

SISTEM IDENTIFIKASI PEMBICARA BERBASIS POWER SPEKTRUM MENGGUNAKAN *HIDDEN MARKOV MODEL*

Agus Buono¹⁾ Wisnu Jatmiko²⁾ Benyamin Kusumoputro³⁾

¹⁾ Departemen Ilmu Komputer FMIPA IPB
Kampus IPB Darmaga-Bogor
email : pudesha@yahoo.co.id

²⁾ Fakultas Ilmu Komputer Universitas Indonesia
Fakultas Ilmu Komputer Kampus UI Depok

³⁾ Fakultas Teknik Universitas Indonesia
Fakultas Teknik Kampus UI Depok
email : nynykusumo@yahoo.com

ABSTRACT

Pada paper ini, disajikan suatu penerapan model HMM sebagai pengenalan pola pada sistem identifikasi pembicara (SIP) dengan ekstraksi ciri menggunakan teknik MFCC yang berbasis nilai power spektrum dari suara. SIP yang dikembangkan bersifat text dependent dan melibatkan 10 pembicara yang mengucapkan frase "pudesha" sebanyak 80 kali tanpa pengkondisian. Untuk melihat efektifitas sistem, dilakukan uji coba baik pada suara tanpa penambahan noise maupun dengan penambahan noise yang bersifat Gaussian pada level 20 hingga 0 dB. Hasil percobaan menunjukkan bahwa teknik HMM dengan jumlah data training yang mencukupi mampu melakukan pengenalan dengan baik (99%) untuk data tanpa penambahan noise. Namun demikian, untuk data dengan penambahan noise (meskipun hanya 20 dB), akurasi sistem drop hingga jauh di bawah 50%. Hasil percobaan menunjukkan bahwa kegagalan ini lebih disebabkan oleh nilai power spektrum sebagai masukan MFCC yang bersifat sensitif terhadap noise, dan bukan pada HMM-nya.

Key words

Sistem identifikasi pembicara, MFCC, HMM, Codebook, SOFM.

1. Pendahuluan

Sistem Pengenalan Pembicara, *Automatic Speaker identification* (ASI), adalah salah satu sistem pengenalan suara yang mengidentifikasi orang atau dari kelompok apa orang tersebut berasal berdasar suara tanpa adanya klaim sebelumnya mengenai orang tersebut, [1]. Dalam perkembangannya, teknik

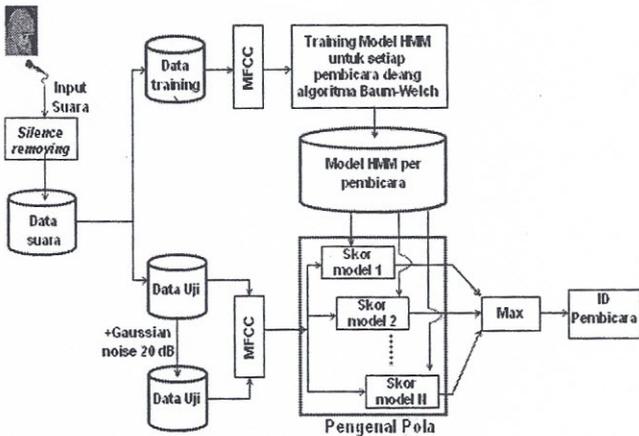
pemodelan suara yang banyak dikaji adalah yang berbasis teori peluang. Satu teknik yang telah menunjukkan efektifitas yang baik dalam merepresentasikan suara adalah HMM (*Hidden Markov Model*), seperti disajikan pada [2].

Dari aspek ekstraksi ciri, *Mel-Frequency Cepstrum Coefficients* (MFCC) merupakan teknik yang telah luas dipakai pada pemrosesan sinyal suara, terutama pada pengenalan pembicara. Penggunaan teknik ini pada sistem pemrosesan sinyal memberikan pengenalan yang lebih baik dibandingkan dengan metode lainnya, Davis and Mermelstein (dalam [3]).

Paper ini disajikan dengan susunan sebagai berikut : Bagian 2 mengenai prinsip sistem identifikasi pembicara. Teknik MFCC untuk ekstraksi ciri pada bagian 3. Bagian 4 menyajikan data, rancangan dan hasil percobaan, dan sebagai penutup adalah kesimpulan dan saran untuk penelitian selanjutnya yang disajikan pada bagian 5.

