developing learning software for Quran verse recitation is aimed to ease people in studying how to recite Quranic verses with correct and proper manner without leaving conventional method, the 'face to face' method. It means the software is just an additional tool to improve their Quranic verse recitation skills by themselves, not as main tool for learning Quran.

Designed learning software for Quranic verse recitation can be achieved by some methods. In this research, we use speech recognition techniques to develop the software. The techniques include cepstral feature and Gaussian Mixture model (GMM). This software will make correction and rectification of reading at each learning session if there is any error in reading. It is expected that people could learn to read the Qur'an easily without feeling doubt about the accuracy of pronunciation and recitation of Quranic verse recitation law (*tajweed*) because they feel have "virtual" mentors in that software and always make corrections in his learning.

1.1 Quranic Verse Recitation Law And Pronunciation

Here are some pronunciation and recitation law which is used as object in the development and the application of the VQL.

1.1.1 Makhraj

Makhraj means place of discharge. *Makhraj* in *hija'iyah* letters means place where *hija'iyah* letters come out from mouth. Considering that Arabic letters differ from Latin letters, hence Arabic letters pronounced in different manner to Malay words pronunciation. Pronunciation of arabic words/letter is determined by *makhraj* of the letter.

1.1.2 Mad

Mad means elongate tone. In tajweed course there are two kind of mad, i.e. *mad ashli/tabi'i* and *mad far'i*. *Ashli* means principal and *far'i* means subsidiary.

▪ *Mad Ashli/Thabi'i*

Mad ashli reading is read in two *harakah*. *Mad ashli* occur when there is *alif* after letters with *fathah*, *ya* consonant after letters with *kasrah*, and *waw* consonant after letters with *dhomah*. Furthermore, long-sign readings can be utilized if *alif*, *waw* consonant or *ya* consonant are not used.

▪ *Mad Far'i*

Mad far'i has many different types with different vowel length. The application of the *Mad* is also different for every *qiraat*.

1.2 Law of Nun Consonant and Tanween

The law of nun consonant and *tanween* can be classified in some types.

▪ *Izhhaar*

Izhhaar means clearly read. *Izhhaar* is readings where *nun* consonant or *tanween* meet *alif*, *hamzah*, ‘*ain*, *ghain*, *ha*, *kha*, and *ha*’ and read with clear sound.

▪ *Idgham*

Idgham means to get into/to change the tone of *nun* when *nun* consonant or *tanween* meet *idgham* letters. Each *idgham* readings read in two *harakah*. *Idgham* Letters are “*ya*”, “*ra*”, “*mim*”, “*lam*”, *waw*” and “*nun*”. There are two kinds of *idgham*:

- *Idgham Bighunnah*

This law occur when *nun* consonant meet “*ya*”, “*nun*”, “*mim*”, dan “*waw*”. *Bighunnah* means to buzz, when reading this kind of *idgham* the tone of the letter is inserted to the nose just like buzzing.

- *Idgham Bilaghunnah*

Idgham bilaghunnah is an inverse of *idgham bughunnah*, where the tone is not inserted into the nose. The letters of *idgham bilaghunnah* are “*lam*” and “*ra*”.

▪ *Iqlab*

Iqlab occur when *nun* consonant or *tanween* meet “*ba*”. The tone of “*ba*” in *Iqlab* readings changed to “*mim*” accompanied with drone. *Iqlab* readings are read in two *harakah*.

▪ *Ikhfa'*

Ikhfa' means to hide/vague. *Ikhfa'* readings read by vague voice of *nun* when *nun* consonant or *tanween* meet *ikhfa'* letters. All the readings with *ikhfa'* read in two *harakah*. *Ikhfa'* letters are the letters except in *izhhaar*, *idgham* and *iqlab*.

Ikhfa' divided into three groups based on the tone of the letter, that is *nun* consonant or *tanween* meet the letters of;

1. “*kaf*” and “*qaf*”, *nun* tone cahenged to buzz “*ng*”.
2. “*tsa*”, “*dal*” and “*tha*”, more vague than first case.
3. Other letters are read vaguer than first and second case.

1.3 Law of *Mimi* Consonant

Law of *mim* consonant is divided into several laws as follows;

▪ *Idgham Mutamatsilayn/Miimi*

idgam mutamatsilayn occurs when *mim* consonant meet “*mim*”, where the tone of *mim* in *mim* consonant is inserted to the next letter tone with buzz tone. This reading is read with two *harakah*.

▪ *Ikhfa Syafawii*

If *mim* consonant meet “*ba*”, then the tone of *mim* in *mim* consonant read with vague with a bit of buzz. It is read with two *harakah*.

▪ *Izhhaar Syafawii*

If *mim* consonant meet letters other than “*mim*” and “*ba*”, then the tone of *mim* in *mim* consonant read obviously. *Izhhaar syafawi* is read with one *harakah*.

1.4 *Ghunnah*

Ghunnah means to buzz. *Ghunnah* occur in two cases, that is when “*mim*” and “*nun*” use tasydid sign. *Ghunnah* is read with two *harakah*.

II. Speech Recognition Techniques

Signal processing is utilized to obtain the characteristics of pronunciation as identifier of faults and correction. Extraction of MFCC coefficient, signal energy, delta MFCC and delta-delta MFCC is conducted to subtract the feature of voices.

2.1 Voice Signal and its Occurrence Process

Voice is signal which is greatly influenced by frequency and a form of discrete signal which is influenced by time. The main component in voice production system is vocal tract. Vocal tract is a resonance tube-shaped object in voice production system which has main part called pharynx, nasal cavity and oral cavity. The vocal tract varies in shapes according to soft plate (*velum*), tongues, lips and jaw which overall called as articulators.

The process of human voice production can be explained as follows; the air flows from lungs to the trachea, a tube composed of cartilage rings, and goes through larynx to the vocal tract. Larynx reacts as a gate between lungs and mouth. Larynx composed of epiglottis, vocal cords and false vocal cords. These three components are closed when human swallow food, so that the food does not enter the lungs, and open again when the human inhale. Phoneme in English can be classified in terminology of “manner of articulation” and “place of articulation”.

Manner of articulation is concentrated on air flow, it means it concerns about track and the level of vocal that goes through. Manner of articulation and voicing divides Phoneme in three big classes. Phoneme that produce employ voicing and solely stimulate the vocal tract on glottis called as *sonorants* (vowels, diphthongs, glides, liquids and nasals). They have continuous, intents and periodic phonemes characteristics.

Voiced sounds is produced by air pressure which flows through vocal cords while vocal cords squeezed to open and close quickly to produce a series of puffs periodic which have fundamental frequency (first harmonics) same as vocal-cord's vibration frequency. The frequency of vocal cords depends on the level of solidity, tension, the length of vocal cords and air flow effect which produced in the glottis, a chamber between vocal cords. The component of this frequency is composed of a number of harmonics from fundamental frequency.

A sound that produced without vibration in vocal cords is called as unvoiced [1].

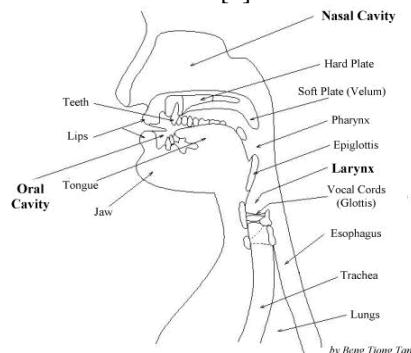


Figure 1. Human's mouth cavity [1]

The following figure illustrates a segment on vowel /ix/. A quasi-periodicity signal on voiced speech can be seen here.

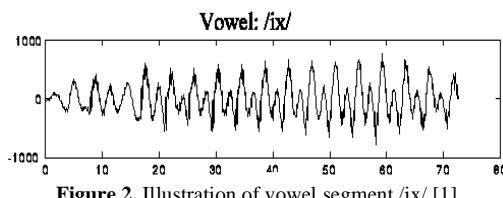


Figure 2. Illustration of vowel segment /ix/ [1]

Fricative sound is generated from the confinement of vocal tract and air flow pressure with high enough velocity to make turbulence. The turbulence is for instance /hh/ or /s/.

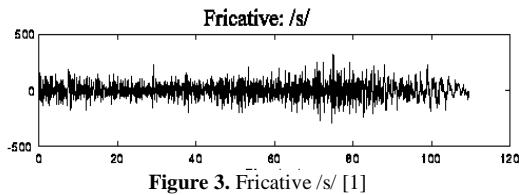


Figure 3. Fricative /s/ [1]

Plosive or stop sounds is generated from vocal tract blocking process by closing the lips and nasal cavity, enabling lateral air pressure, and followed by a beat. This mechanism will generate /p/ and /g/ voices.

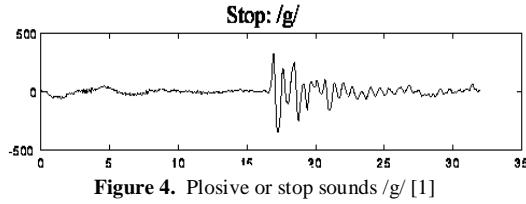


Figure 4. Plosive or stop sounds /g/ [1]

Affricates is a combination of stop and fricative sounds. Stops, fricatives and affricates collectively called as obstruent phonemes which is weak enough and periodic, and basically is a form that generated by blocking stimulus on main vocal tract. Vowel is basically described in terminology of tongue position and lips.

2.2 Voice Signal Processing

Voice signal processing is intended to obtain feature cepstral of human voice. The process of voice signal processing consists of Sampling, Frame Blocking, Windowing, Discrete Fourier Transform (DFT), Filter Power Spectral Subtraction (SS), Filter Bank and Discrete Cosine Transform.

Sampling

Human voices will generate continuous analog signals. Therefore, the analog signal is chopped in certain interval of time. Discrete series sample $x[n]$ is obtained from continuous signal $x(t)$,

$$x[n] = x(nT) \quad (1)$$

where T is sampling period and $i/T = F_s$ is sampling frequency in unit of sample/second. The value of n is the number of samples. According to the Nyquist sampling theory, minimal sampling frequency required is twice of original maximal signal.

$$\text{SamplingFrequency} \geq 2 \times \text{SignalFrequency} \quad (2)$$

Frame Blocking

Frame Blocking is division of voices into several frames where one frame consists of several samples. This process is needed to transform a non-stationer signal to a quasi-stationer signal so it can be transformed from time domain to frequency domain with Fourier transform. Human voice signal indicate quasi-stationer characteristic in range of time from 20 to 40 milliseconds [6]. Hence, on that range of time the Fourier transform can be performed.

Windowing

Sound signal which is chopped into frames will lead to discontinuity in initial and final signal. The discontinuity leads to data error in the process of Fourier transform. Windowing is needed to reduce the effect of discontinuity in chopped signal. If window is defined as $w(n)$, where $0 \leq n \leq N-1$ and N is number of samples in each frame, then the result of windowing process is;

$$w(n) = x(n)W(n), \quad 0 \leq n \leq N-1 \quad (3)$$

Windowing tang that is utilized in this research is Hamming windowing.

$$W_{hamm} = \begin{cases} 0.52 - 0.46 \cos(2\pi n/(N-1)) & 0 \leq n \leq N-1 \\ 0 & others \end{cases} \quad (4)$$

Discrete Fourier Transform (DFT)

Fourier transform is performed to transform from time domain to frequency domain. DFT is a specific form of integral Fourier equation;

$$Y(\omega) = \int w(t)e^{-j\omega dt} \quad (5)$$

The DFT can be obtained by changing the variables time (t) and frequency (w) into discrete form:

$$Y(k\omega_0) = \sum_{n=0}^{N-1} w(nT) e^{-jk\omega_0 nt} \quad (6)$$

Mel Frequency Cepstral Coefficient (MFCC)

The most important information of human voice signal is located at high frequency. This important information determines the characteristic of human voice and Mel Scale is utilized to accommodate this characteristics. The relation between Mel and actual frequency is according to various researches about perception of voice reception by human ear [5].

$$Mel = 1000 \times \log_2(1 + \omega) \quad (7)$$

In the implementation, this scaling is interpreted with Mel Filter Bank where each value of frequency magnitude is filtered by triangle filter series with Mel frequency as middle frequency. The triangular filter represents the process of Mel scaling in the signal.

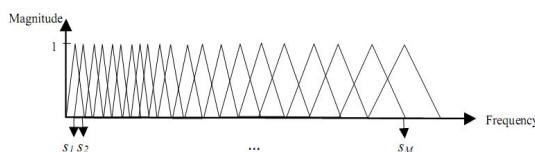


Figure 5. Filter bank construction

After the magnitude of signal spectrum $X[k]$ filtered by Mel Filter Bank, computation of logarithmic value of energy is conducted to each of output band from each filter. Logarithmic signal energy process is utilized to adapt the system just like human ear.

$$s[m] = \ln \left[\sum_{k=0}^{N-1} [Y[k]]^2 H_m[k] \right] \quad 0 \leq m \leq M \quad (8)$$

To obtain the MFCC, the result of energy logarithmic is processed with Discrete Cosine Transform (DCT).

$$c[n] = \sum_{m=0}^{M-1} s[m] \cos \left(\frac{\pi n(m-0.5)}{M} \right) \quad (9)$$

2.3 Gaussian Mixture Model (GMM)

Gaussian Probability density function (pdf) is bell-shaped one dimensional function which is defined by two parameters, that is mean μ and variant σ or covariant Σ . In D dimension it can be formulated as;

$$N(\mathbf{x}; \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \frac{1}{(2\pi)^{D/2} |\boldsymbol{\Sigma}|^{1/2}} \exp \left[-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right] \quad (10)$$

where μ is vector mean and Σ is covariant matrix.

Gaussian mixture model (GMM) is a mixing of several Gaussian distribution or representation of the existence of subclasses in a class. The probability density function of GMM is described as sum of multiplication of weight with Gaussian probability.

$$p(\mathbf{x}; \boldsymbol{\theta}) = \sum_{c=1}^C \alpha_c N(\mathbf{x}; \boldsymbol{\mu}_c, \boldsymbol{\Sigma}_c) \quad (11)$$

where α_c is weight of mixed component c, where $0 < \alpha_c < 1$ for each component and $\sum_{c=1}^C \alpha_c = 1$. whereas, the parameter distribution,

$$\boldsymbol{\theta} = \{\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1, \dots, \boldsymbol{\mu}_C, \boldsymbol{\Sigma}_C\} \quad (12)$$

It is the definition of Gaussian mixture probability density function parameter.

III. Software Design And Implementation

3.1 Software Interface

Virtual Qur'an Lecturer (VQL) software is software to learn how to read Al Qur'an correctly. Commonly, Al Qur'an reading learning book require more competent supervisor. However, VQL is interactive multimedia software accompanied with pronunciation and verse recitation law correction each courses.



Figure 6. Screenshot and Visualization of Virtual Qur'an Lecturer (VQL)

To utilize the software, first, we need to install the software to the computer. After that, the courses can be started. There are several modules or levels in this software, i.e. Pre-Learning, Fundamental, Intermediate and Advance. In fundamental level, the course is in basic level and there is only correction in *makhraj*. In intermediate level, the course includes the law of recitation and also correction in *makhraj*/pronunciation. In this level there is only one kind of recitation law. The advanced level, include courses with combination of more than one law of recitation and still with pronunciation correction. So, the higher level we take the higher difficulty of the course. If we make an error during the learning, the error message will appear on the software which will guide the user to correct the reading and pronunciations.

The software is equipped with a tutorial how to read Al Qur'an recitation correctly as guidance and pre-learning. The tutorial is recommended to

read every time we make error so we know why it is incorrect.

2.4 How the System Works

The basic of the correction and evaluation system is speech recognition where the cepstral feature of the voice of each reading is taken for recognition. The Gaussian Mixture Model is used to model the signals. Formulation and design is done prior to software construction and design. General formulation of the parameters of GMM modeling which is used to recognize the reading is obtained from this stage. The stage of construction and training can be illustrated as follows.

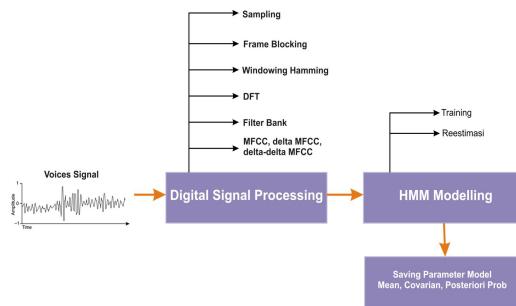


Figure 7. Design And Formulation Of The Software

The stage of design and formulation is composed of digital signal processing (DSP), GMM modeling and saving the GMM parameters to the database. The incoming voice signal is processed to extract the cepstral feature parameters. The incoming voice signal is sampled with sampling frequency of 800 Hz, accordance with Nyquist rule, and then divided into time slots (framing) with frame time 40 ms and overlapping time 20 ms. Each frame is passed through Hamming window to reduce signals discontinuity after chopping. The signal is then transformed to frequency domain by using DFT with $N = 1024$. The signal is passed through the Filter bank of 24 triangle filter. The DCT with MFCC coefficient of 14 is done after filtering process. Other feature calculation like signal energy, delta-MFCC and delta-delta MFCC were done after DCT process.

The four signal feature vector is used as input of the observation data in the GMM. The number of state and model which is used to design the software is not specified to adapt with verses section or the session of each reading in the course. After initialization, the next process is GMM training to obtain the maximum likelihood. After several iterations, if it achieves convergence the iteration can be stopped and all the parameter is stored in the database. Each recitation law in the course has one GMM model for correction.

At the stage of construction and application, correction process is done with similar process at the construction and formulation stage. However, at this stage likelihood estimation with GMM parameter for

each recitation law/pronunciation is done after signal processing process. Furthermore, if the likelihood value is less than the predetermined threshold value then certainly an error occurred in reciting the readings. The error message is appeared in order to warn the user to reread the readings with proper and correct law of recitation.

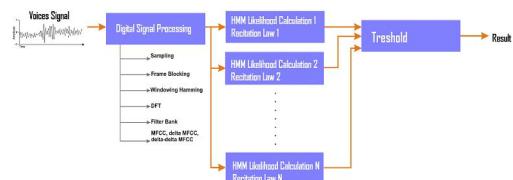


Figure 8. Correction Process In The Software

III. Results

To test the reliability and accuracy of correction, calibration for each recitation law in the software is done with ten speakers to read the readings in wrong and right manner.

1. Analysis of pronunciation correction (*Makhraj*): Fundamental

The accuracy correction for hija'iyah letter pronunciation letter is obtained from the average result for all hija'iyah letters. For example, how much *alif* recited true by the system compared with real true value of pronunciation of *alif* and then other *hija'iyah* letters. The average result from all *hija'iyah* letters is calculated as the correction accuracy which the value is 90%.

