FUZZY HIDDEN MARKOV MODELS FOR INDONESIAN SPEECH CLASSIFICATION

Intan Nurma Yulita[1], Prof. The Houw Liong[2], Adiwijaya[3]

[1]Magister Teknik Informatika, Fakultas Teknik Informatika, Universitas Telkom

Abstrak

Indonesia memiliki banyak suku bangsa sehingga ada banyak dialek. Klasifikasi ucapan sulit dilakukan jika data mengunakan sinyal ucapan dari berbagai orang yang memiliki perbedaan karekateristik karena jenis kelamin dan dialek. Perbedaan ini akan mempengaruhi frekuensi, intonasi, amplitude, dan periode dari suara. Akibatnya, perbedaaan karakteristik membuat system perlu dilatih untuk berbagai referensi template dari sinyal ucapan. Oleh karena itu, penelitian ini telah dilakukan untuk klasifikasi ucapan berbahasa Indonesia. Penelitian ini merancang solusi karakteristik berbeda untuk klasifikasi ucapan berbahasa Indonesia. Solusin mengkombinasikan Fuzzy pada Hidden Markov Models. Desain baru fuzzy Hidden Markov Models ditawarkan pada penelitian ini. Modelnya akan terdiri dari Fuzzy C-Means Clustering yang didesain untuk menggantikan proses vector quantization dan metode baru dari forward and backward untuk menangani derajat keanggotaan data. Hasil menunjukkan FHMM lebih baik dari HMM dengan peningkatan sebesar 3.33 %.

Kata Kunci : Fuzzy, Hidden Markov Models, Bahasa Indonesia, Ucapan, Klasifikasi

Abstract

ndonesia has a lot of tribe, so that there are a lot of 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 requency, intonation, amplitude, and period of the speech. make the system be trained for the various templates reference of speech signal. Therefore, this study has been developed for Indonesian speech classification This study designs the solution of the different characteristics for Indonesian speech classification. The solution combines Fuzzy on Hidden Markov Models. The new design of fuzzy Hidden Markov Models will be proposed in this study. The models will consist of Fuzzy C-Means Clustering which will be designed to substitute the vector quantization process and a new forward and backward method to handle the membership degree of data. The result shows FHMM is better than HMM and the improvement was 3.33 %.

Keywords : Fuzzy, Hidden Markov Models, Indonesian, Speech, Classification

1. INTRODUCTION

This study is dealing with speech classification. This introduction is organized as follows: Chapter one has seven sections; they each discuss the rationale of the study, the theoretical framework approach, the conceptual approach, the problems, hypotheses, assumptions, scopes and delimitations and the significance of the study successively.

2.1 Rationale

Over the past several decades, the speech classification technology has been much done. There are many approaches to Speech Classification for example template-based, knowledge based, and stochastic-based approaches [1, 18]. The successful results were the hidden Markov model (HMM) [10]. Other results were Artificial Neural Network [14], Support Vector Machine [13], Fuzzy [17] and Clustering [20].

Speech classification is a "language-dependent" system. The application of classification in a language cannot be applied into another language because each language has a list of phonemes. A lot of studies have been carried out abroad, but it cannot be applied well in Indonesian. English speech classification is the most speech classification system, it has been developed in references [1, 2, 3, 9, 10, 11, 12, 13, 14, 15, 20]. The number of studies which has done Indonesia speech classification is still few. They have done Indonesian speech classification based on speaker adaptation system [7], and developed the corpus of Indonesian Speech Classification [4,5].

Speech classification is difficult because speech has an unique characteristics. In different time, a same word has different form although has been spoken from same person. So 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 used to classifying speech. However, a method is needed to develop dynamic model of various templates reference speech signal to produce accurate result and for Indonesian Speech classification. This study designs it. The solution combines Fuzzy on Hidden Markov Models. Actually, some study has combined fuzzy on HMM [6, 9, 10, 18] 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 will be designed to substitute the vector quantization process and a new forward and backward method to handle the membership degree of data.

