

# PERBANDINGAN METODE *HIDDEN MARKOV MODEL* DAN *VECTOR QUANTIZATION* UNTUK APLIKASI IDENTIFIKASI SUARA

M. G. J. Harry Khesa S<sup>1</sup>, W. Setiawan<sup>2</sup>, I.G.A.K. Diafari Djuni H<sup>3</sup>

<sup>1,2,3</sup> Jurusan Teknik Elektro dan Komputer, Fakultas Teknik Universitas Udayana

Email: [jayaharrykhesa@yahoo.co.id](mailto:jayaharrykhesa@yahoo.co.id)<sup>1</sup>, [widyadi@ee.unud.ac.id](mailto:widyadi@ee.unud.ac.id)<sup>2</sup>, [igakdiafari@yahoo.com](mailto:igakdiafari@yahoo.com)<sup>3</sup>

## Abstrak

Identifikasi suara merupakan suatu cara yang bisa digunakan untuk mengetahui perbedaan dari masing – masing individu seperti halnya *fingerpint scan*, *retinal scan*, atau *face recognition*. Diperlukan suatu aplikasi untuk memudahkan dalam proses identifikasi. Metode yang biasanya digunakan untuk *recognition* pada Aplikasi Identifikasi Suara adalah *Hidden Markov Model* atau *Vector Quantization*. Kedua metode pengenalan untuk Aplikasi Identifikasi Suara ini akan dibandingkan unjuk kerjanya dalam kondisi ideal maupun tidak ideal. Aplikasi identifikasi suara dengan metode *Vector Quantization* mempunyai nilai unjuk kerja lebih baik dengan hasil unjuk kerja sebesar 93% dibandingkan metode *Hidden Markov Model* yang sebesar 85% pada kondisi ideal dan 78% dibanding 69,5% pada kondisi tidak ideal, hal ini dikarenakan, pada metode *Vector Quantization* menggunakan layer tunggal dan tidak adanya layer tersembunyi seperti pada metode *Hidden Markov Model*.

**Kata Kunci:** Aplikasi Identifikasi Suara, *Vector Quantization*, *Hidden Markov Model*

## Abstract

*Voice identification is a way that can be used to determine the differences of each - each individual as well as fingerprint scan, retinal scan, or face recognition. Needed an application to facilitate the identification process. The method is usually used for the Application Identification Voice recognition is a Hidden Markov Model or Vector quantization. The second method for the introduction of Voice Identification Application will be compared to its performance in ideal conditions and not ideal. Voice identification applications with Vector Quantization method has a better performance value with the performance results of 93% compared to the method of Hidden Markov Models which amounted to 85% in ideal conditions and 78% versus 69.5% in conditions not ideal, this is because, in Vector quantization method uses a single layer and no hidden layer as a method of hidden Markov Model.*

**Keywords:** *Voice Identification Application, Vector Quantization, Hidden Markov Model*

## 1. PENDAHULUAN

Suara merupakan gabungan berbagai sinyal, tetapi suara murni secara teoritis dapat dijelaskan dengan kecepatan osilasi atau frekuensi yang diukur dalam *Hertz* (Hz) dan amplitudo atau kenyaringan bunyi dengan pengukuran dalam *decibel* [1]. Suara dihasilkan melalui proses *Generation* dan *Filtering*. Diproduksinya suara melalui bergetarnya pita suara (*vocal cord*) yang berada di *larynx* untuk menghasilkan bunyi periodik disebut *Generation*. *Vocal tract* yang terdiri dari lidah, gigi, bibir, langit-langit mulut dan lain-lain menghasilkan bunyi periodik yang bersifat konstan kemudian difilterisasi melalui sehingga bunyi tersebut dapat menjadi bunyi keluaran yaitu bunyi vokal

dan/atau bunyi konsonan yang membentuk kalimat yang memiliki arti untuk *speech recognition*.