2. Prinsip Sistem Identifikasi Pembicara

Secara umum, sistem identifikasi pembicara terdiri dari dua subsistem, yaitu subsistem ekstraksi ciri dan subsistem pencocokan pola, seperti disajikan pada **Gambar 1**. Subsistem ekstraksi ciri melakukan proses transformasi sinyal input ke dalam satu set vektor ciri sebagai representasi dari sinyal suara suatu pembicara untuk proses selanjutnya. Subsistem pencocokan pola merupakan bagian untuk melakukan identifikasi suatu pembicara yang belum diketahui dengan cara membandingkan sinyal suaranya yang telah diekstrak ke dalam vektor ciri dengan set vektor ciri dari pembicara yang telah diketahui dan tersimpan dalam sistem, [4].



Gambar 1. Blok diagram sistem identifikasi pembicara dengan HMM sebagai pengenalan pola

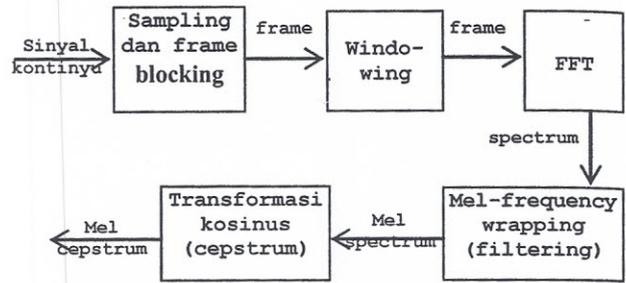
Dari aspek pengembangan sistem, ada dua fase pada sistem identifikasi pembicara. Fase pertama adalah tahap pelatihan. Pada fase ini sistem melakukan pelatihan untuk menentukan parameter model untuk setiap pembicara berdasar data suara pembicara tersebut. Pada penelitian ini suara setiap pembicara dimodelkan dengan menggunakan Hidden Markov Model (HMM) yang dilatih menggunakan algoritma Baum Welch seperti yang disajikan pada [2].

Fase kedua adalah tahapan pengujian, yaitu sinyal input yang diberikan kepada sistem dicocokkan dengan dengan model setiap pembicara yang ada pada sistem. Keputusan untuk menentukan pembicara didasarkan pada skor tertinggi untuk setiap model, yang dihitung dengan algoritma Forward, [2].

3. Mel-Frequency Cepstrum Coefficients

Sinyal suara adalah sinyal yang berubah terhadap waktu secara perlahan (*quasi-stationary*). Untuk rentang waktu yang pendek (antara 5 and 100 milliseconds), karakteristik sinyal ini bisa dianggap stasioner. Namun demikian untuk untuk periode yang cukup panjang (1/5 seconds atau lebih) karakteristik sinyal berubah yang mencerminkan perbedaan bunyi yang diucapkan. Oleh karena itu, *short-time spectral analysis* adalah yang biasa dilakukan untuk mengkarakterisasi sinyal suara, [4].

MFCC adalah teknik ekstraksi ciri yang populer dan paling banyak digunakan. Block diagram teknik MFCC ini disajikan pada Gambar 2, [4]. **Frame blocking:** sinyal suara dibaca per blok (frame) yang terdiri dari N sample. Antara dua frame yang bersisihan terdapat overlap N-M sample, dengan M adalah banyaknya pergeseran antar frame (M<N).



Gambar 2. Block diagram teknik MFCC

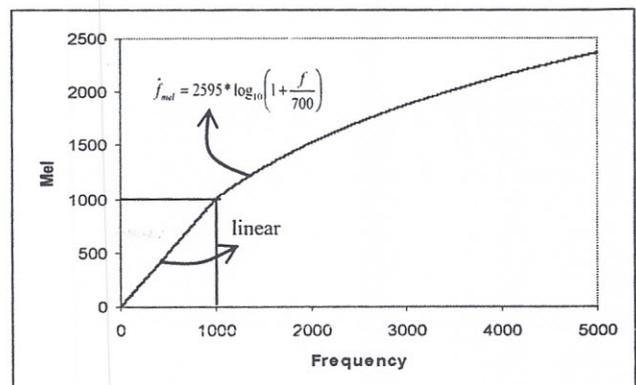
Windowing: proses windowing dilakukan pada setiap frame dengan tujuan untuk meminimumkan diskontinuitas antar sua frame, khususnya pada bagian awal dan akhir.

