

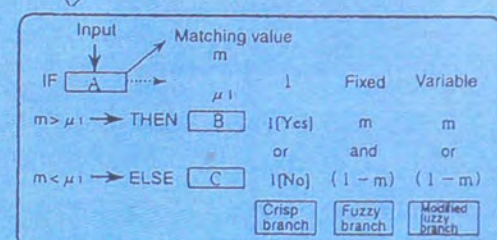
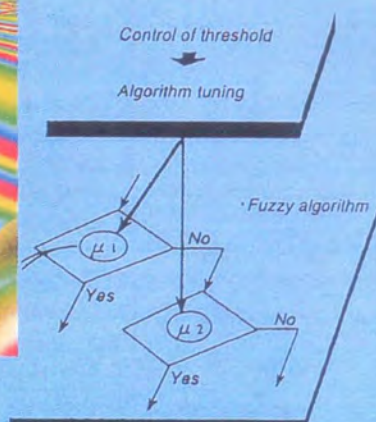
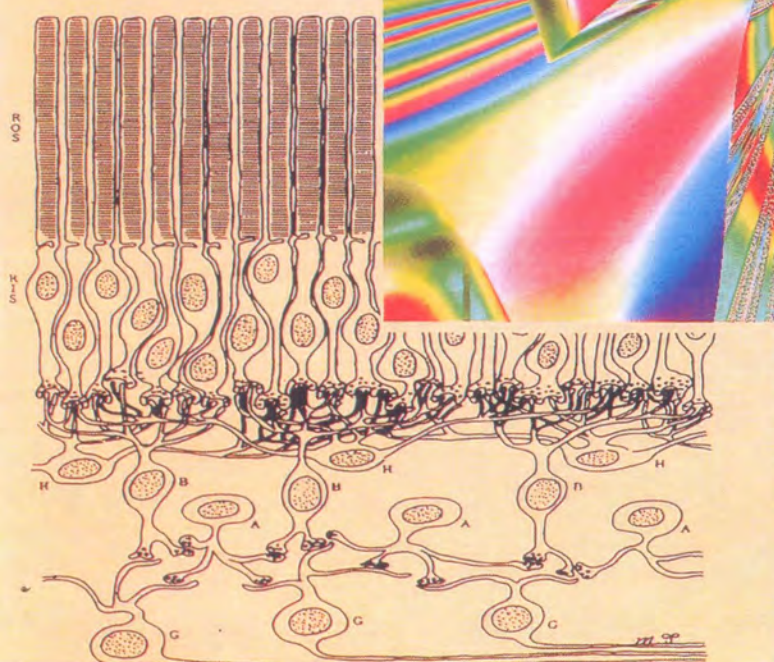
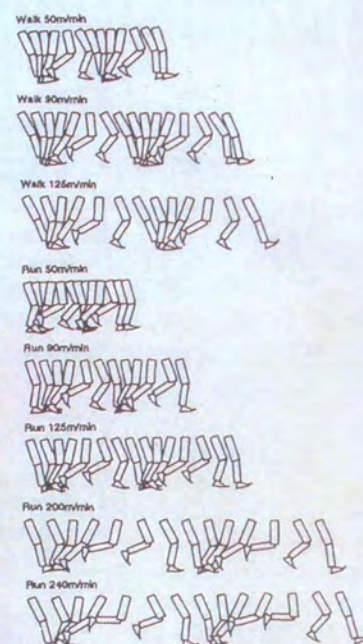
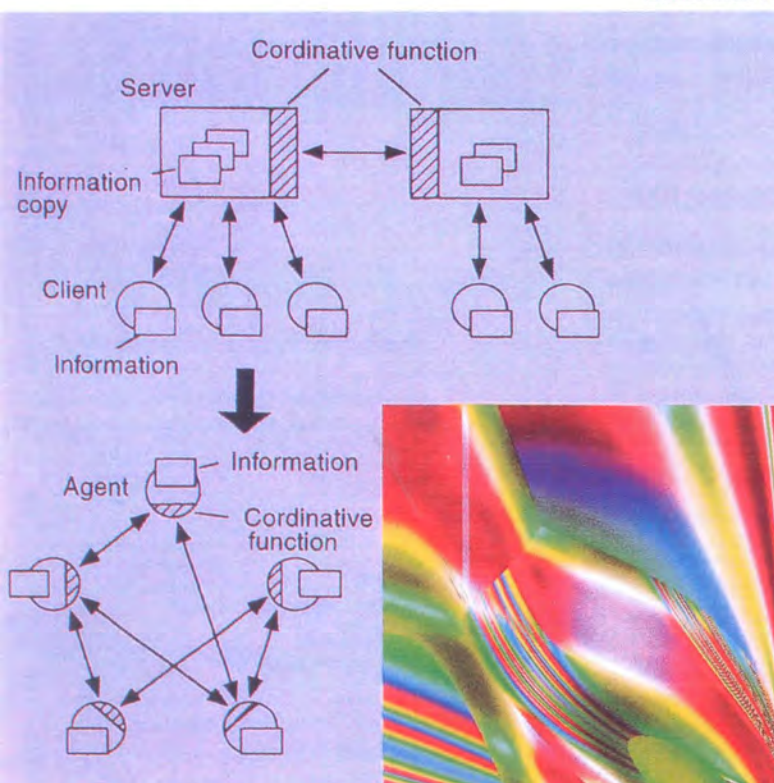
Journal of Advanced Computational Intelligence and Intelligent Informatics

Vol.16 No.3
May 2012

Indexed in Scopus

In Cooperation with

International Fuzzy Systems Association (IFSA),
Japan Society for Fuzzy Theory and Intelligent Informatics (SOFT),
Brazilian Society of Automatics (SBA),
The Society of Instrument and Control Engineers (SICE),
John von Neumann Computer Society (NJSZT),
Vietnamese Fuzzy Systems Society (VFSS),
Fuzzy Systems and Intelligent Technologies Research Society of Thailand (FIRST),
Korean Institute of Intelligent Systems (KIIS), and
Taiwanese Association for Artificial Intelligence (TAAI)



Editorial Board

Honorary Editor

Lotfi A. Zadeh
University of California, U.S.A.

