

## 基于稀疏表示权重张量的音频特征提取算法

林 静<sup>1\*</sup>, 杨继臣<sup>2</sup>, 张雪源<sup>2</sup>, 李新超<sup>2</sup>

(1. 茂名职业技术学院 机电信息系, 广东 茂名 525000; 2. 华南理工大学 电子与信息学院, 广州 510641)

(\* 通信作者电子邮箱 linjing80615@163.com)

**摘要:**为了更好地描述非平稳音频信号的特征,提出了一种基于 Gabor 字典和稀疏表示权重张量的时-频音频特征提取方法。该方法基于 Gabor 字典将音频信号编码为稀疏的权重向量,并进一步将权重向量中的元素重新排列为张量形式,该张量各阶分别刻画了信号的时间、频率以及时长特性,为信号的联合时-频-长表示。通过对该张量进行因子分解,将分解后得到的频率因子和时长因子拼接为音频特征。针对稀疏张量分解时容易产生过拟合的问题,提出一种自调整惩罚参数分解算法并进行了改进。实验结果显示,所提出的特征相对于传统梅尔倒谱系数(MFCC)特征、MFCC 特征及匹配追踪算法(MP)求解的特征联合拼接得到的 MFCC + MP 特征和非均匀尺度-频率图特征对 15 类音效分类效果分别提升了 28.0%、19.8% 和 6.7%。

**关键词:**稀疏表示;张量因子分解;音效分类;时-频特征

**中图分类号:**TN912.3 **文献标志码:**A

### Audio feature extraction algorithm based on weight tensor of sparse representation

LIN Jing<sup>1\*</sup>, YANG Jichen<sup>2</sup>, ZHANG Xueyuan<sup>2</sup>, LI Xinchao<sup>2</sup>

(1. Department of Mechanical and Electrical Information, Maoming Vocational and Technical College, Maoming Guangdong 525000, China;

2. School of Electronic and Information Engineering, South China University of Technology, Guangzhou Guangdong 510641, China)

**Abstract:** A joint time-frequency audio feature extraction algorithm based on Gabor dictionary and weight tensor of sparse representation was proposed to describe the characteristic of non-stationary audio signal. Conventional sparse representation uses a predefined dictionary to encode the audio signal as sparse weight vector. In this paper, the elements in the weight vector were reorganized into tensor format. Each order of the tensor respectively characterized time, frequency and duration property of signal, making it the joint time-frequency-duration representation of the signal. The frequency factors and duration factors were concatenated as audio features through tensor decomposition. To solve the over-fitting problem of sparse tensor factorization, an automatic-adjust-penalty-coefficient factorization algorithm was proposed. The experimental results show that the proposed feature outperforms MFCC (Mel-Frequency Cepstrum Coefficient) feature, MFCC + MP feature concatenated by MFCC and Matching Pursuit (MP) features, and nonuniform scale-frequency map feature by 28.0%, 19.8% and 6.7% respectively, in 15-category audio classification.

**Key words:** sparse representation; tensor factorization; audio effect classification; time-frequency feature

## 0 引言

传统的语音特征通常从频域或者倒谱域对信号进行表示,这些特征均假定在一短帧内信号的统计特性平稳,即信号具有短时平稳性。但是这些特征不适用于刻画非平稳的音频信号,例如闪电声和枪声。为了描述统计特性随时间变化的信号,基于稀疏表示(Sparse Representation)的方法受到了广泛的关注。稀疏表示利用一个过完备的字典,将信号编码为一个稀疏的权重向量序列,每个向量中的非零元素值表示了重建时对应原子的权重。有许多文献使用该权重向量进行信号分类。Zubair 等<sup>[1-2]</sup>直接使用该权重向量作为特征进行音频信号分类,并且通过使用最大投票和平均投票等方法提取更为鲁棒的特征。Chu 等<sup>[3]</sup>使用匹配追踪(Matching Pursuit, MP)算法求解稀疏表示,并且利用分解后原子的频率

和时长的均值和方差信息作为梅尔倒谱系数(Mel-Frequency Cepstrum Coefficient, MFCC)特征的补充特征。Sivasankaran 等<sup>[4]</sup>对 Chu 算法进行了改进,提出使用权重计算加权的频率和时长平均值和方差作为特征。在上述文献中,均使用了原子参数的低阶统计量表征信号的频率和时长分布特性。但是,由于稀疏表示的目的是使用具有区分性的原子表示信号,而计算这些原子的均值和方差则消除了单个原子的表征能力。针对此问题,Wang 等<sup>[5]</sup>提出了非均匀尺度-频率图(Nonuniform Scale-Frequency Map)特征提取方法,他们构建了一个频率-时长图来保存每个原子的频率和时长参数,将该图向量化后利用主成分分析(Principal Component Analysis, PCA)算法和线性判别式分析(Linear Discriminant Analysis, LDA)算法对其进行降维,但该方法的主要不足是向量化丢失了二维的信息,以及没有使用对于时变信号分类具有重要作