FFT: Pada tahap ini setiap frame yang terdiri dari N samples dikonversi dari domain waktu ke domain frekuensi. Output dari proses ini disebut dengan nama spektrum atau periodogram.

Mel-Frequency wrapping: tahap ini merupakan proses pengfilteran dari spektrum setiap frame yang diperoleh dari tahapan sebelumnya. Filter tersebut berupa M filter segitiga sama tinggi dengan tinggi satu. Filter ini dibuat dengan mengikuti persepsi telinga manusia dalam menerima suara. Persepsi ini dinyatakan dalam skala 'mel' (berasal dari Melody) yang mempunyai hubungan tidak linear dengan frekuensi suara, [4]. Dalam hal ini skala mel-frequency adalah linear untuk frekuensi kurang dari 1000 Hz dan logaritmik untuk frekuensi di atas 1000 Hz. Satu relasi antara frekuensi bunyi (dalam Hz) dengan skala mel adalah, [4], [5] :

$$\hat{f}_{mel} = 2595 * \log_{10} \left(1 + \frac{f}{700} \right) \dots \dots \dots (1)$$

Grafik relasi di atas disajikan pada Gambar 3.



Gambar 3. Hubungan antara frekuensi dengan skala mel

Berdasar hubungan tersebut, maka algoritma untuk membuat M filter adalah seperti pada algoritma 1. Sedangkan Gambar 4 menyajikan contoh sebuah filter segitiga ke i.

Algoritma 1: Mengkonstruksi M filter

- a. Pilih jumlah filter yang akan dibuat (M)
- b. Pilih frekuensi terbesar (f_{high}). Dari

nilai ini, maka nilai tertinggi dari \hat{f}_{mel}

$$\text{adalah } \hat{f}_{mel}^{high} = 2595 * \log_{10} \left(1 + \frac{f_{high}}{700} \right)$$

- c. Pusat filter ke i adalah f_i :

c.1. $f_i = \frac{1000}{0.5 * M} * i$ untuk $i=1, 2, 3, \dots, M/2$

c.2. untuk $i=M/2, M/2+1, \dots, M$, maka f_i dihitung dengan prosedur berikut:

- 1. skala mel disekat dengan lebar yang sama, yaitu sebesar Δ , dengan:

$$\Delta = \frac{\hat{f}_{mel}^{high} - 1000}{0.5 * M}$$

Dengan persamaan (1), maka nilai Δ dapat dirumsukan sebagai :

$$\Delta = \frac{5190}{M} \log \left(\frac{700 + f_{high}}{1700} \right)$$

- 2. nilai mel untuk pusat filter ke i adalah:

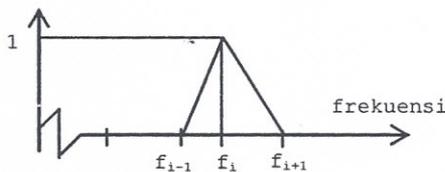
$$a = 1000 + (i - 0.5 * M) * \Delta$$

- 3. pusat dari filter ke i adalah:

$$f_i = 700 * (10^{a/2595} - 1)$$

Nilai koefisien spektrum mel (*mel spectrum coefficients*) adalah jumlah dari hasil pengfilteran sesuai rumus 2, [5]:

$$X_i = \log \left(\sum_{j=0}^{N-1} \text{abs}(X(j)) * H_i(f) \right) \dots \dots \dots (2)$$



Gambar 4. Suatu filter segitiga ke i dengan tinggi 1

Dengan $i=1,2,3,\dots,M$; M adalah banyaknya filter; N adalah banyaknya koefisien FFT; $\text{abs}(X(j))$ adalah magnitude koefisien ke j dari periodogram yang dihasilkan oleh transformasi FFT; dan $H_i(f)$ adalah nilai filter ke i pada titik frekuensi adalah f.

