ISSN 1004-9037,CODEN SCYCE4
Journal of Data Acquisition and Processing Vol. 30,No. 2,Mar. 2015,pp. 399－407
DOI:10.16337/j.1004-9037.2015.02.019

http://sjcj.nuaa.edu.cn
E-mail:sjcj@nuaa.edu.cn
Tel/Fax:＋86-025-84892742

# 一种基于压缩感知的说话人识别参数分析

潘海琦[1]　杨　震[2]　徐珑婷[1]　朱俊华[1]

(1.南京邮电大学通信与信息工程学院,南京,210003; 2.南京邮电大学"宽带无线通信与传感网技术"教育部重点实验室,南京,210003)

**摘　要**:本文为在传统的说话人识别理论研究中"较少的特征参数量不能与较高的识别率共存"的难题找到了一种解决方案。本文基于压缩感知的理论,利用行阶梯观测矩阵进行信号的投影,改变了传统的梅尔频率倒谱系数(Mel-frequency cepstral coefficient, MFCC)参数,从而提出了一种新的识别参数CS-MFCC(Compressed sensing-MFCC)。该参数不仅使得参数存储量降低到少于原存储量的$1/n$($n$为行阶梯观测矩阵的压缩比),而且明显提高了系统的鲁棒性。通过仿真实验证明了当压缩比$n$为4时,平均识别率能够提高到96%以上。

**关键词**:说话人识别;压缩感知;识别率;CS-MFCC;鲁棒性

**中图分类号**:TN912.3　　**文献标志码**:A

## Parameter of Speaker Recognition Based on Compressed Sensing

Pan Haiqi[1], Yang Zhen[2], Xu Longting[1], Zhu Junhua[1]

(1. College of Telecommunication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China; 2. Key Lab of Broadband Wireless Communication and Sensor Network Technology, Ministry of Education, Nanjing University of Posts and Telecommunications, Nanjing, 210003, China)

**Abstract**: A solution is proposed to deal with the problem that "less number of features cannot coexist with higher recognition rate" in the traditional theory of speaker recognition. Ladder observation matrix projection is used to change the traditional Mel-frequency cepstral coefficient (MFCC) parameters based on compressed sensing theory, presenting a new recognition parameters named compressed sensing MFCC (CS-MFCC) parameters. These parameters make storage capacity decrease to less than $1/n$ of the original, here $n$ is the compression ratio of the line ladder matrix, and also greatly increase the robustness of the system. Furthermore simulation results prove that when $n$ is 4, the recognition rate increases to 96% above.

**Key words**: speaker recognition; compressed sensing; recognition rate; CS-MFCC; robustness

## 引　　言

比尔·盖茨认为:"以人类生物特征(指纹,语音,脸相等)进行生物认证的技术,在今后数年内将成

**基金项目**:国家自然科学基金(61271335)资助项目;江苏省普通高校研究生科研创新计划(CXZZ13-0488)资助项目。
**收稿日期**:2013-05-23;**修订日期**:2013-08-08

为 IT 产业内最为重要的技术革命。”与其他生物识别技术相比,说话人识别有着其独有的经济性、便捷性和准确性,因此被广泛的应用在身份鉴定领域。说话人识别是通过提取说话人语音信号中的特征来识别说话人身份的一种技术,说话人的语音特征提取是说话人识别的关键。常用的特征参数有:线性预测倒谱系数,梅尔频率倒谱系数(Mel frequency cepstral coefficient, MFCC)、耳蜗滤波器倒谱系数(Cochlear filter cepstral coefficients, CFCC)等等[1]。说话人识别技术虽然在发展中已经取得了一些成就,但还存在着一些问题亟待解决,例如参数存储量过大、识别系统鲁棒性不高等。

传统说话人识别技术主要包括特征提取技术、模式匹配准则及模式训练技术 3 方面。在奈奎斯特准则下的提取过程,要想实现较高的识别率往往需要提取非常大的数据量的特征参数,使得识别过程中的数据处理量过大。近年来,压缩感知(Compressed sensing, CS)理论以远小于奈奎斯特采样率的速率对信号进行采样,得到数目大大减少的观测序列,给信号采样方法带来一场新的革命。本文提出了一种基于稀疏子空间模型的说话人识别方法[2],在此基础上,提出了一种基于压缩感知理论的新型说话人识别参数,利用行阶梯矩阵观测原始序列,使得参数总存储量明显减少,同时由于行阶梯观测矩阵采用了阵列信号处理中周期信号通过简单叠加提高信噪比的方法,有效地抑制了加性噪声,从而提高了整个系统的鲁棒性,使得识别率提高到 96%以上。

