

Distance Method with PCA based Speaker Identification

Noor Khaled

Department of Computer Science,
Al-Mustansiriyah University, Baghdad, Iraqi

ABSTRACT

This paper is dedicated to present statistic method with dimensionality reduction for speaker identification. LPC is involved to formed set of descriptors for the speaker data. The formed descriptors vectors are evaluated to realize dimensionality reduction by PCA. This set of selected features is recognized by distance method. The experimental result revealed that the proposed work is achieved in better performance than without reduction.

Keywords

Pattern Recognition, Dimensionality Reduction.

1. INTRODUCTION

Oral communication which achieved through speech is the commonest and oldest means of communication [1].several types of information can be transmitted by speech signals covering personal information about different aspect of life [2] .

To discover a person identity based on speaker recognition system, it uses a person's speech properties to find out his/her identity [3].The rest of paper has been organized as follows: Section 2 talks about speaker recognition system and the techniques have been used in our work in these stages i.e. LPC technique and PCA algorithm. Section 3 explains in detail the proposed speaker recognition system along with the block diagram. Finally, the obtained results and graphs are presented in section 4, followed by conclusion in section 5.

2. SPEAKER RECOGNITION

Speaker recognition is the process of distinguishing the person from his/her voice based on information extracted from the speaker vocal tract which it unique for every person. These variances in its structure are an essential property of the speaker; make the voice graduated under biometric identity classification [4]. It's preferred as first stage in accomplishing speaker recognition system to form vocal tract model for the person. This is can be done by forming mathematically a vocal tract model which yielded the human speech, or essentially constructed by involving similar produced characteristics [3].

Depending on task objectives, speaker recognition system can be classified into speaker identification and speaker verification, in which an identification process is to determine which one from enrolled speakers producing the speech act [5]. On the other hand, the purpose of achieving verification process is to confirm the person to be true or imposter [6].

2.1 Pre-emphasis and Windowing

The purpose of pre-emphasis process is to smooth the speech spectrum, in which high frequency components are oppressed through the human speech production methodology; a digital filter is employed as pre-emphasizer to equalize these high frequencies [7]. And it is illustrated as below [4]:

$$Y(n) = x(n) - \alpha x(n-1), \quad 0 \leq \alpha \leq 1 \quad (1)$$

Where: $y(n)$ and $x(n)$ are the output and the input of the filter respectively.

Window technique at each frame is used to reduce signal discontinuity at either end of frame. Mostly used technique is hamming window technique [7]. The Hamming window of length N is given as [4]:

$$W(n) = 0.54 - 0.46 \cos(2\pi n / (N-1)), \quad 0 \leq n \leq N-1 \quad (2)$$

2.2 Feature Extraction

Feature extraction is a methodology for evolving the speech signal into a set of features vectors. The point of this change may be to acquire a new representation which is further compact, less redundant, and most proper for a model built statistically and for the matching purpose [8].

In Linear Predictive Code (LPC) technique the source filter model is based to represent the speech signal, in which the vocal tract transfer function is represented by an all-pole filter as illustrated below [9].

$$H(z) = \frac{1}{1 - \sum_{i=1}^p a_i z^{-i}} \quad (3)$$

The fundamental theory to this method is that the system production is estimated as linear combination of previous introduced samples considering that the input is unidentified, this can be illustrated as in the equation below:

$$\hat{Y}(n) = -\sum_{k=1}^p a(k)x(n-k). \quad (4)$$

where $x(\cdot)$ and $\hat{y}(\cdot)$ are represent the input samples of speech and the outputs of them respectively and $a(k)=[a(1), a(2), \dots, a(p)]^T$ is the LPC coefficients and p is the order of linear predictive [10].

2.3 Features Selection

Features selection is a mechanism of choosing only the most significant features and employs them to build a certain model. It is also named attribute selection or variable selection [11].