2. Accuracy of Quranic Recitation Law

Accuracy of Quranic recitation law can be obtained by system test for Quranic recitation. For a word containing some letters, the law might be *idghom*, *ihkfa'* and *idhar*. How many law from some Arabic words is detected true compared with real true value is obtained as the accuracy of Quranic recitation law. For recitation law correction test, the result suggests that the correction accuracy is poor (70%), hence, the system need a reconfiguration in order to improve the correction of recitation law.

3. Combination of several *Makhraj* and recitation law

The last accuracy tested to the system is the combination of makhraj and Quranic recitation law. It can be obtained by tested both of makhraj and Quranic recitation law from some Arabic words and compared the result with real true value. For combination of several *makhraj* and recitation law test, the result suggests that the correction accuracy is also poor (60%), hence, the system need a reconfiguration back in order to improve combination of *makhraj* and recitation law.

IV. Conclusion and Discussion

The conclusions of this research are as follows;

- a. Virtual Qur'an Lecturer (VQL) has been designed as an interactive multimedia Qur'an recitation learning software using cepstral feature and GMM modeling as the basis of speech recognition technology applied.
- b. The results of the research suggest that the correction accuracy of the VQL software is 70% for pronunciation, 90% for recitation law and 60% for combination of pronunciation and recitation law.

The performance of correction accuracy can be improved by changing the configuration of speech recognition used.

References

- [1] Bardici, Nick (2006). Speech Recognition using Gaussian Mixture Model. PhD Thesis. Blekinge Institute of Technology.
- [2] Moreno, Pedro J. (1996). Speech Recognition in Noisy Environments. PhD Thesis. Carnegie Mellon University.
- [3] Razak, Zaidi (2007). Quranic Verse Recitation Feature Extraction Using Mel-Frequency Cepstral Coefficient (MFCC). Journal University of Malaya.
- [4] Berouti, M., Schwartz, R. And Makhoul, J. (1979). Enhancement Of Speech Corrupted By Acoustic Noise. Proc. Of IEEE ICASSP, Pp. 208-211, Washington DC.
- [5] Bhatnagar, B.E., Mukul (2002). A Modified Spectral Subtraction Method Combined With Perceptual Weighting For Speech Enhancement. MSc Thesis. The University Of Texas At Dallas.

PREDIKSI TINGKAT KEBISINGAN LALU LINTAS DI SURABAYA DENGAN JARINGAN SYARAF TIRUAN STUDI KASUS JALAN EMBONG MALANG

Tutug Dhanardono, Aulia Siti Aisjah, Bambang L.Widjiantoro

Jurusan Teknik Fisika

Fakultas Teknologi Industri - Institut Teknologi Sepuluh Nopember

Kampus ITS Keputih Sukolilo – Surabaya 60111

Telp : +6231-5947188 Fax : +6231-5923626)

Abstrak

Jalan Embong Malang Surabaya merupakan jalan protokol yang banyak dilewati oleh berbagai jenis kendaraan bermotor yang berpotensi menimbulkan kebisingan terhadap penduduk yang bermukim di bagian utara jalan tersebut. Tingkat kebisingan yang terjadi pada setiap saat selalu berubah, tergantung pada tingkat kepadatan, jumlah komposisi kendaraan yang melintas. Pertumbuhan kendaraan di Jawa Timur, terutama kendaraan roda dua, terus meningkat dari tahun ke tahun, sehingga gangguan kebisingan lalu lintas perlu diprediksi agar dapat diketahui seberapa besar tingkat kebisingan yang terjadi seiring dengan pertumbuhan kendaraan tersebut. Dalam penelitian ini, dengan jumlah komposisi kendaraan yang melintas, diprediksi tingkat kebisingan. Jaringan Syaraf Tiruan (JST) adalah salah satu metode yang dapat digunakan untuk memprediksi tingkat kebisingan sehubungan jumlah komposisi kendaraan yang selalu berubah. Metode pembelajaran Levenberg-Marquardt (LM), dengan jumlah hidden layer 1 dan jumlah node 12 untuk prediksi Tingkat Kebisingan Ekuivalen L_{eq} , 11 node untuk prediksi Traffic Noise Index – TNI, 12 node untuk prediksi Noise Pollution Level - L_{NP} . Model rancangan Jaringan Syaraf Tiruan mampu memprediksi dengan baik Leq , TNI, dan L_{NP} , ditandai oleh besarnya eror 0.0477 dBA untuk Leq , 1.3567 dBA untuk TNI, dan 0.2042 dBA untuk L_{NP} .

Kata kunci : Tingkat Kebisingan, L_{eq} , TNI, L_{NP} , JST, LM.

I. Pendahuluan

1.1 Latar Belakang

Jalan Embong Malang Surabaya (JEMS) merupakan jalan utama kota Surabaya dimana lalu lintasnya hanya satu arah dari timur ke barat, dengan ujung-ujung masuk dan keluar terpecah menjadi dua, sehingga tergolong model jalan yang unik, volume transportasi tergolong sedang tetapi kontinyu,,banyak pemukiman di sekitar Jalan Embong Malang Surabaya, tergolong kawasan padat penduduk, aktivitas penduduk di sekitar jalan Embong Malang Surabaya cukup beragam, banyak aktivitas penduduk yang memakan waktu kerja yang cukup lama (<8 jam/hari) di lakukan di tepi jalan Embong Malang Surabaya, sebagai lokasi usaha kecil, tidak ada sumber kebisingan lain yang cukup besar selain dari lalu lintas kendaraan di lokasi pemukiman sekitar jalan Embong Malang Surabaya.

1.2 Perumusan Masalah

Tingkat kebisingan akibat lalu lintas kendaraan sangat tergantung pada jumlah komposisi kendaraan yang melintas di jalan raya tersebut. Apakah jaringan syaraf tiruan dengan struktur narxsp (nonlinear autoregressive network with exogenous inputs series-paralel) dengan algoritma pembelajaran Levenberg-Marquardt mampu memprediksi tingkat kebisingan ekuivalen Leq , Traffic Noise Index TNI, dan Noise Pollution Level L_{NP} , dengan masukan jumlah komposisi kendaraan.

1.3 Tujuan dan Manfaat Penelitian

Memprediksi tingkat kebisingan ekuivalen Leq , Traffic Noise Index TNI, dan Noise Pollution Level L_{NP} akibat lalu lintas kendaraan di Jalan Embong Malang Surabaya terhadap jumlah komposisi kendaraan menggunakan Jaringan Syaraf Tiruan (JST) berstruktur narxsp dengan algoritma pembelajaran Levenberg-Marquardt dan performansi Mean Square Error MSE. Struktur jaringan syaraf tiruan yang terbentuk nantinya dapat digunakan sebagai alat (soft sensor) menentukan tingkat kebisingan bila diketahui jumlah komposisi kendaraan yang melewati suatu jalan.

II. Kajian Pustaka Dan Teori

2.1 Kajian Pustaka

Pemakaian Jaringan Syaraf Tiruan – JST untuk memprediksi tingkat kebisingan lalu lintas jalan raya telah banyak dilakukan (Barry L.Doolan, 2005; Rizky Prasetyo S, 2006; Pichai Pamanikabud, 2006; Dhananjay, 2007) dengan berbagai variasi besaran masukan dan target keluaran yang diinginkan, misal sebagai besaran masukan adalah tingkat kebisingan pada waktu sekarang dan sebelumnya dan sebagai target keluaran adalah waktu mendatang (Rizky Prasetyo S, 2006), atau sebagai masukan adalah jumlah komposisi kendaraan (sepeda motor, mobil/kendaraan ringan, bus/truk) dan sebagai keluaran adalah besaran bising lingkungan, Leq , TNI, L_{NP} , dan sebagainya (Barry

L.Doolan, 2005; Dhananjay,2007). Berbagai metode pembelajaran / pelatihan juga dilakukan, misalnya Backpropagation (Barry L.Doolan, 2005; Dhananjay,2007), Levenberg-Marquardt (Rzky Prasetyo S, 2006). Tetapi pemakaian Jaringan Syaraf Tiruan untuk memprediksi tingkat kebisingan tersebut tidak selalu memberikan hasil yang diharapkan ketika dilakukan validasi dengan memasukkan data masukan yang belum pernah dikenalkan dalam proses training (Rizky Prasetyo S, 2006).

Rumus empiris yang menyatakan tingkat kebisingan lalu lintas kendaraan bermotor di jalan raya sebagai fungsi waktu, dengan menggunakan metode statistik, di jalan Embong Malang Surabaya, memberikan hasil dapat dipergunakan untuk memprediksi tingkat kebisingan pada waktu mendatang (Anik Prasetyowati., 2005). Tetapi rumus empiris tersebut belum divalidasi dengan data tingkat kebisingan di waktu yang akan datang.

Jalan Embong Malang Surabaya merupakan jalan akses yang strategis sebagai penghubung Surabaya bagian timur dan barat, utara dan selatan, serta merupakan lokasi yang mempunyai karakteristik unik sehingga menarik untuk diteliti tingkat kebisingan lalu lintasnya (Anik Parsetiowati, 2005, Aries Purwanto, 2005; Rizky Prasetyo Santosa, 2006).

2.2 Tingkat Kebisingan Ekuivalen, L_{eq} ,

Tingkat Kebisingan dinyatakan dengan persamaan :

$$L_{eq} = 10 \log (\sum P_i 10^{\frac{L_i}{10}}) \text{ dBA} \quad (1)$$

dimana :

t_i = Lamanya waktu dengan tingkat kebisingan L_i

$T = \sum t_i = t_1 + t_2 + t_3 + \dots =$ selang waktu dimana dihitung L_{eq} tersebut

$P_i = t_i/T$ = fraksi waktu

L_i = Tingkat kebisingan selama waktu t_i .

2.3 Noise Pollution Level, L_{NP}

Noise Pollution Level, L_{NP} , didefinisikan sebagai berikut :

$$L_{NP} = L_{eq} + K\delta \text{ PNdB} \quad (2)$$

dengan K adalah konstanta = 2,56. dan δ = deviasi standar.

Noise Pollution Level dapat juga dinyatakan dengan besaran tingkat kebisingan yang dilampaui 10 % dan 50 % dari waktu pengukuran, dengan persamaan sebagai berikut :

$$L_{NP} = L_{50} + d + \frac{d^2}{60} \text{ PNdB} \quad (3)$$

atau

$$L_{NP} = L_{eq} + d \text{ PNdB} \quad (4)$$

dengan :

$$d = L_{10} - L_{90}$$

L_{10} – Tingkat kebisingan yang dilampaui 10% dari waktu pengukuran, dBA
 L_{50} – Tingkat kebisingan yang dilampaui 50% dari waktu pengukuran, dBA
 L_{90} – Tingkat kebisingan yang dilampaui 90% dari waktu pengukuran, dBA
 L_{10} menunjukkan nilai puncak tingkat kebisingan, L_{50} menunjukkan nilai median tingkat kebisingan, dan L_{90} menunjukkan nilai tingkat kebisingan latar belakang.

2.4 Traffic Noise Index

Besaran ini dinyatakan dengan besaran *Percentile Noise Level* dengan hubungan :

$$TNI = 4xL_{10} - 3xL_{90} - 30 \text{ dBA} \quad (5)$$

2.5 Algoritma Levenberg-Marquardt

Secara umum algoritma LM dapat dituliskan sebagai berikut,

1. Pilih suatu vektor parameter awal $\theta^{(i)}$ dan harga awal $\lambda^{(i)}$.
2. Tentukan arah pencarian dari $[R(\theta^{(i)}) + \lambda^{(i)} I]f^{(i)} = -G(\theta^{(i)})$ dengan I adalah matrik identitas.
3. Jika $r^{(i)} > 0,75 \Rightarrow \lambda^{(i)} = \lambda^{(i)}/2$ (Jika pengurangan yang diprediksi mendekati pengurangan yang sebenarnya ambil arah pencarian mendekati arah pencarian Gauss Newton).
4. Jika $r^{(i)} < 0,25 \Rightarrow \lambda^{(i)} = 2\lambda$ (Jika pengurangan yang diprediksi menjauhi pengurangan yang sebenarnya ambil arah pencarian mendekati arah pencarian gradien).
5. Jika $V_N(\theta^{(i)} + f^{(i)}, Z^N) < V_N(\theta^{(i)}, Z^N)$, maka terima $\theta^{(i+1)} = \theta^{(i)} + f^{(i)}$ sebagai suatu iterasi baru dan ambil $\lambda^{(i+1)} = \lambda^{(i)}$ dan $i = i+1$.
6. Jika kriteria penghentian tidak dipenuhi kembali ke langkah 2.

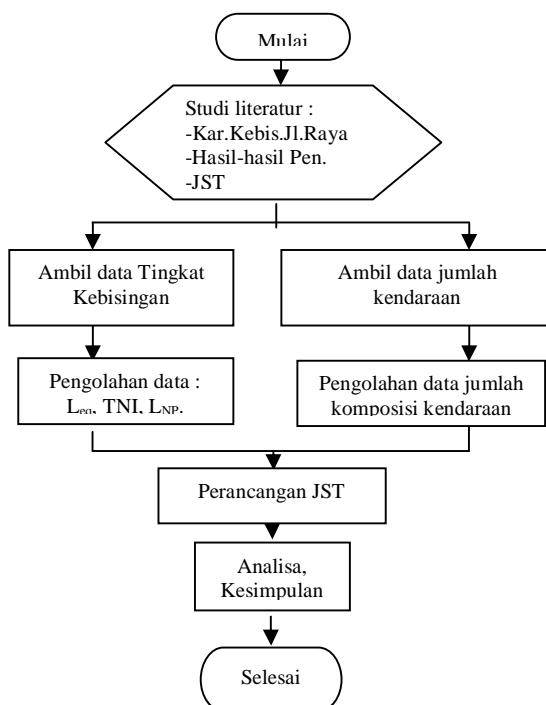
2.6 Performansi Keberhasilan Jaringan

Performansi keberhasilan jaringan untuk memprediksi tingkat kebisingan dinyatakan oleh besaran-besaran *mean square error* (MSE). Yaitu selisih nilai data target dengan nilai data keluaran. Goalnya adalah meminimalkan rata-rata dari jumlah kuadrat selisih tersebut⁽¹²⁾.

$$MSE = \frac{1}{Q} \sum_{k=1}^Q e(k)^2 = \frac{1}{Q} \sum_{k=1}^Q (t(k) - a(k))^2 \quad (6)$$

III. Metode Penelitian

Untuk melihat secara garis besar urutan langkah-langkah penelitian, maka dibuat diagram alir sebagai berikut :



Gambar 1 : Diagram alir penelitian secara garis besar.

3.1 Sampling Data

Pengukuran dilakukan selama satu minggu (hari Senin sampai Minggu), dalam satu hari dilakukan pengukuran tiga kali (pagi, siang, malam), setiap kali pengukuran dilakukan selama 20 menit, dengan pencatatan secara manual setiap 1.67 detik sekali, dengan jumlah data sebanyak 720. Dalam satu hari dilakukan 3 kali pengukuran pagi, siang, dan sore, dengan jumlah data 2160. Dalam satu minggu dilakukan pengukuran 7 kali, sehingga diperoleh jumlah data total sebanyak 15.120. Bersamaan dengan itu dilakukan pengambilan gambar kendaraan yang melintas dengan menggunakan alat perekam gambar (*video recorder*).

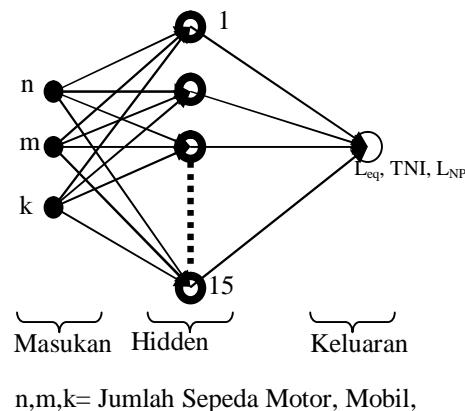
3.2 Perancangan Jaringan Syaraf Tiruan

Data masukan untuk memprediksi tingkat kebisingan lingkungan adalah jumlah komposisi kendaraan (Sepeda Motor, Mobil dan Bus/truk) dan sebagai keluaran atau target adalah tingkat kebisingan lingkungan, yaitu Tingkat Kebisingan Ekuivalen, L_{eq} , Traffic Noise Index, TNI, dan Noise Pollution Level, L_{NP} .

Hasil training dan simulasi dengan menggunakan JST dibahas untuk mengetahui keberhasilan teori tersebut dalam memprediksi tingkat kebisingan dengan perubahan jumlah komposisi jenis kendaraan. Diagram alir untuk prediksi tingkat kebisingan dengan JST dapat dilihat pada gambar 3 di bawah ini.

Dibangun jaringan dengan fungsi *newnarxsp* (*nonlinear autoregressive network with exogenous inputs, seri-paralel*), dimana dalam model jaringan

ini data masukan dan keluaran diinputkan secara seri-paralel, secara berurutan (sequential), data pertama dan kedua dari masukan dan keluaran di tunda (di *delay*). Data masukan dan keluaran (sebagai target) ditraining dengan metode pelatihan Levenberg-Marquardt, dengan jumlah layer hidden sebanyak 1 dan dimulai dengan 1 node.



Gambar 2. Jaringan Syaraf Tiruan Untuk Memprediksi Tingkat Kebisingan. Masukan Adalah Jumlah Komposisi Kendaraan, Target/Keluaran Adalah Tingkat Kebisingan Ekuivalen L_{eq} , Traffic Noise Index TNI, Dan Noise Pollution Level L_{NP} .

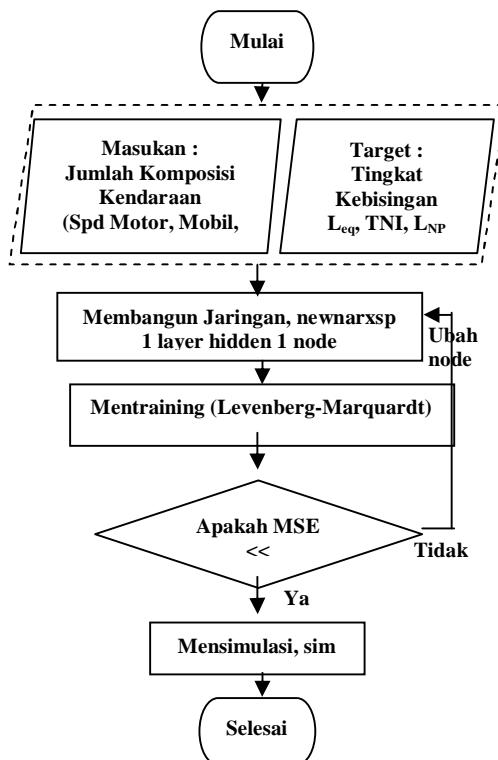
Langkah ini menghasilkan arsitektur jaringan dengan nilai performansi, yang dinyatakan dengan besaran *Mean Square Error* (MSE) tertentu. Hasil latihan ini kemudian disimpan dan diberi suatu nama. Latihan dimulai lagi dengan merubah hidden node menjadi 2 dan seterusnya. Dari hasil training tersebut dipilih jaringan yang mempunyai nilai MSE terkecil untuk dilakukan simulasi/tes dengan menggunakan data baru (data yang belum dipakai dalam training). Proses pengolahan data selesai.