## 1 说话人识别技术

### 1.1 基本原理

说话人识别,又被称为声纹识别,是指通过对说话人发出的语音信号分析和处理,自动确认说话人是否在所记录的话者集合中,以及进一步确认说话人是谁的技术。说话人识别分为说话人辨认(Speaker identification, SI)和说话人确认(Speaker verification, SV)两种。说话人辨认是通过一段语音从若干人中确认出说话人的身份,是一对多的过程。而说话人确认是讲一段语音与某一个说话人的模型比较,以确定是否是该说话人的声音,属于一对一的过程。本文研究的是说话人辨认系统。

说话人识别技术主要包括特征提取技术、模式匹配准则及模式训练技术 3 方面。其基本原理如图 1 所示[1]。

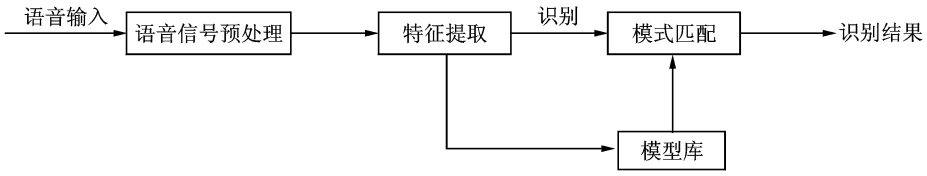



图 1 说话人识别系统原理

Fig. 1 Block diagram of speaker recognition system

### 1.2 特征参数提取

传统的说话人识别首先要对输入的模拟语音信号进行预处理,包括预滤波、采样、量化、加窗、端点检测、预加重等[3]。语音信号预处理之后,接下来很重要的一个环节就是特征参数的提取,即从说话人的语音中提取出反映说话人的身份(发音习惯、心理、生理和行为)的语音特征参数。对特征参数的要求是:

(1)提取的特征参数能有效地代表说话人语音特性,具有很好的区分性;

(2)各阶参数之间有良好的独立性;

(3)特征参数要计算方便,须有高效的计算方法来保证语音识别的实时实现。

最早提出的语音特征参数有线性预测系数(Linear prediction coefficient, LPC)、线性预测倒谱系数(Linear prediction cepstral coefficient, LPCC)等[4]。随后提出的梅尔倒谱参数 MFCC 充分描述了人耳特殊的感知特性,即对频率感知的非线性特性,大量研究表明,MFCC 参数能够比 LPCC 参数更好地提高系统的识别性能[5],因此 MFCC 参数得到了广泛地应用。

### 1.3 系统性能评价

评价说话人识别系统性能好坏的标准有很多,例如识别率、识别时间、识别样本数、鲁棒性以及存储参数量等等。其中识别率为

$$\text{识别率}=\frac{\text{匹配正确的样本数}}{\text{测试总样本数}} \tag{1}$$

## 2 基于压缩感知的说话人识别

### 2.1 压缩感知

压缩感知打破了传统的奈奎斯特采样准则的框架,是一种新型的数据采集和解编码方法。大量的研究证明[2]若有一个长度为 $N$ 的信号 $\boldsymbol{X}$ 是 $K$-稀疏的,那么它经过线性变换后仅用少量的系数便可很好的重构出来,通过压缩传感过程可以直接得到 $M$ 维信号 $\boldsymbol{Y}$,其中 $M<N$,它们之间的关系为

$$\boldsymbol{Y}=\boldsymbol{\Phi X} \tag{2}$$

式中:矩阵 $\boldsymbol{\Phi}$ 为测量矩阵,大小为 $M\times N$。另外,若信号不是绝对稀疏的,但却是可压缩的(即变换域系数虽然不是 0,但是多数很小,可忽略),即通过可逆变换 $\boldsymbol{\Psi}$ 后可以变成稀疏信号,此处 $\boldsymbol{\Psi}$ 被称为稀疏矩阵。稀疏矩阵在信号的重构中至关重要[6],但本文研究的说话人识别系统是非重构系统,所以将不考虑稀疏矩阵的选取。