Salah satu cara yang bisa digunakan untuk mengetahui perbedaan dari masing – masing individu adalah identifikasi suara. Seperti halnya *fingerpint recognition* (identifikasi pola sidik jari pada setiap orang), *retinal scan* (identifikasi berdasarkan pola pembuluh darah pada retina mata), *face recognition* (pengenalan seseorang berdasarkan raut dan ekspresi seseorang dengan kunci utama pada letak mata dan mulut) [1]. Tidak semua pendengaran manusia mampu membedakan suara dari masing-masing individu yang dikenalnya dalam identifikasi terhadap suara manusia,. Kepekaan telinga juga memiliki keterbatasan

dan sensitif terhadap berbagai macam suara. Oleh karena itu dibutuhkan suatu sistem untuk mengenali suara manusia tersebut sehingga mendapatkan hasil yang menunjukkan individu yang dimaksud.

Metode yang biasanya digunakan untuk recognition pada Aplikasi Identifikasi Suara adalah *Hidden Markov Model* (HMM) atau *Vector Quantization* (VQ). Kedua metode recognition ini memiliki algoritma yang berbeda untuk mengenali suara. *Hidden Markov Model* (HMM) adalah sebuah model statistik dari sebuah sistem yang diasumsikan sebuah proses Markov dengan parameter yang tak diketahui, dan tantangannya adalah menentukan parameter-parameter tersembunyi (*hidden*) dari parameter-parameter yang dapat diamati. Sedangkan *Vector Quantization* (VQ) adalah suatu metode untuk melakukan pembelajaran pada lapisan kompetitif yang terawasi. Suatu lapisan kompetitif akan secara otomatis belajar untuk mengklasifikasikan vektor-vektor input. Kelas-kelas yang didapatkan sebagai hasil dari lapisan kompetitif ini hanya tergantung pada jarak antara vektor-vektor input. Jika 2 vektor input mendekati sama, maka lapisan kompetitif akan meletakkan kedua vektor input tersebut kedalam kelas yang sama.

Pada penelitian ini akan mencari hasil perbandingan unjuk kerja dari suatu Aplikasi Identifikasi Suara dengan metode *recognition* yang berbeda yaitu *Hidden Markov Model* (HMM) dan *Vector Quantization* (VQ) dengan parameter persentase keakuratannya dalam mengenali sampel suara yang diujicobakan sehingga didapatkan metode *recognition* mana yang paling cocok digunakan baik dalam kondisi ideal maupun kondisi tidak ideal.

## 2. TINJAUAN PUSTAKA

Tinjauan pustaka pada penelitian ini mengacu kepada beberapa literatur yang terkait dengan penelitian yang dilakukan. Teori yang diambil berupa teori tentang suara, Jaringan Saraf Tiruan, *Mel Frequency Cepstrum Coefficient* (MFCC), *Hidden Markov Model* (HMM), dan *Vector Quantization* (VQ).

### 2.1 Suara

Suara dihasilkan melalui proses *Generation* dan *Filtering*. Pada proses *Generation*, suara pertama kali diproduksi melalui bergetarnya pita suara yang berada di *larynx* untuk menghasilkan bunyi periodik [1]. *Vocal tract* yang terdiri dari lidah, gigi, bibir, langit-langit mulut dan lain-lain menghasilkan bunyi periodik yang bersifat konstan kemudian difilterisasi melalui sehingga bunyi tersebut dapat menjadi bunyi keluaran yaitu bunyi vokal dan/atau bunyi konsonan yang membentuk kalimat yang memiliki arti untuk *speech recognition*. Suara terdiri dari beberapa komponen, yaitu *pitch*, *formant* dan *spectrogram* yang dapat digunakan untuk mengidentifikasi karakteristik suara seseorang untuk kepentingan *voice recognition*.

### 2.2 Jaringan Saraf Tiruan

Sekelompok unit pemroses kecil yang dimodelkan berdasarkan jaringan saraf manusia disebut sebagai Jaringan Saraf Tiruan (JST) atau dalam Bahasa Inggris disebut *Artificial Neural Network* (ANN) atau umumnya hanya disebut *Neural Network* (NN) adalah jaringan dari [2]. JST merupakan sistem adaptif yang dapat mengubah struktur untuk memecahkan masalah berdasarkan informasi eksternal maupun internal yang mengalir melalui jaringan tersebut.

