

Pengenalan Pola Gerak Bibir Dalam Pengucapan Fonem Vokal Bahasa Indonesia

Faridah¹, Nazrul Effendy²

^{1,2} Jurusan Teknik Fisika FT UGM
Jln. Grafika 2 Yogyakarta 55281 INDONESIA

¹faridah@ugm.ac.id

²nazrul@gadjahmada.edu

Intisari— Telah dikembangkan suatu perangkat lunak untuk pengenalan pola gerak bibir dalam pengucapan fonem vokal /a/, /i/, /u/, /e/, dan /o/ bahasa Indonesia. Perangkat lunak ini merupakan studi awal untuk pengembangan alat terapi bicara penderita tuna rungu di Indonesia berbasis metode pembacaan gerak bibir. Secara umum sistem Pengenalan Pola Gerak Bibir ini melibatkan tiga proses, yaitu akuisisi citra, pengolahan citra dan pengenalan pola. Citra video gerak bibir di akuisisi oleh sebuah kamera digital pada pencahayaan ruang 380 – 450 Lux. Komputer akan melakukan segmentasi bibir dan ekstraksi fitur. Pola gerak bibir diwakili oleh pola luas fitur bibir berbentuk persegi panjang pada setiap frame video gerak bibir, dan melalui pola ini akan dikenali fonem vokal yang diucapkan bibir dengan menerapkan algoritma Jaringan Syaraf Tiruan.

Hasil pengujian perangkat lunak pengenalan pola gerak bibir menunjukkan tingkat akurasi sebesar 75,9 %. Kesalahan pengenalan pola oleh perangkat lunak, yang disebabkan oleh gaya menggerakkan bibir dalam berbicara untuk setiap orang berbeda, noise pada citra yang disebabkan oleh warna bibir dan kuat penerangan ruangan yang tidak merata.

Kata kunci— pola pengucapan, fonem vokal bahasa Indonesia, segmentasi bibir, fitur bibir, pola gerak bibir, Jaringan Syaraf Tiruan.

Abstract— It has been developed a system to recognize the patterns of lip motions in the utterance of vowel phoneme /a/, /i/, /u/, /e/, and /o/ in Indonesian language. The system is the results of a beginning study to develop a speech therapy tool for deaf people in Indonesia based on a method of the reading of lip motions. In general, the system consists of three processes, i.e. image acquisition, image processing, and pattern recognition. The video image of lip motion is acquired by a digital camera with room luminance of 380-450 lux. A computer will segment the lip images and extract the features of the images. The patterns of the lip motion are represented by the wide patterns of the lip features in rectangular form of each video frame of the lip motion. The system will identify the vowel phoneme by recognizing the patterns using neural network.

The test results of the software for the recognition of the patterns of the lip motion show that the accuracy of the system is 75.9%. The errors in the recognition are because of the difference of the lip motion style of speakers, noise in the difference of the lip color, and non-uniform room luminance.

Keywords— pattern utterance, Indonesian vowel, lip segmentation, lip feature, lip motion, neural network.

I. PENDAHULUAN

Komunikasi merupakan hal yang paling utama dalam kehidupan, tanpa komunikasi manusia tidak akan berkembang dan banyak informasi yang tidak tersampaikan dengan baik. Manusia tidak akan lepas dari berkomunikasi satu dengan yang lainnya, bahkan untuk penderita cacat, mereka memerlukan komunikasi dengan sesamanya.

Salah satu bentuk komunikasi adalah dengan menggunakan gerakan bibir dari lawan bicara. Bentuk komunikasi ini biasanya digunakan oleh penderita tuna rungu untuk bisa memahami apa yang diucapkan oleh lawan bicaranya. Begitu pula dalam melatih kemampuan bicaranya, penderita tuna rungu juga akan memperhatikan bentuk gerak bibir dan belajar menirukan pengucapan kata atau suku kata/fonem.