收稿日期:2015-10-14;修回日期:2016-01-18。 基金项目:国家自然科学基金资助项目(61301300)。

作者简介:林静(1982—),女,广东茂名人,讲师,硕士,主要研究方向:音视频信号处理; 杨继臣(1980—),男,安徽界首人,副研究员,博士,主要研究方向:音视频信号处理; 张雪源(1987—),男,河北石家庄人,博士研究生,主要研究方向:音视频信号处理; 李新超(1980—),男,河南南阳人,博士研究生,主要研究方向:智能优化、信号处理。

用的时间信息。

本文提出稀疏表示的张量(Tensor)形式,称为权重张量。根据 Gabor 字典的 3 个参数,即时间、频率和时长,将权重排列到 3-阶张量中,权重张量中每一个元素都表示了信号中某个瞬时成分的中心频率、中心时间位置以及持续时长,其中频率、时间以及时长信息通过元素在张量中的索引表示,该瞬时成分的强度通过元素的值表示,因此该张量是信号的联合时-频-长表示。该张量表示相比于直接获取特征向量的方法具有可以联合分析信号成分的优势,具体地,张量中加入了信号的时间信息,该时间信息可以描述信号的时变特性,即瞬时成分出现的不同时间位置,该位置对于瞬时信号分类具有重要作用。但是由于其高维和稀疏特性,该张量不适合直接用作特征,因此通过张量因子分解算法将其分解为若干因子向量,并提出了一种自调整惩罚参数的分解算法以避免因子分解的过拟合。本文提出的张量特征通过时-频-长联合分析保证了张量分析过程中保留了瞬时成分的主要特性,从而得到表征能力更强的频率因子和时长因子。

### 1 权重张量的音频特征提取

本文所提出的特征提取流程如图 1 所示。

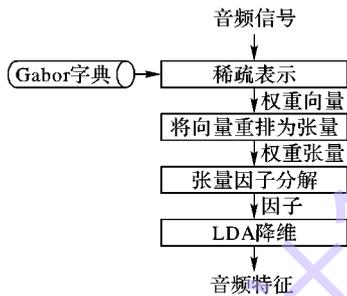


图1 音频特征提取流程

#### 1.1 基于 Gabor 字典的权重向量稀疏表示

Gabor 原子定义为

$$g_{\mu, f, \sigma}(t) = \frac{\lambda_{\mu, f, \sigma}}{\sigma \sqrt{2\pi}} \exp\left\{-\frac{(t-\mu)^2}{2\sigma^2}\right\} \cos[2\pi f(t-\mu)] \quad (1)$$

其中:  $\mu$ 、 $f$  和  $\sigma$  分别为时间、频率和时长参数,  $\mu$  决定原子出现的中心位置,  $f$  表示原子的中心频率,  $\sigma$  控制原子的时长, 每个参数只能从有限的集合中取值, 取值集合记为  $\mu \in \{\mu_1, \mu_2, \dots, \mu_i, \dots, \mu_l\}$ ,  $f \in \{f_1, f_2, \dots, f_j, \dots, f_j\}$  及  $\sigma \in \{\sigma_1, \sigma_2, \dots, \sigma_k, \dots, \sigma_k\}$ ;  $\lambda$  将原子归一化为单位能量。原子定义在短时帧内, 其中心时间位置以帧的起点作为原点, 通常时间位置以线性刻度分布在帧内。字典中原子的个数  $N$  由 Gabor 函数的参数取遍其集合内的所有元素决定, 即共有  $IJK$  个原子, 字典  $C$  可以表示为:

$$C = [g^{(1)}, g^{(2)}, \dots, g^{(n)}, \dots, g^{(N)}] = [g_{\mu_1, f_1, \sigma_1}, g_{\mu_1, f_1, \sigma_2}, \dots, g_{\mu_i, f_j, \sigma_k}, \dots, g_{\mu_l, f_j, \sigma_k}]$$