Cepstrum: Pada tahap ini dilakukan konversi dari koefisien spektrum mel kembali ke domain waktu menggunakan transformasi kosinus sesuai rumus 3.

$$C_j = \sum_{i=1}^M X_i * \cos \left(\frac{j * (i - 0.5) * \pi}{20} \right) \dots \dots \dots (3)$$

Dengan $j=1,2,3,\dots,K$; K adalah banyaknya koefisien; M adalah banyaknya filter segitiga; X_i adalah koefisien spektrum mel yang diperoleh dengan formula (2). Dalam hal ini C_j disebut sebagai mel frequency cepstrum coefficients (MFCC) koefisien ke j.

4. Data, Struktur HMM dan Hasil Percobaan

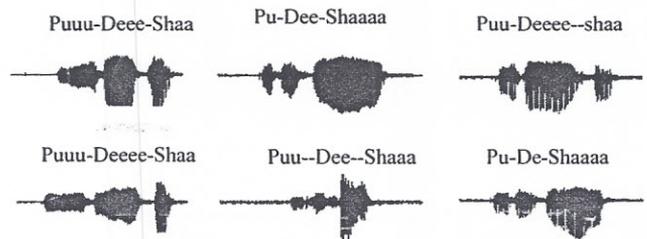
4.1. Data Percobaan

Percobaan pada penelitian menggunakan data yang berasal dari 10 pembicara yang mengucapkan kata "PUDESHA" sebanyak 80 kali tanpa dikontrol cara pengucapannya. Hal ini dengan maksud untuk menunjukkan bahwa model yang dihasilkan mampu bekerja pada suara dengan berbagai cara pengucapan (keras-lemah, panjang-pendek, maupun tinggi-rendah). Karakteristik ke sepuluh pembicara tersebut adalah seperti disajikan pada Tabel 1.

Tabel 1. Daftar 10 Pembicara yang Digunakan dalam Penelitian

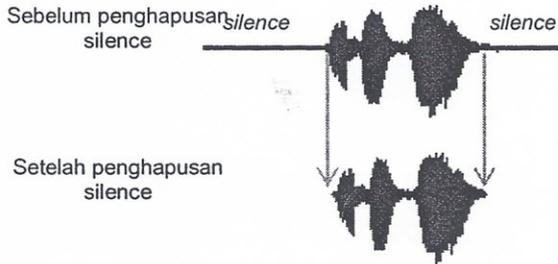
Pembicara	Jenis Kelamin	Umur (th)	Pekerjaan	Suku
1	Pria	40	PNS	Jawa
2	Pria	32	Karyawan	Sunda
3	Pria	23	Mhsw	Jawa
4	Pria	37	PNS	Batak
5	Pria	37	PNS	Jawa
6	Pria	30	PNS	Sunda
7	Pria	23	MHSW	Jawa
8	Wanita	9	Pelajar	Jawa-Minang
9	Wanita	6	Pelajar	Jawa-Minang
10	Pria	23	MHSW	Jawa-Sunda

Setiap suara disampling dengan frekuensi 1.1 kHz dan durasi 1.28 detik (sehingga diperoleh sinyal digital dengan panjang 14080) yang dilakukan dalam ruang kerja yang tertutup. Gambar 5 adalah contoh variasi cara pengucapan dari seorang pembicara tertentu yang digunakan dalam model.



Gambar 5. Beberapa Sinyal Sesuai Cara Pengucapan dari Seorang Pembicara

Pada **Gambar 5** terlihat bahwa dari satu durasi perekaman, ada bagian atau segmen yang tidak ada suara (*silence*), yang umumnya terjadi dibagian depan dan belakang. Untuk meningkatkan akurasi sistem, maka sebelum memasuki tahap ekstraksi, dilakukan proses penghapusan bagian *silence* dengan menggunakan batas sebesar rata-rata absolut bagian *silence* ditambah 1,96 dikali standar deviasinya. **Gambar 6** memberikan input dan output proses penghapusan bagian *silence*.