Editors-in-chief

Toshio Fukuda
Nagoya University, Japan
Kaoru Hirota
Tokyo Institute of Technology, Japan

Advisory Editors

J. Bezdek
Univ. of West Florida, U.S.A.

Z. Bien
KAIST, Korea

D. Fogel
Natural Selection, Inc., U.S.A.

Y. Hatakeyama
Kochi Univ., Japan

N. Kasabov
Auckland Univ. of Technology,
New Zealand

J. Keller
Univ. of Missouri-Columbia, U.S.A.

J. Koza
Stanford Univ., U.S.A.

L. T. Kóczy
Technical Univ. of Budapest, Hungary

V. Kreinovich
The Univ. of Texas at El Paso, U.S.A.

R. J. Marks II
Baylor Univ., U.S.A.

M. Mukaidono
Meiji Univ., Japan

H. P. Nguyen
Ministry of Health of Vietnam,
Vietnam

I. J. Rudas
Óbuda Univ., Hungary

E. Ruspini
SRI International, U.S.A.

P. Sincák
TU Košice, Slovakia

I. B. Turksen
Univ. of Toronto, Canada

R. R. Yager
IONA College, U.S.A.

T. Yamakawa
Kyushu Institute of Technology, Japan

H.-J. Zimmermann
RWTH Aachen, Germany

Editors

R. A. Aliev
Georgia State Univ., U.S.A.

A. Bargiela
Univ. of Nottingham, UK

M. Beldjehem
Sainte-Anne Univ., Canada

D. Berrar
Tokyo Institute of Technology, Japan

S.-B. Cho
Yonsei Univ., Korea

E. P. Dadios
De La Salle Univ., Philippines

Y. Dai
Beijing Institute of Technology, China

L. Ding
National Univ. of Singapore, Singapore

F. Dong
Tokyo Institute of Technology, Japan

J. Ezzine
Ecole Nationale d'Ingénieurs de Tunis,
Tunisia

T. Furuhashi
Nagoya Univ., Japan

Q. Ha
Univ. of Technology, Australia

T. Hashimoto
Saitama Institute of Technology, Japan

Y. Hatakeyama
Kochi Univ., Japan

I. Hayashi
Kansai Univ., Japan

M. Idesawa
Univ. of Electro-Communications, Japan

N. Ikoma
Kyushu Institute of Technology, Japan

A. Inoue
Eastern Washington Univ., U.S.A.

M. Inuiguchi
Osaka Univ., Japan

H. Ishibuchi
Osaka Pref. Univ., Japan

G. Kang
Pukyong National Univ., Korea

D. Katagami
Tokyo Polytechnic Univ., Japan

S. Kawaji
Kumamoto Univ., Japan

K. Kawamoto
Chiba Univ., Japan

D. H. Kim
Hanbat National Univ., Korea

S. Kovács
Univ. of Miskolc, Hungary

N. Kubota
Tokyo Metropolitan Univ., Japan

C.-S. Lee
National Univ. of Tainan, Taiwan

J. Lee
National Central Univ., Taiwan

K. H. Lee
Korea Advanced Institute of Science and
Technology, Korea

S. Mabu
Waseda Univ., Japan

Y. Maeda
Fukui Univ., Japan

K. Miyazaki
National Institution for Academic
Degrees and Univ. Evaluation, Japan

C. K. Morooka
State Univ. of Campinas, Brazil

K. Nagasaka
Tokyo Univ. of Agriculture &
Technology, Japan

Y. Narukawa
Toho Gakuen, Japan

T. Onisawa
Univ. of Tsukuba, Japan

G.-K. Park
Mokpo National Maritime Univ., Korea

W. Pedrycz
Univ. of Alberta, Canada

P. M. Pinto
Univ. of Trás-os-Montes e Alto Douro,
Portugal

C. Pozna
Univ. Transilvania din Brasov, Romania

D. Ralescu
Univ. of Cincinnati, U.S.A.

E. Shimokawara
Tokyo Metropolitan Univ., Japan

K. Takadama
Univ. of Electro-Communications, Japan

H. Takahashi
Shonan Institute of Technology, Japan

Y. Takama
Tokyo Metropolitan Univ., Japan

K. Tenekedjiev
Nikola Y. Vaptsarov Naval Academy,
Bulgaria

V. Torra
IIIA-CSIC, Spain

K. Uehara
Ibaraki Univ., Japan

J. Watada
Waseda Univ., Japan

K. Watanabe
Okayama Univ., Japan

M. R. Widianto
Univ. of Indonesia, Indonesia

K. W. Wong
Murdoch Univ., Australia

J. Yi
Institute of Automation, Chinese
Academy of Sciences, China

G. Zhu
Guilin Univ. of Electronic Technology,
China

Vol.16 No.3
May 2012



Cover Pictures:

Upper left:
A function of Agent

Upper right:
Examples of leg's motion based on data
received for both walking and running during
constant velocity trial
(Dr. Takashi Yokoi, National Institute of
Bioscience and Human Technology)

Center:
2D CG image drawn with "SBART," a
simulated breeding tool.
(Prof. Tatsuo Unemi, Soka University)

Lower left:
The structure of retina

Lower right:
The threshold control method of fuzzy
algorithm in FRASH2
(Prof. Yoichiro Maeda, University of Fukui)

Publishing Staff:

Editor	Keiji Hayashi
Managing Editor	Shinya Wakai
Associate Editors	Reiko Ohta Kunihiko Uchida Shinji Isokawa
Art Director	Prof. Yuji Isa
Publisher	Keiji Hayashi

Published bimonthly by
Fuji Technology Press Ltd.

4F Toranomon Sangyo Bldg.
2-29, Toranomon 1-chome
Minato-ku, Tokyo 105-0001, Japan
Tel: +813-3508-0051
Fax: +813-3592-0648
E-mail: jaciii@fujipress.jp
URL: <http://www.fujipress.jp/JACIII>

Copyright © 2012 by
Fuji Technology Press Ltd.
All rights reserved.

No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in
any form or by any means, electronic, mechan-
ical, photocopying, recording, or otherwise,
without the prior written permission of the
publishers.

Contents

Regular Papers:

- **A Neural Network Model of Students' English Abilities Based on Their Affective Factors in Learning** 375
Fitra A. Bachtiar, Katsuari Kamei, and Eric W. Cooper
- **Fuzzy Hidden Markov Models for Indonesian Speech Classification** 381
Intan Nurma Yulita, Houw Liong The, and Adiwijaya
- **Multiscale Image Aggregation for Dental Radiograph Segmentation** 388
Martin Leonard Tangel, Chastine Fatichah,
Muhammad Rahmat Widyanto, Fangyan Dong,
and Kaoru Hirota
- **Substitute Target Learning Based Control System for Control Knowledge Acquisition Within Constrained Environment** 397
Syafiq Fauzi Kamarulzaman, Takeshi Shibuya,
and Seiji Yasunobu
- **Investigation About Applicability of Personal Values for Recommender System** 404
Shunichi Hattori and Yasufumi Takama
- **Parameter Optimization of Local Fuzzy Patterns Based on Fuzzy Contrast Measure for White Blood Cell Texture Feature Extraction** 412
Chastine Fatichah, Martin Leonard Tangel,
Muhammad Rahmat Widyanto, Fangyan Dong,
and Kaoru Hirota
- **On Network Structure of Stable Strategies in Local Connection Games** 420
Hikaru Iwazaki, Takenori Ujigawa, and Shao-Chin Sung
- **The Analysis of Consumer Purchasing Behavior on Cosmetics** 425
Ya-Ling Wu and You-Shyang Chen
- **Coalition Formation Based Staffing Strategy Development** 430
Mayuko Miyata and Shao-Chin Sung
- **The Effective Technology- and Resource-Management Strategy of Established Japanese Company: A Case Study of Toraya, a Traditional Japanese Confectioner** 436
Kana Sugimoto, Takao Someya, and Shin'ya Nagasawa

Fuzzy Hidden Markov Models for Indonesian Speech Classification

Intan Nurma Yulita, Houw Liong The, and Adiwijaya

Graduate Faculty, Telkom Institute of Technology

Jalan Telekomunikasi No.1, Dayeuhkolot, Jawa Barat 40257, Indonesia

E-mail: intanurma@gmail.com, houwthee@yahoo.co.id, adiwijaya@ittelkom.ac.id

[Received September 15, 2011; accepted November 15, 2011]

Indonesia has many tribes, so that there are many dialects. Speech classification is difficult if the database uses speech signals from various people who have different characteristics because of gender and dialect. The different characteristics will influence frequency, intonation, amplitude, and period of the speech. It makes the system must be trained for the various templates reference of speech signal. Therefore, this study has been developed for Indonesian speech classification. The solution is a new combination of fuzzy on hidden Markov models. The result shows a new version of fuzzy hidden Markov models is better than hidden Markov model.

Keywords: fuzzy logic, hidden Markov models, speech, classification, clustering

1. Introduction

Over the past several decades, speech classification technology has been widely applied. There are many approaches for speech classification for example template-based, knowledge-based, and stochastic-based approaches [1]. The successful results were Hidden Markov Model (HMM) [2]. Other results were artificial neural network [3], support vector machine [4], fuzzy [5] and clustering [6]. Speech classification is a “language-dependent” system. Since each language has many phonemes then the application of classification in a language cannot be applied for all language. Many researches have been carried out abroad, but it cannot be applied well in Indonesian. English speech recognition is the most speech recognition system, it has been developed in references [2–10]. The number of researches which have done Indonesian speech classification is still slightly. They have done Indonesian speech classification based on speaker adaptation system [11], and developed the corpus of Indonesian speech classification [12].