其中:  $n$  为 Gabor 原子在字典中的索引;  $i$ 、 $j$  和  $k$  为参数的索引, 索引之间的映射关系为:

$$n = (i-1)JK + (j-1)K + k \quad (2)$$

由此得到信号  $z$  的稀疏表示:

$$z = Cx + e = \sum_{n=1}^N x_n g^{(n)} + e; \|x\|_0 \leq N_0 \quad (3)$$

其中:  $e$  为残差,  $x$  为权重向量,  $x_n$  是  $x$  中的第  $n$  个元素,  $\|\cdot\|_0$  为  $l_0$ -范数,  $N_0$  定义了权重向量的稀疏度。

#### 1.2 基于权重张量的稀疏表示

首先构建 3-阶张量  $T \in \mathbf{R}^{I \times J \times K}$  将张量元素  $t_{i,j,k}$  的索引记为  $(i, j, k)$ 。之后, 将权重向量中的元素值赋给权重张量元素, 向量和张量元素之间的映射关系由式(2)得:

$$t_{i,j,k} = x_n; n = (i-1)JK + (j-1)K + k \quad (4)$$

式(3)等价于:

$$z = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K t_{i,j,k} g_{\mu_i, f_j, \sigma_k} + e \quad (5)$$

该映射关系如图 2 所示。图中的每个立方体表示一个张量元素。该过程使得张量中元素的索引与原子参数的索引相一致。由图可见, 首先对短时音频帧使用稀疏表示, 以获得 Gabor 原子的权重向量  $x$ , 之后通过对应的 Gabor 原子参数索引, 将权重向量中的元素和张量元素建立一一映射。该过程等价于将向量元素重新排列为张量形式, 同时张量的各阶分别对应时间、频率和时长参数中的一个, 如图 2 中右上角坐标所示。

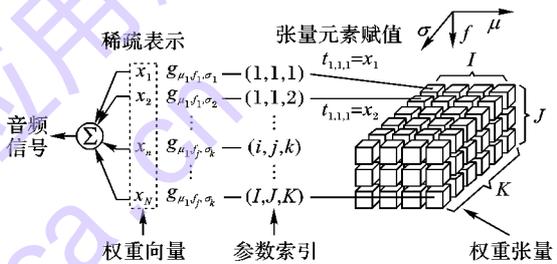


图2 权重张量

权重向量和权重张量均包含了所有原子的权重信息, 只是两者的排列方式不同。张量元素的索引和 Gabor 参数的索引相一致, 每个张量元素需要 3 个索引值来标示其位置, 即时间、频率和时长索引值, 因此张量元素的索引表达了 Gabor 原子参数之间的联合关系。该张量利用其高阶特性来联合地表达了权重与其对应原子的时间、频率和时长参数。此外, 由于权重表示了原子的幅度, 因此张量是信号的联合时-频-长表示。例如, 张量元素  $t_{i,j,k}$  表示信号中在时间  $\mu_i$ , 一个频率为  $f_j$  的分量, 其时长为  $\sigma_k$ , 强度为  $|t_{i,j,k}|$ 。

#### 1.3 张量因子分解

张量元素的绝对值表示了信号中瞬变成分的强度。对张量  $T$  中每个元素取绝对值得到非负张量  $|T|$ 。此外, 权重张量与权重向量一样均满足稀疏条件, 因此  $|T|$  为非负稀疏张量。本节中, 对该非负稀疏张量进行分解。传统的张量分解算法包括 CP (Canonical decomposition Parallel factor analysis) 算法和 Tucker-3 算法, 这些算法针对的是非稀疏的矩阵, 对稀疏矩阵分解时会产生过拟合问题<sup>[6]</sup>。通常算法会使用  $l_1$ -范数惩罚, 通过最小化输出因子的元素的和来避免过拟合<sup>[7]</sup>, 但是该方法没有指明输出因子中的哪些元素应当最小化, 因此需要通过多次随机初始化和更新过程寻找最优解。

在本文中, 张量中各部分的稀疏度被用作惩罚参数, 惩罚它们对应的输出因子元素, 因此对于不同的输出因子的元素有不同的惩罚参数。具体地, 从张量中稀疏部分分解出的因子元素的惩罚参数大于从张量紧致部分分解出的因子元素, 从而使用更多的非零因子元素近似紧致部分, 使用少量的非

零因子元素近似张量中的稀疏部分,使得在保证分解准确性的前提下避免过拟合。同时,与  $l_1$ -范数的方法相比,由于指定了每个输出因子元素的惩罚值,即指明了哪些因子元素应当最小化,因此可以在更少的循环迭代中收敛。