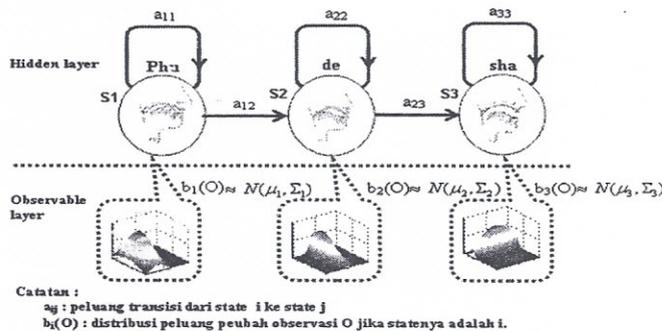


Gambar 6 Ilustrasi Penghapusan Bagian Silence

Setelah dilakukan penghapusan bagian *silence*, dilakukan standarisasi sinyal suara dengan membagi sinyal dengan nilai mutlak simpangan maksimum, sehingga akan diperoleh sinyal dengan simpangan maksimum 1 atau -1. Selanjutnya dilakukan penambahan Gaussian noise dengan level 20 hingga 0 dB. Dengan demikian, pada akhirnya diperoleh beberapa set sinyal yang terdiri 800 data sinyal per set, masing-masing untuk sinyal asli dan sinyal asli dengan penambahan noise.

4.2. Struktur HMM

Pada penelitian ini, identifikasi pembicara didasarkan pada ujaran yang bersifat tetap, yaitu frase "PUDESHA" dan vektor pengamatan pada setiap periode t, O_t , adalah kontinyu yang didekati dengan distribusi Gaussian, maka model left-right HMM dengan distribusi *emitten* bersifat Gaussian akan lebih sesuai. Bentuk model ini disajikan pada **Gambar 7**.



Gambar 7. Model Left-Right HMM dengan Tiga State yang Dipergunakan pada Penelitian

Oleh karenanya, HMM di atas dituliskan sebagai :

$$\lambda = \left(A = \begin{bmatrix} a_{11} & a_{12} & 0 \\ 0 & a_{22} & a_{23} \\ 0 & 0 & a_{33} \end{bmatrix}, B = \begin{bmatrix} N(\mu_1, \Sigma_1) \\ N(\mu_2, \Sigma_2) \\ N(\mu_3, \Sigma_3) \end{bmatrix}, \Pi = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \right)$$

Seperti telah disebutkan sebelumnya, data yang digunakan berasal dari 10 pembicara masing-masing 80 pengucapan. Dalam hal ini dicobakan tiga set perbandingan jumlah data pelatihan dengan jumlah data pengujian, yaitu 20:60, 40:40, dan 60:20. Ada tiga tahapan dalam percobaan ini, yaitu tahap ekstraksi ciri, tahap pelatihan model dan tahapan pengenalan. Pada tahap ekstraksi ciri, semua ucapan dari setiap pembicara dihitung nilai cirinya yang dibaca frame demi frame dengan panjang 256 dan overlap antar frame 156 menggunakan teknik MFCC. Berikutnya data yang telah dihitung cirinya dikelompokkan menjadi dua, satu untuk pelatihan dan sisanya untuk pengujian (pengenalan) sesuai proporsi di atas. Selanjutnya adalah tahap pelatihan. Pada tahap ini dilakukan pendugaan parameter HMM dengan algoritma Baum Welch [2] dari set data pelatihan. Setelah model diperoleh, dilanjutkan tahap pengenalan pembicara untuk set data pengujian yang sudah ditentukan.

4.3. Hasil Percobaan

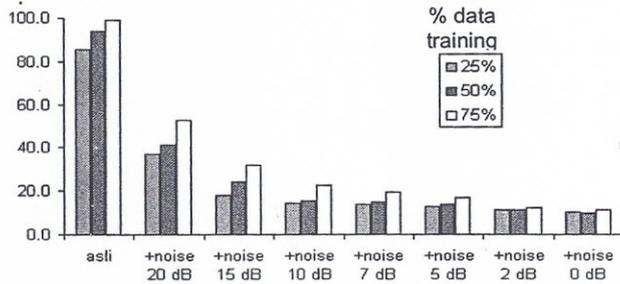
Hasil pengenalan pembicara dengan metode MFCC+HMM untuk sinyal asli disajikan pada **Tabel 2**. Dari tabel tersebut terlihat bahwa teknik yang dipakai berhasil dengan baik mengenali pembicara, yaitu sekitar 99% untuk data asli pada proporsi data training 75%. Sedangkan jika proporsi data training berkurang menjadi 50% dan 25%, maka akurasi turun cukup berarti, yaitu menjadi 93.8% dan 85.5%.