Speech classification is difficult because speech has several unique characteristics. In different time, a same word has different form although it has been spoken from the same person. Hence, speech classification is more difficult if the database uses speech signals from various

people who have different characteristics because of gender and dialect. The different characteristics will influence frequency, intonation, amplitude, and period of the speech. It makes the system must be trained for the various templates reference of speech signal. Therefore, a study still needs to be conducted. Hidden Markov models is a common approach that is used to classify speech. However, a method is needed to develop a solution from the above problem, and for Indonesian speech classification. This study designs it. The solution combines fuzzy on hidden Markov models. Fuzzy handles variant forms of speech more properly than there is no fuzzy. If the number of variant is higher, then the area of each cluster of Fuzzy C-Means clustering is wider. Actually, some study had combined fuzzy on HMM [3, 9, 10, 13] but they were not designed to solve the different characteristics problem in speech dialect and for Indonesian speech. The new design of fuzzy hidden Markov models will be proposed in this study. The models will consist of Fuzzy C-Means clustering which is designed to substitute the vector quantization process and a new forward and backward method to handle the membership degree of data.

2. Materials and Methods

2.1. Obtaining Raw Data

In this study, it was used the speech recognition data from Research and Development Center of Telkom of Indonesia. Data collection was conducted in a soundproof room (it means there is no noise in speech) and the number of involved speakers was 70 people (male and female). Experiments were performed on speech data set with various characteristics dialect and gender. In this data, some dialect of speaker tribes in Indonesia was used, they were Sundanese, Javanese, Batak, Betawi, Balinese but there was no information how much their proportion. The data set was divided into training data (80% of data set) and testing data (20% of data set). The speakers of training data and testing data were different because our speech classification was speaker independent system. **Table 1** lists the used words for the training data and shows the extremely different sounds of each word.

Table 1. Training data.

Words	Sounds	Information
Balaysalasa	'balaysalasa' and 'baleysalasa'	101 files
Lubuklinggaw	'lubuklinggaw' and 'lubuklinggo'	101 files
Prabumulih	'prabumulih' and 'prabulumeh'	101 files
tanjungenim	'tanjungenim' and 'tanjungeenim'	100 files
Tarempa	'tarempa' and 'tareempa'	98 files

2.2. Preprocessing

The purpose of preprocessing is to make all input signals in accordance with the required system specifications [7]. The first step is centering, it aims to shift the location of the discrete amplitude distribution and makes its center locate the axis $y = 0$. Thus, centering makes the average amplitude of the signal to zero. The next step is normalization, a process to equalize the maximum amplitude of the sound signal. Normalization is done by dividing each discrete amplitude values with the maximum amplitude value.

2.3. Feature Extraction

This process aimed at obtaining the characteristics of the voice signal. In this study, MFCC is implemented for feature extraction. It produces 24 parameter values. They are 12 cepstral values and 12 first-order derivative value of these cepstral. The output of this process is that every speech will be divided into a number of frames and each frame will have 24 feature values.

2.4. Vector Quantization (VQ)

Basically, the output of feature extraction is shorter than the original signal. However, in order to process HMM, an observation sequence is needed [7]. The observation represents all variations of existing cepstral. VQ is used for the formation of discrete symbols (codebook) from a series of observations of the HMM model for estimating the vector representation of the shorter term. VQ process is divided into two stages: the formation of codebook and the codebook index determination. When constructing codebook, the input feature vector of the VQ is a whole variety of known voice signal. By using clustering algorithms, feature vector will be grouped into clusters. The cluster center is called codebook. After the codebook is constructed, the next step of VQ can be done by replacing a feature vector with one vector codebook which has the smallest Euclidean distance. The output of VQ is the input of hidden Markov models.

2.5. Hidden Markov Models (HMM)

HMM is a Markov chain which its output symbol describes the chances of output symbol transitions [2, 8]. Observations for each state are described separately by a probability function or density function (probability density function), which is defined as an opportunity to produce a transition between states. Unlike the Observable Markov Model (OMM), HMM consists of a series of double stochastic process that primary process cannot be directly observable (hidden) but only can be observed through another set of stochastic processes that produces a range of observations.

2.5.1. Basic Element

HMM as a discrete observation symbol has the following elements [2, 8]:

1. HMM consists of N states, they are labeled by $1, 2, \dots, N$ and state to- t is given by q_t . N is tested parameter in this study.
2. Number of observation symbols (M). Observation symbol is the output being modeled.
 $V = V_1, \dots, V_m$.
3. Transition probability distribution from one state to another state (A)

$$A = a_{ij}, \quad 1 \leq i, j \leq N. \quad (1)$$

4. Observation probability distribution of k -th symbol in the j -th state (B)

$$B = b_j(V_k), \quad 1 \leq i \leq N, \quad i \leq j \leq N. \quad . (2)$$

5. Initial state probability distribution

$$\pi_i = P(q_1 = i), \quad 1 \leq i \leq N. \quad (3)$$

HMM requires specification of two model parameters N , M , A , B , and, λ are measured. HMM notations are usually written with $\lambda(model) = (A, B, \pi)$.

2.5.2. Basic Problem and Solution

There are three basic problems in HMM to be solved, namely [2, 8]:

1. If a given observation $O = (O_1, O_2, \dots, O_t)$ and model evaluation λ , how to calculate the efficient probability of observations series?
2. If a given observation (O_1, O_2, \dots, O_t) and model evaluation λ , how to choose the optimal states series that represent the observation?
3. How to set the parameters of the model evaluation λ to maximize the probability $P(O|\lambda)$ value?

