

利用抗噪纹理特征的快速鸟鸣声识别

魏静明, 李 应

(福州大学数学与计算机科学学院, 福建福州 350108)

摘 要: 针对非平稳噪声下的鸟鸣声识别问题, 提出一种利用抗噪纹理特征的快速鸟鸣声识别方法, 该方法也解决了纹理特征提取过程中灰度共生矩阵 (GLCM) 占用空间大, 以及计算量大、耗时的问题. 该方法分三个步骤, 首先, 通过短时谱估计算法对鸟鸣声带噪功率谱进行音频增强; 然后, 采用和差统计法 (SDH) 对增强功率谱快速提取纹理特征; 最后, 由随机森林进行分类. 在实验部分, 设计了两组对比实验, 结果表明, 该方法有良好的识别性能、较少的时耗, 且具有噪声鲁棒性.

关键词: 鸟鸣声识别; 抗噪纹理特征; 短时谱估计; 和差统计法

中图分类号: TP391.4 **文献标识码:** A **文章编号:** 0372-2112 (2015)01-0185-06

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2015.01.029

Rapid Bird Sound Recognition Using Anti-Noise Texture Features

WEI Jing-ming, LI Ying

(College of Mathematics and Computer Science, Fuzhou University, Fuzhou, Fujian 350108, China)

Abstract: In order to improve the accuracy of bird sound recognition under non-stationary noise, a rapid bird sound recognition method using anti-noise texture features was proposed. This method also solved the problem of big occupation space, as well as large computation, time-consuming of gray level co-occurrence matrix (GLCM) in extracting texture features. The method contained three steps. Firstly, short-time spectrum estimation was conducted on noise spectrum of bird sound to get the enhanced spectrum. Then, sum and difference histograms (SDH) were applied to calculate texture features rapidly. Finally, random forest (RF) was used to make classification. In the experiment part, two contrast experiments were designed. The results show that this method has good recognition performance, lesser time consumption; meanwhile, it is robust to noise.

Key words: bird sound recognition; anti-noise texture features; short-time spectrum estimation; sum and difference histograms (SDH)

1 引言

生态环境中鸟鸣声包含着丰富的信息, 鸟鸣声识别在了解鸟类行为等特征的同时, 还可以对相关环境做一定程度的分析.

原始鸟鸣声信号经过短时傅里叶变换 (Short-Time Fourier Transform, STFT) 可以得到其时频谱图, 它本质是一幅灰度图像, 因此可以通过图像特征实现鸟鸣声识别. 目前, 借助图像特征, 鸟鸣声识别领域已取得了一些成果. 文献[1]提出一种基于时频谱图有监督的时频分割方法, 利用时频单元方差等特征进行鸟鸣声识别. 文献[2]提出一种基于时频谱图形状特征的方法, 利用角度径向变换特征实现鸟鸣声自动识别.

不同于以上特征, 纹理特征是一种全局统计特征,

在图像识别中具有较大的优越性. 因此, 我们试图利用时频谱图纹理特征进行鸟鸣声识别. 纹理特征提取方法中, 基于统计的灰度共生矩阵 (Gray Level Co-occurrence Matrix, GLCM) 因其原理简单、并能较好反映纹理图像而被广泛采用^[3~5]. 但 GLCM 的二维表示占用存储空间大, 提取二次统计纹理特征的过程计算量大、耗时多^[6]. 在 GLCM 的改进算法中, 和差统计法 (Sum and Difference Histograms, SDH)^[7] 的一维表示能够有效地克服 GLCM 的缺点.

同时真实环境中普遍存在着各种背景噪声, 这使得噪声环境下的鸟鸣声识别具有重要的现实意义^[8]. 文献[9]采用谱减法, 功率谱分析及自相关分析进行噪声环境下的鸟鸣声识别. 文献[10]利用基于帧线性预测编码 (Linear Predictive Coding, LPC) 的聚类, 提取有效音节实

现噪声环境下鸟鸣声识别.但这些方法主要针对特定噪声环境下的鸟鸣声识别,不能普遍适用于真实环境中的非平稳噪声.而短时谱估计^[11]能够有效处理非平稳噪声下的音频增强.