Faridah dkk [1] telah mengembangkan alat terapi berbicara untuk penderita tuna rungu di Indonesia. Alat terapi berbicara ini memberikan terapi berbicara pada penderita tuna rungu dengan cara memantau tingkat kemiripan spektrum suara yang dihasilkan penderita terhadap orang normal. Metode pembacaan gerak bibir, yang menjadi kunci utama

proses belajar berbicara penderita tuna rungu dengan terapis, belum ditemukan pada alat ini.

Pengenalan pola gerak bibir dalam sistem pembacaan gerak bibir otomatis ditemukan di beberapa penelitian. Meier dkk [2] menggunakan fitur berupa informasi dari semua piksel citra sebagai masukan pada mesin pengenalan seperti model Hidden Markov atau Jaringan Syaraf Tiruan. Untuk mengurangi jumlah vektor masukan, analisis hanya dilakukan pada komponen-komponen yang utama, misalnya dengan menggunakan teknik aliran optik pada Mase dan Pentland [3], atau metode integral proyeksi oleh Petajan [4].

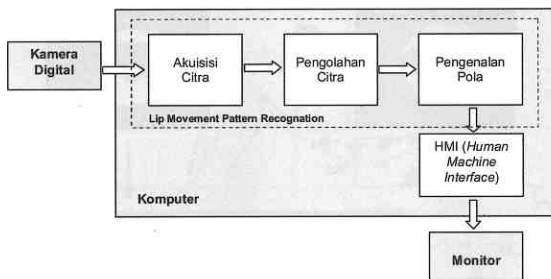
Segmentasi dan ekstraksi fitur merupakan tahapan penting dalam pengenalan pola gerak bibir. Kaas dkk [5] mengembangkan metode 'snakes' yang telah diaplikasikan untuk segmentasi bibir [6][7]. Zhang [8] menggunakan informasi warna *hue* dan tepi bibir untuk melakukan segmentasi dan lokalisasi area bibir. Perbedaan warna *hue* bibir dan kulit disekitarnya relatif tinggi meskipun pada kondisi bibir bergerak pada saat berbicara. Metode ini dikembangkan oleh Eveno [9] yang menghasilkan transformasi warna baru berbasis RGB yaitu peta kurva kromatik. Transformasi ini memungkinkan untuk membedakan warna bibir dan kulit disekitarnya meskipun

keadaan pencahayaan sekitar berubah dan kondisi wajah tanpa make-up. Transformasi yang dihasilkan telah diaplikasikan untuk mengembangkan model parametris untuk segmentasi bibir [10].

Dalam paper ini, dikembangkan suatu perangkat lunak untuk pengenalan pola gerak bibir dalam pengucapan fonem vokal bahasa Indonesia. Perangkat lunak ini merupakan studi awal untuk pengembangan alat terapi bicara penderita tuna rungu di Indonesia berbasis metode pembacaan gerak bibir. Dengan menggunakan perangkat lunak ini, tingkat kemiripan gerak bibir penderita tuna rungu dengan orang normal dalam proses belajar berbicara dalam bahasa Indonesia bisa dianalisis.

II. METODE PENELITIAN

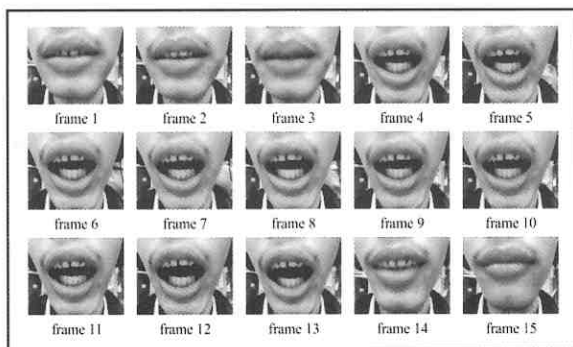
Sistem pengenalan pola gerak bibir yang dikembangkan dalam paper ini ditunjukkan oleh Gambar 1. Secara umum, sistem terdiri dari 3 bagian utama, yaitu akuisisi citra, pengolahan citra, dan pengenalan pola gerak bibir.