### 2.2 观测矩阵

投影后的序列,根据不同投影矩阵,波形呈现不同特征,取一帧浊音信号如图 2 所示,在高斯随机矩阵观测后的波形如图 3 所示[7]。

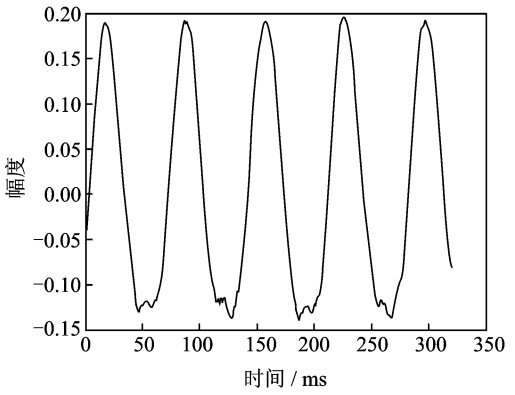



图 2 一帧浊音信号的波形图

Fig. 2 Waveform of frame voice signal

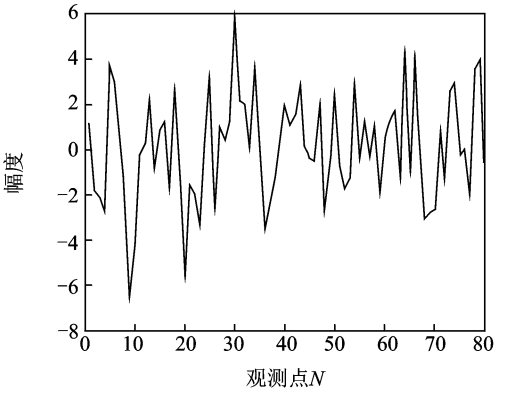



图 3 一帧浊音经高斯随机矩阵观测后的波形

Fig. 3 Waveform of frame voice signal observed by Gaussian random matrix

在文献[3]中提出了一种新的观测矩阵——行阶梯观测矩阵,压缩比为 4 的观测矩阵形如下

$$\boldsymbol{\Phi}=\begin{bmatrix} 111100000 & \cdots & 0 \\ 000011110 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 011110000 \\ 0 & \cdots & 000001111 \end{bmatrix} \tag{3}$$

图 2 为一帧浊音信号，经压缩比为 4∶1 的行阶梯矩阵观测后的波形如图 4 所示。

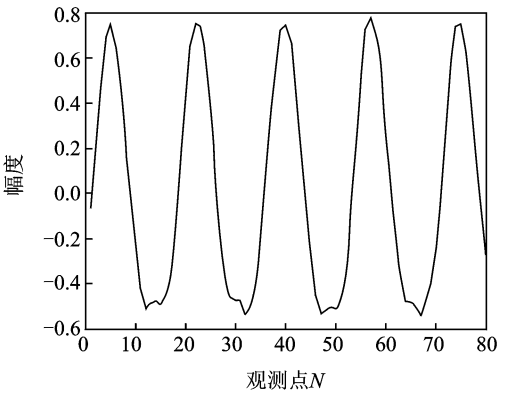



图 4 一帧浊音经行阶梯矩阵观测后的波形

Fig. 4 Waveform of frame voice signal observed by ladder matrix

通过对图 3 和图 4 的比较，发现文献[3]所提出的行阶梯观测矩阵更好的保留了原始信号波形特征。因此在本文所提出的说话人识别系统中，选用了行阶梯矩阵。

在压缩比为 4 的行阶梯矩阵下观测值 $\boldsymbol{x}$ 与原始信号 $s_0$ 关系为

$$\begin{cases} x_1 = s_{01} + s_{02} + s_{03} + s_{04} \\ x_2 = s_{05} + s_{06} + s_{07} + s_{08} \\ \vdots \qquad \vdots \\ x_i = s_{0(4i-3)} + s_{0(4i-2)} + s_{0(4i-1)} + s_{0(4i)} \end{cases} \qquad i = 1, 2, \cdots, 64 \tag{4}$$