由于  $|T|$  的稀疏性,将其分解为因子矩阵会导致过拟合,因此将其分解为因子向量的外积,即

$$|T| \approx u \circ w \circ s \tag{6}$$

其中:  $u \in R^I, w \in R^J, s \in R^K$ , 分别为刻画时间、频率和时长的因子向量,符号  $\circ$  表示外积<sup>[8]</sup>。式(6)的元素形式表示为  $|T|_{i,j,k} \approx u_i w_j s_k$ , 其中  $u_i, w_j$  和  $s_k$  分别表示  $u, w$  和  $s$  中第  $i, j$  和  $k$  个元素。

张量  $|T|$  沿第 1-模展开为矩阵形式,记为  $|T|^{(1)}$ , 因此式(6)的矩阵形式表示为:

$$|T|^{(1)} \approx u(w \otimes s)^T \tag{7}$$

其中:  $\text{tr}$  表示转置运算,  $\otimes$  表示 Kronecker 积, 定义为  $w \otimes s = [(w_1 s)^T, (w_2 s)^T, \dots, (w_J s)^T]^T$ <sup>[8]</sup>。矩阵  $|T|^{(1)}$  的第  $i$  行记为  $|T|_{i:}^{(1)}$ , 因此式(6)的向量形式表示为:

$$|T|_{i:}^{(1)} \approx u_i (w \otimes s)^T \tag{8}$$

为求解式(6), 首先定义代价函数  $D(|T|, u, w, s)$ , 描述张量和因子之间的不相似度:

$$D(|T|, u, w, s) = \| |T| - u \circ w \circ s \|_F^2 + \sum_{i=1}^I \alpha S(|T|_{i:}^{(1)}) u_i + \sum_{j=1}^J \beta S(|T|_{:j}^{(2)}) w_j + \sum_{k=1}^K \gamma S(|T|_{:k}^{(3)}) s_k \tag{9}$$

其中:  $\alpha \geq 0, \beta \geq 0$  和  $\gamma \geq 0$  是固定的惩罚参数, 控制因子的整体稀疏程度。稀疏度测度  $S(\cdot)$  表示零元素的比例, 其值由张量中各部分的稀疏度决定。沿 1-模展开后的矩阵和代价函数的构成如图 3 所示, 可以看出, 代价函数  $D$  中包含了近似失真和稀疏惩罚两部分, 其中近似失真由 Frobenius 范数<sup>[8]</sup>度量。向量  $|T|_{i:}^{(1)}$  的稀疏度用来惩罚  $u_i$ , 如果  $|T|_{i:}^{(1)}$  的稀疏度较大, 则惩罚参数将较大, 从而产生较小的  $u_i$ ; 反之亦然。

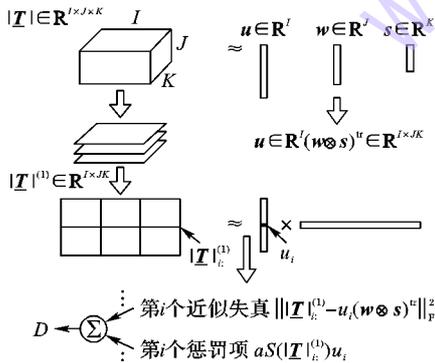


图 3 张量分解的代价函数示意图

为了最小化代价函数  $D$ , 本文采用 ALS (Alternating Least Squares) 算法<sup>[8]</sup>:

- 1) (随机) 初始化  $u, w$  和  $s$ 。
- 2) 固定  $w$  和  $s$ , 求  $u$  从而最小化  $D$ 。
- 3) 固定  $u$  和  $s$ , 求  $w$  从而最小化  $D$ 。
- 4) 固定  $u$  和  $w$ , 求  $s$  从而最小化  $D$ 。
- 5) 如果  $(D_{old} - D_{new})/D_{old} > \epsilon$ , 转到步骤 1); 否则停止。其中当代价函数值在相邻两次迭代中的相对变化小于一个门

限  $\epsilon$  时停止迭代。

因子  $w$  刻画了数据中显著的频率值,  $s$  表示它们的持续时间,  $u$  指出其出现的位置, 即以帧起始位置为时间原点频率成分出现在帧内的时间位置, 但是受到分帧的影响, 不同帧之间的  $u$  因子不具有共同时间原点, 因此在完成联合时-频-长分解后丢弃该因子, 不作为特征。将  $w$  和  $s$  拼接起来, 通过 LDA 降维后作为特征。