2.2 Theoretical Framework

Input data Fuzzy Hidden Markov Models (FHMM) have degree of membership for each cluster. To apply it, Fuzzy C-Means Clustering is applied in the preprocessing step. In this process, FCM produces codebook. Input data is changed by codebook and produces the degree of membership of the frame for each cluster. The data input is processed forward and backward algorithm in FHMM. The forward and backward of this study is called fuzzy forward and backward algorithm.

2.3 Conseptual Framework

There are three variables applied to conduct measurement on this study, namely:

Table 1.1 Variables

Variable	Variable's information
The number of cluster	The variable indicates the number of codebook or recognized sound. It is used in Fuzzy C-Means Clustering.
The power (w)	The variable indicates range of each cluster in Fuzzy C-Means Clustering.
The number of state	The number of state to construct Fuzzy Hidden Markov Models

2.4 Identification Problems

The main problems discussed in this study are:

a. Applying speech classification for the Indonesian language.
b. Applying Fuzzy C-Means Clustering to get codebook
c. Applying a new forward and backward algorithm to process the data with membership degree.
d. Classifying the speech to the class label. The input data format is file of *.wav and the output is text.
e. Comparing the FHMM's performance with HMM. The HMM's program in this study is the program of other study [8].

2.5 Hypotheses

HMM observation data is the quantization results of feature extraction data. The observation data of HMM does not represent the real condition. The real condition

observation data sometimes belong to some clusters. Therefore this study proposed using fuzzy method. The proposed method will construct with fuzzy clustering. The output of fuzzy clustering is membership degree of the clusters for each data. To implement this output, a new forward and backward algorithm is needed. This algorithm is called fuzzy forward and backward algorithm and the system is called Fuzzy Hidden Markov Models (FHMM).

From these premises above, FHMM is more accurate to classify speech than HMM because the data input will be suitable to represent data with different characteristic for each speech.

2.6 Assumptions

The assumptions of this study are:

a. To implement Indonesian speech, the data contains a list of city in Indonesia and in small size (630 files of .wav).
b. There will be no noise in speech.
c. The normalization and feature extraction of system will be taken from other study. It is the thesis of bachelor degree's student which has been tested [8]. Therefore, this study will not discuss both processes.

2.7 Scopes and Delimitations

The scopes of this study are:

a. Speech classification for Indonesian.
b. Speech classification for isolated word.
c. Speech classification for speaker independent system.
d. Speech classification for matching the level of speech
e. MFCC is implemented for feature extraction.
f. Frequency of sampling is 8000 hertz.

2.8 Importance of the study

This study has four objectives, namely:

a. Developing speech classification for Indonesian
b. Applying Fuzzy C-Means Clustering to substitute vector quantization
c. Developing a new Fuzzy Hidden Markov Models, especially a new fuzzy forward and backward algorithm which implement data with membership degree for each cluster.

5. CONCLUSIONS AND RECOMMENDATIONS

This chapter discusses the conclusion of the study that proves that most of the numerical results of this study are stated in chapter 4 and the methods applied in this study was shown to be reasonable accurate. Based on the results of the study some recommendations are provided.

5.1 Conclussions

From the analysis of the performance of FHMM by using the data in this study, it can be concluded that the optimal condition of FHMM to obtain a good accuracy in this study are the number of cluster is 32, the number of state is 7, and the power (w) is 1.05. With this optimal condition, the FHMM's accuracy is 92.50 % and it is better than the HMM's accuracy, the improvement is 3.33 %.

5.2 Recommendations for future works

Since our method is an effective way to Indonesian Speech Classification, it is strongly recommend the use of FHMM for a bigger database, and the implementation of more efficient time complexity on FHMM. The proposed method needs longer time than HMM. Other recommendation is the use of bigger frequency on FHMM.

REFERENCES

[1] B.H. Juang and Lawrence R. Rabiner: Automatic Speech Recognition – A Brief History of The Technology Development. Georgia Institute of Technology and Rutgers University and The University Of California, 2004.

[2] B.H. Juang and Lawrence R. Rabiner: Fundamentals of Speech Recognition, Prentice-Hall International, Inc, 1993.

[3] B. H. Juang; L. R. Rabiner, Hidden Markov Models for Speech Recognition, Technometrics, Vol. 33, No. 3., pp. 251-272. 1991.