In this work Principal Component Analysis Algorithm (PCA) is used for features selection process. The steps of performing PCA algorithm are [12]:

Let x_1, x_2, \dots, x_M are $M \times 1$ vectors

Step1: Calculate mean:

$$\bar{x} = \frac{1}{M} \sum_{i=1}^M x_i \quad (5)$$

Step2: Subtract the mean:

$$\phi_i = x_i - \bar{x} \quad (6)$$

Step3: From the matrix $[\phi_1, \phi_2 \dots \phi_M]$ ($N * M$ matrix), then compute:

$$C = \frac{1}{M} \sum_{n=1}^M \phi_n \phi_n^T = AA^T \quad (7)$$

Step4: Find the values of the eigenvalues of c: $\lambda_1 > \lambda_2 > \dots > \lambda_N$

Step5: Find the values of the eigenvectors of c: $u_1, u_2 \dots u_N$

Step6: (dimensionality reduction step) maintain only the values meet to the K largest eigenvalues.

To select K, use the following equation

$$\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^N \lambda_i} > 0.9 \quad (8)$$

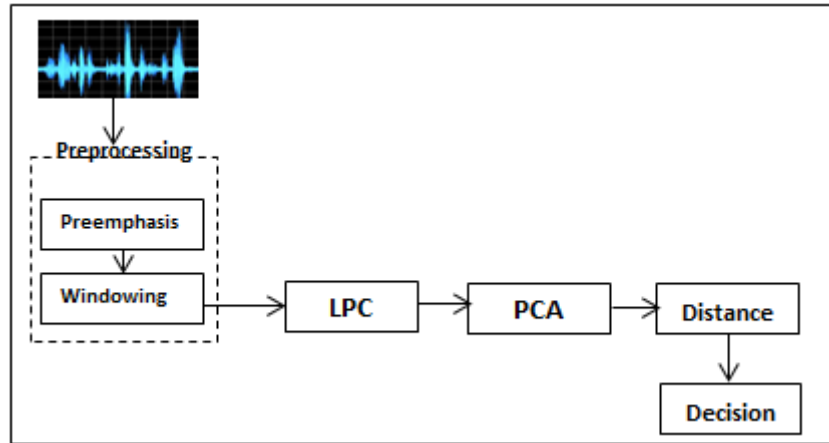


Fig 1: Layout of proposed system

Pre-emphasis process is achieved on input signal by applying a high pass filter which is given in the equation (1). The value of α in this equation is setting to 0.97. Speech signal is segmented into fixed duration which named frames, each with of N samples. The purpose of this segmentation is to achieve stationary. This fixed duration value which equivalent to frame length is tested to find its suitable value which leads to the best recognition rate; experimentally 32ms is adopted in this system. The energy value of each frame is tested according to certain threshold value to fix silent region for removing it. Hamming window technique is applied as given in the equation (2) on overlapped frames with ratio of 50% of frame size. Then the features vector is extracted by LPC technique, the multidimensional feature parameters are reduced to a smaller vector and maintained the features which are more distinctive between speakers by PCA algorithm. Finally the decision is made by distance method.

4. SYSTEM PERFORMANCE

Sampling frequency of 16 KHz, sampling resolution of 16-bits, mono recording channel and recorded file format is wav have been considered to capture the speech utterances. A set of tests have been conducted on (10) speakers, (7 male and 3 female).

Fig2 presents the selection process effect on accurate rate. The accuracy rate by using PCA algorithm which satisfy reduction process is better than when using all features.

3. PROPOSED SYSTEM

Fig 1 presents the main layout of this proposed system.

As shown in the figure this work conducts the following main stages:

1. Preprocessing speech signal stage has been done on input signal.
2. Features extraction stage has been applied by LPC technique.
3. Analysis and selection process has been done by PCA.
4. Distance method is used in recognition phase.