IV. Hasil Penelitian dan Pembahasan

4.1 Penyiapan Data Untuk Training

Prediksi Tingkat Kebisingan dengan menggunakan Jaringan Syaraf Tiruan dilakukan dengan mengambil masukan adalah data jumlah komposisi kendaraan dan sebagai data keluaran adalah Tingkat Kebisingan Lingkungan, seperti yang disusun seperti tabel 4.1 di bawah ini :

Data masukan pada JST yaitu : jumlah komposisi kendaraan dan keluaran pada JST adalah nilai L_{eq} , TNI dan L_{NP} yang telah diuraikan diatas, diuji atau ditraining dengan menggunakan struktur jaringan 1 hidden layer dan berbagai jumlah hidden node mulai dari 1 hingga 15, dengan struktur seperti digambarkan pada gambar 3.



Gambar 3. Giagram Alir Penelitian Prediksi Tingkat Kebisingan dengan JST.

Tabel 1. Data Masukan Dan Data Keluaran Untuk Prediksi Dengan JST

NO	Data Masukan			Data Keluaran		
	Jumlah Jenis Kendaraan		Bis/truk	Tingkat Kebisingan, dBA		
	Spd Mtr	Mobil		L _{eq}	TNI	L _{NP}
1	674	138	3	75.4	63	83
2	798	457	5	75	67	82
3	838	433	3	74.6	64	83
4	411	134	0	73.7	64	82
5	587	414	3	74.3	64	81
6	874	418	1	75.3	65	81
7	1112	173	3	76.6	66	82
8	872	173	9	75.1	63	84
9	1516	428	1	75.6	64	86
10	896	163	3	76.5	66	83
11	702	445	15	73.8	62	81
12	974	451	3	74.7	67	82
13	1006	180	3	76.5	65	83
14	783	409	3	75	67	85
15	1147	451	1	75.6	65	83
16	989	200	5	78	68	84
17	874	462	7	76	66	85
18	1139	497	5	77.4	64	85
19	940	192	4	79.3	68	86
20	730	455	8	75.9	60	95
21	1258	452	2	76.9	63	102

4.2 Hasil Training dan Simulasi Tingkat Kebisingan Ekuivalen, L_{eq}

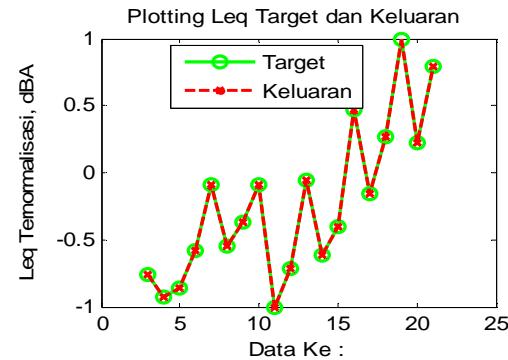
Hasil training dan simulasi untuk besaran Leq dinyatakan dengan nilai MSE, variasi eror, plot regresi dan plot mapping antara nilai target dan nilai hasil simulasi, serta variasi eror terhadap pasangan data masukan - keluaran. Nilai MSE yang dimiliki untuk setiap jumlah hidden node dapat dilihat pada tabel 2.

Tabel 2. Nilai Mean Square Error (MSE) dari Leq untuk Berbagai Jumlah Node Dalam 1 Hidden Layer,

No	Jumlah Hidden Node	Nilai MSE
1	1	3.72 x 10 ⁻²
2	2	1.34 x 10 ⁻²
3	3	8.0998 x 10 ⁻⁰⁴
4	4	1.3104 x 10 ⁻³¹
5	5	3.4787 x 10 ⁻²¹
6	6	1.5453 x 10 ⁻²⁹
7	7	3.5415 x 10 ⁻²⁷
8	8	2.8803 x 10 ⁻³¹
9	9	9.9891 x 10 ⁻²⁴
10	10	9.7310 x 10 ⁻³²
11	11	3.5349 x 10 ⁻²⁹
12	12	8.4335 x 10⁻³²
13	13	1.5894 x 10 ⁻³¹
14	14	5.0212 x 10 ⁻³¹
15	15	2.6860 x 10 ⁻²⁷

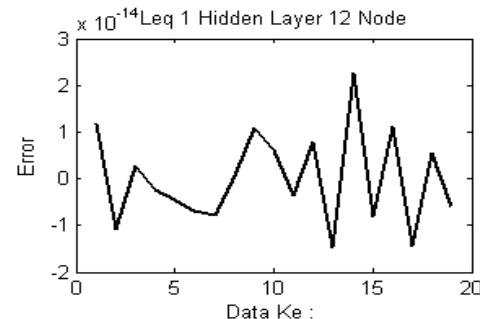
Hasil training tentang kecepatan jaringan dengan 1 hidden layer dan 12 node diperlihatkan pada gambar-gambar dibawah ini.

Plotting data antar hasil simulasi dan target dapat dilihat pada gambar 4.. Terlihat bahwa nilai Tingkat Kebisingan Ekuivalen, L_{eq}, sebagai keluaran dari JST sangat berimpit dengan nilai L_{eq} yang ditargetkan. Hal ini menunjukkan bahwa proses training berhasil.



Gambar 4. Ploting Data Hasil Simulasi dan Target Nilai L_{eq} untuk Jaringan dengan 1 Hidden Layer Dan 12 Node.

Hal ini juga sesuai dengan hasil simulasi yang ditandai dengan eror antara nilai Leq hasil simulasi dan nilai Leq yang ditargetkan, seperti ditunjukkan pada gambar 4.2 di bawah ini.



Gambar 5. Variasi Eror terhadap Pasangan Data Target Dan Keluaran untuk Besaran L_{eq}, dengan 1 Hidden Layer dan 12 Node.

Pada gambar 5 di atas terlihat bahwa variasi nilai eror pada setiap pasangan data masukan dan keluaran adalah berorde 10^{-14} . Nilai ini merupakan harga yang sangat kecil sehingga perbedaan antara nilai keluaran dari JST dengan nilai yang ditargetkan tidak signifikan.

4.3 Hasil Training dan Simulasi Traffic Noise Index, TNI

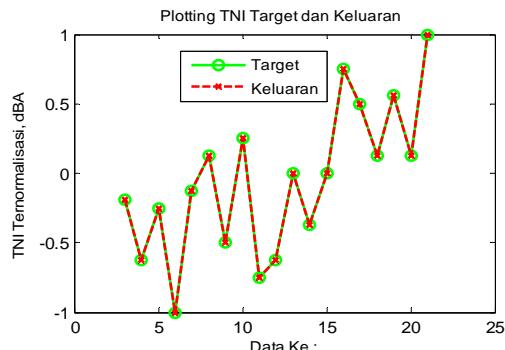
Hasil training dan simulasi untuk besaran TNI dinyatakan dengan nilai MSE yang dimiliki untuk setiap jumlah hidden node dapat dilihat pada tabel 3.

Dari tabel 3 di bawah ini dapat dilihat bahwa nilai MSE terkecil dimiliki oleh jaringan dengan jumlah hidden node 11.

Tabel 3. Nilai Mean Square Error (MSE) dari TNI untuk Berbagai Jumlah Node dalam 1 Hidden Layer

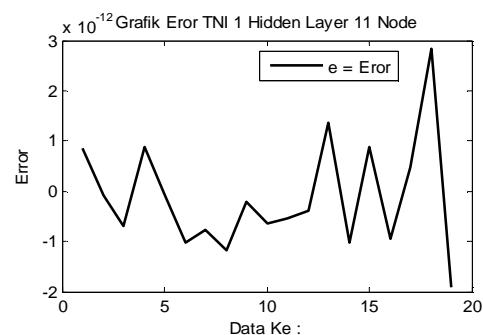
No	Jumlah Hidden Node	Nilai MSE
1	1	7.63×10^{-2}
2	2	5.25×10^{-2}
3	3	1.307×10^{-1}
4	4	8.9400×10^{-21}
5	5	1.1991×10^{-29}
6	6	2.9877×10^{-26}
7	7	1.4681×10^{-21}
8	8	7.1501×10^{-22}
9	9	1.5485×10^{-30}
10	10	2.3794×10^{-39}
11	11	1.8035×10^{-31}
12	12	2.9842×10^{-31}
13	13	3.1983×10^{-31}
14	14	1.1007×10^{-28}
15	15	8.2425×10^{-22}

Berikut ini disajikan hasil simulasi terbaik, yaitu jaringan dengan 1 hidden layer dan 12 node, Plotting data antar hasil simulasi dan target dapat dilihat pada gambar 6 di bawah ini.



Gambar 6. Ploting Data Hasil Simulasi dan Target Nilai TNI untuk Jaringan dengan 1 Hidden Layer Dan 11 Node

Pada gambar 6 terlihat bahwa nilai Tingkat Kebisingan Ekuivalen, TNI, hasil keluaran dari JST sangat berimpit dengan nilai TNI yang ditargetkan. Hal ini berarti bahwa proses training berhasil dengan baik. Hal ini juga sesuai dengan hasil simulasi yang ditandai dengan eror antara nilai TNI hasil simulasi dan nilai TNI yang ditargetkan. Besarnya eror tersebut ditunjukkan pada gambar 4.4 di bawah ini.



Gambar 7. Variasi Eror Terhadap Pasangan Data Masukan dan Keluaran untuk Besaran TNI, Dengan 1 Hidden Layer dan 11 Node

Dari gambar 7 di atas terlihat bahwa variasi nilai eror pada setiap pasangan data masukan dan keluaran adalah berorde 10^{-12} . Ini merupakan nilai yang sangat kecil, sehingga perbedaan antara nilai target dan nilai keluaran JST tidak signifikan.

4.4 Hasil Training dan Simulasi Noise Pollution Level, L_{NP}

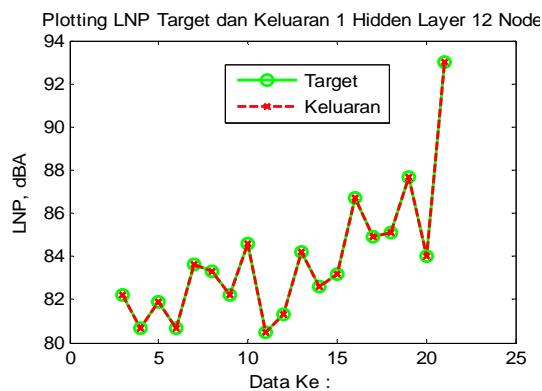
Hasil training dan simulasi untuk besaran LNP dinyatakan dengan nilai MSE yang dimiliki untuk setiap jumlah hidden node dapat dilihat pada tabel 4.

Tabel 4. Nilai Mean Square Error (MSE) Dari L_{NP} untuk Berbagai Jumlah Node Dalam 1 Hidden Layer,

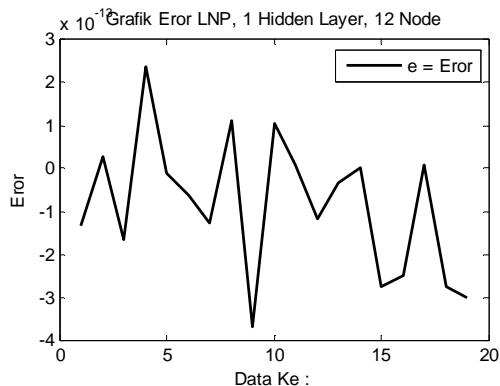
No	Jumlah Hidden Node	Nilai MSE
1	1	6.30×10^{-2}
2	2	1.346×10^{-1}
3	3	2.3440×10^{-23}
4	4	6.7283×10^{-24}
5	5	6.2359×10^{-27}
6	6	9.2472×10^{-28}
7	7	4.8120×10^{-27}
8	8	3.7237×10^{-31}
9	9	5.7666×10^{-30}
10	10	4.7175×10^{-22}
11	11	5.9062×10^{-28}
12	12	2.7766×10^{-31}
13	13	2.0996×10^{-28}
14	14	9.8672×10^{-25}
15	15	1.5688×10^{-25}

Dari tabel 4 tersebut dapat dilihat bahwa nilai MSE terkecil dimiliki oleh jaringan dengan jumlah hidden node 12.

Pada gambar 4.5 di bawah terlihat bahwa nilai Noise Pollution Level, L_{NP} , hasil keluaran dari JST sangat berimpit dengan nilai L_{NP} yang ditargetkan. Hal ini berarti bahwa proses training berhasil dengan baik. Hal ini juga sesuai dengan hasil simulasi yang ditandai dengan eror antara nilai L_{NP} hasil simulasi dan nilai L_{NP} yang ditargetkan, seperti ditunjukkan pada gambar 4.6 berikut :



Gambar 8. Ploting Data Hasil Simulasi dan Target Nilai L_{NP} untuk Jaringan Dengan 1 Hidden Layer Dan 11 Node



Gambar 9. Variasi Eror Terhadap Pasangan Data Masukan dan Keluaran untuk Besaran L_{NP} , dengan 1 Hidden Layer dan 12 Node.

Dari gambar 9 terlihat bahwa variasi nilai eror pada setiap pasangan data masukan dan keluaran dari Noise Pollution Level L_{NP} adalah berorde 10^{-12} . Ini berarti bahwa nilai data target dan nilai keluaran JST tidak berbeda.

4.5 Validasi

Dalam melakukan validasi terhadap jaringan yang telah rancang, dengan struktur 1 hidden layer 12 node untuk Tingkat Kebisingan Ekuivalen L_{eq} , 1 hidden layer 11 node untuk Traffic Noise Index TNI, dan 1 hidden layer 12 node untuk Pollution Noise Level L_{NP} , diberikan data masukan baru dari hasil pengukuran, yaitu Jumlah Komposisi Kendaraan (Sepeda motor, Mobil, Bus/Truk) dan data target keluaran baru dari ketiga besaran tingkat kebisingan tersebut. Data tersebut diperlihatkan pada tabel 4.5.

Tabel 5. Data untuk Validasi Model Jaringan.

Jumlah Kendaraan			Tingkat Kebisingan, dBA		
Sepeda Motor	Mobil	Bus/Truk	L_{EQ}	TNI	L_{NP}
760	154	2	80.604	84.5	90.6
606	337	8	77.148	68	83.1
1139	455	4	78.251	78	86.75

Nilai komposisi kendaraan yang ditunjukkan pada tabel 5 di atas digunakan untuk validasi pada

model jaringan syaraf tiruan yang berfungsi sebagai prediktor terbaik (saat menghasilkan nilai MSE terkecil) dari masing-masing besaran tingkat kebisingan yang diperoleh pada proses training. Hasil validasi dapat dilihat pada tabel 6.

Tabel 6. Hasil Validasi

Tingkat kebisingan, dBA	Target, dBA	Hasil Prediksi, dBA	Eror (Target-Prediksi), dBA
L_{eq}	78.25126	77.62311	0.62814
TNI	78	70.26538	7.7346
L_{NP}	86.75	84.17167	2.5796

Dari tabel 6 tersebut terlihat bahwa nilai eror, yaitu selisih antara hasil prediksi JST terhadap data target yang diharapkan sangat bervariasi untuk berbagai besaran tingkat kebisingan. Dalam tabel tersebut ditunjukkan bahwa rancangan model Jaringan Syaraf Tiruan yang dibangun mampu memprediksi Tingkat Kebisingan Ekuivalen L_{eq} , dengan nilai eror 0.62814 dBA. Kemampuan untuk memprediksi besaran Pollution Level L_{NP} , dengan nilai eror 2.5796 dBA, cukup baik, karena telinga manusia baru dapat membedakan perubahan tingkat kebisingan bila selisihnya > 3 dBA. Sedangkan untuk validasi terhadap TNI menunjukkan eror sebesar 7.7346 dBA, berarti model jaringan tersebut tidak mampu memprediksi besaran Traffic Noise Index TNI dengan baik.

Hasil eror prediksi yang dihasilkan dari usaha mentraining kembali dapat dilihat pada lampiran VII, sedangkan hasil terbaik, yaitu nilai eror terkecil, dapat dilihat pada tabel 7 berikut.

Tabel 7. Hasil Validasi dengan Mentraining Kembali

Tingkat kebisingan, dBA	Target, dBA	Hasil Prediksi, dBA	Eror (Target-Prediksi), dBA
L_{eq}	78.25126	78.2035	0.0477
TNI	78	76.6433	1.3567
L_{NP}	86.75	86.9556	-0.2042

Pada tabel 4.7, dapat dilihat bahwa eror antara data target dan hasil prediksi untuk Tingkat Kebisingan Ekuivalen adalah 0.0477, untuk Traffic Noise Index adalah 1.3567, dan untuk Noise Pollution Level adalah 0.2042 merupakan harga yang relatif sangat kecil, sehingga tidak signifikan untuk mempengaruhi penambahan tingkat kebisingan. Ini berarti bahwa Jaringan Syaraf Tiruan mampu memprediksi ketiga besaran tingkat kebisingan tersebut.

V. Kesimpulan Dan Saran

Dari hasil pembahasan dapat diberikan kesimpulan dan saran sebagai berikut :

5.1 Kesimpulan

- Model Jaringan Syaraf tiruan terbaik untuk prediksi Tingkat Kebisingan Ekuivalen L_{eq}

- dan *Noise Pollution Level* L_{NP} pada struktur 1 Hidden Layer dengan jumlah node 12 dan untuk prediksi *Traffic Noise Index* TNI pada struktur 11 node.
- Model rancangan Jaringan Syaraf Tiruan mampu memprediksi dengan baik Tingkat Kebisingan Ekuivalen Leq, Traffic Noise Index TNI, dan Noise Pollution Level L_{NP} , ditandai dengan eror 0.0477 dBA untuk Leq, 1.3567 dBA untuk TNI, dan 0.2042 dBA untuk L_{NP} .
- [8] Rizky Prasetyo Santoso, Tutug Dhanardono, Bambang Lelono Widjiantoro (2006), *Implementasi Jaringan Syaraf Tiruan Untuk Memprediksi Tingkat Kebisingan Jalan Embong Malang Surabaya*, Tugas akhir Jurusan Teknik Fisika,
- Copyright 1992–2009 by The MathWorks, Inc.,
www.mathworks.com

5.2 Saran

Dari hasil penelitian tersebut terbukti bahwa model rancangan JST yang terbentuk mampu memprediksi Tingkat Kebisingan Ekuivalen L_{eq} , Traffic Noise Index TNI, dan Noise Pollution Level L_{NP} , tetapi perlu di teliti lagi apakah model JST tersebut mampu memprediksi untuk kondisi jalan lainnya di Surabaya. Oleh karena itu disarankan untuk menguji kemampuan JST tersebut terhadap jalan lain selain jalan Embong Malang Surabaya.