基于以上分析,本文提出一种利用抗噪纹理特征的快速鸟鸣声识别方法.该方法首先通过短时谱估计对鸟鸣声做增强处理,然后由 SDH 对得到的增强功率谱快速提取纹理特征,再结合随机森林(Random Forest, RF)进行识别和分类.实验结果表明,该方法提取的抗噪纹理特征具有良好的识别性能,在纹理特征提取环节减少了时耗,同时能有效提高非平稳噪声环境下鸟鸣声识别性能.

2 抗噪纹理特征的提取

抗噪纹理特征的提取分为前端处理和纹理特征提取两个环节,如图 1 所示.从真实环境中采集的原始鸟鸣声信号 $y(t)$ 可表示为:

$$y(t) = s(t) + n(t) \quad (1)$$

式中, $s(t)$ 为纯净鸟鸣声, $n(t)$ 为加性环境噪声.经过预处理后,对 $y(t)$ 进行 STFT 可得到其幅度谱 $Y(k, l)$, 其中 k 为帧索引, l 为频率索引.接下来在前端处理环节对 $Y(k, l)$ 进行音频增强.

2.1 前端处理

前端处理环节采用的短时谱估计是一种基于噪声估计的音频增强算法,具体实现如下.

(1) 对鸟鸣声幅度谱 $Y(k, l)$ 进行一个固定因子平滑,得到平滑后带噪声功率谱:

$$S(k, l) = \alpha S(k-1, l) + (1-\alpha) |Y(k, l)|^2 \quad (2)$$

式中, $\alpha = 0.7$ 为固定的平滑因子.

(2) 通过前向-后向相结合的双向搜索算法寻找 $S(k, l)$ 的最小频谱分量:

$$S_{\min}(k, l) = \max\{S_{\min1}(k, l), S_{\min2}(k, l)\} \quad (3)$$

式中, $S_{\min1}(k, l)$ 表示前向搜索出来的最小值, $S_{\min2}(k, l)$ 表示后向搜索出来的最小值.

(3) 计算有效信号存在的概率:

$$P(k, l) = \alpha_1 P(k-1, l) + (1-\alpha_1) H(k, l) \quad (4)$$

式中, $\alpha_1 = 0.2$ 为固定的平滑因子, $H(k, l)$ 是有效信号存在性的判别准则,表示当前频点判定为音频信号还是噪声.可描述如下:

$$H(k, l) = \begin{cases} 1, & \text{存在有效信号, } S(k, l)/S_{\min}(k, l) > \phi(k) \\ 0, & \text{不存在有效信号, 其他} \end{cases} \quad (5)$$

式中, $\phi(k)$ 是依赖于具体频率的判别阈值,当 k 小于 1kHz 或界于 1 到 3kHz 时,取 $\phi(k) = 2$,当 k 界于 3kHz 到信号采样频率的一半时,取 $\phi(k) = 5$.

(4) 计算时-频平滑因子 $\eta(k, l)$:

$$\eta(k, l) = \alpha_2 + (1-\alpha_2) P(k, l) \quad (6)$$

式中, $\alpha_2 = 0.95$ 为根据实际情况设定的常量,显然, $\alpha_2 \leq \eta(k, l) \leq 1$.由 $\eta(k, l)$ 可以进行噪声功率谱的估计: $N(k, l) = \eta(k, l) N(k-1, l) + (1-\eta(k, l)) |Y(k, l)|^2$

(5) 得到上述结果后,便可以进行频谱增益因子的计算:

$$G(k, l) = C(k, l) / (C(k, l) + \sigma N(k, l)) \quad (8)$$

式中, $C(k, l) = |Y(k, l)|^2 - N(k, l)$ 表示纯净信号功率谱, σ 为过减因子,其值为:

$$\sigma = 8 - \frac{9}{25} \times 10 \lg \frac{|Y(k, l)|^2}{N(k, l)} \quad (9)$$

(6) 最后,得到增强后的音频信号功率谱:

$$X(k, l) = G(k, l) \times |Y(k, l)|^2 \quad (10)$$

算法中相关参数值,都是我们经过多次实验对比后确定的,适用于非平稳噪声环境下鸟鸣声的增强处理.