## 2 实验和结果

### 2.1 权重张量

本文采用正交匹配追踪算法 (Orthogonal Matching Pursuit, OMP)<sup>[9]</sup> 进行信号的稀疏表示分解。字典原子参数集设置如下  $\mu \in \{1 + 32m; 0 \leq m \leq 15, m \in \mathbf{Z}\}, f \in \{0.5 \times (m/25)^{2.6}; 1 \leq m \leq 25, m \in \mathbf{Z}\}, \sigma \in \{2^m; 3 \leq m \leq 11, m \in \mathbf{Z}\}$ , 因此权重张量的维数是  $16 \times 25 \times 9$ , 经过 LDA 降维后的特征向量维数为 13。

一短时帧的女性语音和河流流水声分别如图 4(a) 和 (b) 所示, 其权重张量分别如图 4(c) 和 (d) 所示, 其中稀疏表示原子数目选择为 32。图中尺寸较大的点表示权重较大的值, 反之, 尺寸较小的点表示较小的权重值, 权重为 0 的元素没有显示。如图 4(c) 中可见一个尺寸较大的点, 该点的索引为 (9, 12, 6), 根据原子参数取值该点的中心位置为 257 个采样点, 频率为 0.074 (即 1186 Hz), 时长为 256 个采样点, 因此该点表示了在该语音帧中的一个能量集中的频率分量。与之相比, 图 4(d) 中的点能量均相似, 且能量普遍较小, 与河流流水无显著频率成分的事实相一致。对比图 4(e)、(g) 与图 4(f)、(h) 可以看出, 语音的中心频率集中在一个因子元素上, 而河流流水声相对较为分散; 此外, 两者时长因子的分布模式也有较大不同。

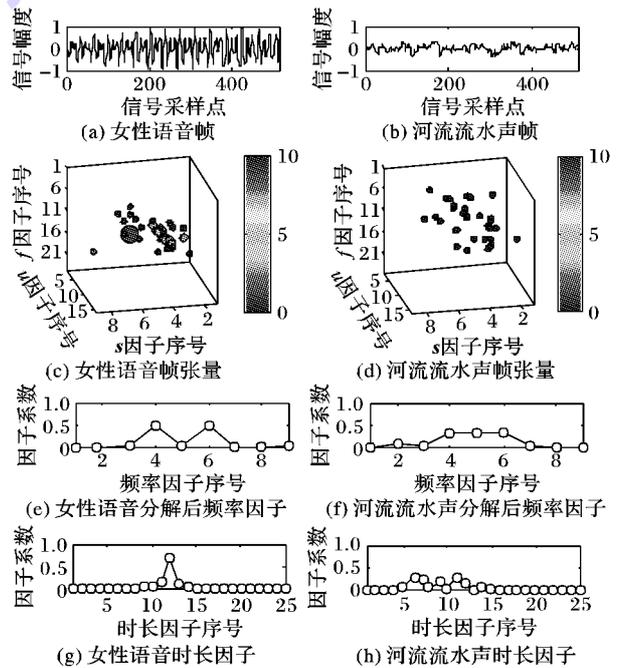


图 4 两种不同信号的张量分解对比

### 2.2 音效分类实验结果

本文使用 15 类音效分类任务评估所提出的特征, 音效包括男性语音、女性语音、婴儿语音、掌声、脚步声、鸟叫声、猫叫

声、河流流水声、雷声、引擎声、枪声、警报声、钢琴声、小提琴声和鼓声。所有音效都是从 Digital Juice 音效库 I 和 II<sup>[10]</sup> 以及 BBC (British Broadcasting Corporation) 音效库<sup>[11]</sup> 中收集的,所有音效样本时长介于 3 ~ 8 s,单声道,16 kHz 采样,使用 32 ms 帧和 8 ms 帧移。

高斯混合模型 (Gaussian Mixture Model, GMM) 对于该多分类问题有最好并且最稳定的分类结果<sup>[3]</sup>,因此本文统一采用该分类器对各组特征进行分类实验比较。GMM 的混合度为 8,惩罚参数  $\alpha$ 、 $\beta$  和  $\gamma$  通过最优化和男性语音的二分任务设定为 0.132、0.109 和 0.097。门限  $\varepsilon$  根据经验设定为 0.01。在 Matlab 7.14.10 仿真平台下将本文所提出的 13 维张量特征 (Tensor) 与 13 维 MFCC 特征、MFCC 特征和匹配追踪算法稀疏分解的 MP 特征拼接得到的 17 维联合 MFCC + MP 特征<sup>[3]</sup>、16 维非均匀尺度-频率图特征 (Map)<sup>[5]</sup> 比较分类效果。此外,为了验证所提出的自调整惩罚参数分解算法的有效性,Tensor 特征还与 CP 分解算法 (CP) 和  $l_1$ -范数惩罚的 CP 算法 ( $l_1$ -CP) 进行比较。