Daftar Pustaka

- [1] Anik Prasetyowati dan Tutug Dhanardono (2005), “Penentuan Tingkat Kebisingan Lalu-Lintas Sebagai Fungsi Waktu Di Jalan Embong Malang Surabaya”. Tugas Akhir, Jurusan Teknik Fisika FTI-ITS.
- [2] Pichai Pamanikabud, Niran Janchom, and Marupong Tansatcha (2006), “Application Of Neural Network Vehicle Noise In Prediction Of Highway Traffic Noise”, *International Conference on Green and Sustainable Innovation*, November 29th – December 1st, Amora Tapae Hotel, Chiangmai Thailand.
- [3] Dragan Cvetković, Momir Praščević, Violeta Stojanović (1997), “Naiss - Model For Traffic Noise Prediction”, *University of NIŠ The scientific journal Facta Universitatis Series: Working and Living Environmental Protection*, Vol. 1, No 2, pp. 73 – 81.
- [4] Barry L. Doolan and Steven J.B. Carter (2005), “Road Traffic Noise Prediction An Artificial Intelligence Approach”, *Proceeding of Acoustics*.
- [5] Dhananjay K. and Prashant B. Nagarnaik (2007), “Assessment And Ann Modeling Of Noise Levels At Major Road Intersections In An Indian Intermediate City”, *Journal of Research in Science, Computing, and Engineering* 4:3, pp.39-49.
- [6] Erawati Tri Utami, Tutug Dhanardono (2009), *Analisa Tingkat Kebisingan Lingkungan Akibat Lalu Lintas Pada Jalan Raya Tandes - Surabaya*, Tugas Akhir, Jurusan Teknik Fisika FTI-ITS,.
- [7] Howard Demuth, Mark Beale, Martin Hagan; *Neural Network Toolbox™ 6 User's Guide*, ©

APLIKASI GELOMBANG ULTRASONIK PADA REAKSI DEGRADASI GLISEROL

Yuyun Yuniat¹, Sumarno, Mahfud

¹Jurusan Teknik Kimia, Institut Teknologi Sepuluh Nopember Surabaya
Kampus ITS Keputih Sukolilo Surabaya 60111
*Email : yuniati73@gmail.com

Abstrak

Pengaruh kavitas dalam gelombang ultrasonik dalam reaksi kimia dipelajari dengan mendegradasi senyawa gliserol. Senyawa gliserol banyak dihasilkan sebagai hasil samping pembuatan biodiesel sehingga dengan memanfaatkannya kembali dapat meningkatkan daya jual minyak diesel nabati yang relatif cukup tinggi di pasaran. Pada penelitian ini digunakan peralatan ultrasonik batch jenis S 100 H dengan frekuensi 37 kHz. Parameter proses yang dipelajari antara lain suhu cairan (bulk), yaitu 40, 50, dan 60°C dan waktu sonifikasi antara 10-60 menit untuk reaktan gliserol pada perbandingan massa 1:10 (gliserol-air). Hasil penelitian ini membuktikan bahwa gliserol dapat didegrasi menggunakan gelombang ultrasonik. Salahsatu produk degradasi ini adalah metanol yang dikenal sebagai bahan bakar dan pelarut organik.

Kata kunci : ultrasonik, gliserol, degradasi, metanol

I. Pendahuluan

Memburuknya kualitas udara di perkotaan dan ketergantungan pada sumber energi minyak bumi mendorong para ilmuwan untuk selalu mencari sumber-sumber energi alternatif yang lebih ramah lingkungan dan bersifat berkelanjutan. Hal ini disebabkan karena populasi dunia yang semakin tinggi juga menyebabkan lonjakan kebutuhan energi. Salahsatu sumber energi alternatif yang banyak diminati dan dikembangkan adalah biodiesel. Permasalahan lain muncul ketika investasi industri biodiesel sangat tinggi membuat harga biodiesel di pasaran masih kurang bersaing dengan minyak diesel fosil. Sebagai upaya untuk meningkatkan daya saing biodiesel maka dilakukan dengan memanfaatkan hasil samping dalam proses pembuatan biodiesel yang berupa gliserol. Melalui reaksi transesterifikasi gliserol yang dihasilkan adalah 10% dari berat biodiesel yang diproduksi (Dasari dkk, 2005), hal ini merupakan peluang agar dilakukan penelitian tentang gliserol lebih lanjut. Salah satu usaha untuk memanfaatkan gliserol adalah dengan mentransformasi gliserol menjadi bahan bakar lain berupa metanol dan etanol melalui teknik degradasi atau pemutusan ikatan molekul dalam media air subkritis hingga superkritis.

Penelitian degradasi gliserol menjadi sangat penting apalagi jika dihasilkan bahan bakar lain berupa metanol, mengingat Indonesia mempunyai tingkat penduduk yang sangat padat sehingga jelas diperlukan sumber-sumber energi yang bisa bertahan lama dan berkelanjutan. Potensi biodiesel yang telah banyak dikembangkan di Indonesia juga menjadi sumber bagi ketersediaan gliserol sebagai hasil sampingnya. Dari sisi kebijakan pemerintah diharapkan mendukung produksi biodiesel di tengah krisis energi dan memberikan kesempatan terbukanya lapangan pekerjaan.

Aplikasi gelombang ultrasonik untuk berbagai keperluan reaksi kimia telah banyak digunakan terutama dalam pengolahan air limbah. Ultrasonik dipakai untuk mendegradasi senyawa organik yang bersifat polutan. Melalui teknologi ini akan diperoleh yield yang cukup tinggi dengan energi input yang cukup rendah.

Penelitian ini bertujuan untuk mempelajari pengaruh gelombang ultrasonik pada reaksi degradasi gliserol dalam menghasilkan bahan kimia lain serta membandingkan proses ini dengan teknologi subkritis/ superkritis air

II. Tinjauan Pustaka

Penggunaan ultrasonik untuk keperluan reaksi kimia disebut sebagai teknologi sonokimia, dimana digunakan frekuensi antara 20-40 kHz sebab pada batasan inilah peralatan laboratorium biasa digunakan. Medan ultrasonik akan menghasilkan efek kimia dan fisika yang diakibatkan oleh meledaknya gelembung kavitas mikro yang disebabkan getaran ultrasonik, sehingga efek tersebut yang berdampak terjadinya proses sonokimia (Mason, 1997).

Ultrasonik dapat meningkatkan kereaktifan kimia pada suatu sistem yang secara efektif bertindak sebagai katalis untuk lebih mereaktifkan atom-atom dan molekul dalam sistem. Pada reaksi yang menggunakan bahan padat, ultrasonik dapat memecah padatan dari energi yang ditimbulkan oleh runtuhnya kavitas yang melaluinya. Efeknya adalah memberikan komponen reaktan padat luas permukaan yang lebih besar untuk meningkatkan laju reaksi. (Suslick, 1994).

Berbagai penelitian yang telah dilakukan dengan menggunakan iradiasi ultrasonik dalam proses kimia, antara lain adalah : Matouq dkk (2007) melakukan penelitian degradasi diazinon yaitu sejenis komponen pencemar dalam air menggunakan

gelombang ultrasonik berfrekuensi tinggi. Pada penelitian ini dipakai konsentrasi awal diazinon yang berbeda-beda, yaitu 800,1200 dan 1800 ppm, volume cairan berbeda-beda, serta proses berlangsung dalam waktu yang tetap, yaitu 300 detik. Peralatan ultrasonik yang dipakai mempunyai diameter tranduser 20 mm yang dilengkapi dengan piezoceramics. Frekuensi yang diterapkan sebesar 1,7 MHz dengan power input 9,5 W. Kondisi terbaik dicapai saat degradasi dilakukan pada konsentrasi awal 1200 ppm dan volume 50 mL. Kemampuan degradasi akan menurun apabila volume cairan bertambah besar. Proses degradasi diazinon mengikuti model reaksi order satu semu dengan nilai konstanta sebesar 0,01/ detik. Behnajady dkk (2008) melakukan penelitian degradasi Malachite Green (MG). Pada penelitian tersebut menggunakan ultrasonik bath pada frekuensi 35 kHz untuk mengamati pengaruh konsentrasi MG, power, suhu, faktor pengadukan, penambahan etil alkohol dan iso butanol selama proses degradasi dengan ultrasonik. Proses ini menghasilkan energi aktivasi sebesar 30,95 kJ/mol pada batasan suhu antara 21-34°C dan diyakini bahwa reaksi dikendalikan oleh difusi. Dengan penambahan etil alkohol dan iso butanol dimaksudkan untuk menimbulkan radikal hidroksil, dimana pengaruh iso butanol lebih besar dengan diamatinya komponen N-demetilasi sebagai salah satu produk intermediate. Prado dkk (2008) mempelajari pengaruh sonokimia terhadap degradasi triclosan yang merupakan polutan mikro dalam limbah domestik dan rumah tangga. Peralatan ultrasonik yang dipakai berjenis tanduk dioperasikan dengan sistem kontinyu dan pulsa berfrekuensi 80 kHz. Reaksi dijalankan dalam peralatan gelas bervolume 250 mL dilengkapi jaket pendingin dan peralatan pengendali suhu. Untuk semua kasus percobaan, larutan bervolume 200 mL mengandung 5 mg/ Lt triclosan dikenai iradiasi ultrasonik dengan input power 135 W dan reaksi dipertahankan pada suhu 20°C. Hasil penelitian dilaporkan bahwa frekuensi rendah iradiasi ultrasonik mampu mendegradasi polutan ini hingga seratus persen dengan reaksi degradasi mengikuti kinetika reaksi order satu. Mojtahedi dkk, (2008) melakukan penelitian tentang kondensasi turunan hydrazine dapat dilakukan dengan lebih cepat menggunakan ultrasonik.

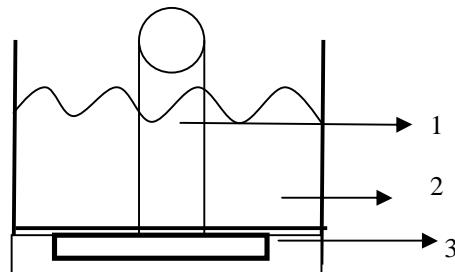
III. Metodologi

Penelitian ini menggunakan bahan-bahan seperti gliserol p.a dan HPLC water sebagai pereaksi. Variabel yang diterapkan pada penelitian ini meliputi :

- a. Waktu : 10-60 menit
- b. Rasio gliserol-air (massa) 1:10
- c. Suhu bulk (cairan) : 40, 50, dan 60°C
- d. Frekuensi peralatan ultrasonik : 37 kHz

Peralatan

Peralatan yang digunakan pada penelitian ini adalah ultrasonik jenis S 100 H dan labu ukur sebagai tempat reaksi (reaktor). Skema peralatan dapat dilihat pada gambar di bawah ini.



Gambar 1. Skema peralatan penelitian

Keterangan gambar :

1. Reaktor
2. Peralatan ultrasonik tipe S100 H
3. Tranducer

Prosedur percobaan

Mula-mula dibuat terlebih dahulu campuran gliserol-air dengan perbandingan massa 1:10 sebagai pereaksi (bahan yang hendak direaksikan). Sebelum melakukan percobaan, media air yang terdapat pada peralatan ultrasonik dipanaskan hingga tercapai suhu yang diinginkan yaitu 40, 50, dan 60°C. Setelah suhu proses tercapai, campuran reaksi yang sudah dituang pada wadah bervolume 10 ml diletakkan pada peralatan ultrasonik untuk dilakukan degradasi. Proses dihentikan jika lamanya proses sudah mencapai waktu yang diinginkan. Hasil degradasi dikeluarkan dan dianalisis menggunakan metode GC-FID.

IV. Hasil Dan Pembahasan

Dua parameter yang dipelajari dalam penelitian ini adalah temperatur cairan (bulk) dan waktu sonifikasi, dengan kondisi percobaan yang tetap adalah frekuensi gelombang yaitu 37 kHz dan perbandingan massa gliserol-air adalah 1:10.

Hasil degradasi gliserol menggunakan gelombang ultrasonik menghasilkan senyawa metanol dan beberapa komponen yang belum teridentifikasi. Pada penelitian degradasi gliserol yang telah dilakukan sebelumnya menggunakan teknologi subkritis/ superkritis air dihasilkan senyawa kimia berupa asetaldehid, metanol dan etanol serta masih ada beberapa komponen lain yang masih belum bisa teridentifikasi. (Yuniati dkk, 2010).

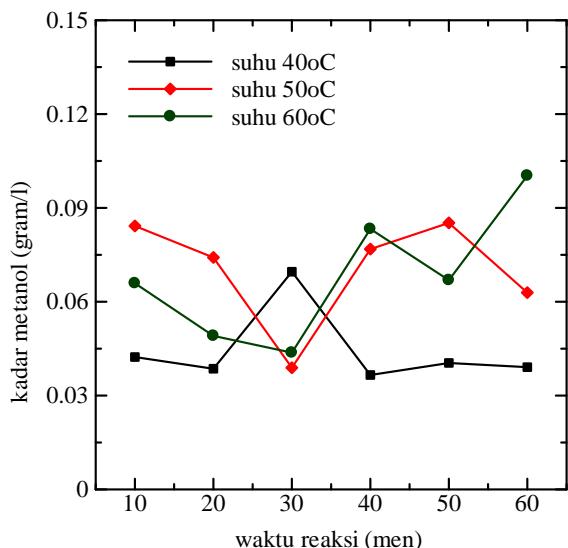
Pada kondisi subkritis/ superkritis air, media air yang digunakan mempunyai sifat-sifat yang sangat berbeda dengan air pada kondisi biasa. Air pada kondisi subkritis/ superkritis mempunyai range temperatur antara 100°C hingga di atas 374°C serta tekanannya adalah di atas 22,1 MPa.

Perbedaan jenis senyawa yang dihasilkan disebabkan karena pada pemakaian gelombang ultrasonik menyebabkan properties air juga berubah. Efek kavitas yang disebabkan karena runtuhnya

gelembung cairan menjadikan kondisi ekstrem pada titik-titik air sehingga menimbulkan terbentuknya radikal bebas yang lebih banyak. Akibatnya adalah pemutusan ikatan rantai organik senyawa mempunyai arah mekanisme yang berbeda, meskipun bisa jadi masih bisa memungkinkan ada senyawa/ komponen yang sama.

Pengaruh waktu reaksi

Hubungan antara konsentrasi metanol dengan waktu reaksi dapat dilihat pada gambar di bawah ini.



Gambar.2. Hubungan konsentrasi metanol dan waktu reaksi

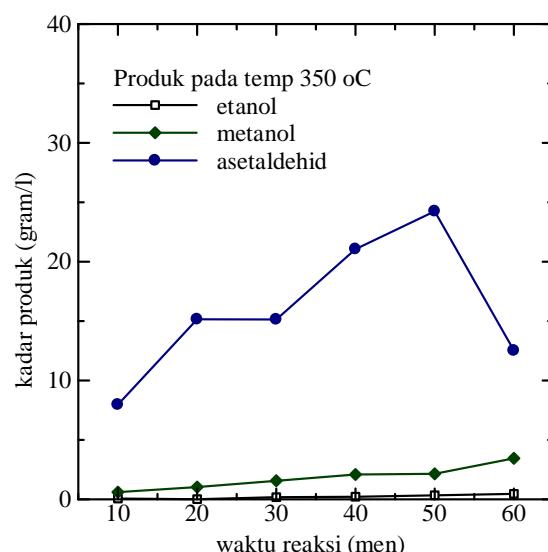
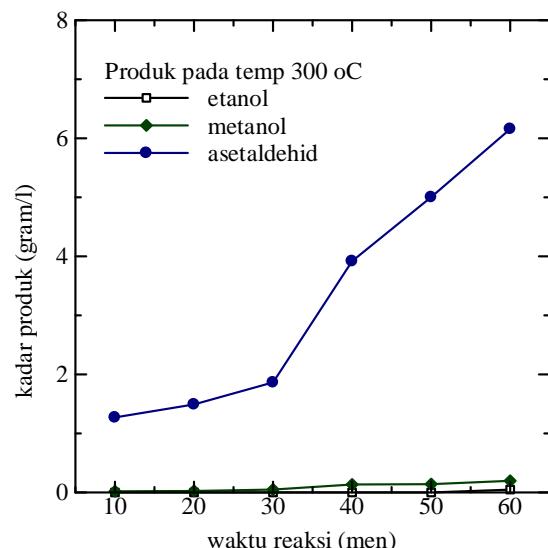
Semakin lama proses sonifikasi dapat meningkatkan kadar metanol. Kadar metanol tertinggi diperoleh pada waktu sonifikasi 60 menit yaitu 0,100255 gram/lt. Dari gambar itu terlihat bahwa glicerol yang terdegradasi semakin banyak sebagai efek dari proses sonokimia yang semakin lama.

Penelitian sejenis yang menggunakan teknologi subkritis/ superkritis air pada tekanan 250 kgf/cm² dan temperatur reaksi 300°C dan 350°C menghasilkan produk seperti gambar di bawah ini.

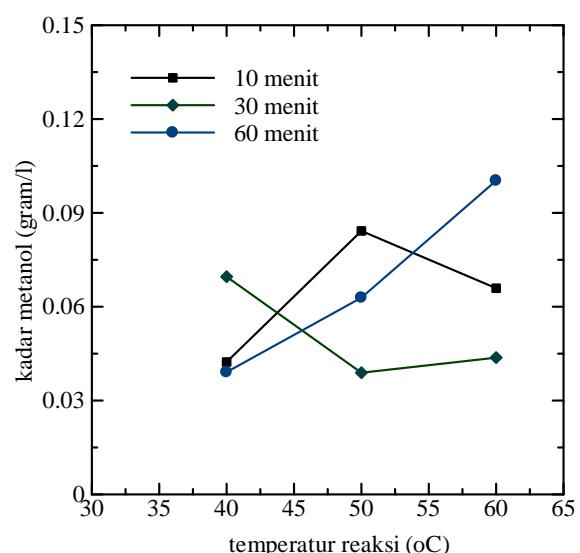
Jika dibandingkan proses menggunakan gelombang ultrasonik dengan teknologi subkritis/ superkritis air ternyata sama-sama menghasilkan produk metanol. Kadar metanol yang dihasilkan melalui proses teknologi air subkritis/ superkritis tertinggi adalah 3,4489 gr/l yang dicapai ketika temperatur reaksi 350°C dan kadar metanol terendah yang dapat dihasilkan adalah 0,01902 gr/l pada temperatur reaksi 300°C.

Pengaruh temperatur cairan (bulk)

Hubungan antara konsentrasi metanol dengan temperatur reaksi dapat dilihat pada gambar di bawah ini.