2.2 纹理特征提取

2.2.1 计算和向量与差向量

增强后的音频信号功率谱 $X(k, l)$ 是一幅灰度图像,其大小为 $M \times N$,灰度级为 L .图像中两像素点相对位移 $(\Delta k, \Delta l)$ 一定的情况下,我们分别定义和与差为:

$$S(k, l) = X(k, l) + X(k + \Delta k, l + \Delta l) \quad (11)$$

$$D(k, l) = X(k, l) - X(k + \Delta k, l + \Delta l) \quad (12)$$

当 $((k, l), (k + \Delta k, l + \Delta l)) \in (M \times N) \times (M \times N)$ 时,和的统计定义为:

$$C(i | \Delta k, \Delta l) = C_S(i) = \text{Card}\{S(k, l) = i\} \quad (13)$$

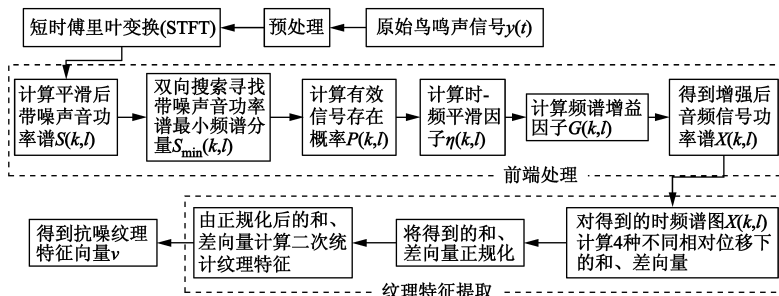


图1 抗噪纹理特征的提取流程

式中, $i = 0, 1, \dots, 2(L-1)$. 差的统计定义为:

$$C(j|\Delta k, \Delta l) = C_D(j) = \text{Card}\{D(k, l) = j\} \quad (14)$$

式中, $j = -(L-1), -(L-2), \dots, 0, \dots, L-2, L-1$. 对式(14)进一步分析可知, 当 $0 < d \leq L-1$, 有 $C_D(d) = C_D(-d)$. 因此对差向量在 $j=0$ 处进行折叠, 可以得到简单化的新差向量:

$$C_D^*(j) = \begin{cases} C_D(j), & j=0 \\ 2C_D(j), & 0 < j \leq L-1 \end{cases} \quad (15)$$

由于正规化后的特征值能够使纹理分辨率更高, 因此, 实际应用中, 需要对和向量与差向量进行正规化:

$$S(i) = C_S(i)/N_H, D(j) = C_D^*(j)/N_H \quad (16)$$

$$\text{式中}, N_H = \sum_{i=0}^{2(L-1)} C_S(i) = \sum_{j=0}^{L-1} C_D^*(j).$$

2.2.2 纹理特征的选择与计算

目前, 基于纹理的二次统计特征有 14 种, 其中有且仅有 5 种不相关且具有良好的识别效果, 我们选择这 5 种二次统计特征, 并用正规化后的和向量与差向量计算如下:

(1) 角二阶距(UNI):

$$\text{UNI} = \sum_{i=0}^{2(L-1)} S^2(i) \sum_{j=0}^{L-1} D^2(j) \quad (17)$$

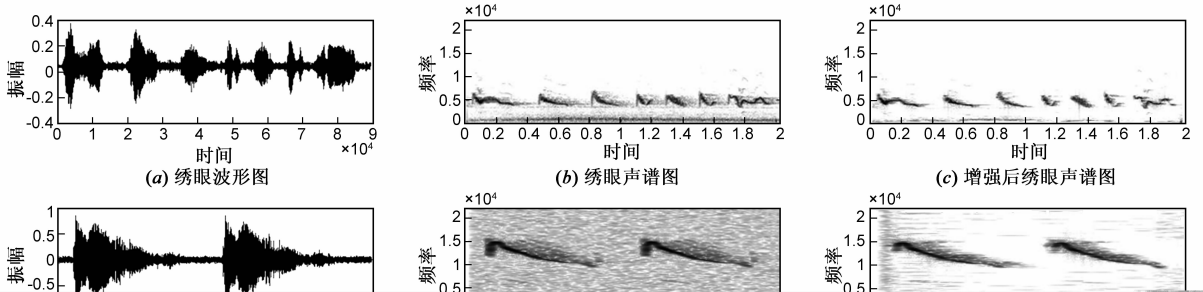
(2) 对比度(CON):

$$\text{CON} = \sum_{j=0}^{L-1} j^2 D(j) \quad (18)$$

(3) 相关(COR):

$$\text{COR} = \left(\sum_{i=0}^{2(L-1)} (i - 2\mu)^2 S(i) - \sum_{j=0}^{L-1} j^2 D(j) \right) / 2 \quad (19)$$

式中, $\mu = \sum_{i=0}^{2(L-1)} i S(i)$.



3.2 纹理特征提取性能分析

为了便于用实例分析纹理特征提取过程中和向量、差向量的计算, 以及 SDH 相对 GLCM 在空间占用方面的优越性, 并计算最终的纹理特征向量. 我们从上面

(4) 熵(ENT):

$$\text{ENT} = - \sum_{i=0}^{2(L-1)} S(i) \log S(i) - \sum_{j=0}^{L-1} D(j) \log D(j) \quad (20)$$

(5) 局部平稳(HOM):

$$\text{HOM} = \sum_{j=0}^{L-1} \frac{D(j)}{(1+j)^2} \quad (21)$$

分别用 a, b, c, d 表示图像 $X(k, l)$ 中两像素点相对位移 $(\Delta k, \Delta l)$ 为 $(1, 0), (1, 1), (0, 1), (-1, 1)$ 的四种情况, 通过以上过程计算每种情况下的 5 个二次统计纹理特征值, 并按顺序排在一起, 最终可以得到一个维数为 1、长度为 20 的抗噪纹理特征向量 $\mathbf{v} = (v_a, v_b, v_c, v_d)$, 记 $i \in \{a, b, c, d\}$, 有 $v_i = (i. uni, i. con, i. cor, i. ent, i. hom)$, $i. uni$ 表示在情况 i 下得到的角二阶距值, 其他 4 个意义类同.

3 方法性能分析

这部分, 我们实例分析抗噪纹理特征提取过程中前端处理和纹理特征提取两个环节的性能.

3.1 前端处理性能分析

真实环境中采集的绣眼和秃鹰两种鸟鸣声如图 2 所示.

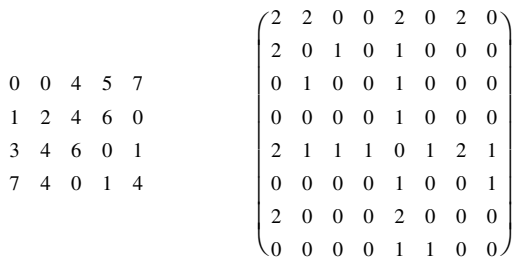
在图 2 中, (a)、(d) 分别为绣眼和秃鹰的波形图, (b)、(e) 分别为增强前绣眼和秃鹰的声谱图, (c)、(f) 分别为经短时谱估计增强后绣眼和秃鹰的声谱图. 通过 (b) 和 (c) 以及 (e) 和 (f) 的对比, 可以看出, (c)、(f) 两幅声谱图包含的背景噪声的纹理明显少于 (b)、(e) 两幅声谱图, 同时, (c)、(f) 两幅声谱图中有效声音信号的纹理更加突出. 这说明了在前端处理环节采用的短时谱估计算法能够有效地降低真实环境中非平稳噪声对鸟鸣声信号的影响.