Gambar 1. Sistem pengenalan pola gerak bibir.

A. Akuisisi Citra

Gerak bibir seseorang dalam mengucapkan suatu fonem vokal bahasa Indonesia direkam oleh kamera digital dalam bentuk data citra video untuk selanjutnya dijadikan 15 buah *frame* citra yang masing-masing berukuran 120x160 piksel (Gambar 2). Bentuk bibir akan berubah dari *frame* pertama ke *frame-frame* berikutnya dan membentuk pola tertentu sesuai dengan fonem vokal yang diucapkan.



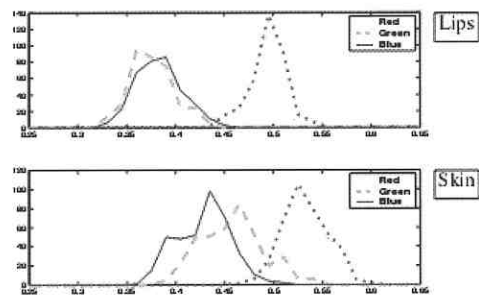
Gambar 2. Contoh pengambilan data dalam pengucapan fonem vokal /a/.

B. Pengolahan Citra

Untuk mendapatkan pola perubahan bentuk bibir dari *frame-frame*, dilakukan proses pengolahan citra pada masing-masing *frame*. Algoritma utama dari pengolahan citra adalah segmentasi bibir dan ekstraksi ciri atau fitur yang akan membentuk pola saat bibir digunakan untuk berbicara.

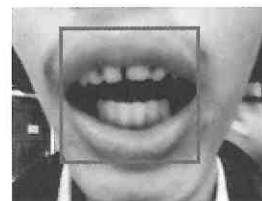
Pada penelitian ini segmentasi bibir dilakukan berdasarkan perbedaan komposisi warna antara bibir dan kulit. Warna kulit lebih dicirikan dengan komposisi warna dari pada kecerahan, bahkan pada orang yang berbeda-beda, komposisi warna pada kulit sangat konstan meskipun banyak terjadi iluminasi. Grafik histogram yang menggambarkan komposisi warna RGB pada bibir dan kulit dapat dilihat pada Gambar 3 [9]. Dapat dilihat bahwa perbedaan warna antara merah dan hijau lebih besar pada bibir dibandingkan pada kulit. Hulbert and Poggio [11] mendefinisikan nilai *pseudo hue* untuk menggambarkan perbedaan ini, yaitu :

$$h(x, y) = \frac{R(x, y)}{G(x, y) + R(x, y)} \quad (1)$$



Gambar 3. Perbandingan histogram kulit (*skin*) dan bibir (*lips*) [9].

Fitur bentuk adalah fitur dasar dalam *visual-content* pada citra. Dengan fitur ini setiap obyek gambar dapat dibedakan berdasarkan bentuk dari obyek tersebut. Bentuk dasar dalam geometri adalah bujur sangkar, persegi panjang, segitiga, lingkaran dan ellips. Pada penelitian ini bentuk bibir dalam pengucapan huruf vokal bahasa Indonesia diwakili oleh fitur bentuk persegi panjang (Gambar 4).



Gambar 4. Fitur bentuk persegi panjang yang mewakili bentuk bibir.