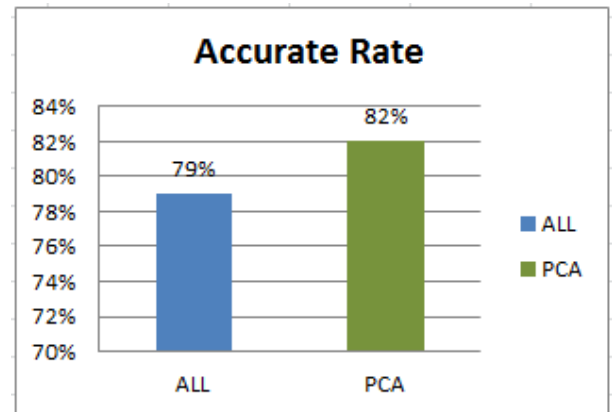


Fig 2: Comparison Accurate Rate with Selection Process and with all features

5. CONCLUSIONS

Some conclusions have been drawn from the obtained test results:

- 1- The results of test conducted on recognition based on feature analysis reveal that there are an exact number of features that lead to high accuracy rate.
- 2- The work with exact number of features which satisfies the dimensionality reduction has led to reduce the computation complexity and finally the time consuming.

6. REFERENCES

- [1] Bansod N.S., Seema Kawathekar, Dabhade S.B., "Review of Different Techniques for Speaker Recognition System", *Advances in Computational Research*, Vol.4, No.1, ISSN: 0975-3273, pp. 57-60, 2012.
- [2] Qin Jin," Robust Speaker Recognition", PhD thesis, Carnegie Mellon University, 2007.
- [3] Jucheng Yang, "Biometrics", In Tech, London, 2011.
- [4] Sejal Shah, Archana Bhise, "Fast Speaker Recognition using Efficient Feature Extraction Technique", *International Journal of Computer Science and Network*, Vol.2, No.1, ISSN: 2277-5420, pp.357-362, 2013.
- [5] Arun Kumar G, Harish H.M, Jagadisha N, Srinivasa Raoudara, Spoorti Jainar, "Design and Development of Text-independent Speaker Identification System", *IJASCSE*, Vol.2, No.3,pp.7-14, 2013.
- [6] Nakhat Fatima , Thomas Fang Zheng, "Short Utterance Speaker Recognition", In Proceeding of the 2012International Conference on Systems and Informatics, pp.1746-1750, 2012.
- [7] Rajiv Chechi,Reetu," Performance Analysis of MFCC And LPCC Techniques In Automatic Speech Recognition", *International Journal of Engineering Research & Technology (IJERT)*, Vol.2, No.9, ISSN: 2278-0181, pp.3142-3146, 2013.
- [8] Nisha.V.S, M. Jayasheela, "Survey on Feature Extraction and Matching Techniques for Speaker Recognition System", *International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE)*, Vol.2, No.3, pp.357-362, 2013.
- [9] Thapanee Seehapoch, Sartra Wongthanavas, "Speech Emotion Recognition Using Support Vector Machines", In proceeding of the 5th international conference on knowledge and smart technology,pp.86-91, 2013.
- [10] Yusnita MA, Paulraj MP, Sazali Yaacob, Shahrman AB ,Nor Fadzilah Mokhtar, " Statistical Formant Descriptors with Linear Predictive Coefficients for Accent Classification" , In proceeding of the 8th IEEE Conference on Industrial Electronics and Applications, pp. 906 – 911, 2013.
- [11] P.Abinaya, Dr.J.Sutha," Effective Feature Selection For High Dimensional Data using Fast Algorithm", *International Journal of Advanced Research in Computer Science & Technology (IJARCST)*, Vol. 2, No.6, ISSN : 2347 – 9817, pp.1790-1799, 2014.
- [12] Shilpa lakhina, Sini Joseph , Bhupendra verma," Feature Reduction using Principal Component Analysis for Effective Anomaly-Based Intrusion Detection on NSL-KDD", *International Journal of Engineering Science and Technology*, Vol. 2, No.6, pp.1790-1799, 2010.