采集的绣眼声谱图中截取一个大小为 4×5 、灰度级为 8 的灰度图像, 并由其分别计算 GLCM 与和向量、差向量, 所得结果如图 3 所示.

在图 3 中, (b) 为 (a) 在距离 $d = 1$ 和方向 $\theta = 0^\circ$ 时

计算得到的 GLCM, 矩阵中每一个元素值都是对 (a) 进行遍历统计得到的, 第 i 行 j 列的元素值表示 (a) 中水平方向距离为 1 的像素对中一个像素值为 i 、另一个像素值为 j 的个数, $i, j = 0, 1, \dots, 7$, 正向和逆向重复计数, 可知, GLCM 为对称矩阵. 如像素对 $(1, 2)$ 或 $(2, 1)$ 只在 (a) 中第 2 行出现一次, 故 (b) 中第 1 行 2 列以及第 2 行 1 列元素值均为 1.

而 (c) 为 (a) 在相对位移取 $(1, 0)$ 的情况下得到的和向量与差向量. $C_S(i)$ 表示 (a) 中水平方向距离为 1 的像素对和为 i 的个数, $i = 0, 1, \dots, 14$. 如, (a) 中和为 6 的组合中, 像素对 $(6, 0)$ 出现 2 次, 像素对 $(2, 4)$ 出现 1 次, 其他均为 0, 因此 $C_S(6) = 6$. $C_D^*(j)$ 表示 (a) 中水平方向距离为 1 的像素对差的绝对值为 j 的个数, $j = 0, 1, \dots, 7$. 如, (a) 中差的绝对值为 1 的组合中, 像素对 $(0, 1)$ 出现 2 次, 像素对 $(1, 2)$ 出现 1 次, 像素对 $(3, 4)$ 出现 1 次, 像素对 $(4, 5)$ 出现 1 次, 因此 $C_D^*(1) = 10$.



(a) 4×5 的灰度图像

(b) 对应的 GLCM

i	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
$C_S(i)$	2	4	0	2	4	2	6	2	0	2	4	2	2	0	0

j	0	1	2	3	4	5	6	7
$C_D^*(j)$	2	10	8	4	4	0	4	0

(c) 和向量与差向量
图 3

通过 (b) 与 (c) 的对比, 不难看出, SDH 的一维表示占用空间为 $15 + 8$, GLCM 的二维表示占用空间为 8×8 , 用灰度级来表示分别为 $2(L - 1) + 1 + L$ 与 L^2 , 其中 $L = 8$. 当 $L = 2$ 时, $L^2 = 4$ 小于 $2(L - 1) + 1 + L = 5$, 当 $L = 3$ 时, $L^2 = 9$ 大于 $2(L - 1) + 1 + L = 8$, 当 $L > 3$ 时, 随着 L 的增大, L^2 呈指数增长, $2(L - 1) + 1 + L$ 则呈线性增长, 两者的差值也越来越大. 可见, SDH 在空间占用方面有明显的优越性.

对图 3(c) 中和向量与差向量进行正规化后, 由式 (17) ~ (21), 可得 $v_a = (0.0231, 1.8125, 23.1250, 3.9698, 0.0764)$. 接下来, 用同样的方式计算相对位移为 $(1, 1)$, $(0, 1)$, $(-1, 1)$ 三种情况下的和向量、差向量, 并计算相应纹理特征, 得到其纹理特征向量 $\mathbf{v} = (v_a, v_b, v_c, v_d) = (0.0231, 1.8125, 23.1250, 3.9698, 0.0764, 0.0311,$

$4.5833, 35.6667, 3.6061, 0.1550, 0.0263, 17.1333, 19.4667, 3.8676, 0.0871, 0.0478, 20.5000, 13.7083, 3.0582, 0.0374)$.