Tabel 2. Akurasi (%) Sistem dengan Teknik MFCC+HMM pada Berbagai Jenis Data Uji

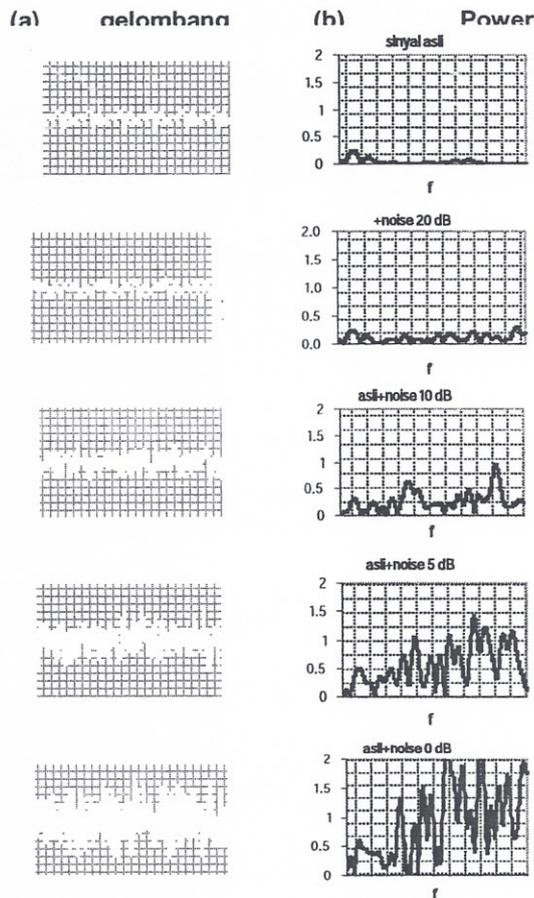
Jenis Data Uji	% data training		
	25%	50%	75%
Asli	85.5	93.8	99.0
Asli+noise 20 dB	37.0	41.1	52.8
Asli+noise 15 dB	18.0	24.3	32.0
Asli+noise 10 dB	14.4	15.4	22.5
Asli+noise 7 dB	13.9	14.9	19.5
Asli+noise 5 dB	12.7	13.8	17.3
Asli+noise 2 dB	11.2	11.2	12.3
Asli+noise 0 dB	10.4	10.0	11.3

Hal ini menunjukkan bahwa untuk memberikan hasil yang optimum, maka jumlah teknik HMM memerlukan jumlah data pelatihan yang mencukupi. Dengan 60 data pelatihan dengan setiap suara mempunyai panjang rata-rata sekitar 50 frame, maka panjang total barisan yang dipergunakan

untuk pelatihan adalah sekitar 3000 frame. Dari tabel tersebut juga terlihat bahwa dengan bertambahnya noise, akurasi turun secara drastis, yaitu menjadi 52% untuk noise 20 dB, dan untuk noise yang lebih tinggi, akurasi jauh di bawah 50%. Hal tersebut terlihat jelas secara visual seperti yang ditunjukkan pada Gambar 8.



Gambar 8. Akurasi Sistem (%) Dengan Teknik MFCC+HMM untuk Berbagai Data Uji pada Berbagai Noise



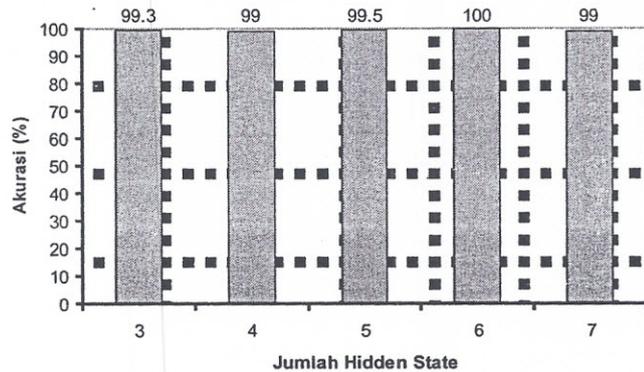
Gambar 9. Perbandingan Pola Power Spektrum antara Sinyal Asli dengan Sinyal Asli yang Ditambah Noise