The solutions to the problem above are [2, 8]:

1. Evaluation (evaluation of opportunities)

The commonly used method is to examine each possible sequence of N states along the T (the number of observations). It is not efficient. Another simple procedure is forward and backward procedures.

a. Forward procedure

Forward variable ($\alpha_t(i)$) at t -time and i -state is defined by $\alpha_t(i) = P(O_1, O_2, \dots, O_t, q_t = i | \lambda)$. The forward opportunities function can be solved for N -state and T -symbol inductively with the following steps:

i. Initialization:

$$\alpha_t(i) = \pi_i b_i(O_1), \quad 1 \leq i \leq N \quad . . \quad (4)$$

ii. Induction:

$$\alpha_{t+1}(j) = \left| \sum_{i=1}^N (\alpha_t(i) \alpha_{ij}) \right| b_j(O_{t+1}) \quad (5)$$

$$1 \leq (i, j) \leq N, \quad 1 \leq t \leq T-1$$

iii. Termination:

$$P(O|\lambda) = \sum_{i=1}^N (\alpha_t(i)) \quad (6)$$

Forward probability is calculated based on the Trellis diagram pattern. There are n points each time slot in the pattern. All possible sequence is combined to N states.

b. Backward procedure

Backward variable $\beta_t(i)$ in time to t and i -state is defined by $\beta_t(i) = P(O_{t+1}, O_{t+2}, \dots, O_T, q_t = i | \lambda)$. Step backward procedure is as follows:

i. Initialization :

$$\beta_t(i) = 1, \quad 1 \leq i \leq N \quad (7)$$

ii. Induction

$$\beta_t(j) = \sum_{i=1}^N \alpha_{ji} b_j(O_{t+1}) \beta_{t+1}(i) \quad . . \quad (8)$$

$$1 \leq (i, j) \leq N, \quad t = T-1, T-2, \dots, 1$$

To obtain the state to the i -th time t and the rows of observations at time $t+1$, then it is assumed that the possible j -state at time $t+1$, to obtain a transition from i to j , and rows of observation on the j -th state. Then it calculates the observation of the j -state.

c. Forward-Backward procedure

The combination of forward and backward procedure can be used to obtain the values of $P(O|\lambda)$. Opportunity in the state at t -time of the N state before time $t-1$ can be calculated with the function of the forward opportunities $\alpha_t(i)$. Backward probability function is used to calculate the probability of observation symbol sequence that starts from time $t+1$ to T . By mathematical calculation, using a forward-backward procedure is illustrated

as the following formula:

$$P(O|\lambda) = \sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) \alpha_{ij} \beta_{t+1}(j)$$

$$= \sum_{i=1}^N \alpha_t(i) \beta_t(i) \quad (9)$$

2. Decoding

The second problem is looking for the hidden state sequence (hidden) for a sequence of generated observations from model. The solution is used to find the optimal state sequence. It is Viterbi algorithm (dynamic programming). Viterbi algorithm maximizes the probability value $P(Q|O, \lambda)$ so it will produce the optimal state sequence. Based on the Bayes rule, mathematically it is expressed as this formula:

$$P(Q|O, \lambda) = \frac{P(Q, O|\lambda)}{P(O|\lambda)} \quad (10)$$

3. The third problem solution is to adjust the (training) parameters based on certain optimal criterion. The usual method to solve this third problem is the Baum-Welch algorithm. This algorithm is an iterative method that works to find the values of local maximum of the probability function. This training process continues until a critical state is met. The model result should be better training than the previous model.

2.6. Fuzzy Hidden Markov Models (FHMM)

The proposed FHMM does not implement vector quantization. The substituted process is Fuzzy C-Means clustering. Fuzzy C-Means clustering has two functions. First, it obtains the codebook by clustering processing, the codebook is a cluster center. Second, it changes the feature extraction output to be the data with membership degree for each cluster. The data is used to be the fuzzy hidden Markov models input.

Figure 1 can be elaborated that the system is designed to have 2 ways (training and testing). Both ways have the same number of stages. They are preprocessing, feature extraction and Fuzzy C-Means clustering. The system input is speech. The speech is normalized. The normalized speech is extracted by feature extraction processing. The training of Fuzzy C-Means clustering is done to get codebook. After the codebook is constructed, the next step of vector quantization can be done by replacing a feature vector with a row of frame membership degree for each cluster. The testing of vector quantization only replaces a feature vector with a row of frame membership degree for each cluster. After Fuzzy C-Means clustering, the training does re-estimation process for FHMM and the testing process decided the most similar reference model. The system output is text.

2.6.1. Fuzzy C-Means Clustering

The steps of Fuzzy C-Means clustering will be shown in the following steps [14]:

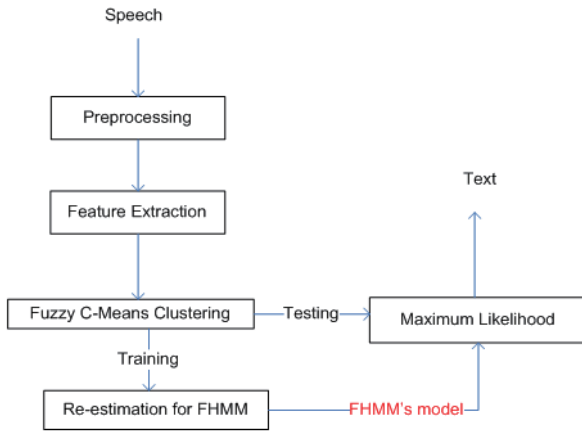


Fig. 1. Speech classification using FHMM.