至于 SDH 与 GLCM 时耗方面的对比, 这里只分析其时间复杂度, 具体实例数据可从实验部分获知. 表 1 列出了由 GLCM 计算本文选择的 5 个纹理特征的公式^[3-5]. 可见, GLCM 的时间复杂度为 $O(L^2)$, 由式 (17) ~ (21) 可知, SDH 的时间复杂度仅为 $O(L)$, 显然, 比起 GLCM, SDH 在时间复杂度方面也有一个数量级的提升.

表 1 GLCM 下 5 种特征计算公式

纹理特征名称	GLCM
UNI	$\text{UNI} = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P^2(i, j)$
CON	$\text{CON} = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i - j)^2 P(i, j)$
COR	$\text{COR} = \frac{1}{\sigma_x \sigma_y} \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i - \mu_x)(j - \mu_y) P(i, j)$
ENT	$\text{ENT} = - \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P(i, j) \log P(i, j)$
HOM	$\text{HOM} = \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{P(i, j)}{1 + (i - j)^2}$

4 实验结果与分析

4.1 声音样本集

实验用到的 10 种纯净鸟鸣声来自 Freesound^[12] 声音数据库, 分别为白面鸡、朱颈斑鸠、董鸡、红毛鸡、山斑鸠、水秧鸡、绣眼、野山鸡、竹鸡、鹧鸪. 每种鸟鸣声有 40 个样本, 共 400 个样本. 实验用到的三种加性背景噪声分别为高斯白噪声、汽车噪声和流水噪声. 其中, 高斯白噪声是由计算机随机产生的均值为 0、功率谱密度为 1 且均匀分布的随机信号经高斯化得到的, 汽车噪声和流水噪声是以 44.1kHz 的采样频率分别在交通主干线附近和山间小溪处录取的. 为规范以上声音文件的编码格式和长度, 我们将它们统一转换成采样频率为 8kHz、采样精度为 16b、长度为 2s 的单声道 WAV 格式声音片段. 在预处理部分, 对声音片段分帧加 Hanning 窗, 帧长为 16ms (128 个采样点), 帧移为 8ms (64 个采样点).

4.2 实验设计与结果分析

为了检验利用抗噪纹理特征的快速鸟鸣声识别方法的性能, 设计了两组对比实验, 每组均进行 20 次. 每次实验, 无放回地从每类鸟鸣声中随机抽取 10 个样本作为训练数据, 剩下 30 个样本作为测试数据.

第 1 组实验, 对比由 SDH 与 GLCM 提取纹理特征的识别率和时耗. 对纯净的鸟鸣声片段, 做预处理和 STFT 后, 分别使用 GLCM 和 SDH 来提取纹理特征, 然后由 10 类共 100 个训练样本的纹理特征向量来训练 RF 分类

器,最后用训练好的 RF 分类器对 10 类共 300 个测试样本进行识别和分类.每次实验,记录 GLCM 和 SDH 对每类鸟鸣声提取的纹理特征的识别率,得到的 20 个值分别是两种方法下每类鸟鸣声 30 个测试样本识别率的均值;记录 GLCM 和 SDH 在每类鸟鸣声纹理特征提取过程的时耗,得到的 20 个值分别是两种方法下每类鸟鸣声 10 个训练样本和 30 个测试样本相应时耗的均值.