每种特征对所有音效类的平均分类效果如图 5 所示,每一类的识别率定义为正确识别为该类的样本数与该类总样本数的比值,所有类的平均识别率通过 10-折交叉验证得到。由于稀疏表示中稀疏度的不同会导致识别率不同,因此选择以 2 为底的对数尺度的原子数目进行稀疏表示,分别测定不同原子数目下的识别效果。Tensor 特征、非均匀尺度-频率图特征和 MFCC + MP 特征对每一音效类的识别率如图 6 所示。注意,图 6 中每种特征对应的原子数是根据图 5 中该种特征达到最高识别率时的原子数决定的。

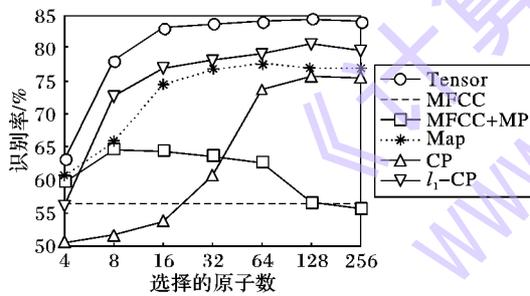


图 5 不同特征的平均分类效果

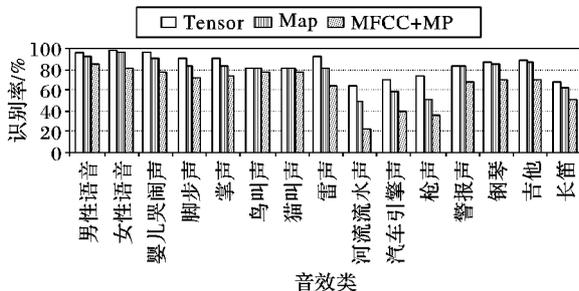


图 6 不同特征对各音效类的分类效果

最好的平均识别率时由 Tensor 特征在原子数为 128 时得到的 84.8%,非均匀尺度-频率图 Map 特征的平均识别率在原子数为 64 时取得最高,为 77.6%。此外,Tensor 特征在原子数目为 16 ~ 256 时识别率均高于 83%,因此该特征对原子数目具有鲁棒性。在各原子数目下,Tensor 特征分类效果均优于非均匀尺度-频率图特征,原因在于 Tensor 特征相比于非均匀尺度-频率图特征多使用了时间信息进行联合分析,而

时间信息对于瞬变信号分类具有重要作用。该结论也可以从图 6 中看出,其中,与非均匀尺度-频率图特征相比,Tensor 特征显著提升了河流流水声,雷声,引擎声和枪声的分类效果,此外,对人类语音、掌声和脚步声也均有一定提升。MFCC + MP 特征最高的识别率时 8 个原子数对应的 64.6%。Tensor 特征较非均匀尺度-频率图特征、MFCC 特征和 MFCC + MP 特征识别率分别高出 6.7%、28.0% 和 19.8%。

### 3 结语

本文以音效分类为应用背景,提出了一种基于稀疏表示权重张量的特征提取方法,刻画了音频信号中瞬变成分的特性。Gabor 原子的时间、频率以及时长特性用来建立一个张量,描述了信号的联合时-频-长分布特点。该张量刻画了信号更多具有区分力的特性,并且利用张量分解得到描述信号短帧内频率分布的频率因子和时长分布的时长因子。该张量结构和因子分解过程相比于直接进行向量特征提取的方法描述了更多信号的区分特性。此外针对稀疏张量分解还提出了一种自调整惩罚参数分解算法,利用张量的稀疏度动态作为惩罚参数,有效避免了过拟合。实验结果显示,所提出的张量特征对 15 类音效的分类效果显著优于其他特征。