1. Initializing data input, matrix X , with size $n \times m$, (n = number of frames, m = number of features)
2. Determining the parameters:
 - a. Number of clusters (k)
 - b. Maximum iterations (t)
 - c. The expected smallest error (ξ)
 - d. 1st iteration
 - e. Power (w)

The number of cluster indicates the variation of recognized sound. If the number of cluster is 16 then there are 16 variation of recognized sound. The power of Fuzzy C-Means clustering indicates range of each cluster. If the power is 2 then the cluster range is wider than that if the power is 1.3. It means if the power is 2 then membership degree of data is higher than that if the power is 1.3.

3. Generating random values from the matrix U which is a matrix number of frames, and the number of clusters, to make the matrix elements of the initial partition U . Calculating the partition matrix (μ_{ik}):

$$\mu_{ik} = \frac{\mu_{ik}}{Q_j} \quad \dots \quad (11)$$

4. Calculating the k -th cluster center (V_{kj}):

$$V_{kj} = \frac{\sum_{i=1}^n ((\mu_{ik})^w X_{ij})}{\sum_{i=1}^n (\mu_{ik})^w} \quad \dots \quad (12)$$

5. Calculating the objective function (P_t) at iteration t :

$$P_t = \sum_{i=1}^N \sum_{k=1}^c \left(\sum_{j=1}^m (X_{ij} - V_{kj})^2 (\mu_{ik})^w \right) \quad \dots \quad (13)$$

6. Doing iteration and at each iteration the partition matrix (μ_{ik}) will be updated:

$$\mu_{kj} = \frac{\left(\sum_{i=1}^m (X_{ij} - V_{kj})^2 \right)^{\frac{-1}{w-1}}}{\sum_{k=1}^c \left(\sum_{i=1}^m (X_{ij} - V_{kj})^2 \right)^{\frac{-1}{w-1}}} \quad \dots \quad (14)$$

7. Checking the stop condition:

- a. If new objective function value less the same old objective function value is less than the expected error value, or more than the maximum t value iteration, ($|P_t - P_{t-1}| < \xi$) or ($t > MaxIter$), then stop.
- b. Step 4 will be repeated if the condition has not stopped and $t = t + 1$.

Fuzzy C-Means clustering is done to obtain the cluster center (codebook). After the codebook is constructed, the next step can be done by replacing a feature vector with a row membership degree of frame for each cluster. After the codebook is obtained, then calculate membership degree of data for each cluster (κ_{xz}) using the following equation:

$$\kappa_{xz} = \frac{\left(\sum_{y=1}^m (X_{xy} - V_{zy})^2 \right)^{\frac{-1}{w-1}}}{\sum_{z=1}^c \left(\sum_{y=1}^m (X_{xy} - V_{zy})^2 \right)^{\frac{-1}{w-1}}} \quad \dots \quad (15)$$

Note:

1. x : number of frames of observation data
2. y : number of features
3. z : number of clusters

2.6.2. Fuzzy Forward-Backward

The difference between HMM and FHMM is that for each observation HMM refers to one codebook value of one frame while in FHMM, observation refers to a frame value but it has all the values in each codebook with different membership degree. Therefore, a new framework of forward and backward calculation needs to be conducted. In this sub-section, the other forward and backward calculation is also shown [13].

Initialization of forward calculation ($t = 1$):

1. HMM

$$\alpha_1(i) = \pi_1 b_i(O_1) \quad \dots \quad (16)$$

2. The proposed FHMM

$$\alpha_1(i) = \pi_1 b_i(\Theta_1) \quad \dots \quad (17)$$

3. Other FHMM

$$\alpha_1(i) = \pi_1 \left| \sum_{m=1}^M u(m,1) b_i(O_1) \right| \quad (18)$$

Induction of forward calculation ($t = 2, \dots, T$):

1. HMM

$$\alpha_{t+1}(j) = \left| \sum_{i=1}^N (\alpha_t(i) \alpha_{ij}) \right| b_j(O_{t+1}) \quad . . . (19)$$

2. The proposed FHMM

$$\alpha_{t+1}(j) = \left| \sum_{i=1}^N (\alpha_t(i) \alpha_{ij}) \right| b_j(\Theta_{t+1}) \quad . . . (20)$$

3. Other FHMM

$$\alpha_{t+1}(j) = \left| \sum_{i=1}^N (\alpha_t(i) \alpha_{ij}) \right| \left| \sum_{m=1}^M u(m,t) b_j(m) \right| \quad (21)$$

Induction of backward calculation ($t = T$):

1. HMM

$$\beta_t(j) = \sum_{j=1}^N a_{ji} b_j(O_{t+1}) B_{t+1}(j) \quad (22)$$

2. The proposed FHMM

$$\beta_t(j) = \sum_{j=1}^N a_{ji} b_j(\Theta_{t+1}) B_{t+1}(j) \quad (23)$$

3. Other FHMM

$$\beta_t(j) = \left| \sum_{j=1}^N a_{ji} B_{t+1}(j) \right| \left| \sum_{m=1}^M u(m,t) b_j(m) \right| \quad (24)$$

Calculation of forward-backward:

1. HMM

$$P(O|\lambda) = \sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) a_{ji} b_j(O_{t+1}) B_{t+1}(j) \quad (25)$$

2. The proposed FHMM

$$P(O|\lambda) = \sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) a_{ji} b_j(\Theta_{t+1}) B_{t+1}(j) \quad (26)$$

3. Other FHMM

$$P(O|\lambda) = \sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) a_{ji} B_{t+1}(j) \rho \quad . . . (27)$$

$$\rho = \sum_{m=1}^M u(m,t) b_j(m) \quad (28)$$

Note:

$$1. \quad b_i(\Theta_x) = \sum_{z=1}^c B_{zi} \kappa_{xz} \quad (29)$$

This formula means that the input data is observation data which has membership degree for each cluster, and the output data is the observation probability distribution of x -th symbol in the i -th state (B).

$$2. \quad u(m,t) = \text{similarity}(cb(m), O_t) \quad (30)$$

$cb(m)$ is a cluster center vector for index m .