20 次实验取平均后,结果如表 2 所示.其中,为方便对比时耗,统一记 GLCM 的相应时耗比为 100,并按比例计算得到 SDH 的相应时耗比.可见,识别率方面,纹理特征具有良好的识别性能,同时 SDH 在白面鸡、董鸡、水秧鸡这 3 类鸟鸣声的识别率略小于 GLCM,而另外 7 类鸟鸣声的识别率则大于等于 GLCM,并且,平均识别率还有微小的提升.更重要的是,时耗比方面,从数据上看 SDH 明显全优于 GLCM,平均时耗比也减少了 20.2%.显然,比起 GLCM,采用 SDH 提取的纹理特征能够在保证良好识别率的前提下,有效地减少时耗.

表 2 两种纹理特征提取方法用于 10 类鸟鸣声分类的识别率和时耗比

音频类型	识别率 (%)		时耗比 (%)	
	GLCM	SDH	GLCM	SDH
白面鸡	97	96	100	79
朱颈斑鸠	94	95	100	81
董鸡	98	97	100	77
红毛鸡	96	96	100	80
山斑鸠	92	93	100	87
水秧鸡	97	96	100	76
绣眼	94	95	100	81
野山鸡	96	97	100	83
竹鸡	93	94	100	78
鹧鸪	98	98	100	76
平均	95.5	95.7	100	79.8

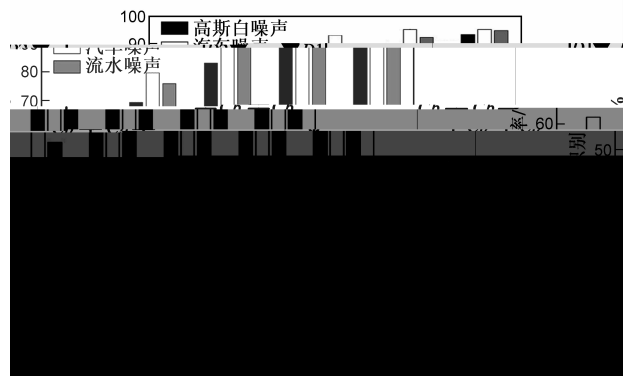
第 2 组实验,对比利用抗噪纹理特征快速鸟鸣声识别方法在不同噪声环境和不同信噪比情况下的识别率.为了模拟真实环境中的鸟鸣声,实验用到 3 种常用加性背景噪声,分别是高斯白噪声、汽车噪声和流水噪声.对 10 类纯净鸟鸣声共 400 个声音片段中的每一个声音片段按信噪比为 0dB、10dB、20dB、30dB 和 40dB 分别添加 3 种背景噪声,经过预处理和 STFT 后,实验甲不做增强,实验乙用短时谱估计进行音频增强,然后由 SDH 对相应功率谱提取纹理特征,接下来,采用 RF 分类器进行识别和分类,其训练和测试阶段设计同前一组实验.每次实验,分别记录实验甲和实验乙在不同噪声环境和不同信噪比下鸟鸣声的平均识别率,得到的 30 个值分别是不同条件下 10 类鸟鸣声 300 个测试样本识别率的均值.

图 4 展示了实验甲在不同噪声环境和不同信噪比

下的平均值别率.从中可以看出,没有经过音频增强的前端处理时,该方法对带噪声音频数据的识别性能并不良好,尤其是信噪比小于 30dB 时,三种不同噪声环境下,平均识别率均不到 80%.这表明,该方法在未经音频增强的前端处理时,并不能取得噪声环境下令人满意的识别性能,并且,不同的噪声环境对方法性能的影响也是不一样的,相比起来,汽车噪声的影响最小,高斯白噪声的影响最大.



图 5 是实验乙在不同噪声环境和不同信噪比下的平均识别率.具体分析两幅图中 30 个数据可知,经过增强后,平均识别率的涨幅最小为 6.5%,对应着汽车噪声环境下信噪比为 40dB 的情况;涨幅最大为 43%,对应着高斯白噪声环境下信噪比为 10dB 的情况.在信噪比介于 0dB 到 20dB 之间时,平均识别率整体有相当大的提高.显然,经过音频增强的前端处理后,全部三种不同噪声环境下,该方法的平均识别率都得到了可观的提升.但是对于信噪比接近 0dB 时,增强后的该方法性能依然不佳.这可能是由于噪声信号的能量大于等于有效音频信号的能量,在增强处理过程中,噪声功率谱的估计比较困难.