C. Pengenalan Pola

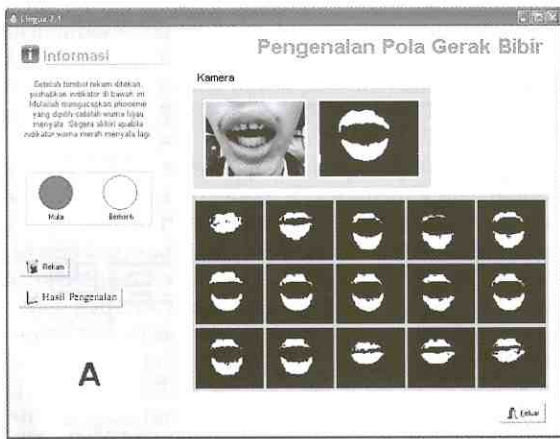
Perangkat lunak pengenalan pola gerak bibir dibangun menggunakan algoritma dasar Jaringan Syaraf Tiruan *Back Propagation* [2]. Pemodelan jaringan syaraf tiruan dilakukan dengan menggunakan data latih berupa video

citra gerak bibir pengucapan fonem vokal /a/, /i/, /u/, /e/, dan /o/, yang diucapkan oleh 8 orang dengan masing-masing orang mengucapkan tiap fonem sebanyak 10 kali.

Sistem yang dibuat selanjutnya diuji untuk mendapatkan tingkat akurasi dalam melakukan pengenalan pola gerak bibir. Tingkat akurasi didapatkan dengan cara membandingkan jumlah keberhasilan sistem dalam mengenal huruf vokal yang diucapkan oleh suatu rekaman pola gerak bibir terhadap jumlah total pengujian yang dilakukan.

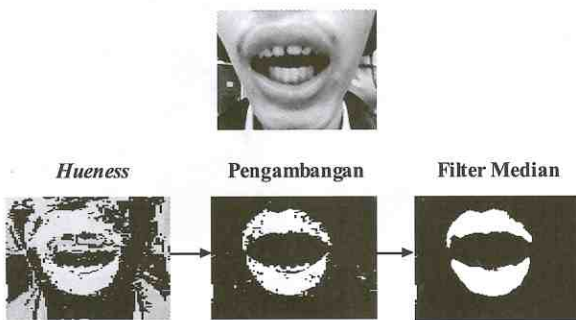
III. HASIL DAN PEMBAHASAN

Tampilan utama dari perangkat lunak pengenalan pola gerak bibir ditunjukkan oleh Gambar 5. Data yang diambil dalam bentuk data citra video dijadikan 15 buah *frame* citra yang masing-masing berukuran 120x160 piksel.



Gambar 5. Tampilan utama dari perangkat lunak pengenalan pola gerak bibir.

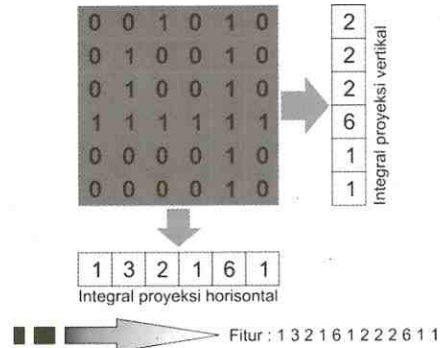
Segmentasi bibir dilakukan dengan pengembangan nilai *hueness* dengan nilai ambang lebih besar atau sama dengan 217 pada pencahayaan ruang 380 – 450 Lux dan dilanjutkan dengan penggunaan filter median untuk menghilangkan *noise* citra (Gambar 6).



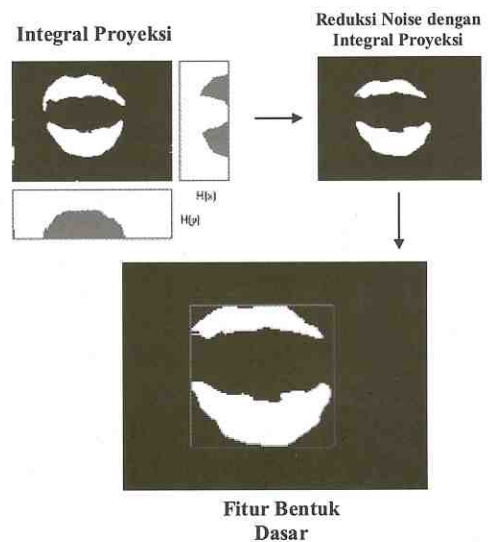
Gambar 6. Proses segmentasi bagian bibir.