Aplikasi Identifikasi Suara adalah salah satu contoh dari JST. Aplikasi Identifikasi Suara berfungsi untuk mengenali suara atau bunyi. Suatu Aplikasi Identifikasi Suara mempunyai keakuratan yang berbeda-beda dalam mengenali atau mengidentifikasi suara tergantung dari beberapa faktor seperti metode ekstraksi ciri, metode *recognition*, *noise*, dan lain-lain.

### 2.3 Mel Frequency Cepstrum Coefficient (MFCC)

Ekstraksi fitur pada ASR (*Automatic Speech Recognition*) merupakan proses perhitungan urutan dari fitur vektor yang mampu merepresentasikan sinyal wicara yang ada secara optimal [3]. Fitur yang biasa digunakan adalah *cepstral coefficient*. MFCC merupakan metode ekstraksi fitur yang menghitung koefisien *cepstral* yang didasarkan pada variasi dari frekuensi kritis pada telinga manusia.

Tahapan-tahapan dalam MFCC adalah *frame blocking*, *windowing*, *fast fourier transform*, *mel frequency-wrapping*, dan *cepstrum*. sebagai berikut. Hasil dari proses ini dinamakan *Mel-Frequency Cepstrum Coefficients* (MFCC) yang nantinya akan digunakan dalam proses *recognition* pada Aplikasi Identifikasi Suara.

### 2.4 Hidden Markov Model

Hidden Markov Model merupakan sebuah model statistik dari sebuah sistem yang diasumsikan sebuah proses Markov dengan adanya parameter yang tak diketahui, lalu akan ditentukan parameter-parameter tersembunyi (*hidden*) dari parameter-parameter yang dapat diamati [4]. Setiap kondisi memiliki distribusi kemungkinan disetiap output yang berbeda. Walaupun *parameter model* diketahui, model tersebut tetap tersembunyi. HMM dapat digunakan untuk aplikasi dibidang *temporal pattern recognition* seperti pengenalan suara, tulisan, gestur, bio-informatika, kompresi kalimat, *computer vision*, ekonomi, finansial, dan pengenalan not balok.

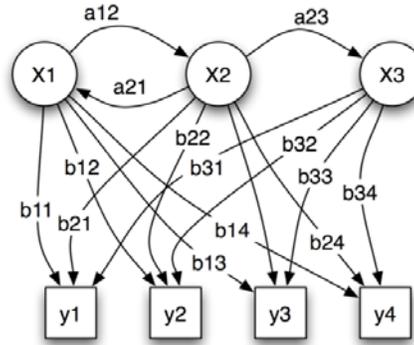
Dalam HMM, hanya urutan observasi saja yang dapat teramati, sedangkan urutan statenya tersembunyi.

Secara umum, HMM terdiri atas elemen-elemen berikut :

1. Himpunan nilai *output* observasi  $O = \{o_1, o_2, \dots, o_M\}$ , dimana  $M$  adalah jumlah simbol observasi.
2. Himpunan *state*  $S = \{1, 2, \dots, N\}$ . Dimana  $N$  menyatakan jumlah *state* yang terdapat pada HMM.
3. Himpunan probabilitas transisi antar *state*. Pada dasarnya, diasumsikan bahwa *state* berikutnya tergantung pada *state* pada saat ini. Asumsi ini menyebabkan proses perhitungan menjadi lebih mudah dan efisien untuk dilakukan. Probabilitas transaksi dapat dinyatakan dengan sebuah matriks  $A = \{a_{ij}\}$ , dimana  $a_{ij}$  adalah probabilitas transaksi dari *state*  $i$  ke *state*  $j$ .
4. Himpunan probabilitas *output*  $B = \{b_i(k)\}$  pada setiap *state*. Yang juga disebut probabilitas emisi

$b_i(k)$  adalah probabilitas dari simbol *output*  $o_k$  pada *state*  $i$ .

5. Himpunan *state* awal  $\pi = \{\pi_i\}$ , dimana  $\pi_i$  adalah probabilitas *state*  $i$  menjadi *state* awal pada urutan *state* HMM.