图 5 为一段语音信号在预处理之后未通过观测矩阵的波形，以及预处理之后在压缩比为 4 的行阶梯矩阵下的观测序列波形。

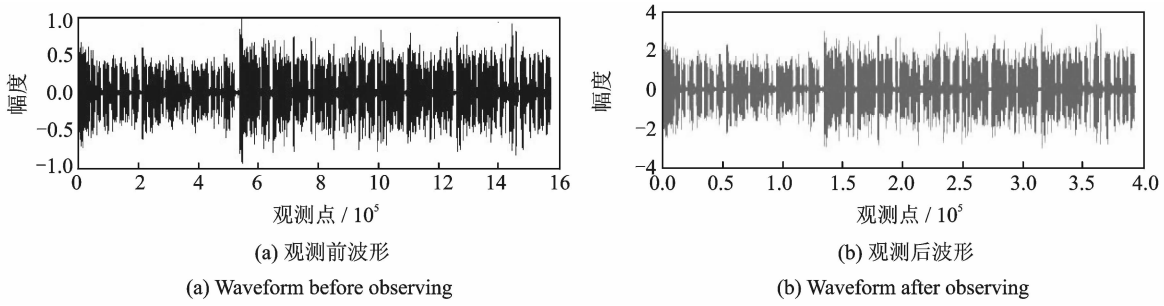



图 5 观测前后波形对比

Fig. 5 Comparison of waveform before and after observing

由图 5 可以看出,在压缩比为 4 的行阶梯矩阵观测前后,信号的波形的总体包络形状没有明显的变化,对比后发现两者区别是:

(1)观测后的采样点数是观测前的 1/4,即压缩比是 4∶1;

(2)投影矩阵实现了序列的压缩叠加,观测后的信号幅度是观测前的 4 倍。

基于压缩感知理论,本文提出了一种新的说话人识别参数 CS-MFCC。

### 2.3 CS-MFCC 参数的提取

CS-MFCC 参数在传统 MFCC 参数的基础上引入了行阶梯观测矩阵,不仅使得识别参数量明显减少,同时加入的行阶梯矩阵相当于一个"隐能量算子"[8-9]。因为语音浊音具有比较明显的周期性,通过叠加,得到进一步增强[10-11],而加性随机噪声,没有这个特性,叠加后相互抵消,因此改进后的参数对加性噪声有了抑制作用,提高了信噪比,增强了整个识别系统的鲁棒性,从而识别率大大提高[12]。

与传统的说话人识别系统[13]相似,在提取 CS-MFCC 参数之前,需要对输入的模拟语音信号进行预处理,包括预滤波、采样、量化、加窗、端点检测等,提取 CS-MFCC 的流程为:

(1)对采样后的离散信号进行首次分帧,本文取每帧样点数 $N=256$ 点;

(2)用行阶梯矩阵对分帧后的信号矩阵进行观测,得到观测矩阵,将观测矩阵一维化得到观测序列;

(3)对观测序列进行二次分帧,对每帧序列进行预加重处理后再经过离散傅里叶变换,取模的平方得到离散功率谱 $S(n)$。

(4)通过 $M$ 个 $H_m(n)$后所得的功率值,计算 $S(n)$和 $H_m(n)$在各离散频率点上的乘积之和,得到 $M$ 个参数 $P_m$,$m=0,1,\cdots,M-1$,即

$$P_m=\sum_{w=0}^{N/2}\mid S(n)\mid H_m\left(w\frac{2\pi}{N}\right) \tag{5}$$

(5)计算 $P_m$ 的自然对数,得到 $L_m$,$m=0,1,\cdots,M-1$,即

$$L_m=\log P_m \tag{6}$$

(6)对 $L_0,L_1,\cdots,L_{m-1}$计算其离散余弦变换,得到 $D_m$,$m=0,1,\cdots,M-1$。

(7)舍去代表直流成分的 $D_0$,取 $D_1,D_2,\cdots,D_k$ 作为 CS-MFCC 参数,其中 $k$ 为参数的阶数。

其中,滤波器的维数 $M$ 和 CS-MFCC 参数的阶数都要通过实验确定其最优取值。本文经大量实验验证得知,当 $M=30$,$K=11$ 时,系统识别率最佳。

### 2.4 参数提取及存储量的压缩

传统的 MFCC 的提取流程如图 6 所示[2],而改进后的 CS-MFCC 参数的提取流程如图 7 所示。当压缩比 $n$ 远远小于每帧的分帧数时,实验证明压缩后的观测序列仍然具有短时平稳特性。由前面的叙述中可以知道,信号在压缩比为 $n$ 的行阶梯矩阵观测下,采样点数是观测前的 $1/n$;若原信号采样点数为 $H$ 个,则观测后的采样点数为 $H/n$ 个。