Fitur bibir yang diwakili oleh bentuk persegi panjang didapatkan dengan menggunakan integral proyeksi dari citra (Gambar 7), di mana panjang dan lebar bibir didapatkan

dengan mencari letak piksel bernilai putih terluar dengan mengabaikan nilai putih yang dikarenakan oleh *noise*. Hasil segmentasi bibir yang dihasilkan oleh perangkat lunak pengolahan citra ditunjukkan oleh Gambar 8.



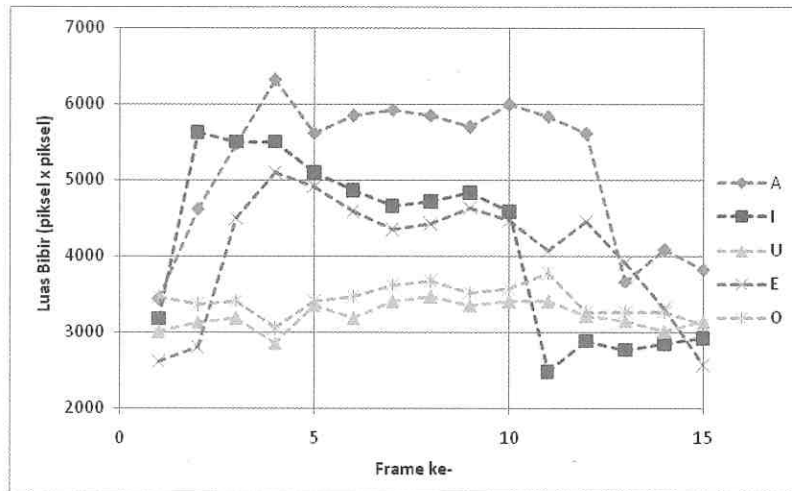
Gambar 7. Pola integral proyeksi horisontal dan vertikal pada sebuah citra [4].



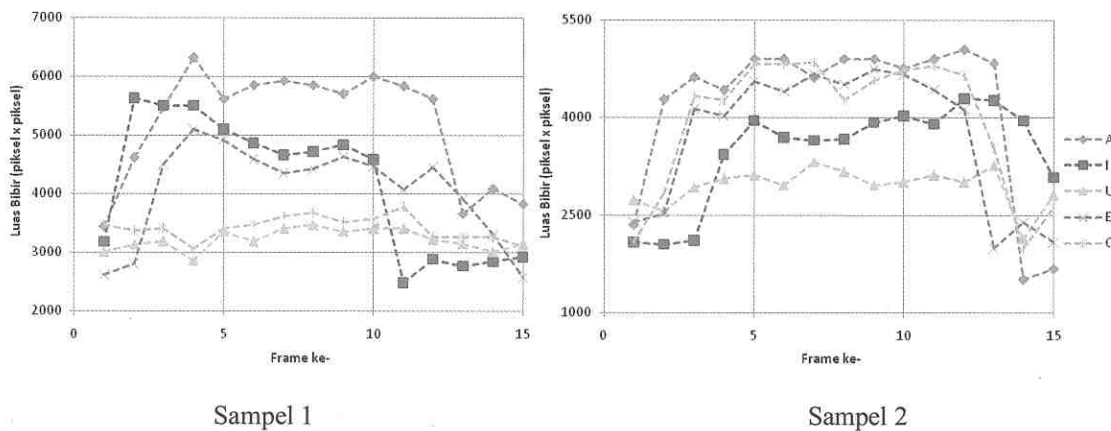
Gambar 8. Fitur bentuk bibir yang diwakili oleh bentuk persegi panjang.