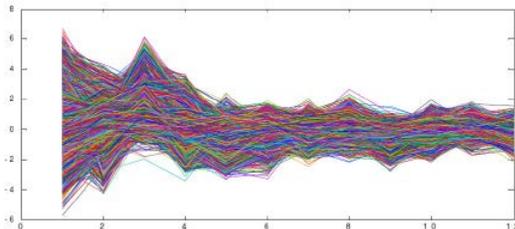


Gambar 1. Parameter Probabilistik pada Hidden Markov Model

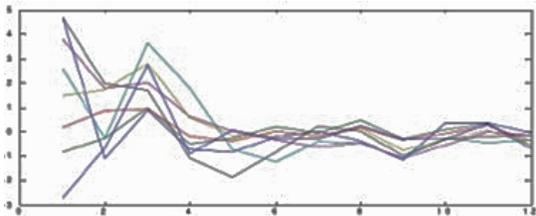
### 2.5 Vector Quantization

*Vector Quantization* adalah suatu metode untuk melakukan pembelajaran pada lapisan kompetitif yang terawasi [5]. Suatu lapisan kompetitif akan secara otomatis belajar untuk mengklasifikasikan vektor-vektor masukan. Kelas-kelas yang didapatkan sebagai hasil dari lapisan kompetitif ini hanya tergantung pada jarak antara vektor-vektor input. Jika 2 vektor input mendekati sama, maka lapisan kompetitif akan meletakkan kedua vektor masukan tersebut kedalam kelas yang sama.

Teknik VQ terdiri dari mengekstraksi sebagian kecil vektor corak sebagai contoh untuk menandai karakter spesifik pembicara agar lebih efisien. Dengan penggunaan corak data suara pelatihan yang dikelompokkan untuk membentuk suatu *codebook* untuk masing-masing pembicara. Dalam langkah pengenalan, data dari pembicara yang diuji dibandingkan kepada *codebook* dari tiap pembicaraan mengukur perbedaannya. Perbedaan ini kemudian digunakan untuk membuat keputusan pengenalan suara dari pembicara tersebut.



Gambar 2. Vektor Sebelum Mengalami Proses Vector Quantization



Gambar 3. Vektor Setelah Mengalami Proses Vector Quantization

### 3. METODOLOGI PENELITIAN

Metode yang digunakan pada penelitian ini yaitu membandingkan hasil unjuk kerja dari Aplikasi Identifikasi Suara yang menggunakan dua metode *recognition* yang berbeda. Aplikasi Identifikasi Suara ini diujicobakan dalam dua kondisi, yaitu kondisi ideal dan kondisi tidak ideal dengan kosakata "Selamat Pagi". Yang dimaksudkan dengan kondisi ideal adalah kondisi dimana tidak adanya noise pada rekaman yang diujicobakan. Sedangkan kondisi tidak ideal adalah kondisi dimana adanya noise pada rekaman yang diujicobakan dan diujicobakan dengan kosakata yang berbeda.

Tahapan tahapan pada penelitian ini sebagai berikut,

1. Studi Literatur
2. Membuat aplikasi identifikasi suara menggunakan perangkat lunak Matlab dengan menggunakan metode ekstraksi ciri *Mel Frequency Cepstrum Coefficient*, metode *recognition Hidden Markov Model* dan *Vector Quantization*.
3. Merekam suara dari 10 orang mahasiswa dan 10 orang mahasiswi sebanyak 10 kali dalam 10 hari yang berbeda dalam kondisi sepi, kondisi ramai, dan nama masing-masing responden,

4. Hasil rekaman hari ke-1 sampai hari ke-5 digunakan sebagai *data training* dan hasil rekaman hari ke-6 sampai hari ke-10 digunakan sebagai *data testing*.
5. Mencatat hasil identifikasi rekaman suara dari hasil uji coba identifikasi suara yang menggunakan metode *Hidden Markov Model* dan *Vector Quantization*.
6. Menganalisis perbandingan unjuk kerja Aplikasi Identifikasi Suara dengan metode *Hidden Markov Model* dan *Vector Quantization*.
7. Membuat kesimpulan dari hasil perbandingan unjuk kerja Aplikasi Identifikasi Suara dengan metode *Hidden Markov Model* dan *Vector Quantization*.