Gambar 3. Hubungan Kadar Produk dan Waktu Reaksi



Gambar 4. Hubungan Antara Konsentrasi Metanol dan Waktu Reaksi

Semakin tinggi temperatur reaksi menunjukkan fenomena kadar metanol yang semakin tinggi. Adanya temperatur reaksi yang tinggi akan menyebabkan energi aktivasi reaksi menurun sehingga proses reaksi berlangsung lebih mudah. Dengan demikian, kadar metanol yang dihasilkan juga makin tinggi, meskipun pada waktu sonikasi 30 menit terjadi anomali, dimana semakin tinggi temperatur bulk (cairan) yang semakin tinggi justru akan menurunkan kadar metanol yang dihasilkan.

Secara keseluruhan, penggunaan gelombang ultrasonik mampu menghasilkan metanol dengan kadar yang hampir sama ketika digunakan media air subkritis/ superkritis pada temperatur antara 300°C hingga 350°C dengan tekanan yang cukup tinggi, yaitu 25 Mpa.

V. Kesimpulan

Gliserol dapat didegradasi menggunakan gelombang ultrasonik pada frekuensi gelombang 37 kHz dengan parameter proses yang berpengaruh adalah temperatur bulk dan waktu sonikasi. Salah satu produk degradasi yang dihasilkan adalah metanol dengan kadar tertinggi adalah 0,100255 gram/lt ketika proses sonikasi selama 60 menit dan temperatur bulk (cairan) adalah 60°C. Produk asetaldehid dan etanol tidak ditemukan saat degradasi gliserol menggunakan gelombang ultrasonik.

Penggunaan gelombang ultrasonik dalam mendegradasi senyawa organik dapat memberikan efek reduksi energi aktivasi tanpa input energi thermal yang tinggi.

Daftar Pustaka

- [1] Behnajady, M.A., Nasser, M., Mohammad, S., Behrouz, „Effect of operational parameters on degradation of Malachite Green by ultrasonic irradiation”, Ultrasonic sonochemistry Journal. 15. pp.1009-1014, 2008.
- [2] Dasari, M.A, P.P. Kiatsimkui, W.R. Sutterlin, G.J. Suppes, “ Low-pressure hydrogenolysis of glycerol to propylene glycol”, *Applied Catalysis A: General* 281(1-2):225-231, 2005.
- [3] Mason, T. J. Ultrasound in synthetic organic chemistry. Chemical Society Reviews, vol 26,1997.
- [4] Matouq, M.A., Zaid, A.A., Tomohiko, T., Salah, A., Moh.A.Shannag, ”Degradation of dissolved diazinon pestiside in water using the high frequency of ultrasound wave”, Ultrasonic Sonochemistry Journal. 18. pp. 869-874, 2007.
- [5] Mojtabedi, M., Mashal, J., M. Saeed, A., ”Convenient ultrasound mediated synthesis of substituted pyrazolones under solvent-free conditions”, Ultrasonic Sonochemistry Journal. 15. pp. 828-832, 2008.
- [6] Prado, L.S., Ruth B., Carmen G.J., Maria, L., Marta, L., Christos, P., Nicolas, K., Dionissios, M., Elefteria, P. 2008, ”Sonochemical degradation of triclosan in water and wastewater”, Ultrasonic Sonochemistry Journal. 15. pp. 689-694, 2008.
- [7] Suslick, K.S., The Chemistry of ultrasound. Review. The Yearbook of Science & Future 1994, Encyclopedia Britannica, Chicago, pp.138-155, 1994.
- [8] Yuniati,Y., Mahfud, Sumarno, “Pemanfaatan gliserol sebagai hasil samping biodiesel menjadi produk kimia lain dalam media air subkritis hingga superkritis “, Prosiding Seminar nasional Rekayasa dan Proses 2010. Jurusan Teknik Kimia, Universitas Diponegoro, Semarang, 2010.

DESIGN OF VOICE BIOMETRIC SECURITY SYSTEM AT WEB LOGIN PAGE USING IDENTIFICATION OF MEL FREQUENCY CEPSTRAL COEFFICIENT-GAUSSIAN MIXTURE MODEL (MFCC-GMM)

Budiman Putra AR, Suyanto

Jurusan Teknik Fisika

Fakultas Teknologi Industri - Institut Teknologi Sepuluh Nopember

Kampus ITS Keputih Sukolilo – Surabaya 60111

e-mail : suyanto@ep.its.ac.id & giganggu_2@yahoo.co.id

Abstract

The need of security for web page was increased as the development of online activity especially trading or banking. Speaker recognition can be used to secure the web page that need high security level. In this research, a speaker recognition system at web page was successfully built for login authentication security. For enrollment and verification need, speech signal from clients was recorded in 35 seconds for enrollment and 10 seconds for verification and then transferred to server by network. Then this signal will be processed with Sampling, Frame Blocking, Windowing Hamming and then Discrete Fourier Transform. The signal in frequency domain will be filtered by Nonlinear Power Spectral Subtraction to reduce background noise. For identification, the system extract the feature of Mel Frequency Cepstral Coefficient (MFCC), and to build the model of these features using Gaussian Mixture Model (GMM). To improve the security level, the system use Secure Socket Layer (SSL) with 1024 bits RSA encryption. From this research, we succeed to optimizing the signal quality up to 5 dB SNR, the mean error recognition level of FAR is about 23.3% and FRR 27.5 and the maximum accuracy of recognition system is about 88% when the quality of speech signal is clean. The computation time for enrollment is about 552573,5 miliseconds and for verification is about 129062,6 miliseconds.

Keywords - Security, Nonlinear Power Spectral Subtraction (SS), Mel Frequency Cepstral Coefficient (MFCC), Gaussian Mixture Model (GMM), Secure Socket Layer (SSL), RSA

I. Introduction

The development of information and communication technology especially internet based application change and ease human activities. Now, trading and banking can be done by internet. This activity usually called as e-banking and e-commerce. Consequences of this change is the need for higher level of security verification than the conventional one such as PIN or password. Now, this verification is still used although this verification is very mature from the cyber crime attacks such as hacking, cracking and carding.

Increases of security needs at internet application make biometric be one of the effective solution [3]. Biometric is the method to identify physical characteristic of human such as face, fingerprint, voiceprint, gait, etc. This method more reliable than conventional method because of it's characteristics or feature is specific for each human. And also this biometric can not be easily imitated by other who is not legitimate like PIN number or password.

From above backgrounds we make this research. The problems which is solved of this research are:

- How to design and make a web application with high level security use voice biometric to verify the user/client

- How performances of system designed based on signal quality, error recognition level, and time consumption for enrollment and verification process are

The aim of the research is to design and implement the human identity verification using voiceprint at web page for login security. Then, this design will be analyzed about it's performance.

II. Theory Review

A. Speech Signal

Excitation powers the speech production process. It is produced by the airflow from lungs, and then carried by trachea through the vocal folds, see Fig. 1 below. During inspiration, air is filled into lungs, and during expiration the energy will be spontaneously released. The trachea conveys the resulting air stream to the larynx.

Larynx refers as an energy provider to serve inputs to the vocal tract, and the volume of air determines the amplitude of the sound. The vocal folds at the base of larynx, and glottis triangular-shaped space between the vocal folds are the critical parts from speech production point of view. They separate the trachea from the base of vocal tract. The types of sounds are determined by the action of vocal folds, and we call it

excitation. Normally excitations are characterized as phonation, whispering, friction, compression, vibration, or a combination of these.

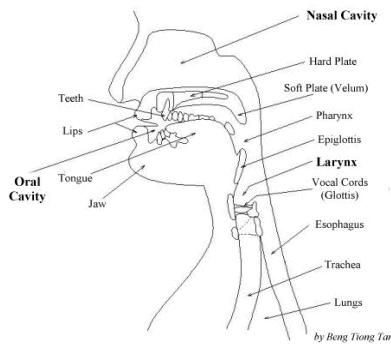


Fig 1. Anatomical structure of human vocal system [8]

B. Digital Signal Processing

Digital signal processing used in this research are Sampling, Frame Blocking, Windowing Hamming, Discrete Fourier Transform, Nonlinear Power Spectral Subtraction, Mel Filter Bank and Discrete Cosine Transform.

Sampling

The continuous time signal (speech) is sampled at discrete time point and then, the samples are quantized due to obtain a digital signal. The sequence of samples $x[n]$ is obtained from the continuous time signal $x(t)$, which stated clearly in the relationship below:

$$x[n] = x(nT) \quad (1)$$

where, T is the sampling period and $1/T = fs$ is the sampling frequency in samples/second, n is represent as the number of samples. The above equation is mainly used to obtain a discrete time representation of a continuous time signal through periodic sampling. The size of the sample for digital signal is determined by the sampling frequency and the length of the speech signal in seconds.

Frame Blocking

Process of segmenting the speech samples into the small frames with the time length within the range of 20-40msec known as Framing blocking. Framing blocking enables the non-stationary speech signal to be segmented into quasi-stationary frames, and enables Fourier Transformation of the speech signal. It is because, speech signal is known to exhibit quasi-stationary behavior within the short time period of 20-40msec. Thus, it is rationale if the Fourier Transformation of the speech signal was enable, because a single Fourier Transform of the entire speech signal cannot capture the time varying frequency due to the non-stationary behavior of the speech signal.

Windowing Hamming

Windowing step is a meant to window each individual frame, in order to minimize the signal discontinuities at the beginning and the end of each frame. If we define the window as $w(n)$, $0 \leq n \leq N-1$, where N is the number of samples in each frame. Thus, the result of the windowing can be shown based on equation below:

$$w(n) = x(n)W(n), \quad 0 \leq n \leq N-1 \quad (2)$$

Here, hamming window most commonly used as window shape in speech recognition technology, by considering the next block in the feature extraction processing chain, integrates all the closest frequency lines. Impulse response of the Hamming window was shown, based on the equation (3) below:

$$W_{hamm} = \begin{cases} 0.52 - 0.46\cos(2\pi n/(N-1)) & 0 \leq n \leq N-1 \\ 0 & lainnya \end{cases} \quad (3)$$

For this reasons, Hamming window is commonly used in MFCC extraction, which shrinks the values of the signal toward zero at the window boundaries and avoiding discontinuities.

Discrete Fourier Transform (DFT)

Discrete Fourier Transform (DFT) is normally computed via Fast Fourier Transform (FFT) algorithm, which is widely used for evaluating the frequency spectrum of speech. FFT is a fast algorithm which exploits the inherent redundancy in the DFT and reduces the number of calculations. It provides exactly the same result as the direct calculation. The Discrete Fourier Transform is represented by the equation below, where $Y(k)$ is the Fourier Transform of widowed signal each frame $w(n)$.

$$Y(k\omega_0) = \sum_{n=0}^{N-1} w(nT)e^{-jk\omega_0 nt} \quad (4)$$

Nonlinear Power Spectral Subtraction

Nonlinear Power Spectral Subtraction (SS) filter is speech enhancement which works in frequency domain signal. This filter has two phases, there are *Voice Activity Detection* (VAD) and *Spectral Subtracting*.

Voice Activity Detection (VAD)

The purpose of Voice Activity Detection (VAD) is to determine whether a frame of the captured signal represents voiced, unvoiced, or silent data. Voiced sounds are periodic in nature and tend to contain more energy than unvoiced sounds, while unvoiced sounds are more noise-like and have more energy than silence. Silence has the least amount of energy and is a representation of the background noise of the environment. The VAD plays a central role in spectral subtraction algorithms because its

accuracy dramatically affects the noise suppression level and amount of speech distortion that occurs. The noise estimate in spectral subtraction uses the VAD to decide when to update the noise reference in the absence of speech.

The energy level detection VAD algorithm used in this paper is described below. The initial noise spectrum, mean, and variance are calculated assuming the first 10 frames are noise only. Threshold are calculated for speech and noise decisions and all statistics are gradually updated when a noise frame is detected. The update factors α and β can be tuned and have been set to 0.95. Other research has extended the energy calculation VAD to dual and multiple spectral subbands within each frame.

The first step of the algorithm is to buffer the data into the k^{th} frame, $x(n, k)$, and transform it into the frequency domain.

$$Y(\omega, K) = DFT(x(n, k)) \quad (6)$$

Next, the noise spectrum and noise mean for $k=1$ are initialized.

$$No(\omega) = Y(\omega, k) \quad (7)$$

$$\mu_N = \frac{1}{L} \sum_{\omega=0}^{L-1} No(\omega) \quad (8)$$

If $VAD = 0$, then the noise spectrum, mean, and standard deviation for frame are all updated. Frames 2 through 10 are assumed to be noise in order to get a good initial average of the stationary noise in the environment.

$$No(\omega) = \alpha_N No(\omega) + (1 - \alpha_N) Y(\omega, k) \quad (9)$$

$$\mu_{No}(k) = \frac{1}{L} \sum_{\omega=0}^{L-1} No(\omega) \quad (10)$$

$$\mu_{No} = \beta_N \mu_{No} + (1 - \beta_N) \mu_{No}(k) \quad (11)$$

$$\sigma_{No} = (\beta_N \sigma_{No}^2 + (1 - \beta_N) \mu_{No}(k)^2)^{1/2} \quad (12)$$

Thresholds are updated if a frame does not contain speech, using the mean and variance of the noise estimate, where threshold settings are adjusted using the multipliers α_S and α_N , which can be adapted and set experimentally. Optimally adapting these VAD thresholds has been the subject of recent research, but was not attempted in this paper because sensitivity to the thresholds was reduced by the iteration of the algorithm.

$$Tresh_S = \mu_{No} + \alpha_S \sigma_{No} \quad (13)$$

$$Tresh_{No} = \mu_{No} + \alpha_N \sigma_{No} \quad (14)$$

Spectral Subtracting

In this research, subtracting phase use constant of oversubtraction dan spectral floor based on SNR posteriori. The general equation of spectral subtracting as below,

$$|\hat{s}(\omega)|^2 = \begin{cases} |\hat{y}(\omega)|^2 - \alpha |\hat{d}(\omega)|^2 & \text{jika } |\hat{y}(\omega)|^2 > |\hat{d}(\omega)|^2 \\ \beta |\hat{d}(\omega)|^2 & \text{lainnya} \end{cases} \quad (15)$$

Constanta of α oversubstraction on purpose to overestimate speech dan constant β spectral floor. The value of α is function of posteriori signal-to-noise ratio (SNR)as equation as below,

$$\alpha = \alpha_0 - \frac{3}{20} SNR_{posterior} \quad -5 \text{dB} \leq SNR_{posterior} \leq 20 \text{ dB} \quad (16)$$

$$SNR_{posterior}(\omega) = \frac{|Y(\omega)|^2}{|\hat{D}(\omega)|^2} \quad (17)$$

Mel Frequency Cepstral Coefficient (MFCC)

Mel scale is a unit of special measure or scale of perceived pitch of a tone. It does not correspond linearly to the normal frequency, but behaves linearly below 1 kHz and logarithmically above 1 kHz. This frequency range is based on the studies of the human perception of the frequency perception of the frequency content of sound. Therefore we can use the following formula to compute the mels for a given frequency f in Hz :

$$Mel = 1000 \times \log_2(1 + \omega) \quad (18)$$

MFCC is the most commonly used set of features in speaker recognition. It is the parametric representation of the speech signal based on the Fourier spectrum. Theprocess to compute it is:

1. Division of signal into overlapping frames.
2. Pre-emphasis of the signal, which consists on rising the level of higher frequencies .
3. Every frame is multiplied by a window function, such as Hamming window.
4. Determination of the amplitude spectrum using FFT.
5. Conversion to Mel scale by applying a filter bank to the magnitude spectrum.
6. Application of the discrete cosine transform (DCT) to the logarithm of the filter bank output

$$c[n] = \sum_{m=0}^{M-1} s[m] \cos\left(\frac{\pi n(m-0.5)}{M}\right) \quad (19)$$

The first coefficient is a measure of the energy in the signal and depends on the intensity. The rest of the coefficients contains the information of the vocal tract filter and are fairly uncorrelated. MFCC is known to deconvolve the source and the vocal tract. However, it is noticed that in practice the cepstrum coefficients are affected by high pitched voices. For more details about mel-scaled cepstral coefficients.

C. Gaussian Mixture Model (GMM)

Gaussian mixture model uses multi-modal Gaussian distribution to represent speaker's voice and vocal tract configurations. For a feature vector denoted as \mathbf{C} , the mixture density for speaker s is defined as

$$p(\mathbf{x}; \boldsymbol{\theta}) = \sum_{c=1}^C \alpha_c N(\mathbf{x}; \boldsymbol{\mu}_c, \Sigma_c) \quad (20)$$

The density is a weighted linear combination of M component uni-modal Gaussian densities each parameterized by a mean vector and covariance matrix. Collectively, the parameters of a speaker's density model are denoted as

$$\boldsymbol{\theta} = \{\boldsymbol{\alpha}_1, \boldsymbol{\mu}_1, \Sigma_1, \dots, \boldsymbol{\alpha}_C, \boldsymbol{\mu}_C, \Sigma_C\} \quad (21)$$

D. Client-Server Architecture

At computer network configuration, a computer used to save database and accessed by many people called as server then other computer used to access database from remote area and has lower capacity than server called client. These server and client are connected by network. These structure called as client server architecture. This model has been applied at world network now.

E. Secure Socket Layer RSA 1024 Bit

Secure Sockets Layer (SSL) is the most widely used protocol for implementing cryptography on the Web. SSL uses a combination of cryptographic processes to provide secure communication over a network. SSL provides a secure enhancement to the standard TCP/IP sockets protocol used for internet communications. The secure sockets layer is added between the transport layer and the application layer in the standard TCP/IP protocol stack.

RSA is an asymmetry cryptography algorithm. This algorithm has two key, there are public key which published and just for encrypting the plaintext; and private key which is secret and just for decryption of ciphertext.

II. System Configuration

For designing of system overall, the configuration is divided into five sections of designing. There are designing *Voice Biometric Security* system, designing digital signal processing, modelling of GMM, activate SSL RSA 1024 bit and then integration of system at web page.

A. Design of Voice Biometric Security System

Voice Biometric Security System is a speaker verification at web application for login security. This system generally can be divided to be some process as figure below,

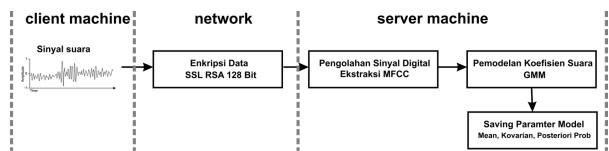


Fig 3. Block Diagram of Voice Biometric Security System

System design at these figure is a process when client enroll to the system. Speech signal from client enter and transferred to server by encryption SSL RSA 1024 bit. After speech signal was transferred to server, this signal was processed to extract the feature MFCC about 14 coefficients. These coefficients then was modelled by GMM with 12 mixtures. After training of Expectation Maximization (EM) is converged, the parameters of model is saved to database.