Grafik luas bibir, yang merupakan perkalian antara panjang dan lebar bibir, pada tiap *frame* untuk pengucapan fonem vokal /a/, /i/, /u/, /e/, dan /o/ dalam bahasa Indonesia ditunjukkan oleh Gambar 9. Dapat dilihat bahwa pola grafik luas bibir untuk tiap-tiap fonem vokal berbeda-beda sehingga pola grafik luas bibir ini akan dijadikan pola ciri yang dikenali dalam perangkat lunak pengenalan pola gerak bibir.

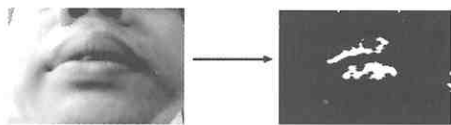
Hasil pengujian dengan menggunakan 60 data uji menghasilkan tingkat akurasi pengenalan pola gerak bibir sebesar 75,9 %. Ada beberapa hal yang menyebabkan kesalahan pengenalan pola oleh perangkat lunak, yaitu (1) cara menggerakkan bibir setiap orang berbeda (Gambar 10) dan (2) *noise* pada citra yang disebabkan oleh warna bibir (Gambar 11) dan kuat penerangan ruangan yang tidak merata (Gambar 12).



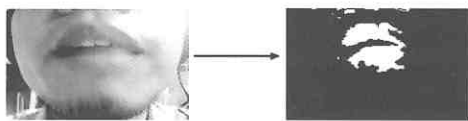
Gambar 9. Grafik luas bibir pada tiap *frame* untuk pengucapan fonem vokal /a/, /i/, /u/, /e/, dan /o/ dalam bahasa Indonesia.



Gambar 10. Grafik luas bibir pada tiap *frame* untuk pengucapan fonem vokal /a/, /i/, /u/, /e/, dan /o/ dalam bahasa Indonesia oleh 2 orang yang berbeda.



Gambar 11. *Noise* karena warna bibir yang mendekati warna kulit.



Gambar 12. *Dark corner* pada kedua ujung bibir karena pencahayaan ruang yang tidak merata.

IV. KESIMPULAN DAN SARAN

Telah dikembangkan perangkat lunak pengenalan pola gerak bibir dalam pengucapan fonem vokal /a/, /i/, /u/, /e/, dan /o/ dalam bahasa Indonesia. Pola gerak bibir diwakili oleh pola luas fitur bibir berbentuk persegi panjang pada setiap

frame video gerak bibir. Hasil pengujian perangkat lunak pengenalan pola gerak bibir menunjukkan tingkat akurasi sebesar 75,9 %. Kesalahan pengenalan pola oleh perangkat lunak, yaitu gaya menggerakkan bibir dalam berbicara untuk setiap orang berbeda, *noise* pada citra yang disebabkan oleh warna bibir dan kuat penerangan ruangan yang tidak merata.

Fitur yang digunakan untuk mewakili citra dinamis yang berubah karena perubahan pola gerak bibir dari satu *frame* ke *frame* yang lain untuk selanjutnya harus diteliti sehingga diketahui seberapa signifikan fitur tersebut bisa membedakan pola saat mengucapkan fonem yang berbeda-beda. Perangkat lunak pengenalan pola gerak bibir juga bisa dikembangkan dengan melakukan penelitian tentang metode segmentasi bibir yang lebih handal terhadap *noise* yang timbul.

UCAPAN TERIMA KASIH

Penelitian ini dibiayai oleh DIKS, Fakultas Teknik, Universitas Gadjah Mada. Ucapan terimakasih kepada Denish

Zuli Anugra, ST. dan Dian Nandiwardhana, ST. untuk bantuan dalam penelitian.