3. Similarity measure m (represents the number of features) using one of the following equations:

a. Cosine similarity:

$$(x_i, x_j) = \frac{\sum_{k=1}^m X_{ik} X_{jk}}{\sqrt{\sum_{k=1}^m (X_{ik})^2 \sum_{k=1}^m (X_{jk})^2}} \quad . . . (31)$$

b. Manhattan distance:

$$(x_i, x_j) = \sum_{k=1}^m |X_{ik} - X_{jk}| \quad (32)$$

c. Euclidean distance:

$$(x_i, x_j) = \sqrt{\sum_{k=1}^m |X_{ik} - X_{jk}|^2} \quad (33)$$

4. Four formulas of the proposed forward and backward calculation are changed because every value $b_j(O_t)$ refers to all codebook with different degrees of membership.

3. Experimental Results

3.1. Compare HMM and FHMM if the Number of Cluster was Altered

The purpose of experiment was to obtain the optimal number of cluster. The static variables were the number of states and the power (w). In this experiment, the number of state was 7 (seven) and the power (w) was 1.1.

Table 2 shows 32 clusters is better than 16 clusters. It means that the system required 32 variant of recognized sound to obtain a good accuracy. The experiment did not try if the number of cluster is more than 32, for example 64 clusters since this study had only five recognized words, all words had few phonemes. If the number of cluster was 64, it would cause the overspecialization system.

3.2. Compare FHMM for Each Power (w)

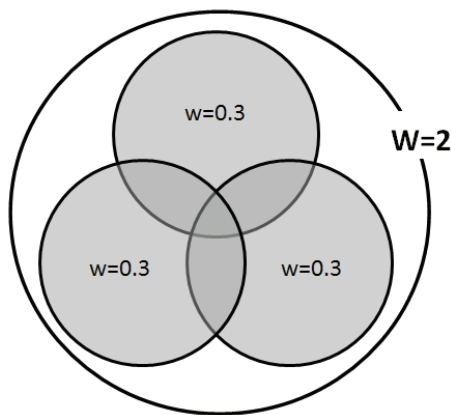
The purpose of experiment was to obtain the optimal power (w) of FHMM. The static variables were the number of states and the number of cluster. In this experiment,

Table 2. If the number of cluster was altered.

The number of cluster	HMM	FHMM
16	66.67 %	84.17%
32	80 %	88.33%

Table 3. If the power (w) was altered.

w	Accuracy
1.05	92.5%
1.1	88.33%
1.3	83.33%
1.5	65%
1.7	46.67%

**Fig. 2.** The influence of w .

the number of state was 7 (seven) and the number of cluster was 32. The number of cluster was 32 because it was the optimal number which was obtained from experiment of **Table 2**.

Table 3 shows FHMM accuracy if power (w) was ranging from 1.05–1.7. The optimal power (w) was 1.05 and if w increased then FHMM accuracies decreased. The explanation of the result will be shown in **Fig. 2**.

Figure 2 shows an illustration that if the power (w) was 1.3, each data had different degrees of membership for each cluster. Otherwise, if the power (w) was 2, three clusters have the same region and each data has the same degrees of membership for each cluster. It means that there is no different among observation data and the system will only recognize one label.

3.3. Compare FHMM for Each State

The purpose of experiment was to obtain the optimal number of states. The static variables were the number of cluster and the power (w). In this experiment, the number cluster was 32 and the power was 1.05. The parameter values were the optimal values which were obtained from experiment of **Tables 2** and **3**.

Table 4. If the number of cluster was altered.

The number of states	HMM	FHMM
5	80%	90%
6	86.67%	90%
7	80%	92.5%
8	86.67%	90%
9	89.17%	91.6%
10	85.83%	91.67%

Table 5. HMM and FHMM for their optimal condition.

Method	Accuracy
HMM	89.17%
FHMM	92.5%

From **Table 4**, the optimal number of states of HMM was 9 and the optimal number of state of FHMM was 7. The accuracy was not influenced by the number of states. It is shown that even if the number of states were increased the accuracy is not always increased.

3.4. Compare HMM and FHMM

The purpose of experiment was to compare HMM and FHMM if they have the optimal condition (the best accuracy). The optimal condition of HMM was if the number of cluster was 32 and the number of state was 9. The optimal condition of FHMM was if the number of cluster was 3, the power (w) was 1.05, and the number of state was 7.

From **Table 5**, FHMM was better than HMM. FHMM could improve HMM accuracy and its improvement was 3.33%. The improvement is small, but we have to remember that it is not effective when a real application tries to obtain the optimal condition because it means try to alter number of state and of course, it is waste of time. Hence FHMM is a good solution for this problem, just try once and the accuracy is equal to or greater than 90%.

4. Conclusion and Recommendation

4.1. Conclusion

From the analysis of the performance of FHMM, by using the data in this study, it can be concluded that a new version of FHMM can handle several problems of variant speech because of gender and dialect. It is better than HMM.