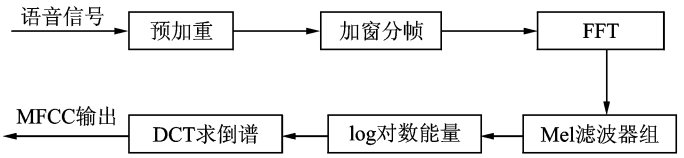



图 6 MFCC 的提取流程

Fig. 6 Process of extracting MFCC

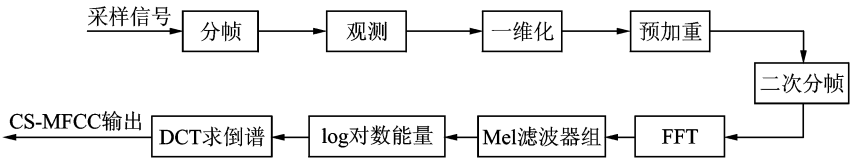



图 7 CS-MFCC 参数的提取流程

Fig. 7 Process of extracting CS-MFCC

若 CS-MFCC 参数提取中二次分帧的每帧样点数与传统 MFCC 的分帧每帧样点数相同，且每帧的 CS-MFCC 参数减少至 12 个(原始 MFCC 参数为 13 个)，则提取出来的 CS-MFCC 参数总数将小于原先 MFCC 参数的 $1/n$。从而实现了对参数提取及存储量降低的目的。图 8 描述了对一段语音提取的传统 MFCC 参数和改进后的 CS-MFCC 参数，可以看出前者数据更密集，即提取参数数据量更大。

## 2.5 识别系统的鲁棒性

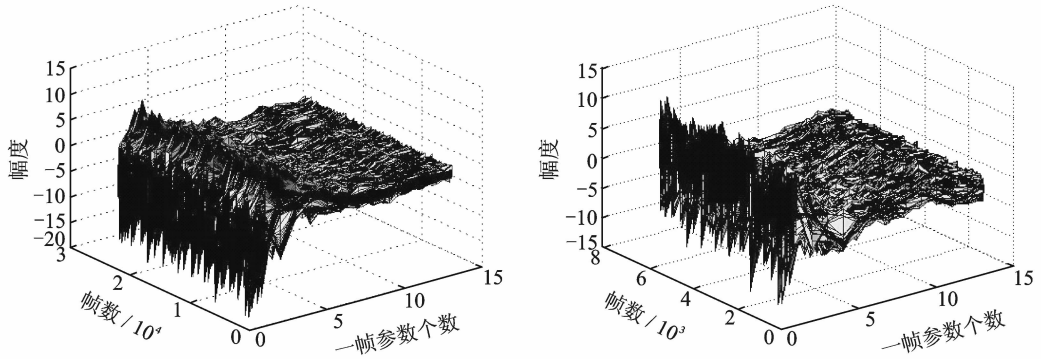
在研究传统的说话人识别过程中，通常有一个难题，即倘若减少提取参数的个数，则识别率会大大下降。但是，本文提出的 CS-MFCC 参数为这个难题找到了两全之策。

导致识别率降低的原因有很多，其中最重要的是加性噪声的存在。设原信号为 $S_0$，加入白噪声 $N$ 后得到

$$\boldsymbol{S}=\boldsymbol{S}_0+\boldsymbol{N} \tag{7}$$

此时信噪比 SNR 可以表示为

$$\mathrm{SNR}=10\lg\left(\frac{\|\boldsymbol{S}_0\|^2}{\|\boldsymbol{S}-\boldsymbol{S}_0\|^2}\right) \tag{8}$$




(a) 从语音中提取出的MFCC参数
(a) MFCC extracting from voice signal

(b) 从语音中提取出的CS-MFCC参数
(b) CS-MFCC extracting from voice signal

图 8 参数量对比

Fig. 8 Compare the number of parameters

通过上文的说明可以知道若加入压缩比为 4 的行阶梯矩阵，则提取 CS-MFCC 参数的个数可以压缩为原先的 1/4。由于原始语音浊音具有明显的周期性，根据阵列信号处理中常用的提高信噪比方法，含噪的周期性信号经过简单的叠加信噪比会得到一定程度的提高，所以经过观测后信号的信噪比大于观测前的信噪比，进而增强了系统的鲁棒性。图 9 描述了一段语音信噪比随着行阶梯矩阵的压缩比变化的趋势。其中，压缩比为 1 时即为未加观测矩阵的情况。由图 9 可以看出经过行阶梯矩阵的投影信