#### 参考文献:

- [1] ZUBAIR S, WANG W. Audio classification based on sparse coefficients[C]// Sensor Signal Processing for Defence (SSPD 2011). London, UK: The Institution of Engineering and Technology Press, 2011: 1 - 5.
- [2] ZUBAIR S, YAN F, WANG W. Dictionary learning based sparse coefficients for audio classification with max and average pooling[J]. Digital Signal Processing, 2013, 23(3): 960 - 970.
- [3] CHU S, NARAYANAN S, KUO C C J. Environmental sound recognition with time-frequency audio features[J]. IEEE Transactions on Audio, Speech, and Language Processing, 2009, 17(6): 1142 - 1158.
- [4] SIVASANKARAN S, PRABHU K M M. Robust features for environmental sound classification[C]// Proceedings of the 2013 IEEE International Conference on Electronics, Computing and Communication Technologies. Piscataway, NJ: IEEE, 2013: 1 - 6.
- [5] WANG J C, LIN C H, CHEN B W, et al. Gabor-based nonuniform scale-frequency map for environmental sound classification in home automation[J]. IEEE Transactions on Automation Science and Engineering, 2014, 11(2): 607 - 613.
- [6] TAKEUCHI K, ISHIGURO K, KIMURA A, et al. Non-negative multiple matrix factorization[C]// Proceedings of the 23rd International Joint Conference on Artificial Intelligence. Beijing: AAAI, 2013: 1713 - 1720.
- [7] LIU J, LIU J, WONKA P, et al. Sparse non-negative tensor factorization using columnwise coordinate descent[J]. Pattern Recognition, 2012, 45(1): 649 - 656.
- [8] CICHOCKI A, ZDUNEK R, PHAN A H, et al. Nonnegative Matrix and Tensor Factorizations: Applications to Exploratory Multi-way Data Analysis and Blind Source Separation[M]. New York: John Wiley & Sons, 2009: 35 - 37.
- [9] CHANG L H, WU J Y. An improved RIP-based performance guarantee for sparse signal recovery via orthogonal matching pursuit[J]. IEEE Transactions on Information Theory, 2014, 60(9): 5702 - 5715.