For verification process, these phase is relatively same as enrollment process. After coefficient MFCC is extracted then system calculate the loglikelihood with model who he claimed to. Then, this loglikelihood is compared to threshold value for make decision.

B. Design of Digital Signal Processing (DSP)

To extract the MFCC coefficient, speech signal have to processed with some steps. These steps are Sampling, Frame Blocking, Windowing, Discrete Fourier Transform (DFT), Nonlinear Power Spectral Subtraction (SS), Mel Filter Bank, and Discrete Cosine Transform (DCT), see figure below.

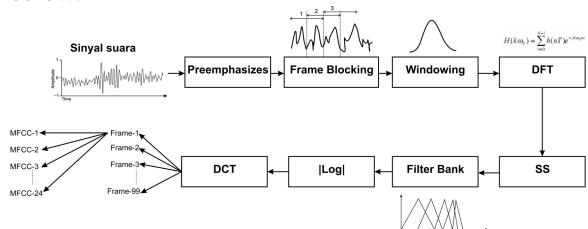


Fig 4. Digital signal processing of Voice Biometrics Security

Speech signal was recorded in 35 or 10 seconds, then processed step by step with sampling rate 8000 Hz, Frame Blocking 25 ms or 320 samples with overlapping 50%, Widowing Hamming, DFT 1024 points. These framed signal on frequency domain is filtered by Nonlinear Power Spectral Subtraction (SS) with constantas of oversubtraction $\alpha_0=5$ and spectral floor $\beta=0.001$. After filtering, signal enter to mel filter bank, absolut logarithm calculating and last is processed by DCT to get 14 coefficients of MFCC.

C. Design of Gaussian Mixture Model

Initialization of GMM is done by K-Means with cluster centre number is about 12 that selected randomly. K-means process at 50 iterations to get convergence of state. After initialization step, then

training GMM to maximize loglikelihood value. Terminate of this training is parameterized by convergence of state. The convergence of state is maximum iteration is about 1000, error 1×10^{-4} with minimum repetition of error up to 100 iterations.

D. Activation of SSL RSA 1024 Bit

To activate SSL process on the web application is did by sets the configuration server files. Encryption at SSL use socket number 8443. This encryption of SSL active as long as client communicate with server.

E. Integration on the Web Application Page

Figure in below is the visualization of web application for verification and enrollment process.



Fig 5. Visualization Of Enrollment And Verification Page

Visualization on the top is enrollment and verification web page. To process or access, client have to fill up the username form blank. Then they record their speech by push the record button. After 35 seconds record their speech, this signal send to server with push the send button. The web page will show the process status at number 9. After the speech signal was sent, then client push the submit button. The enrollment or verification will be processed by system. Process of enrollment start by digital signal processing to extract MFCC feature, training GMM and saving these parameters to database. Then verification process begin with digital signal processing to extract MFCC and then calculate of loglikelihood value from the model who he claimed to.

F. Construction of Biovoice Database

This system database is saved by My SQL named “*biovoice*”. Database consist of 14 tables consist of 12 tables GMM parameters each mixture, one table identity of enrolled client and one temporary table to save .wav filename and client session id of web application.

III. Simulation and Analysis

This design then is analysed to get its performance. Performance analysis is consist of three views, there are signal quality after filtering, recognition rate and time consumption.

A. Signal Quality

From simulation at 5 speech signal samples get the result of quality signal as figure below.

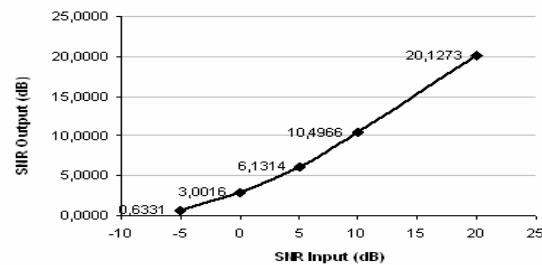


Fig 6. SNR Input And Output After Filtering

Filter is use constantas of oversubtraction $\alpha = 5$ and spectral floor $\beta = 0.001$. From simulation result, this filter configuration can increases signal quality about 5 dB SNR at quality input is -5 dB SNR. But this improvement degrade step by step as increases as quality input.

B. Recognition Rate

The recognition rate of this system is analyzed from three views. There are Match-Mismatch Language views, filtering correlation and then recognition rate of sytem overall. For Match-Mismatch Language analysis and filtering correlation use speech 20 speaker in duration 2 second for enrollment and verification process. Then for the recognition rate system analysis use 10 speakers with speech signal in 35 seconds for enrollment and 10 seconds for verification process.

▪ Match-Mismatched Language

Table 1. Correlation of Match-Mismatching Language –EER (%)

No	Language		% EER
	Enrollment	Verification	
1	Indonesia	Indonesia	34.8
2	Indonesia	Multilingual*	38.2
3	Multilingual*	Indonesia	40.1

*) With variation of language, consist of Inggris, Arab, Jawa

From simulation result, there is no correlation between match or mismatched of language when enroll and verify with error recognition EER. So this application can accept whatever client language use.

▪ Nonlinear Power Spectral Subtraction

Table 2. Correlation Filter SS – EER (%)

No	SNR dB		EER (%)		
	Enrollment	Verification	No SS	SS	% Optimize
1	Clean*	-5	53,0	47,2	10,94
2	Clean	0	48,4	40,1	17,15
3	Clean	5	49,0	36,2	26,12
4	Clean	10	42,5	32,3	24,0
5	Clean	20	45,0	31,6	29,78
Rata-rata			21,60		

*) Without noise

With use filter SS, from table in the top, the recognition rate get improved about 20 % higher than without filter SS. And the maximum optimization of recognition rate is when signal input high is 20 dB SNR.

▪ Recognition Rate of System

Table 3. FAR and FRR System

No	Signal Quality		FAR (%)	FRR (%)
	Enrollment	Verification		
1	Clean	0	46,7	40
2	Clean	10	18,9	30
3	Clean	20	13,3	30
4	Clean	Clean	14,4	10
Mean		23,325	27,5	

*) Without noise

Error recognition rate FAR is error when an impostor is accepted as a legitimate. If this error so big so in one identity can accept more than one client. This condition can make the system unsecure from impostor. In this system has mean of FAR for quality signal input verification 0,10,20 dB up to clean is about 23,3. This value show that in this system, one identity can be accessed by 2-3 impostor at 10 access client. So this system is not secure enough for web application and need upgrading on its biometric and system configuration.

And error recognition rate FRR is error system when a legitimate user is rejected. If this value so big, the client/user comfort will decreases and it is not match with aimof this biometric system made. In simulation, the FRR of system is about 27,5%. With this value can be described when legitimate users access, more than 2 acess process is rejected at 10 times access. So from this system properties, this system is not comfortable enough for clients.

C. Time Consumption

At enrollment phase, computation process consist of three, there are digital signal processing (DSP) to extract feature MFCC, modelling GMM and saving model paramters GMM on data base. From the simulation, we get mean of time

consumption each process, DSP about 487093,7miliseconds, GMM 113778,2 miliseconds and saying to database is about 411 miliseconds then total time consumption for *enrollment* phase is 552573,5 miliseconds.

For verification phase, computation process can devide to two process, there are digital signal processing and access to database with calculate the loglikelihood value. Time consumption for DSP is about 128290,7miliseconds, and second process is about 771,9 miliseconds so total time consumption for verification phase is about 129062,6 miliseconds.

This mean of time consumption is small enough for web application especially for login page. With ignore the transfer time consumption of data, this process can be applied into web application easily.

V. Conclusion

After designing, simulation and analysis of system, there are some conclusion about this research,

- *Voice Biometric Security* at web page using identification of MFCC-GMM with number of feature is about 14 and number of mixtures is about 12, SSL with RSA algorithm of encryption and 1024 bit keylength has been designed.
- *Voice Biometric Security System* have performances, there are; optimize signal quality up to 5 dB SNR; mean of error recognition rate for quality signal when verification from -5 dB SNR to clean are FAR sebesar 23,3% and FRR 27,5% with minimum error FAR 14% and FRR 10% or maximum acuratu of recognition about 88%; mean of time consumption for *enrollment* phase is about 552573,5 miliseconds and verification is about 129062,6 miliseconds .

For further research, this design can be reconfigured especially for speech duration and number mixture of GMM but still make the system comfortable for clients.

References

- [1] Besacier, Laurent dkk. 2004. "Voice Biometric Over The Internet In The Framework Of COST Action 275" . Eurasip Journal on Applied Signal processing 2004
- [2] Jain, K Anil, dkk. 2008. "Biometric Template Security, Review Article". Eurasip Journal on Applied Signal processing 2008
- [3] Gilmore, William, dkk. 2008. " The Future Of Online Internet Marketing: A Solution To Behavioral Marketing using Biometrics". Journal of Bussines & Economic Research February 2008.
- [4] Wayman, James. 2005."Biometrics Systems,

- Technology Design and Performance Evaluation". Italy, USA: Springer
- [5] Bimbot, Frederic dkk.2004. "Tutorial on Text Independent Speaker Verification. EURASIP Journal on Applied Signal Processing". Hindawi Publishing Corporation.
 - [6] Kinnunen, Tomi H.2005. "Optimizing Spectral Feature Based Text Independent Speaker Recognition", Dissertation. Department Computer Science and Engineering, University of Jouensuu, Jepang.
 - [7] Houtamaki, Rosa Emilia Gonzales. 2005. "Fundamental Frequency Estimation and Modeling for Speaker Recognition", Master's thesis. Department of Computer Science, University of Joensuu, Jepang.
 - [8] Kurnia, Indah Sari. 2006. "Aplikasi Pengenalan Penutur Untuk Sistem Absensi Dengan Metode Penutur Dependent (Fase Pembelajaran)". Tugas Akhir Jurusan Teknik Telekomunikasi PENS ITS
 - [9] Berouti, M., Schwartz, R. And Makhoul, J., "Enhancement Of Speech Corrupted By Acoustic Noise", Proc. IEEE ICASSP, Pp. 208-211, Washington DC, April 1979.
 - [10] Bhatnagar, B.E., Mukul. 2002. "A Modified Spectral Subtraction Method Combined With Perceptual Weighting For Speech Enhancement", Thesis. Master Of Science In Electrical Engineering, The University Of Texas Dallas
 - [11] Rittscher, Jens dan Chuck Stewart. _____. "Statistical and Learning Techniques in Computer Vision Lecture 4: Gaussian Mixture Models and the EM Algorithm".
 - [12] Bishop, Christopher M. 2006. "Mixture Models and The EM Algorithm, Advances Tutorial ". Microsoft Research: Cambridge.
 - [13] Paalanen, Pekka. 2005. "Feature Representation and Discrimination Based on Gaussian Mixture Model Probability Densities – Practices and Algorithms". Department of Information Technology, Lappeenranta University of Technology, Finlandia.

TEKNIK DETEKSI DINI PENDERITA KELAINAN PITA SUARA MENGGUNAKAN ANALISA SINYAL AKUSTIK

Hertiana Bethaningtyas, Dyah Kusumaningrum, Sekartedjo, Dhany Arifianto

Jurusan Teknik Fisika

Fakultas Teknologi Industri - Institut Teknologi Sepuluh Nopember

Kampus ITS Keputih Sukolilo – Surabaya 60111

email: {hertie_ana, sekar, dhany}@ep.its.ac.id

Abstrak

Makalah ini merupakan perkembangan metode diagnosa tak sentuh pada pasien penderita kelainan pita suara. Gejala awal kelainan pita suara biasanya ditandai dengan suara parau atau adanya perubahan suara dari kondisi normal. Penentuan tingkat keparahan dilakukan dengan menggunakan analisa estimasi kontur frekuensi dasar F_0 . Dimana dari kondisi keparahan tersebut dilakukan analisa sinyal menggunakan transformasi wavelet untuk mencari feature yang terbaik dalam penentuan jenis kelainan pita suara yang diderita.

Kata Kunci: suara parau, analisa sinyal, tingkat keparahan, transformasi wavelet

I. Pendahuluan

Pita Suara adalah organ yang memproduksi suara melalui pergerakan bolak-balik otot pita suara dan interaksi dengan organ lain. Gangguan pada fungsi pita suara akan menyebabkan kehilangan sementara atau permanen kemampuan berbicara dan sering menimbulkan suara parau. Deteksi dini pada stadium awal terjadinya gangguan pada laring memperbesar peluang pulihnya organ ini berfungsi normal kembali. Namun, bila baru diketahui pada stadium lanjut, seseorang dapat kehilangan kemampuan berbicara, bahkan pada beberapa kasus, dapat menyebabkan kematian, misalnya pada kasus kanker nasofaring[1].

Mencari ciri (feature) yang dapat dijadikan tolok ukur akurat untuk menentukan tipe dan tingkat parah-tidaknya kelainan pada pita suara, merupakan tantangan tersendiri mengingat sifat sinyal suara manusia yang berubah secara waktu (nonlinear time-varying). Teknik yang dipakai adalah pasien diminta mengucapkan vokal (misalnya /a/) secara kontinyu dalam satu nafas[2].

Penelitian menggunakan metode non-invasive telah dilakukan oleh group akustik Teknik Fisika FTI-ITS sejak 1998 bekerja sama dengan Fakultas Kedokteran Universitas Airlangga dan RSUD dr. Soetomo Unit Pelaksana Teknis-Telinga Hidung Tenggorok (UPF-THT). Ide penelitian ini berawal dari kesulitan penegakan diagnosa bagi dokter spesialis THT yang harus melakukan dengan memasukkan kabel optis elastis (laringoskopi) ke tenggorok karena bersifat invasif sehingga menimbulkan ketidaknyamanan pada pasien[1].

Kelanjutan penelitian yang dilakukan saat ini, berdasarkan tingkat keparahan kelainan pita suara yang diderita oleh pasien, dilakukan kelanjutan analisa sinyal suara menggunakan

transformasi wavelet dengan mother wavelet “Daubechies (db)”. Pemilihan mother wavelet “db” dilakukan karena algoritma sinyal pada “db” sangat padat (ompactly support), *effective band*, dan memiliki impuls respon yang hamper flat mendekati filter sempurna yang diharapkan[5].

II. Metode Analisa Sinyal Akustik

Sinyal akustik (suara) merupakan besaran fisis analog yang dapat berubah menurut waktu, ruang dan variabel bebas lainnya. Suara manusia terbentuk akibat proses fonasi yang dilakukan oleh organ articulator. Salah satu organ yang terpenting dalam proses fonasi adalah pita suara yang berfungsi sebagai generator. Pita suara digerakkan oleh otot-otot intrinsic laring yang dapat dikendalikan (volunter), sehingga dapat dilatih untuk menghasilkan suara yang diinginkan. Bila terjadi kelainan pada pita suara, sering ditandai dengan gejala suara parau. Suara parau adalah suatu istilah umum untuk setiap gangguan yang menyebabkan perubahan suara. Ketika parau, suara dapat terdengar serak, kasar dengan nada lebih rendah daripada biasanya, suara lemah, hilang suara, suara tegang dan susah keluar, suara terdiri dari beberapa nada, nyeri saat bersuara, atau ketidakmampuan mencapai nada atau intensitas tertentu.

Adanya perubahan suara yang terjadi, juga menimbulkan perubahan sinyal suara. Dimana sinyal suara tersebut dapat diolah dengan melakukan analisa sinyal akustik, agar detail dan penyebab perubahan yang terjadi dapat diketahui cirinya berdasarkan domain waktu, frekuensi, dan amplitudo. Pada makalah ini digunakan 2 metode dalam melakukan analisa sinyal akustik, yaitu Metode estimasi frekuensi dasar (F_0) untuk mengetahui tingkat keparahan penyakit kemudian menggunakan metode *wavelet transform* untuk mengetahui jenis kelainan pita suara yang

diderita.

2.1. Estimasi Frekuensi Dasar

Frekuensi dasar (fundamental frequency/ F_0) adalah ciri utama dari suara manusia yang bukan hanya membedakan satu individu dan individu yang lain tetapi juga usia dari seseorang tersebut. Secara persepsi besaran fisis frekuensi dasar suara manusia disebut pula pitch. Perbedaannya adalah F_0 memiliki satuan yaitu Hertz, karena merupakan jumlah gerak bolak-balik getaran pita suara dalam satuan waktu sedangkan pitch tidak memiliki besaran hanya memiliki kualitas secara kualitatif. Contohnya penyanyi selalu menyesuaikan dengan lagu yang akan dibawakan dengan nada dasar lagu tersebut agar tidak sumbang (over-pitch). Selain itu, suara anak-anak pria akan memiliki pitch yang tinggi yang kemudian sesuai dengan usianya yang beranjak dewasa, akan memiliki pitch yang rendah. Pada anak-anak, baik pria dan wanita, dan wanita dewasa F_0 berkisar antara 250-300 Hz, sedangkan pria dewasa memiliki F_0 berada disekitar 150 Hz.

2.2. Transformasi Wavelet

Wavelet mulai diperkenalkan sejak tahun 1980-an sebagai salah satu alternatif pendekatan baru analisa dan pemrosesan sinyal. Kemampuan untuk menganalisa suatu data dalam domain waktu dan frekuensi secara simultan dengan skala resolusi yang berbeda pada tiap kondisi yang dimiliki, menjadikan salah satu kelebihan dari transformasi wavelet. Fungsi matematika dari wavelet melakukan pemotongan data menjadi kumpulan-kumpulan frekuensi yang berbeda, sehingga masing-masing komponen tersebut dapat dipelajari dengan menggunakan skala resolusi yang berbeda. Dengan fungsi dilatasi dan translasi transformasi wavelet menurut Daubechies sebagai berikut: [5]

$$\Psi_{j,k} = \alpha^{j/2} \Psi(2^j t - k); j, k \in \mathbb{Z} \quad (3)$$

dengan α merupakan parameter dilatasi, 2^j adalah dilatasi parameter frekuensi atau skala, k parameter waktu atau lokasi ruang, dan \mathbb{Z} mengkondisikan nilai j dan k bernilai integer. Berdasarkan persamaan (3) diketahui bahwa wavelet memiliki karakteristik antara lain translasi, dilatasi serta berosilasi singkat.

Dalam melakukan pemrosesan sinyal, transformasi wavelet menggunakan 2 komponen penting. Yaitu *scaling function* dan *wavelet function*. kedua komponen tersebut dapat disebut sebagai *mother wavelet* yang harus memenuhi kondisi:[5]

$$\int_{-\infty}^{\infty} \Psi(x) dx = 0 \quad (4)$$

yang menjamin terpenuhinya sifat orthogonalitas vektor [5].