噪比有了明显的提高。

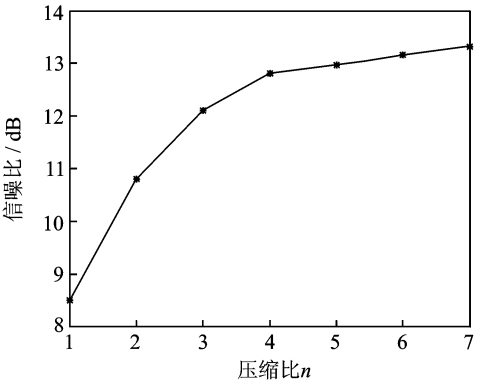



图 9 信噪比与压缩比的关系

Fig. 9 Relationship of SNR and compression ratio

## 3 实验设计及结果分析

### 3.1 CS-MFCC 参数设计

上文已经提到,带通滤波器组所含滤波器的个数 $M$ 和 CS-MFCC 参数的阶数 $K$ 要通过实验确定最优值。本文实验中选择了 PKU-SRSC 语音库中的 50 个训练样本(25 名男性,25 名女性),共 3 000 段语音,每段语音时间为 10～15 s,测试数据有纯净语音和 NOISE-92 数据库中的各种噪声在不同的信噪比下混合而成。测试组包括 2 000 段语音。通过 20 次重复试验得出了在不同的 $M$ 和 $K$ 下识别率的变化的平均结果,总结如表 1 所示。

**表 1 不同的滤波器个数和不同的 MFCC 阶数的正确识别率**

**Table 1 Recognition rate using different number of filters and order of MFCC**

| 滤波器个数 $M$/个 | CS-MFCC 参数的阶数 $K$/% | | | |
|---|---|---|---|---|
| | 10 | 11 | 12 | 13 |
| 20 | 93.44 | 95.87 | 96.11 | 96.10 |
| 25 | 93.18 | 95.92 | 96.43 | 96.42 |
| 30 | 93.33 | 96.68 | 96.65 | 96.61 |
| 35 | 93.89 | 96.12 | 96.27 | 96.31 |
| 40 | 93.79 | 96.03 | 96.32 | 96.35 |

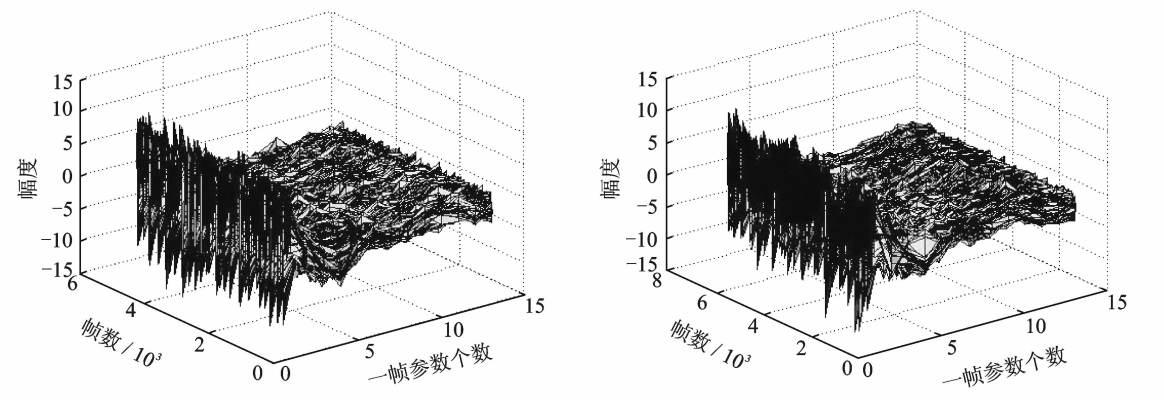
由表 1 可以看出,当固定参数的阶数 $K$ 时,增加滤波器的个数并不能对识别率有明显的提高作用;当固定滤波器的个数 $M$ 时,增加参数的阶数语音识别率提高了。这是因为,截去的部分越少,损失的原特征信息也越少。当阶数 $K$ 增加到 11 时,识别率没有了实质性的提高,这是因为 CS-MFCC 的大部分信息都保留在前 12 个参数中了,图 10 所示为从一位女测试语音和一位男测试语音提取出来的 CS-MFCC 参数图,可以看出参数波形超过 12 以后非常平缓,主要信息集中在前 12 个参数中。综上所述,在后续的识别实验中选取滤波器个数 $M$ 为 30,CS-MFCC 参数的阶数 $K$ 取为 11。