[4] Dessi Puji Lestari,Koji Iwano, Sadaoki Furui: A Larger Vocabulary Continuous Speech Recognition System for Indonesian Languange, 15th Indonesian Scientific Conference in Japan Proceedings, ISSN: 1881-4034, 2006.

[5] Fiarni, Cut: Analisis Efektifitas Pengembangan Korpus Speech Recognition Bahasa Indonesia Menggunakan Pendekatan Interatif, Master's Thesis Of Institut Teknologi Bandung, 2007.

[6] Harun Uguz , Ali Ozturk, Rıdvan Saracoglu, and Ahmet Arslan: A Biomedical System Based on Fuzzy Discrete Hidden Markov Model for The Diagnosis of The Brain Diseases, Expert Systems With Applications 35 1104–1114, 2008.

[7] Hammam Riza, and Oskar Riandi: Toward Asian Speech Translation System: Developing Speech Recognition and Machine Translation for Indonesian Language, International Joint Conference on Natural Language Processing, 2008.

[8] Jarwadi: Speech to Text Menggunakan Database Diphone dalam Bahasa Indonesia dengan Metode Pendekatan Hybrid Hidden Markov Models dan Algotitma Genetika, Institut Teknologi Telkom, 2008.

[9] Jia Zeng, and Zhi Qiang Liu: Interval Type-2 Fuzzy Hidden Markov Models, Center for Media Technology, School of Creative Media, City University of Hong Kong, 2004.

[10] Jia Zeng And Zhi-Qiang Liu: Type-2 Fuzzy Hidden Markov Models to Phoneme Recognition, Center for Media Technology, School of Creative Media, City University Of Hong Kong, IEEE, 2004.

[11] Julianto, Rexcon: Implementasi Model Hidden Markov pada Perangkat Lunak Pengenalan Suara Berbahasa Indonesia dengan Kata Terisolasi and Ketergantungan pada Pembicara, Jurusan Teknik Informatika Institut Teknologi Telkom, 2004.

[12] Journal: Lawrence R. Rabiner: A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, Proceedings of the IEEE, Vol.77, No.2, 1989.

[13] Lei Chen, Sule Gunduz, and M. Tamer Ozsu: Mixed Type Audio Classification with Support Vector Machine, IEEE, 2006.

[14] Peter Merkx And Jadrian Miles: Automatic Vowel Classification in Speech, An Artificial Neural Network Approach Using Cepstral Feature Analysis, Final Project For Math 196s: Dr. Garrett Mitchener, 2005.

[15] Rabiner, Lawrence R.: A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, Procedings Of The IEEE,, Vol. 77, No.2, February 1989.

[16] Rahadian, Dilla: Impelementasi Teknologi Speech pada Aplikasi Asisten Pribadi Komputer dengan Speech Recognition Engine Berbasis Hidden Markov Model, Departemen Teknik Informatika Institut Teknologi Telkom, 2009.

[17] Ramin Halavati, Saeed Bagheri Shouraki, Mahsa Eshraghi, Milad Alemzadeh: A Novel Fuzzy Approach to Speech Processing, 5th Hybrid Intelligent Systems Conference, 2004.

[18] Sinout D. Shenouda, Fayez W. zaki, Amr Goneid: Hybrid Fuzzy HMm System for Arabic Connectionist Speech Recognition, Proceedings of the 5th WSEAS International Conference on Signal Processing, robotics and Automation, pp 64-69, 2006.

[19] Sri Kusumadewi, and Hari Purnomo: Aplikasi Logika Fuzzy untuk Pendukung Keputusan, Penerbit Graha Ilmu, 2004.

[20] Stephen E. Levinson, Lawrence R. Rabiner, Aaron E. Rosenberg, and Jay G. Wilpon: Interactive Clustering Techniques for Selecting Speaker-Independent Reference Templates For Isolated Word Recognition. IEEETransactions on Acoustics, Speech, and Signal Processing, Vol. Assp-27, 1979.

[21] Xianghua Xu, Jie Zhu, and Qiang Guo: Fuzzy Clustering Based Parameter Clustering Method in Mandarin Speech Recognition, Journal of The Chinese Institute of Engineers, Vol. 28, No. 5, Pp. 887-891, 2005.