REFERENSI

- [1] Faridah, Utami SS, Wibowo S, Wijaya E, "Alat Terapi Berbicara untuk Penderita Tuna Rungu di Indonesia", *Media Teknik*, vol. 2, Mei 2008.
- [2] Meier U, Huerst W, Duchnowski P, "Adaptive Bimodal Sensor Fusion for Automatic Speechreading", *Proceeding of the International Conference on Acoustics, Speech, and Signal Processing ICASSP'96*, 1996.
- [3] Mase K, Pentland A, "Automatic Lipreading by Optical Flow Analysis", *Systems and Computer in Japan*, 22(6):67-76,1991
- [4] Petajan ED, *Automatic Lipreading to Enhance Speech Recognition*, PhD thesis, University of Illinois at Urbana-Champaign, 1984.
- [5] Kass M, Witkin A, Terzopoulos D, "Snakes : Active Contour Models", *International Journal on Computer Vision*, 1(4):321-331, 1988.
- [6] Delmas P, Coulon PY, Fristot V, "Automatic Snakes for Robust Lip Boundaries Extraction", *IEEE International Conference on Acoustic, Speech, and Signal Processing (ICASSP'99)*, Phoenix, USA, 1999.
- [7] Terzopoulos D, Waters K, "Analysis and Synthesis of Facial Image Sequences Using Physical and Anatomical Models", *IEEE Trans Pattern Analysis and Machine Intelligence*, 15(6):569-579, 1993.
- [8] Zhang X, Mersereau RM, "Lip Feature Extraction Towards an Automatic Speechreading System", *International Conference on Image Processing (ICIP'00)*, Vancouver, September 2000.
- [9] Eveno N, Caplier A, Coulon PY, "A New Color Transformation For Lips Segmentation", *IEEE Workshop on Multimedia Signal Processing (MMSp'01)*, Cannes, France, 2001.
- [10] Eveno N, Caplier A, Coulon PY, "A Parametric Model for Realistic Lip Segmentation", *International Conference on Control, Automation, Robotics and Vision (ICARCV)*, Singapore, 2002
- [11] Hulbert A, Poggio T, "Synthesizing a Colour Algorithm from Examples", *Science*, 239:482-485, 1998.

ISSN 2089-7154

TEKNOFISIKA

JURNAL ILMIAH BIDANG TEKNIK FISIKA

Vol.1, No.2, September 2012

PENGUKURAN KEKUATAN DAN SUDUT ELEVASI PANCARAN ANTENA *OMNIDIRECTIONAL*

Sunarno

RANCANG BANGUN SISTEM SIMULASI DENYUT NADI PERGELANGAN TANGAN PADA *VITAL SIGN SIMULATOR* SEBAGAI MEDIA PEMBELAJARAN KETERAMPILAN MEDIK MAHASISWA KEDOKTERAN

A. Wiraswan, Sunarno , R. Wijaya , R. N. Hidayah

RANCANG BANGUN SISTEM SIMULASI PEMERIKSAAN SUHU TUBUH PADA *VITAL SIGN SIMULATOR* UNTUK PELATIHAN KETERAMPILAN MEDIK MAHASISWA KEDOKTERAN

D. Hardiyanto, Sunarno , R. Wijaya , R. N. Hidayah

DESAIN KONSEPTUAL *SAFETY INSTRUMENTED SYSTEM* PADA PROSES PENGGILINGAN BATUBARA

R. Astuti, K. Suryopratomo, A. N. I. Wardana

PENGARUH BENTUK DAN CACAH KUBU SIRIP PADA ALIRAN DI DALAM PIPA

A. R. Ritonga, K. Suryopratomo, Sihana

PENGENALAN POLA GERAK BIBIR DALAM PENGUCAPAN FONEM VOKAL BAHASA INDONESIA

Faridah, N. Effendy

PROYEKSI PERMINTAAN LISTRIK SEKTOR RUMAH TANGGA DENGAN MENGGUNAKAN *MODEL END-USE*

A. Darmawan, M. K. Ridwan, A. A. Setiawan

TEKNOFISIKA

Vol.1

No.2

September
2012

HALAMAN
57-104

JURUSAN
TEKNIK FISIKA
UGM

ISSN
2089-7154

TEKNOFISIKA

JURNAL ILMIAH BIDANG TEKNIK FISIKA

ISSN 2089-7154

Vol.1, No.2, September 2012

Periode Terbit

Januari, Mei, September

Terbit Pertama

Mei 2012

DEWAN REDAKSI

Penanggung Jawab

Prof. Ir. Sunarno, M.Eng., Ph.D.