Adanya penemuan yang mendukung orthonormal wavelet pada penggunaan *mother wavelet Daubechies*, menunjukkan analisa wavelet diskrit dapat dipraktekkan. Filter Wavelet *Daubechies* memiliki karakteristik umum yang secara lengkap didukung oleh wavelet wavelet dengan fasa ekstremal dan memiliki jumlah *vanishing moment* paling tinggi untuk lebar yang ditentukan. *Vanishing moment* menunjukkan kemampuan wavelet dalam merepresentasikan sifat polinomial.

Karakteristik Karakteristik umum dari filter wavelet *daubechies* adalah secara lengkap didukung oleh wavelet dengan fasa ekstremal dan memiliki jumlah *vanishing moment* paling tinggi untuk lebar yang ditentukan. *Vanishing moment* menunjukkan kemampuan wavelet dalam merepresentasikan sifat polinomial. Filter skala yang dihubungkan adalah filter fasa minimum.

III. Hasil Dan Pembahasan

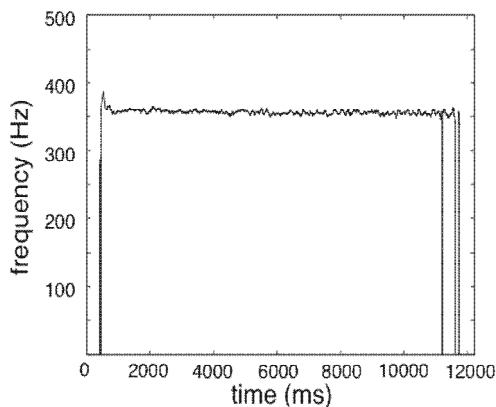
Pada bab ini dijelaskan mengenai prosedur pengambilan data, kriteria pasien dan peralatan yang dipakai serta kondisi eksperimen (experimental set-up). Kemudian dilanjutkan dengan diskusi mengenai hasil eksperimen yang diperoleh dengan menggunakan teknik wavelet.

Pengambilan data dapat dilakukan dengan kerjasama penelitian antara Departemen Telinga, Hidung, Tenggorok dan Bedah Kepala-Leher (THT-KL) RSUD dr. Soetomo Surabaya. Dengan perekaman suara parau pasien di dalam ruang audiologi THT-KL (*double-walled-sound-attenuated booth*). Pasien diminta untuk melaflakan fonem /a/ secara terus-menerus dalam satu tarikan nafas sesuai dengan kapasitas paru-paru yang dimiliki. Lama Makalah harus mengandung hasil-hasil simulasi atau pengukuran sebagai validasi metode. Lama fonasi pada orang sehat berkisar antara 12-14 detik. Namun, pada penderita kelainan pita suara pada umumnya memiliki kapasitas paru-paru yang lebih pendek. Semakin tinggi stadium penyakit yang diderita, semakin pendek durasidari fonasi pada saat pelafalan.

Sinyal suara yang diambil dengan menggunakan mikrofon Shure-58 yang memiliki respon frekuensi rata pada rentang frekuensi suara manusia (speech). Mikrofon ini dihubungkan dengan analog-to-digital converter (ADC) 24-bit Creative EMU 0404. Akuisisi data dilakukan pada laju pencuplikan (sampling rate) 16 kHz dan disimpan dalam format .wav dan menggunakan kode PCM. Pada saat perekaman dan penyimpanan data suara, identitas pasien diganti dengan kombinasi huruf dan angka serta diberi label penyakit yang telah didiagnosa oleh dokter

spesialis THT-KL menggunakan teknik gold-standard sebelum dilakukan perekaman.

Pada data suara dari 62 responden yang berpartisipasi, dilakukan analisa estimasi frekuensi dasar. Dimana ciri utama suara ini yaitu F_0 dipilih sebagai karakteristik utama dalam teknik diagnosa tanpa sentuh. Sifat utama dari F_0 adalah tidak berubah secara waktu karena merupakan parameter dasar dari fungsi fisiologis pita suara yang merupakan frekuensi alami dari getaran keduanya. Dari pengamatan lanjut pada F_0 akan diperoleh ukuran kuantitatif tambahan (*yard-stick*) dalam satu satuan waktu untuk mengkuantifikasi besar dari gangguan osilasi akibat gangguan getaran pita suara. Perbedaan nilai kuantitatif ini yang akan dipakai untuk menentukan persamaan pola, pengklasifikasian jenis kelainan fisiologis dan tingkat keparahan yang tengah diderita oleh pasien dibandingkan dengan yang sehat. Pada Gambar 2 ditunjukkan penderita A yang memiliki kelainan polip selama 1 bulan dan Gambar 3 adalah penderita polip B setelah 6 bulan. Dari Gambar 2 menunjukkan bahwa pada penderita kelainan laring polip pada stadium awal memiliki profil F_0 mirip dengan suara orang normal. Hampir tidak terlihat distorsi, kecuali pada daerah di sekitar 2000 msec. Terjadi sedikit penurunan (dip) F_0 dari 190 Hz ke 160 Hz.

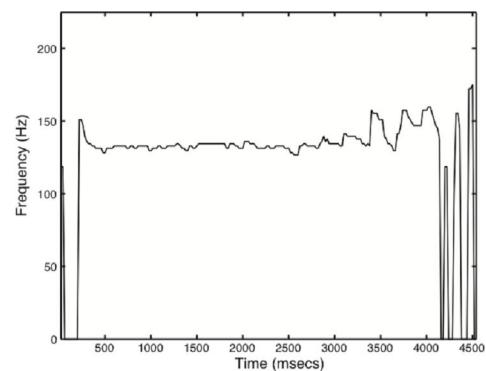


Gambar 1. F_0 dari Partisipan Normal

Gambar 2: F_0 dari penderita polip satu bulan

Sedangkan pada Gambar 3 terlihat dengan jelas semakin besarnya distorsi pada kontur F_0 penderita. Pada awal (0 – 10 mdetik), dan diakhiri fonasi pada posisi 3000 mdetik hingga akhir fonasi 4500 mdetik. Distorsi pada penderita stadium menengah ini pada akhir fonasi berupa osilasi F_0 dari 3000 mdetik hingga 4100 mdetik kemudian diikuti dengan diskontinuitas pada kontur F_0 . Diskontinuitas F_0 ini merupakan daerah yang menunjukkan parah tidaknya dari ke-parau-an suara pasien, padahal pasien melafalkan vokal /a/ secara kontinu. Durasi fonasi penderita juga tinggal hampir separuh (4,5 detik) dibanding penderita polip

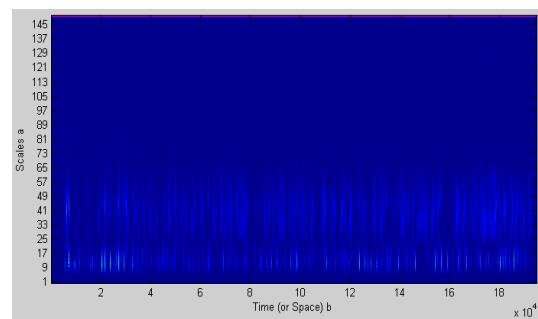
stadium awal. Pola distorsi dan laju distorsi per satuan waktu menjadi salah satu ukuran untuk menentukan tingkat keparahan kelainan laring yang diderita. Namun dari distorsi F_0 saja masih belum cukup untuk menentukan jenis kelainan pita suara yang tengah dialami oleh penderita. Diperlukan pembanding dari penderita dengan jenis kelainan pita suara yang berbeda.



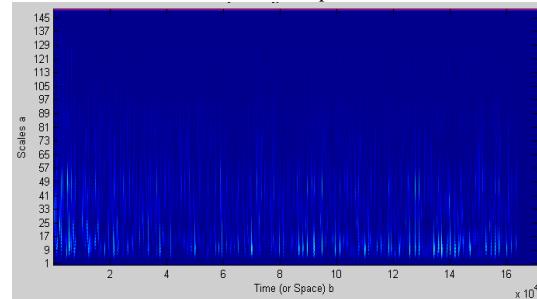
Gambar 3. F_0 Dari Penderita Polip Enam Bulan

Salah satu teknik pengolahan sinyal yang telah dikembangkan sejak awal abad 20 adalah wavelet yang merupakan teknik alternatif selain transformasi Fourier. Pada penelitian ini difokuskan penggunaan wavelet Daubechies yang memiliki kelebihan dibanding wavelet yang lain.

Pada Gambar 4 dan 5 ditunjukkan suara normal pria dengan dua usia yang berbeda dengan menggunakan wavelet Daubechies. Sumbu mendatar adalah fungsi waktu yaitu durasi fonasi dan sumbu tegak adalah skala dari koefisien wavelet Daubechies, sedangkan intensitas warna adalah amplitudo (strength) dari sinyal suara.

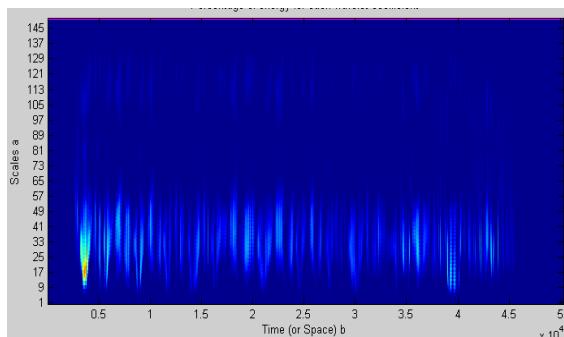


Gambar 4. Suara normal pria usia 54 tahun

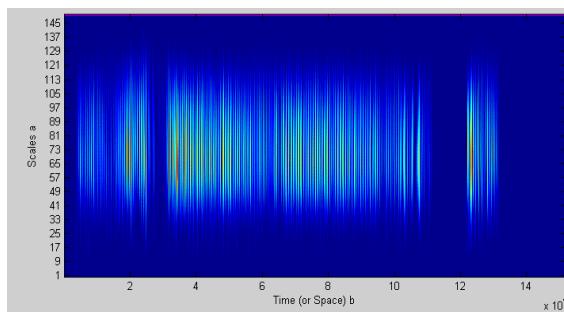


Gambar 5. Suara normal pria usia 24 tahun

Dari kedua gambar tersebut, terlihat bahwa sebaran koefisien wavelet Daubechies adalah merata (homogen) dibidang skala-waktu.



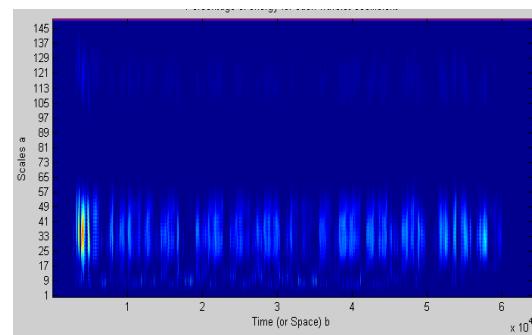
Gambar 6. Skala-Waktu Daubechies Penderita Polip 1 Bulan



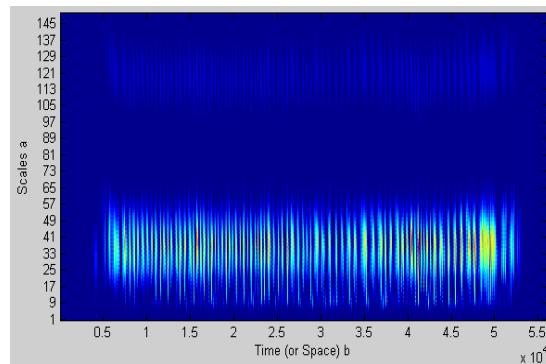
Gambar 7. Skala-Waktu Daubechies Penderita Polip 6 Bulan

Pada Gambar 6 merupakan tampilan dari penderita polip 1 bulan dan (pasien yang sama pada Gambar 2) dan gambar 7 adalah tampilan penderita polip 6 bulan (pasien yang sama pada Gambar 3). Ditunjukkan dari intensitas dari koefisien Daubechies pada penderita polip satu bulan mulai menunjukkan peningkatan keterangan (brightness) dibanding suara normal. Peningkatan intensitas yang tidak merata ini juga diartikan distorsi dari frekuensi dari sinyal suara. Dari hasil ini konsisten dengan keunggulan wavelet Daubechies dibanding transformasi Fourier yang dipakai untuk mengekstraksi F_0 . Distorsi pada penderita polip 1 bulan ini terjadi pada sekitar 0,5 detik dari awal fonasi. Pada Gambar 7 menunjukkan skala-waktu dari penderita polip 6 bulan. Terlihat lokalisasi dari intensitas skala wavelet pada bidang skala-waktu. Selain itu, intensitas dari skala juga menunjukkan semakin tinggi amplitudo dari distorsi pada sinyal suara masukan bila dibandingkan suara normal.

Yang menjadi keunggulan teknik ini adalah posisi dan intensitas dari skala wavelet Daubechies dapat dijadikan ukuran untuk menentukan jenis dari kelainan laring yang tengah dialami oleh pasien. Dari hasil ini diperoleh cara untuk menentukan tingkat keparahan dan jenis kelainan laring dari pasien dengan memanfaatkan ciri yang diperoleh dari dua teknik analisis sinyal yang berbeda, yaitu distorsi F_0 dan lokasi skala wavelet.



Gambar 8. Skala-Waktu Daubechies Penderita Nodul Vocal Stadium Dini



Gambar 9. Skala-Waktu Daubechies Penderita

Yang signifikan terjadi pada daerah skala-waktu setelah sekitar 0,5 detik dari awal fonasi dimana intensitas skala lebih jelas, yang memiliki arti intensitas keparauan lebih jelas pada penderita nodul vokal. Perbedaan lain yang terlihat juga adalah perbedaan lokasi skala di awal fonasi. Pada penderita nodul vokal, intensitas tinggi diawal fonasi merata dari skala rendah sekitar 9 hingga 49, sedangkan pada polip sekitar 9 hingga 20. Ini memiliki arti bahwa penderita polip memiliki karakteristik suara dominan pada frekuensi rendah sedangkan penderita nodul vokal lebih merata dari frekuensi rendah hingga frekuensi tinggi.

Dari Gambar 9 diatas ditunjukkan pola sebaran skala wavelet Daubechies untuk penderita kelainan nodul vokal selama 6 bulan. Terlihat jelas keparauan yang semakin parah melalui intensitas skala wavelet Daubechies. Intensitas ini tidak lagi terletak hanya pada awal fonasi tapi sudah menyebar merata pada bidang skala-waktu wavelet. Bila dibandingkan dengan penderita polip 6 bulan seperti yang ditunjukkan pada Gambar 7, kedua penderita kelainan laring dari jenis yang berbeda memiliki kesamaan yaitu sebaran merata dengan intensitas tinggi pada skala wavelet. Perbedaannya adalah pada penderita polip, dominan pada frekuensi tinggi yang ditunjukkan oleh posisi skala yang berada pada rentang 40 – 105. Pada penderita nodul vokal stadium lanjut skala wavelet tetap pada posisi rentang 9 – 49.

IV. Kesimpulan

Telah dilakukan analisa untuk memperoleh kontur F_0 dan analisa skala-waktu wavelet Daubechies untuk semua penyakit dari 62

responden yang berpartisipasi pada penelitian ini. Dari jumlah tersebut 4 orang adalah normal sebagai kelompok kendali (baseline). Dari hasil percobaan dengan menggunakan dua teknik analisa sinyal yaitu estimasi kontur F_0 menggunakan transformasi Fourier dan skala-waktu (time-scale) wavelet Daubechies, diperoleh gabungan besaran yang dapat menunjukkan tingkat ke-parahan dan jenis dari kelainan pita suara yang sedang dialami oleh pasien.

Daftar Pustaka

- [1] Soedjak, S., "Analisa Suara Penyakit pada Pita Suara", Disertasi Doktor, Fakultas Kedokteran, Universitas Airlangga, 1997.
- [2] Koike, Yasuo, "Vowel Amplitude Modulation in Patients with Laryngeal Diseases", J. Acoust. Soc. Amer., vol 45, no.4, pp. 839-844, 1969.
- [3] Murphy, P.J., "Perturbation-free Measurement of the Harmonics-to-Noise Ratio in Voice Signals Using Pitch Synchronous Harmonic Analysis", J. Acoust. Soc. Amer., vol.105, no. 5, pp. 2866-2881, May 1999.
- [4] Arifianto D., Sekartedjo, "Speech Disorder Analysis using Time-Varying Autoregressive", Proc. IEEE-MWSCAS 2004, pp.III-191-III-194, 2004, Hiroshima, Japan.
- [5] Daubechies, Inggrid, "Orthonormal Bases of Compactly Supported Wavelet", Commun. On Pure and Applied Math., vol. 41, pp.909-996., November 1988.
- [6] Michaelis, D., Gramss, T., Strube, H. W., "Glottal-to-Noise Exitation Ratio-a New Measure for Describing Pathological Noise", J. Acustica. acta acustica, vol. 83, pp.700-706.,1997.
- [7] Moran, R. J., Reilly, R. B., Chazal, P., Lacy, P. D., "Telephony - Based Voice Pathology Assessment Using Automated Speech Analysis", IEEE Transaction on Biomedical Engineering, vol. 53, no. 3, March 2006.

KEBUTUHAN CAHAYA BUAATAN PADA SISTEM PENCAHAYAAN TERPADU DALAM RUANG KULIAH P-103 TEKNIK FISIKA ITS

Heri Justiono

Jurusan Teknik Fisika

Fakultas Teknologi Industri - Institut Teknologi Sepuluh Nopember
Kampus ITS Keputih Sukolilo – Surabaya 60111

Abstrak

Pencahayaan dalam ruang yang membutuhkan energi listrik relatif kecil adalah pencahayaan terpadu yaitu pencahayaan yang diperoleh dari paduan antara cahaya alami dan cahaya buatan. Pada pencahayaan terpadu, kontribusi cahaya buatan berperan penting pada saat kondisi langit tidak memberikan pencahayaan yang mencukupi, sehingga pada saat seperti itu perlu diketahui berapa besar cahaya buatan yang minimal dibutuhkan untuk menjaga kuat pencahayaan pada bidang kerja agar tetap sesuai dengan persyaratan yang berlaku. Sebaliknya bila kondisi langit memberikan pencahayaan yang mencukupi, maka pencahayaan buatan dalam ruang tidak diperlukan. Dalam penelitian yang telah dilakukan pada ruang kuliah P-103 Teknik Fisika, diperoleh bahwa pada kondisi kuat pencahayaan dilapangan terbuka sebesar 10.000 lux dibutuhkan tambahan cahaya buatan minimal sebesar 13.904 lumen dari alat penerangan yang terpasang.

Kata kunci : pencahayaan terpadu, kontribusi cahaya buatan

I. Pendahuluan

Pemakaian energi listrik oleh masyarakat semakin hari semakin meningkat seiring dengan kemajuan teknologi, sedangkan jumlah energi listrik yang tersedia peningkatannya relatif kecil bila dibandingkan dengan peningkatan pemakaian, oleh karena itu perlu adanya usaha penghematan pemakaian energi listrik. Salah satu usaha yang dapat dilakukan yaitu merancang pencahayaan dalam ruang dengan memadukan pencahayaan alami dengan pencahayaan buatan sedemikian rupa sehingga diperoleh pencahayaan ruang yang tetap memenuhi standar dengan konsumsi cahaya buatan yang minimal.

