Journal of Fiber Bioengineering and Informatics 8:1 (2015) 179–193
doi:10.3993/jfbi03201518

A Preliminary Study on the Feature Distribution of Deceptive Speech Signals ⋆

Xinyu Pan [a,b,*], Heming Zhao [a], Yan Zhou [a], Cheng Fan [a]
Wei Zou [a], Zhiqiang Ren [a], Xueqin Chen [a]

[a] *School of Electronics and Information Engineering, Soochow University, Suzhou 215006, China*

[b] *School of Electronics & Information Engineering, Suzhou University of Science and Technology Suzhou 215009, China*

Abstract

A preliminary study is conducted to compare the feature distribution between normal and deceptive speech, and the results are reported in this paper. The objective of this research is to show that deceptive speech may be recognized through the acoustic parameters of general speech characteristics. Six speech parameters, i.e., Mel-frequency Cepstral Coefficients (MFCC), Relative Spectral Filter Perceptual Linear Prediction (RASTA-PLP), pitch frequency, time-domain samples, zero-crossing rate and fractal dimension are used in the statistics. The distributions of these parameters indicate clear differences between the two speech styles. The lowest average degree of difference for these features was 4.74%, and the highest degree was over 20%. Therefore, the distribution demonstrates that there is significant distinction between speech relating the truth and speech relating falsehoods. Linear Discriminant Analysis (LDA) and the Gaussian Mixture Model (GMM) are used to recognize the two psychological states of people's pronunciation, with accuracy above 50%. The results show that there is in fact deceptive information in speech signals and that it can be detected by pattern recognition. These findings provide the theoretical basis for detecting deception in speech signals.

Keywords: Deceptive Speech; Feature Distribution; MFCC; RASTA-PLP; LDA; GMM

1 Introduction

Speech signals include information regarding gender, identity, language and age. A speaker's mental status can be the subject of speculation by listeners throughout the conversation. However, mental status is very difficult to discern by means of computer identification because of limited

⋆This work is supported by the National Natural Science Foundation of China (No. 61071215, 61372146, 61340004), the Postgraduate Research Innovation Project of Jiangsu Province, China (No. CXZZ12_0815, CXZZ13_0812), the Natural Science Foundation of Jiangsu Province, China (No. BK20131196, BK20130324), Specialized Research Fund for the Doctoral Program of Higher Education (SRFDP) (No. 20123201120009), and Natural Science Foundation of the Jiangsu Higher Education Institutions of China (No. 12KJB510029).

*Corresponding author.
Email address: panxy@mail.usts.edu.cn (Xinyu Pan).

information. A number of successful studies show that emotions and even personality traits can be perceived by a computer through speech signals [1-3]. An instrument that might distinguish between lies and the truth would be a great contribution to many security applications, among other types of applications. Thus, it is of interest to distinguish between truth and deception through speech signals.

Traditional measurements have been used to judge the truth or falsity of voices, such as Psychological Stress Evaluators (PSE) and Voice Stress Analyzers (VSA). Additionally, Layered Voice Analysis (LVA) based on brain activity can also provide evidence regarding lying. However, the results from experiments utilizing these measures indicate that there is a great deal of room for improvement [4-7].

Only a limited amount research has evaluated the importance of different speech characteristics in detecting deceptive speech. Although the F1, F2 and F3 vowel formants cannot establish a correlation between truthfulness and deceptiveness, there may yet be a way to detect deception using acoustic parameters [8].

There is no adequate theoretical basis for lie detection using only voice signals [9], but feature selection can determine the performance of speech-processing systems [10]. Thus, it is important to establish the scientific support that digital speech signal features contain adequate information for detecting deception. It is not clear whether all or only some features of speech signals change when people tell lies, and there is no appropriate method for measuring the degree of the mutation. The characteristics of normal and deceptive speech including Mel-frequency Cepstral Coefficients (MFCC) [11, 12], Perceptual Linear Prediction (PLP) [13], pitch frequency, time-domain samples, zero crossing rate and fractal dimensions are used in the statistical analysis of this paper. The results of the experiment show that the distribution of these features in truthful and deceptive speech acts is clearly different. To a certain extent, the findings provide a theoretical basis for lie detection using only voice signals.

This paper is organized as follows: The next section describes the speech database and briefly introduces the related speech characteristics. Section 3 presents the statistical results of all of the speech features and the analyses of the differences calculated from the distribution of these features. The deception detection results are presented in section 4. Finally, the conclusions are provided in the last section.

2 Description of the Corpus and Speech Features

The speech database and six speech features (MFCC, PLP, pitch frequency, time-domain samples, zero crossing rate and fractal dimensions) are briefly described in this section.

2.1 Corpus

The speech corpus was obtained from a game. In this game, each actor (or actress) tells a complicated story, and another participant can ask the actor (or actress) any type of question. The participant then decides on the truth or falsehood of the statement based on the answers to the questions. The deceitful actors are chosen for our speech database. Each actor's voice is selected manually. We take the actor's self-introduction as the truth and the description of the story as the deception. In this way, 30 actors, 15 males and 15 females are selected. All speech is