- [6] 何维国, 谢伟, 张健. 上海电网电能质量监测系统[J]. 华东电力, 2010, 38(4): 497-499. (HE W G, XIE W, ZHANG J. Power quality monitoring system of Shanghai grid[J]. East China Electric Power, 2010, 38(4): 497-499.)
- [7] 全国电压电流等级和频率标准化技术委员会. GB/T 12325-2008, 电能质量供电电压偏差[S]. 北京: 中国标准出版社, 2008: 3. (Standard voltages, current ratings and frequencies. GB/T 12325-2008, Power quality-deviation of supply voltage[S]. Beijing: China Standard Press, 2008: 3.)
- [8] IEEE Standards Coordinating Committee 22 on Power Quality. IEEE Std1159.3-2003 IEEE recommended practice for the transfer of power quality data[S]. New York: IEEE, 2004: 1-200.
- [9] 马友忠, 慈祥, 孟小峰. 海量高维向量的并行 Top-k 连接查询[J]. 计算机学报, 2015, 38(1): 86-98. (MA Y Z, CI X, MENG X F. Parallel Top-k join on massive high-dimensional vectors[J]. Chinese Journal of Computers, 2015, 38(1): 86-98.)
- [10] 王有为, 王伟平, 孟丹. 基于统计方法的 Hive 数据仓库查询优化实现[J]. 计算机研究与发展, 2015, 52(6): 1452-1462. (WANG Y W, WANG W P, MENG D. Query optimization by statistical approach for Hive data warehouse[J]. Journal of Computer Research and Development, 2015, 52(6): 1452-1462.)
- [11] 刘亚秋, 李海涛, 景维鹏. 基于 Hadoop 的海量嘈杂数据决策树算法的实现[J]. 计算机应用, 2015, 35(4): 1143-1147. (LIU Y Q, LI H T, JING W P. Implementation of decision tree algorithm dealing with massive noisy data based on Hadoop[J]. Journal of Computer Applications, 2015, 35(4): 1143-1147.)
- [12] 张逸, 杨洪耕. 海量电能质量数据交换格式文件快速解析方案[J]. 电力自动化设备, 2013, 33(12): 116-121. (ZHANG Y, YANG H G. Fast parsing of massive PQDIF files[J]. Electric Power Automation Equipment, 2013, 33(12): 116-121.)
- [13] 张逸, 杨洪耕, 叶茂清. 基于分布式文件系统海量电能质量监测数据管理方案[J]. 电力系统自动化, 2014, 38(2): 102-108. (ZHANG Y, YANG H G, YE M Q. A data management scheme of massive power quality monitoring data based on distributed file system[J]. Automation of Electric Power Systems, 2014, 38(2): 102-108.)
- [14] 魏雪琴. 电能质量数据压缩研究[D]. 合肥: 安徽大学, 2013: 5-25. (WEI X Q. The research on power quality data compression [D]. Hefei: Anhui University, 2013: 5-25.)
- [15] 赵东旭. 电能质量数据中心建设关键问题研究[D]. 北京: 华北电力大学, 2012: 10-20. (ZHAO D X. Research on key issues in power quality data center construction[D]. Beijing: North China Electric Power University, 2012: 10-20.)
- [16] 王淑祥, 马素霞. 基于 Hadoop 的海量电能质量监测数据分析平台研究[J]. 中国科技信息, 2013, 25(13): 79-80. (WANG S X, MA S X. Study on massive power quality monitoring data analysis platform based on Hadoop[J]. China Science and Technology Information, 2013, 25(13): 79-80.)
- [17] 李斌, 郭雅娟, 陈锦铭, 等. 电能质量监测系统 95 概率大值的 top-k 优化研究[J]. 电力信息化, 2013, 11(1): 20-24. (LI B, GUO Y J, CHEN J M, et al. Research on the top-k optimization method of 95% probability values for the power quality monitoring system[J]. Electric Power Information Technology, 2013, 11(1): 20-24.)
- [18] 白宇, 郭显娥. 深度优先稳定原地归并排序的高效算法[J]. 计算机应用, 2013, 33(4): 1039-1042. (BAI Y, GUO X E. Efficient algorithm of depth-first stable in-place merge sort[J]. Journal of Computer Applications, 2013, 33(4): 1039-1042.)
- [19] 孙琳琳, 侯秀萍, 朱波, 等. 基于多线程归并排序算法设计[J]. 吉林大学学报(信息科学版), 2015, 33(1): 105-110. (SUN L L, HOU X P, ZHU B, et al. Merge sort algorithm design based on multi thread[J]. Journal of Jilin University (Information Science Edition), 2015, 33(1): 105-110.)
- [20] 姜忠华, 徐文丽, 刘家文, 等. 智能归并排序[J]. 电子设计工程, 2011, 19(21): 53-55. (JIANG Z H, XU W L, LIU J W, et al. Intelligently merge sort[J]. Electronic Design Engineering, 2011, 19(21): 53-55.)
- [21] HOSSAIN N, ALMA M G R, AMIRUZZAMAN M, et al. An efficient merge sort technique that reduces both times and comparisons [C]// 2004 ICTTA: Proceedings of the 2004 International Conference on Information and Communication Technologies: from Theory to Applications. Piscataway, NJ: IEEE, 2004: 857-858.
- [22] KAZMI S Z H, KAZMI S S, KAZMI S S H, et al. Chunk sort[J]. International Journal of Computer Science and Information Security, 2010, 8(4): 162.
- [23] 赵霞, 林天华. 分布式电能质量信息系统[J]. 科技导报, 2013, 31(19): 58-61. (ZHAO X, LIN T H. Design and implementation of distributed electric power quality information system[J]. Science & Technology Review, 2013, 31(19): 58-61.)

### Background

This work is partly supported by the Science Technology Research Projects for the High school of Hebei Education Department (YQ2013038), the Natural Science Foundation of Hebei Province (F2015207009) and the Science Research Project of Hebei University of Economics and Business(2013KYY17).

**ZHAO Xia**, born in 1979, Ph. D., lecturer. Her research interests include software engineering, nonlinear control technology.

**LIN Tianhua**, born in 1979, M. S., associate professor. His research interests include software engineering, software architecture, information system.

**MA Suxia**, born in 1964, M. S., professor. Her research interests include software engineering, software architecture, component technology.

**QI Linhai**, born in 1964, associate professor. His research interests include database, information system, business intelligence.

(上接第 1429 页)

- [10] Digital Juice, Incorporated. The digital juice sound FX library [DB/OL]. [2015-05-20]. <http://www.digitaljuice.com>.
- [11] British Broadcasting Corporation (BBC). BBC sound effects library [DB/OL]. [2015-05-20] <http://www.sound-ideas.com/bbc.html>.

### Background

This work is supported by the National Natural Science Foundation of China (61301300).

**LIN Jing**, born in 1982, M. S., lecture. Her research interests include audio and vedio signal processing.

**YANG Jichen**, born in 1980, Ph. D., associate research fellow. His research interests include audio and vedio signal processing.

**ZHANG Xueyuan**, born in 1987, Ph. D. candidate. His research interests include audio and vedio signal processing.

**LI Xinchao**, born in 1980, Ph. D. candidate. His research interests include intelligent optimization, signal processing.