### 3.2 识别实验

#### 3.2.1 不同压缩比对识别率的影响(实验 1)

为了研究行阶梯矩阵的压缩比对系统识别率的影响,本文在信噪比为 30 dB 的情况下,对 26 个训练样本(其中男声录音和女声录音各 13 个)以及 52 个测试样本,其中每个样本时间为 10～15 s,固定每

帧观测个数为 64,实验结果如图 11 所示。




(a) 从男音中提取出的CS-MFCC参数
(a) CS-MFCC parameters from a male voice

(b) 从女音中提取出的CS-MFCC参数
(b) CS-MFCC parameters from a female voice

图 10 男音、女音 CS-MFCC 参数图

Fig. 10 CS-MFCC parameters from female and male voices

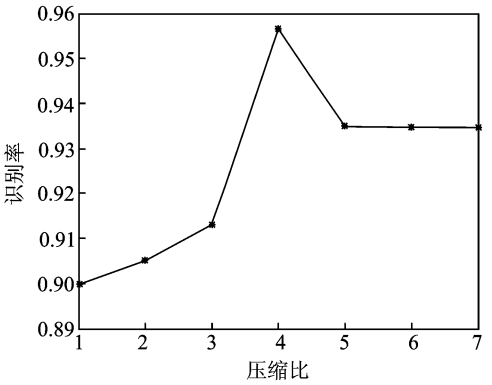



图 11 不同压缩比对识别率的影响

Fig. 11 Recognition rate after different compression ratio

由图 11 可以看出,随着行阶梯矩阵的压缩比的增大,语音识别率先提高后降低,这是因为压缩比的增大使得呈周期性的信号进一步增强,而噪声叠加后相互抵消,系统抗噪性能增强,然而压缩比增加的同时,也使得观测序列与原始序列差别越大,信号渐渐失去了短时平稳特征。综上所述,在本文所研究的系统中,最佳行阶梯矩阵压缩比为 4∶1。

另外本文还进行了无噪情况下,采用了 50 个训练样本(其中男声录音和女声录音各 25 个)以及 150 个测试样本提取 CS-MFCC 参数的识别实验,其中每个样本时间为 10～15 s,样本内容为标准英式英语的语段,经过 30 次以上的实验证明,识别率维持在 96%以上。

3.2.2 有噪情况下与 MFCC 参数的比较(实验 2)

为了探究 CS-MFCC 参数性能,本文进行了下面的识别实验,识别模型均采用高斯混合模型(Gaussian mixture model, GMM)。实验中采用了噪声数据库中的高斯白噪声,并且设置信噪比为15 dB,测试集是训练集的 1,3,5 倍时,考察采用 CS-MFCC 参数的系统识别率,参数采用压缩比为 4 的行阶梯矩阵观测,并与传统 MFCC 参数的识别率对比。实验结果是由多次实验后的平均值统计所得,具体如表 2 所示。

表 2 两种参数识别率对比

Table 2 Comparison of two parameters %

| 参数 | 测试集/训练集 | | |
|---|---|---|---|
| | 1/1 | 3/1 | 5/1 |
| MFCC | 94.22 | 92.51 | 90.28 |
| CS-MFCC | 98.68 | 97.76 | 96.70 |

由表 2 可以看出,CS-MFCC 参数的识别率明显高于 MFCC 参数的识别率。这是因为行阶梯矩阵使得信号叠加,根据阵列信号处理基本理论,原始语音浊音具有明显的周期性,而加性随机噪声并没有明显的周期性,而具有白噪声特性,因此通过叠加后信噪比得到了提高,系统鲁棒性进一步的增强。