4.2. Recommendations for Future Works

Since our method is an effective way for Indonesian speech classification, it is highly recommended for a bigger database and other languages.

References:

- [1] S. D. Shenouda, F. W. Zaki, and A. Goneid, "Hybrid Fuzzy HMm System for Arabic Connectionist Speech Recognition," Proc. of the 5th WSEAS Int. Conf. on Signal Processing, robotics and Automation, pp. 64-69, 2006.
- [2] L. R. Rabiner, "A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," Proc. of the IEEE, Vol.77, No.2, 1989.
- [3] P. Melin, J. Urias, D. Solano, et al., "Voice Recognition with Neural Networks, Type-2 Fuzzy Logic and Genetic Algorithms," Engineering Letters, Vol.13, No.2, 2006.
- [4] L. Chen, S. Gunduz, and M. T. Ozsu, "Mixed Type Audio Classification with Support Vector Machine," Proc. of the IEEE Int. Conf. on Multimedia and Expo, 2006.
- [5] R. Halavati, S. B. Shouraki, M. Eshraghi, and M. Alemzadeh, "A Novel Fuzzy Approach to Speech Processing," 5th Hybrid Intelligent Systems Conf., 2004.
- [6] S. E. Levinson, L. R. Rabiner, A. E. Rosenberg, and J. G. Wilpon, "Interactive Clustering Techniques for Selecting Speaker-Independent Reference Templates For Isolated Word Recognition," IEEE Trans. on Acoustics, Speech, and Signal Processing, Vol.Assp-27, 1979.
- [7] B. H. Juang and L. R. Rabiner, "Fundamentals of Speech Recognition," Prentice-Hall, 1993.
- [8] B. H. Juang and L. R. Rabiner, "Hidden Markov Models for Speech Recognition," Technometrics, Vol.33, No.3, pp. 251-272, 1991.
- [9] J. Zeng and Z.-Q. Liu, "Interval Type-2 Fuzzy Hidden Markov Models," Proc. of Int. Conf. on Fuzzy Systems, Vol.2, pp. 1123-1128, 2004.
- [10] J. Zeng and Z.-Q. Liu, "Type-2 Fuzzy Hidden Markov Models to Phoneme Recognition," Proc. of the 17th Int. Conf. on Pattern Recognition, 2004.
- [11] H. Riza and O. Riandi, "Toward Asian Speech Translation System: Developing Speech Recognition and Machine Translation for Indonesian Language," Int. Joint Conf. on Natural Language Processing, 2008.
- [12] D. P. Lestari, K. Iwano, and S. Furui, "A Larger Vocabulary Continuous Speech Recognition System for Indonesian Language," 15th Indonesian Scientific Conf. in Japan Proceedings, 2006.
- [13] H. Uguz, A. Ozturk, R. Saracoglu, and A. Arslan, "A Biomedical System Based on Fuzzy Discrete Hidden Markov Model for The Diagnosis of The Brain Diseases," Expert Systems With Applications, Vol.35, pp. 1104-1114, 2008.
- [14] S. Kusumadewi, H. Purnomo, and A. Logika, "Fuzzy untuk Pendukung Keputusan," Penerbit Graha Ilmu, pp. 84-85, 2004.

**Name:**

Intan Nurma Yulita

Affiliation:

Lecturer, Faculty of Informatics, Telkom Institute of Technology
 Graduate Faculty, Telkom Institute of Technology

Address:

Jalan Telekomunikasi No.1, DayeuhKolot, Jawa Barat 40257, Indonesia

Brief Biographical History:

2004-2008 S.T. in Informatics, Telkom Institute of Technology
 2009-2011 M.T. in Informatics, Telkom Institute of Technology

Main Works:

- applying fuzzy logic on data mining

Membership in Academic Societies:

- Soft Computing – Indonesia (SC-INA)
 - Data Mining – Indonesia (INDO-DM)
-

**Name:**

Houw Liong The

Affiliation:

Professor, Telkom Institute of Technology
 Professor, School of Telematics, Institut Teknologi Harapan Bangsa
 Professor and Researcher, Bandung Institute of Technology (ITB)
 Professor, Artificial Intelligence at ST Inten, Bandung

Address:

Jalan Telekomunikasi No.1, DayeuhKolot, Jawa Barat 40257, Indonesia

Brief Biographical History:

1959-1964 Bachelor degree in Physics from Bandung Institute of Technology (ITB)

1964-1968 Ph.D. in Physics from University of Kentucky

Main Works:

- researching complex systems, climate change, climate predictions, flood predictions and space science

Membership in Academic Societies:

- Indonesia Global Network
 - Climate Change Network
 - Interactive Physics
-

**Name:**

Adi Wijaya

Affiliation:

Lecturer, Telkom Institute of Technology

Address:

Jalan Telekomunikasi No.1, DayeuhKolot, Jawa Barat 40257, Indonesia

Brief Biographical History:

1999 B.Sc. in Mathematics from Bandung Institute of Technology (ITB)

2004 M.Sc. in Mathematics from Bandung Institute of Technology (ITB)

2012 Ph.D. in Mathematics from Bandung Institute of Technology (ITB)

Main Works:

- graph theory, information theory, and optimization on image processing

Membership in Academic Societies:

- Indonesian Mathematics Society (Indo-MS)
 - Indonesian Combinatorial Society (InaCombS)
-