Dewan Penyunting

Dr.-Ing. Awang N. I. Wardana, ST., MT., M.Sc. (Ketua)

Dr. Gea O. F. Parikesit, ST., M.Sc.

Ir. Kutut Suryopratomo, MT., Msc.

Dr.-Ing. Sihana

Mitra Bestari

Dr. Khasani ST., M.Eng. (Teknik Mesin FT UGM)

drg. Rinaldi Budi Utomo, M.S., SpKGA(K) (Fakultas Kedokteran Gigi UGM)

Dr. Ir. Risanuri Hidayat, M.Sc. (Teknik Elektro dan Teknologi Informasi FT UGM)

Dr. Ir. Sri Ratna Sulistiyanti, M.T. (Universitas Lampung)

Dr. Ir. T. Haryono, M. Sc., CEng MIET (Teknik Elektro dan Teknologi Informasi FT UGM)

dr. Widyandana, MHPE, Ph.D. (Fakultas Kedokteran UGM)

Sekretariat Redaksi

Memory M. Waruwu, ST.

Alamat Redaksi :

JURUSAN TEKNIK FISIKA

FAKULTAS TEKNIK - UNIVERSITAS GADJAH MADA

Jl. Grafika No. 2 Yogyakarta 55281

Telp./Fax. : 0274-580882

Website : <http://tf.ugm.ac.id/teknofisika/>

e-Mail : teknofisika@yahoo.com

DAFTAR ISI

	halaman
1. Pengukuran Kekuatan dan Sudut Elevasi Pancaran Antena <i>Omnidirectional</i> <i>Sunarno</i>	(57-64)
2. Rancang Bangun Sistem Simulasi Denyut Nadi Pergelangan Tangan pada <i>Vital Sign Simulator</i> Sebagai Media Pembelajaran Keterampilan Medik Mahasiswa Kedokteran <i>Aditya Wiraswan, Sunarno, Rony Wijaya, Rachmadya Nur Hidayah</i>	(65-72)
3. Rancang Bangun Sistem Simulasi Pemeriksaan Suhu Tubuh pada <i>Vital Sign Simulator</i> untuk Pelatihan Keterampilan Medik Mahasiswa Kedokteran <i>Denny Hardiyanto, Sunarno, Rony Wijaya, Rachmadya Nur Hidayah</i>	(73-79)
4. Desain Konseptual <i>Safety Instrumented System</i> pada Proses Penggilingan Batubara <i>Rini Astuti, Kutut Suryopratomo, Awang N. I. Wardana</i>	(80-89)
5. Pengaruh Bentuk dan Cacah Kubu Sirip pada Aliran di Dalam Pipa <i>Anwar Ramadhan Ritonga, Kutut Suryopratomo, Sihana</i>	(90-95)
6. Pengenalan Pola Gerak Bibir Dalam Pengucapan Fonem Vokal Bahasa Indonesia <i>Faridah, Nazrul Effendy</i>	(96-100)
7. Proyeksi Permintaan Listrik Sektor Rumah Tangga dengan Menggunakan <i>Model End-Use</i> <i>Arif Darmawan, M. Kholid Ridwan, Ahmad Agus Setiawan</i>	(101-104)

Diterbitkan oleh :

JURUSAN TEKNIK FISIKA
FAKULTAS TEKNIK - UNIVERSITAS GADJAH MADA
Jl. Grafika No. 2 Yogyakarta 55281
Telp./Fax. : 0274-580882
Website : <http://tf.ugm.ac.id/teknofisika/>
e-Mail : teknofisika@yahoo.com