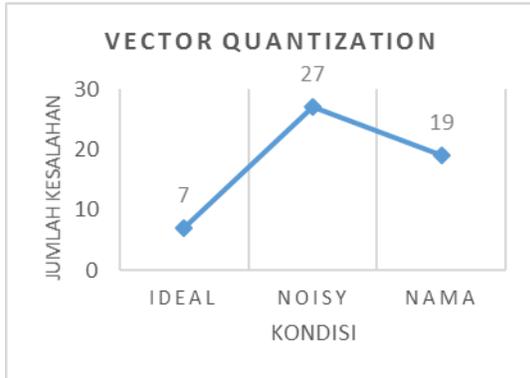
### 4. HASIL DAN PEMBAHASAN

Penelitian perbandingan unjuk kerja metode *Hidden Markov Model* dan *Vector Quantization* untuk Aplikasi Identifikasi Suara ini menggunakan data berupa rekaman suara dari 10 orang mahasiswa dan 10 orang mahasiswi. Ada 3 jenis rekaman suara yang digunakan yaitu rekaman suara kata "Selamat Pagi" tanpa adanya noise, rekaman suara kata "Selamat Pagi" dengan adanya noise, dan rekaman suara dari nama masing-masing responden. Rekaman suara untuk masing-masing jenis rekaman berjumlah 10 yang nantinya dibedakan menjadi 5 rekaman untuk *data training* dan 5 rekaman untuk *data testing*. Rekaman suara ini akan diujicobakan pada Aplikasi Identifikasi Suara dan nantinya akan didapatkan hasil unjuk kerja metode mana yang lebih baik untuk Aplikasi Identifikasi Suara.

#### 4.1 Hasil Unjuk Kerja Aplikasi Identifikasi Suara Metode Vector Quantization

Berdasarkan uji coba yang sudah dilakukan dengan Aplikasi Identifikasi Suara Metode *Vector Quantization* pada kondisi ideal, dari 100 data testing yang diujicobakan, sebanyak 93 sampel berhasil diidentifikasi dengan benar. Hasil unjuk kerjanya sebesar 93%. Sedangkan pada kondisi tidak ideal, dari 200 data testing yang diujicobakan, sebanyak 152 sampel berhasil diidentifikasi

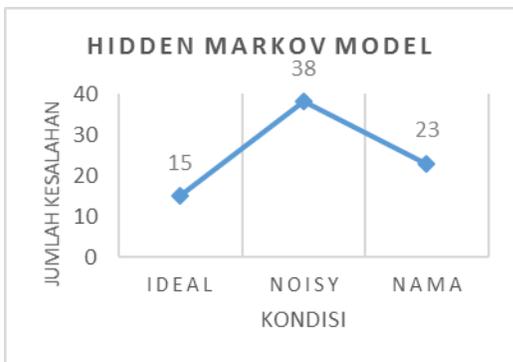
dengan benar. Hasil unjuk kerjanya sebesar 78%.



Gambar 5. Grafik Jumlah Kesalahan Aplikasi Identifikasi Suara Metode *Vector Quantization*

#### 4.2 Hasil Unjuk Kerja Aplikasi Identifikasi Suara Metode *Hidden Markov Model*

Berdasarkan ujicoba yang sudah dilakukan dengan Aplikasi Identifikasi Suara Metode *Hidden Markov Model* pada kondisi ideal, dari 100 data testing yang diujicobakan, sebanyak 85 sampel berhasil diidentifikasi dengan benar. Hasil unjuk kerjanya sebesar 85%. Sedangkan pada kondisi tidak ideal, dari 200 data testing yang diujicobakan, sebanyak 139 sampel berhasil diidentifikasi dengan benar. Hasil unjuk kerjanya sebesar 69,5%.