Gagalnya sistem ini disebabkan oleh nilai power spektrum yang sensitif terhadap gangguan noise, seperti ditunjukkan pada Gambar 9. Bagian kiri, (a), menunjukkan dengan jelas bahwa bentuk gelombang

sinyal suara mengalami perubahan dengan adanya tambahan noise mulai dari terkecil 20 dB hingga terbesar 0 dB. Meskipun, secara alamiah ciri sinyal asli dengan sinyal asli yang sudah ditambah noise tidaklah berubah, namun pola power spektrumnya telah mengalami perubahan. Perubahan ini sudah mulai terlihat pada noise 20 dB, dan secara nyata pada noise 10 dB

Fakta empiris yang ditunjukkan pada Gambar 9 mengatakan bahwa power spektrum sebagai penentu ciri sinyal suara sangat sensitif terhadap gangguan noise. Dengan demikian, maka nilai feature yang diperoleh dari ekstraksi ciri telah mengalami distorsi yang cukup besar dari sinyal aslinya, yang pada akhirnya akan menurunkan akurasi sistem.

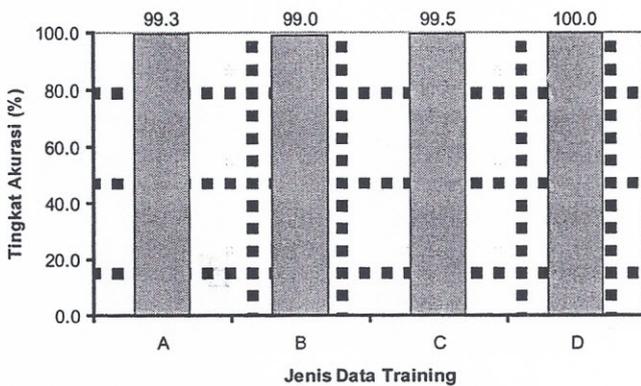
Sedangkan kalau diperhatikan dari jumlah hidden state dari model HMM, terlihat adanya perbedaan tingkat akurasi dari jumlah hidden state 3 hingga 7, dengan nilai tertinggi pada jumlah hidden state 6, yaitu mencapai 100% (untuk data asli), seperti ditunjukkan pada Gambar 10. Meskipun demikian, secara absolut, nilai akurasi dari setiap jumlah hidden state tersebut tidaklah terlalu berbeda nyata (semuanya di atas 99%). Hal ini menunjukkan bahwa pemilihan jumlah hidden state pada HMM tidaklah memberikan pengaruh yang nyata terhadap hasil akurasi sistem.



Gambar 10. Perbandingan Akurasi pada Berbagai Jumlah Hidden State

Dari aspek pemilihan data training, ada hal yang cukup menarik adalah adanya fakta bahwa pemilihan ucapan sebagai data pelatihan secara acak akan memberikan akurasi sistem yang lebih baik, seperti disajikan pada Gambar 11. Dari gambar tersebut terlihat bahwa sistem yang dikembangkan dengan data training yang dipilih dari suara yang secara relatif berada jauh dari pusat distribusinya akan memberikan akurasi yang rendah, (B). Untuk data training berasal dari suara yang menyebar di sekitar pusat distribusi, (A), juga kurang memberikan hasil yang bagus. Kalau dipilih suara secara sistematis dari yang dekat hingga yang terjauh dari pusat distribusi suara (C), maka akurasi lebih tinggi dibanding (A) dan (B). Hasil tertinggi adalah kalau data training dipilih secara

acak, seperti yang ditunjukkan pada (D), yaitu hampir mencapai 100%. Meskipun demikian, dari nilai absolut akurasi tersebut masih cukup baik, yaitu untuk semuanya tidak kurang dari 99%.