Kontribusi cahaya alami pada pencahayaan terpadu dalam ruang, besarnya selalu berubah-ubah, hal ini disebabkan kondisi terang langit pada setiap saat terjadi perubahan.

Pada saat kondisi terang langit tidak memberikan pencahayaan yang mencukupi, maka dibutuhkan jumlah cahaya buatan yang minimal sedemikian besar hingga menghasilkan kuat pencahayaan di dalam ruang yang masih memenuhi persyaratan. Sebaliknya, bila terang langit memberikan pencahayaan yang cukup dalam ruang, maka pencahayaan buatan tidak dibutuhkan.

Pada penelitian ini pokok bahasananya dibatasi untuk ruang kuliah P-103 Teknik Fisika yang telah diketahui ukuran ruangnya, faktor pemanfaatan dari permukaan-permukaan ruang, tinggi luminer terhadap bidang kerja dan faktor langit di titik ukur.

II. Landasan Teori

2.1. Pencahayaan Alami

Pencahayaan yang cahayanya berasal dari cahaya langit secara langsung maupun melalui pantulan atau transmisi disebut pencahayaan alami. Pada pencahayaan alami terdapat beberapa istilah yang digunakan seperti langit perencanaan, faktor pencahayaan siang hari, titik ukur utama dan samping.

• Langit Perencanaan

Keadaan langit yang dijumpai selalu menunjukkan variabilitas yang besar, sehingga perlu ditetapkan keadaan langit yang sering dijumpai pada selang waktu antara pukul 8.00 sampai pukul 16.00 dengan nilai kegagalan kurang dari 10 % atau nilai keberhasilan paling rendah 90 %. Keadaan langit yang demikian disebut langit perencanaan yang nantinya digunakan untuk keperluan perancangan pencahayaan alami di dalam ruang.

Keadaan langit yang dipilih dan ditetapkan sebagai langit perencanaan adalah :

- Langit biru jernih tanpa awan atau langit yang seluruhnya tertutup awan abu-abu putih
- Keadaan langit yang memberikan kuat pencahayaan

Pada bidang horizontal di lapangan terbuka sebesar 10.000 lux

• Faktor Pencahayaan Siang Hari

Didefinisikan bahwa besarnya faktor pencahayaan siang hari adalah perbandingan kuat pencahayaan di suatu titik pada bidang kerja di dalam ruang terhadap kuat pencahayaan pada bidang horizontal di lapangan terbuka pada saat yang sama. Kedua kuat pencahayaan tersebut, tidak memperhatikan cahaya matahari, sedangkan bidang

kerja yang dimaksud di atas adalah bidang horizontal yang letaknya 0,75 meter di atas lantai ruangan. Dalam mempermudah perhitungan besarnya faktor pencahayaan siang hari, diasumsikan langit mempunyai luminansi atau terang yang *uniform* atau merata.

Faktor pencahayaan siang hari terdiri dari 3 komponen, yaitu :

1. Komponen langit.
2. Komponen refleksi luar.
3. Komponen refleksi dalam

Besarnya faktor pencahayaan siang hari dapat dihitung dengan menjumlahkan 3 macam komponen tersebut.

$$FP = KL + KRL + KRD$$

Dimana :

FP : adalah besarnya faktor pencahayaan siang hari

KL : adalah besarnya komponen langit

KRL : adalah besarnya komponen refleksi luar

KRD: adalah besarnya komponen refleksi dalam

• Titik Ukur Utama dan Samping

Adalah titik-titik ukur di dalam ruang yang kuat pencahayaannya dipilih sebagai indikator untuk keadaan pencahayaan seluruh ruang.

Titik-titik ukur tersebut terletak pada bidang kerja atau bidang horizontal yang letaknya setinggi 0,75 meter di atas lantai.

Ada dua macam titik ukur yaitu : titik ukur utama dan titik ukur samping

a. Titik Ukur Utama (TUU)

Yaitu titik ukur yang letaknya di tengah-tengah kedua dinding samping ruangan dan berjarak 1/3 d dari bidang lubang cahaya.

b. Titik Ukur Samping (TUS)

Yaitu titik ukur yang letaknya 0,5 meter dari dinding samping ruangan dan berjarak 1/3 d dari bidang lubang cahaya.

Yang dimaksud bidang lubang cahaya adalah bidang vertikal sebelah dalam dari lubang cahaya, sedangkan d adalah jarak antara bidang lubang cahaya dengan dinding seberangnya. Bila $d \leq 6$ meter, maka ketentuan 1/3 d diganti dengan 2 meter. Jarak antara dua titik ukur tidak boleh lebih dari 3 meter. Jadi bila ruangan mempunyai bidang lubang cahaya dengan panjang lebih besar dari 7 meter, maka jumlah titik ukur utama harus ditambah.

2.2. Pencahayaan Buatan

Pencahayaan yang sumber cahayanya berasal dari lampu atau luminer disebut pencahayaan buatan.

Berdasarkan besarnya fluks cahaya yang dipancarkan ke atas dan ke bawah oleh sumber

cahaya atau luminer, pencahayaan buatan di dalam ruang dibedakan menjadi 5 macam, yaitu :

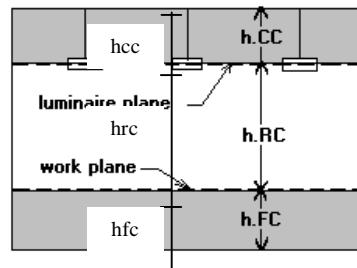
- Pencahayaan tak langsung
- Pencahayaan semi tak langsung
- Pencahayaan difus
- Pencahayaan semi langsung
- Pencahayaan langsung

Besarnya kuat pencahayaan rata-rata pada bidang kerja dapat dicari dengan metoda *zonal cavity*

Metoda Zonal Cavity

Pada metoda ini ruangan dibagi menjadi 3 daerah ruang atau rongga, yaitu : daerah ruang langit-langit, daerah ruang kamar, dan daerah ruang lantai.

Pembagian daerah ruang seperti pada gambar di bawah ini



Gambar 1. Pembagian Daerah Ruang⁽⁴⁾

Dimana : h_{cc} adalah tinggi ruang langit-langit.
 h_{rc} adalah tinggi ruang kamar
 h_{fc} adalah tinggi ruang lantai.

Dalam metoda ini, kuat pencahayaan rata-rata pada bidang kerja dapat dicari melalui persamaan di bawah ini.

$$E_{av} = \frac{F \cdot Cu}{A}$$

Dimana :

E_{av} adalah kuat pencahayaan rata-rata pada bidang kerja

F adalah besarnya fluks cahaya total yang dipancarkan oleh luminer

Cu adalah koefisien penggunaan (*Coefficient of utilization*)

A adalah lantai ruangan

Yang dimaksud dengan koefisien penggunaan (*Coefficient of utilization*) adalah perbandingan antara besarnya fluks cahaya yang sampai pada bidang kerja dengan fluks cahaya yang dipancarkan oleh luminer. Besarnya koefisien penggunaan tergantung dari distribusi cahaya yang dipancarkan oleh luminer, ukuran ruangan, faktor pemantulan permukaan-permukaan ruang, tinggi bidang kerja

terhadap lantai dan tinggi luminer terhadap bidang kerja.

Untuk memperoleh koefisien penggunaan, terlebih dulu menentukan perbandingan ruang langit-langit, kamar, dan lantai dengan menggunakan persamaan dibawah ini .

$$\begin{aligned} \text{FCR} &= 5 \cdot h_{fc} \cdot [(p+l) / (p \cdot l)] \\ \text{RCR} &= 5 \cdot h_{rc} \cdot [(p+l) / (p \cdot l)] \\ \text{CCR} &= 5 \cdot h_{cc} \cdot [(p+l) / (p \cdot l)] \end{aligned}$$

Dimana,

FCR = Rasio perbandingan ruang lantai
 RCR = Rasio perbandingan ruang kamar
 CCR = Rasio perbandingan ruang langit-langit

FCR dan CCR digunakan untuk mencari refleksi ruang lantai efektif (ρ_{fc}) dan refleksi ruang langit-langit efektif (ρ_{cc}) dengan menggunakan tabel *Persentasi Koefisien Pemantulan Permukaan Efektif untuk Rongga Langit-langit dan Rongga Lantai*. Setelah itu RCR, ρ_w , ρ_f , dan ρ_{cc} digunakan untuk mencari koefisien penggunaan pada tabel *Koefisien Penggunaan* sesuai dengan tipe alat penerangan yang digunakan. Untuk alat penerangan yang dipasang menempel pada langit-langit, ρ_{cc} sama dengan nol.

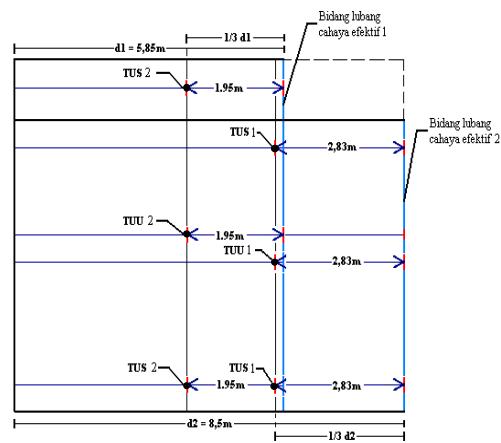
III. Ruang Kuliah P-103

Ruang kuliah P-103 Teknik Fisika mempunyai data-data sebagai berikut.

Panjang (p)	= 8,5
Lebar (l)	= 7 m
Tinggi langit-lagit (h)	= 3,75 m
Tinggi ruang lantai (h_{fc})	= 0,75 m
Tinggi ruang kamar (h_{rc})	= 2,75 m
Tinggi ruang langit-langit (h_{cc})	= 0,25 m
Koefisien pantulan langit-langit	= 80 %
Koefisien pantulan dinding	= 50 %
Koefisien pantulan lantai	= 30 %

Ruang kelas P-103 memiliki dua buah bidang lubang cahaya efektif. Kedalaman bidang lubang cahaya efektif pertama (d_1) sama dengan 583 cm sedangkan kedalaman bidang lubang cahaya efektif kedua (d_2) sama dengan 850 cm, sehingga terdapat dua buah TUU dan TUS, yaitu TUU-1 dan TUU-2 untuk bidang lubang cahaya efektif 1, dan TUU-2 dan TUS-2 untuk bidang lubang cahaya efektif 2.

TUU dan TUS terletak pada jarak $1/3 d$. Untuk $d_1 = 5,85$ m, TUU dan TUS terletak pada jarak 1,95 m dari bidang lubang cahaya efektif 1. Untuk $d_2 = 8,5$ m, TUU dan TUS terletak pada jarak 2,83 m dari bidang lubang cahaya efektif 2. **Gambar 2** menunjukkan letak titik ukur kedua bidang lubang cahaya efektif. Pengukuran tingkat pencahayaan dilakukan pada bidang kerja sepanjang TUU dan TUS untuk mengetahui karakteristik faktor langit (fl) ruangan, disamping nilai fl pada TUU dan TUS kedua bidang lubang cahaya efektif.



Gambar 2. Letak Titik Ukur Utama dan Samping

IV. Hasil Pengukuran Dan Analisa

Pengukuran dilakukan pada selang waktu antara pukul 07.00 WIB – 18.00 WIB.

Hasil pengukuran kuat pencahayaan di titik-titik ukur utama dan samping untuk bermacam-macam kondisi kuat pencahayaan di lapangan terbuka dapat dilihat pada tabel 1 dan 2 di bawah ini.

Tabel 4.1. Data Pengukuran Lubang Cahaya Efektif 1

E _{lap. terbuka}	E _{S1} (lux)	E _u (lux)	E _{S2} (lux)
15000 Lux	141	60	82,5
21000 Lux	255	108,5	143,5
29000 Lux	285	160	198,5
34000 Lux	298,5	190	221,5
34500 Lux	333	210	240
39500 Lux	350	227	277
35500 Lux	300	190	230
31000 Lux	190	130	183
17000 Lux	206	94,5	117
14000 Lux	122	55	72
10000 Lux	91	50	63
3000 Lux	17	10	17

Setelah tingkat pencahayaan di titik ukur dan di lapangan terbuka diketahui, maka nilai faktor langit (fl) dapat dihitung dengan persamaan:

$$fl = E_{di \text{ titik ukur}} / E_{lap. terbuka}$$

Hasil perhitungan faktor langit dapat dilihat pada Tabel 3 dan Tabel 4 di bawah ini.

Tabel 2. Data Pengukuran Lubang Cahaya Efektif 2

E _{lap. terbuka}	E _{S1} (lux)	E _u (lux)	E _{S3} (lux)
15000 Lux	82.5	87.5	109.5
21000 Lux	153	151.5	184.5
29000 Lux	170.5	236	260
34000 Lux	252.5	266.5	285
34500 Lux	275	295	317
39500 Lux	308	300	300
35500 Lux	277	263	303
31000 Lux	123	210	267
17000 Lux	133.5	129.5	171.5
14000 Lux	77	77	117
10000 Lux	64	70	86
3000 Lux	10	15	27

Tabel 3. Faktor Langit di Titik Ukur Utama dan Samping Lubang Cahaya Efektif 1

E _{lap. terbuka}	fl di TUU	fl di TUS ₁	Fl di TUS ₂
15000 Lux	0,0094	0,004	0,0055
21000 Lux	0,01214286	0,0051667	0,0068333
29000 Lux	0,00982759	0,0055172	0,0068448
34000 Lux	0,00877941	0,0055882	0,0065147
34500 Lux	0,00965217	0,006087	0,0069565
39500 Lux	0,00886076	0,0057468	0,0070127
35500 Lux	0,0084507	0,0053521	0,0064789
31000 Lux	0,00612903	0,0041935	0,0059032
17000 Lux	0,01211765	0,0055588	0,0068824
14000 Lux	0,00871429	0,0039286	0,0051429
10000 Lux	0,00906737	0,004952	0,0063396
3000 Lux	0,00566667	0,0033333	0,0056667

Tabel 4. Faktor Langit di Titik Ukur Utama dan Samping Lubang Cahaya Efektif 2

E _{lap. terbuka}	fl di TUU	fl di TUS ₁	Fl di TUS ₂
15000 Lux	0,0055	0,0058333	0,0073
21000 Lux	0,0072857	0,0072143	0,0087857
29000 Lux	0,0058793	0,0081379	0,0089655
34000 Lux	0,0074265	0,0078382	0,0083824
34500 Lux	0,007971	0,0085507	0,0091884
39500 Lux	0,0077975	0,0075949	0,0075949
35500 Lux	0,0078028	0,0074085	0,0085352
31000 Lux	0,0039677	0,0067742	0,0086129
17000 Lux	0,0078529	0,0076176	0,0100882
14000 Lux	0,0055	0,0055	0,0083571
10000 Lux	0,00638144	0,0070427	0,0086191
3000 Lux	0,0033333	0,005	0,009

Harga rata-rata faktor langit di titik ukur utama dan samping dari Tabel 3 dan 4 adalah sebagai berikut.

Di lubang cahaya efektif 1:

fl rata-rata di titik ukur utama = 0,00906737
 fl rata-rata di titik ukur samping = 0,004952

fl rata-rata di titik ukur samping 2 = 0,0063396

Di lubang cahaya efektif 2 :

fl rata-rata di titik ukur utama = 0,00639244

fl rata-rata di titik ukur samping 1 = 0,0070427

fl rata-rata di titik ukur samping 2 = 0,0086191

Sesuai dengan standar nasional Indonesia, fl minimal yang harus dipenuhi adalah sebagai berikut.

Lubang cahaya efektif 1:

Di titik ukur utama = 2,0475 % atau 0,020475

Di titik ukur samping = 1,1700 % atau 0,011700

Lubang cahaya efektif 2:

Di titik ukur utama = 2,9750 % atau 0,029750

Di titik ukur samping = 1,7000 % atau 0,01700

Dengan membandingkan faktor langit yang diperoleh dari hasil pengukuran dengan faktor langit yang dianjurkan (sesuai SNI), maka ruang P-103 tidak memenuhi standar kelayakan pencahayaan siang hari. Dengan demikian maka pencahayaan dalam ruang P-103 harus dipadukan dengan pencahayaan buatan.

Pada saat kuat pencahayaan dilapangan terbuka sebesar 10.000 lux, faktor langit di titik ukur utama dari lubang cahaya efektif 2 sebesar 0,006381 sedangkan faktor langit standar SNI sebesar 0,029750, sehingga selisihnya sebesar 0,023369. Jadi kuat pencahayaan dari cahaya buatan yang dibutuhkan agar pencahayaan dalam ruang memenuhi standar sebesar $0,023369 \times 10.000 = 233,69$ lux atau dibutuhkan cahaya buatan sebesar $233,69 \times 8,5 \times 7 = 13.904$ lumen.

V. Kesimpulan

Dari pembahasan yang telah diuraikan, dapat disimpulkan bahwa:

- Ruang kuliah P-103 tidak memenuhi standar kelayakan pencahayaan siang hari.
- Pada saat kuat pencahayaan di lapangan terbuka sebesar 10000 Lux, ruang kuliah P-103 membutuhkan tambahan cahaya buatan sebesar 13.548 lumen.

Daftar Pustaka

- [1] Julian, Warren G, "LIGHTING : BASIC CONCEPTS", Departemen of Architectural Science University of Sydney, 1986.
- [2] Kaufman, J, "IES LIGHTING HANDBOOK, THE STANDARD LIGHTING GUIDE FOURTH EDITION", New York, 1966.
- [3] Muhammin, "TEKNOLOGI PENCAHAYAAN", Refika Aditama, Bandung, 2001.
- [4] Seno, Bambang Darmanto, "POKOK-POKOK PEMIKIRAN TENTANG PENGUKURAN DAYA SERAP CAHAYA DARI SUATU PERMUKAAN YANG BERWARNA", Study

Literature, Teknik Fisika, Fakultas Teknologi Industri, ITS, 1983.

- [5] Soegijanto, "BANGUNAN DI INDONESIA DENGAN IKLIM TROPIS LEMBAB DITINJAU DARI ASPEK FISIKA BANGUNAN", Dirjen Pendidikan Tinggi, Departemen Pendidikan Dan Kebudayaan, 1998.
- [6] SNI 03-2396-2001, "TATA CARA PERANCANGAN SISTEM PENCAHAYAAN ALAMI PADA BANGUNAN GEDUNG", BSN, 2001.
- [7] SNI 03-6197-2000, "KONSERVASI ENERGI SISTEM PENCAHAYAAN PADA BANGUNAN GEDUNG", BSN, 2000.