Gambar 6. Grafik Jumlah Kesalahan Aplikasi Identifikasi Suara Metode *Hidden Markov Model*

#### 4.3 Perbandingan Hasil Unjuk Kerja Aplikasi Identifikasi Suara Metode *Hidden Markov Model* dan *Vector Quantization*

Dari hasil ujicoba, aplikasi identifikasi suara dengan metode *Vector Quantization* mempunyai nilai unjuk kerja sebesar 93% dibandingkan dengan aplikasi identifikasi suara dengan metode *Hidden Markov Model* yang mempunyai nilai unjuk kerja sebesar 85% pada kondisi ideal. Pada kondisi tidak ideal, hasil unjuk kerja aplikasi identifikasi suara dengan metode *Vector Quantization* 78% sedangkan hasil unjuk kerja aplikasi identifikasi suara dengan metode *Hidden Markov Model* adalah 69,5%.

Aplikasi identifikasi suara dengan metode *Vector Quantization* mempunyai nilai unjuk kerja lebih baik, hal ini dikarenakan, pada metode *Vector Quantization* menggunakan layer tunggal dan tidak adanya layer tersembunyi seperti pada metode *Hidden Markov Model*.



Gambar 7. Grafik Perbandingan Hasil *Recognition* pada Kondisi Ideal



Gambar 8. Grafik Perbandingan Hasil *Recognition* pada Kondisi Tidak Ideal (*Noise*)



Gambar 9. Grafik Perbandingan Hasil *Recognition* pada Kondisi Tidak Ideal (Nama)

## 5. SIMPULAN

Berdasarkan penelitian yang sudah dilakukan, dapat ditarik beberapa simpulan sebagai berikut :

1. Metode *Vector Quantization* dan *Hidden Markov Model* dapat digunakan sebagai metode *recognition* pada Aplikasi Identifikasi Suara.
2. Hasil unjuk kerja Aplikasi Identifikasi Suara dengan metode *Vector Quantization* pada kondisi ideal adalah sebesar 93% sedangkan pada kondisi tidak ideal didapatkan hasil unjuk kerja sebesar 85%.
3. Hasil unjuk kerja Aplikasi Identifikasi Suara dengan metode *Hidden Markov Model* pada kondisi ideal adalah sebesar 78% sedangkan pada kondisi tidak ideal didapatkan hasil unjuk kerja sebesar 69.5%.
4. Aplikasi identifikasi suara dengan metode *Vector Quantization* mempunyai nilai unjuk kerja lebih baik dibandingkan aplikasi identifikasi suara yang menggunakan metode *Hidden Markov Model*, hal ini dikarenakan, pada metode *Vector Quantization* menggunakan layer tunggal dan tidak adanya layer tersembunyi seperti pada metode *Hidden Markov Model*.

5. Metode *Vector Quantization* direkomendasikan digunakan sebagai metode *recognition* pada Aplikasi Identifikasi Suara karena memiliki hasil unjuk kerja yang lebih baik dibandingkan dengan metode *Hidden Markov Model* baik dalam kondisi ideal maupun tidak ideal.

## 6. DAFTAR PUSTAKA

- [1] Nuh Al-Azhar, Muhammad. Audio Forensic: Theory and Analysis. Pusat Laboratorium Forensik POLRI.
- [2] Adler, John., Azhar, Muhammad., Supatmi, Sri. (Januari 2013). "Identifikasi Suara dengan MATLAB sebagai Aplikasi Jaringan Saraf Tiruan". Teknik Komputer, Universitas Komputer Indonesia.
- [3] Dave, Namrata. (2013). *Feature Extraction Methods LPC, PLP and MFCC In Speech Recognition. International Journal for Advance Research in Engineering and Technology*, (Volume 1, Issue VI, July 2013).
- [4] Srinivasan, A (2011). "Speech Recognition Using Hidden Markov Model". Department of Electronics and Communication Engineering Srinivasa Ramanujan Centre. SASTRA University.
- [5] Srinivasa, A., Ramaiah, Ch (Desember 2012). "Speech Samples Recognition based on MFCC and Vector Quantization". Department Of Computer Science and Engineering V.R.Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India