Gambar 11. Perbandingan Akurasi dari Empat Jenis Data Training

5. Kesimpulan dan Saran

Beberapa fakta yang diperoleh pada bagian ini adalah :

1. Teknik MFCC dan HMM dapat diterapkan untuk pengenalan pola suara dengan baik (akurasi di atas 99%) untuk data tanpa penambahan noise.
2. Secara empiris terlihat bahwa perbedaan jumlah hidden state pada HMM tidak memberikan pengaruh yang berarti bagi kinerja sistem. Jumlah hidden state sebanyak 3 telah mampu memberikan akurasi di atas 99%.
3. Akurasi sistem dipengaruhi oleh jenis data suara yang dipakai sebagai data pelatihan. Pemilihan data untuk pelatihan secara acak akan memberikan hasil pengenalan yang lebih baik dibanding kalau data pelatihan dipilih tidak acak. Namun demikian, perbedaan nilai akurasi yang diberikan secara absolut tidak berarti, semuanya di atas 99%.
4. Jumlah data training memberikan perbedaan akurasi yang signifikan, yaitu 99%, 92% dan 89% untuk sinyal tanpa penambahan noise, masing-masing untuk jumlah data training 60 suara (75%), 40 suara (50%), dan 20 suara (25%).
5. Akurasi sistem turun menjadi sekitar 50% untuk data suara bernoise 20 dB. Untuk noise yang lebih besar (10 dB hingga 0 dB), akurasi turun drastis di bawah 50% hingga mencapai sekitar 10%.
6. Kegagalan sistem MFCC+HMM untuk sinyal dengan penambahan noise lebih dikarenakan nilai power spektrum sebagai penciri sinyal yang merupakan input bagi ekstraksi ciri MFCC bersifat sangat sensitif terhadap noise.

Dari hasil yang diperoleh di atas, maka untuk penelitian selanjutnya perlu dilakukan kajian yang lebih mendalam terkait dengan noise. Solusi terhadap permasalahan noise dapat dilakukan dari 3 pendekatan, yaitu pada bagian praproses melalui *denoising*, pada tahapan ekstraksi dengan menggunakan parameter yang bersifat *robust* terhadap noise, dan pada bagian pengenalan dengan mengembangkan model HMM atau model lainnya yang mengakomodasi masalah noise.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih yang setinggi-tingginya kepada Departemen Ilmu Komputer IPB atas ijin penggunaan laboratorium komputer untuk melaksanakan percobaan ini.

REFERENSI

- [1] J. Campbell, "Speaker Recognition: A Tutorial", Proc. of the IEEE, Vol 85, No. 9, pp 1437-1462, 1997.
- [2] L.R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition", Proceeding IEEE, Vol 77 No. 2, pp 257-289, 1989.
- [3] Todor D. Ganchev. *Speaker Recognition*. PhD Dissertation, Wire Communications Laboratory, Department of Computer and Electrical Engineering, University of Patras Greece. 2005
- [4] Cornaz, C. dan U. Hunkeler. An Automatic Speaker Recognition System. Mini-Project. http://www.ifp.uiuc.edu/~minhdo/teaching/speaker_recognition, access : August, 15, 2005.
- [5] M. Nilsson dan M. Ejarsson. *Speech Recognition using Hidden Markov Model : Performance Evaluation in Noisy Environment*. Master Thesis, Departement of Telecommunications and Signal Processing, Blekinge Institute of Technology, Maret 2002.

Agus Buono, memperoleh gelar Sarjana dan Master bidang statistik di IPB pada tahun 1992 dan 1996. Gelar Master dan Doktor bidang Ilmu Komputer diperoleh dari Universitas Indonesia pada tahun 2000 dan 2009. Saat ini sebagai Staf Pengajar Departemen Ilmu Komputer Institut Pertanian Bogor.

Benyamin Kusumoputro, memperoleh gelar Sarjana bidang fisika dari Institut Teknologi Bandung dan Doktor Optoelektronika dari Tokyo Institute of Technology-Jepang. Gelar Profesor diperoleh pada tahun 2002 dari Universitas Indonesia. Saat ini sebagai Staf Pengajar Fakultas Teknik Universitas Indonesia.

Wisnu Jatmiko, memperoleh gelar Sarjana Elektro dan Magister Ilmu Komputer dari Universitas Indonesia. Ph.D bidang komputer diperoleh dari Jepang pada tahun 2008. Saat ini sebagai Dosen Fakultas Ilmu Komputer Universitas Indonesia.