## 4 结束语

本文基于压缩感知理论提出了一种新的说话人识别参数 CS-MFCC,该参数采用行阶梯观测矩阵,在使提取的参数量大大降低的同时,增强了系统的鲁棒性,进而提高了系统的识别率。作者通过仿真实验证明在采用同样的识别模型,并且对同样的测试人群,传统的说话人识别系统识别率达到 90%,然而采用压缩比为 4 的行阶梯矩阵提取的 CS-MFCC 参数,使得系统识别率达到了 96%以上。

**参考文献:**


[1] 何强,何英. MATLAB 扩展编码[M]. 北京:清华大学出版社,2002.
He Qiang, He yin. MATLAB expand programming[M]. Beijing: Tsinghua University Publishing House, 2002.

[2] Xu Longting, Yang Zhen. Speaker identification based on sparse subspace model[C]//APCC2013, 19th Asia-Pacific Conference on Communications. Bali, Indonesia: IEEE,2013:37-41.

[3] 叶蕾,杨震,王天荆,等. 行阶梯观测矩阵、对偶仿射尺度内点重构算法下的语音压缩感知[J]. 电子学报,2012,40(3):429-434.
Ye Lei, Yang Zhen, Wang Tianjin, et al. Compressed sensing of speech signal based on row echel on measurement matrix and dual affine scaling interior point reconstruction method[J]. ACTA Electronic a Sinica,2012,40(3):429-434.

[4] Zhang X, Guo Y, Hou X. A speech recognition method of isolated words based on modified LPC cepstrum[C]//IEEE International Conference on Granular Computing. San Jose, California, USA: [s. n.], 2007:481-485.

[5] Hosseinzadeh D, Krishnan S. Combining vocal source and MFCC features for enhanced speaker recognition performance using GMMs[C]//9th IEEE Workshop on Multimedia Signal Processing. Chania, Greece: [s. n.], 2007:365-368.

[6] 张贤达. 现代信号处理[M]. 2 版. 北京:清华大学出版社,2002:188-206.
Zhang Xianda . Modern signal processing [M]. 2nd Ed. Beijing:Tsinghua University Publishing House, 2002:188-206.

[7] Donoho D. Compressed sensing[J]. IEEE Trans Information Theory,2006,52(4):1289-1306.

[8] Prabhakar S, Pankanti S, Jain A K. Biometric recognition: Security and privacy concerns[J]. IEEE Security & Privacy Magazine, 2003(1):33-42.

[9] Borgstrom B J, Alwan A. Utilizing compressibility in reconstructing spectrographic date, with applications to noise robust ASR[J]. IEEE Signal Processing Letters, 2009,16(5):398-401.

[10] Bouzid M. Robust quantization of LPC parameters for speech communication over noisy channel[C]//Second International Conference on the Applications of Digital Information and Web Technologies. London, UK: [s. n.],2009:713-718.

[11] Gemmeke J F, Cranen B. Using sparse representations for missing data imputation in noise robust speech recognition[C]// European Signal Processing Conf (EUSIPCO). Lausanne, Switzerland: [s. n.], 2008:987-991.

[12] 张弓,杨萌,张劲东,等. 压缩感知在雷达目标探测与识别中的研究进展[J]. 数据采集与处理,2012,27(1):1-12.
Zhang Gong, Yang Meng, Zhang Jingdong, et al. Advances in theory and application of compressed sensing in radar target detection and recongnition[J]. Journal of Data Acquisition and Processing,2012,27(1):1-12.

[13] 马蜂,张宁,戴礼荣. 基于语音信号稀疏性的 FDICA 初始化和后处理方法[J]. 数据采集与处理,2012,27(2):210-217.
Ma feng, Zhang ning, Dai Lirong. FDICA initialization and post-processing method based on sparseness of speech[J]. Journal of Data Acquisition and Processing,2012,27(2):210-217.


作者简介:**潘海琦**(1989-),女,硕士研究生,研究方向:语音信号处理,E-mail:1011010416@njupt.edu.cn;**杨震**(1961-),男,博士生导师,研究方向:现代网络通信、语音处理与现代语音通信;**徐珑婷**(1989-),女,博士研究生,研究方向:基于压缩感知的语音信号处理;**朱俊华**(1989-),男,硕士研究生,研究方向:基于压缩感知的语音信号处理。