5 结论

本文实现了一种利用抗噪纹理特征快速鸟鸣声识别方法.该方法首先使用了一个短时谱估计的音频增强作为前端处理,接着由 SDH 对得到的增强功率谱

快速提取抗噪纹理特征并结合 RF 进行分类. 实验结果表明, 该方法在纹理特征提取阶段能够在保证良好识别率的前提下有效地减少时耗, 并且具有较好的抗噪性.

参考文献

- [1] Neal L, Briggs F, Raich R, et al. Time-frequency segmentation of bird song in noisy acoustic environments[A]. Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing[C]. Prague, Czech Republic: IEEE press, 2011. 2012 – 2015.
- [2] Lee C H, Hsu SB, Shih J L, et al. Continuous birdsong recognition using Gaussian mixture modeling of image shape features[J]. IEEE Transactions on Multimedia, 2013, 15(2): 454 – 464.
- [3] Sahar J, Zahra S, Mehdi A. A robust brain MRI classification with GLCM features[J]. International Journal of Computer Applications, 2012, 37(12): 1 – 5.
- [4] Benazir K K, Vijayakumar. Fingerprint matching by extracting GLCM features[A]. Proceedings of International Conference and Workshop on Emerging Trends in Technology[C]. New York, USA: Foundation of Computer Science press, 2012. 30 – 34.
- [5] RodeK N, Patil S S. Analysis of MS using GLCM[J]. International Journal of Science, Engineering and Technology Research, 2012, 1(5): 121 – 126.
- [6] 薄华, 马缚龙, 焦李成. 图像纹理的灰度共生矩阵计算问题的分析[J]. 电子学报, 2006, 34(1): 155 – 158.
Bo Hua, Ma Fulong, Jiao Licheng. Research on computation of GLCM of image texture[J]. Acta Electronica Sinica, 2006, 34(1): 155 – 158. (in Chinese)
- [7] Unser M. Sum and difference histograms for texture classification[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8(1): 118 – 125.
- [8] 冯霞, 龚晓峰, 张利丹, 武瑞娟. 基于纹理特征的背景噪声提取的应用研究[J]. 电子学报, 2009, 37(9): 2092 – 2095.

Feng Xia, Gong Xiaofeng, Zhang Lidan, Wu Ruijuan. Research of background noise extraction based on texture feature[J]. Acta Electronica Sinica, 2009, 37(9): 2092 – 2095. (in Chinese)

- [9] Bardeli R, Wolff D, Kurth F, et al. Detecting bird sounds in a complex acoustic environment and application to bioacoustics monitoring[J]. Pattern Recognition Letters, 2010, 31(12): 1524 – 1534.
- [10] Chu W, et al. Noise robust bird song detection using syllable pattern-based hidden Markov models[A]. Proceedings of International Conference on Acoustics, Speech, and Signal Processing[C]. Prague, Czech Republic: IEEE press, 2011. 345 – 348.
- [11] 刘翔, 高勇. 一种引入延迟的语音增强算法[J]. 现代电子技术, 2011, 34(5): 85 – 88.
- [12] Universitat Pompeu Fabra. Repository of Sound Under the Creative Commons License [DB/OL]. <http://www.freesound.org>, 2012-5-14.

作者简介



魏静明 男, 1988 年 11 月出生, 河南新乡人. 2014 年 3 月毕业于福州大学数学与计算机科学学院, 工学硕士, 读研期间主要研究方向为声音识别.

E-mail: wjm3219@yeah.net



李应 男, 1964 年 8 月出生, 福建福州闽清人. 现为福州大学数学与计算机科学学院教授, 主要研究方向为环境声音识别、信息安全.

E-mail: fj_liying@fzu.edu.cn