

# 基于自适应最优核时频分布的鸟类识别

孙斌 万鹏威 陶达 赵玉晓

(中国计量学院计量测试工程学院, 杭州, 310018)

**摘要:** 针对鸟鸣声信号的非稳态特性, 提出了一种基于自适应最优核时频分布 (Adaptive optimal kernel, AOK) 的鸟类识别方法。首先对采集的鸟鸣声信号进行预处理, 通过 AOK 时频分析方法得到时频谱图, 分析不同鸟类声音信号在不同时间和不同频率下的能量分布。然后, 将时频谱图转化成灰度图像, 求取灰度共生矩阵, 提取基于灰度共生矩阵不同角度的图像特征参数作为鸟类识别的特征值。最后选取已知鸟种的图像纹理特征训练生成训练模板, 将待识别的鸟种的图像纹理特征参数生成测试模板, 利用动态规整 (Dynamic time warping, DTW) 算法进行模板的匹配, 将匹配值进行大小比较, 找到最小匹配值对应的模板, 从而实现鸟类的识别。通过对 40 种常见鸟类的实验表明, 总体识别率达到 96%。

**关键词:** 自适应最优核时频分布; 灰度共生矩阵; 图像纹理特征; 动态规整

**中图分类号:** TP317.4      **文献标志码:** A

## Identification of Birds Based on Adaptive Optimal Kernel Time-Frequency Distribution

Sun Bin, Wan Pengwei, Tao Da, Zhao Yuxiao

(Institute of Measurement and Testing Engineering, China Jiliang University, Hangzhou, 310018, China)

**Abstract:** A bird identification method for the transient characteristics of birdsong signal based on adaptive optimal kernel (AOK) time-frequency distribution identification is proposed. The collected birdsong signal is preprocessed and the spectrum is obtained through the AOK time-frequency analysis method. Different energy distribution of birds sound signal at different time and different frequency are also analyzed. Then diagram spectrum is turned into gray image, the gray level co-occurrence matrix is calculated, image features is extracted as the eigenvalues of birds identification based on gray co-occurrence matrix parameters at different angles. Finally, the image texture of the known species is selected to generate training template and the image texture characteristic parameters of the species for identifying is used to generate the test template, Template matching is achieved using dynamic time warping (DTW) algorithm. The matching value are compared to find the minimum matching value corresponding templates, therefore the recognition of birds are realized. Finally, 40 kinds of common birds experiments demonstrate that the overall recognition rate reaches 96%.

**Key words:** AOK time-frequency distribution; gray level co-occurrence matrix; image texture characteristic parameter; dynamic time warping

## 引言

鸟声识别和说话人识别一样,都属于生物识别的一种,都是基于物种以及物间的差异性来实现的。现有的鸟声识别技术大都是利用鸟鸣声的短时平稳特性,对鸟鸣声进行分帧处理,然后提取梅尔倒谱系数(Mel cepstral coefficients, MFCC)或者线性预测倒谱系数(Linear prediction cepstral coefficients, LPCC)作为鸟鸣声信号的特征值点集。最后建立高斯混合模型(Gaussian mixture model, GMM)或者隐马尔可夫模型(Hidden Markov model, HMM)来进行识别。虽然现有技术都能取得很好的识别率,但存在着计算量偏大、数据量大以及识别耗时长等缺点<sup>[1]</sup>。针对现有技术的缺点,本文提出了一种基于自适应最优核时频分布(Adaptive optimal kernel, AOK)的鸟声识别方法。

时频分析方法是在单一的时域和频域分析不能满足实际处理的需要的前提下应运而生的,随着多年的发展,时频分析方法已经发展成为信号处理领域的主要研究热点。研究表明现阶段主要的时频分析方法如短时傅里叶变换、S变换、小波变换和 HHT 变换等方法普遍存在着较强的交叉项干扰和时频分辨率不高的弊端<sup>[2-3]</sup>。基于上述考虑,本文采用了 AOK 时频分析方法。该方法不仅具有时频分辨率高的特性,还具有很强的抗交叉干扰项的能力,相对其他常规的时频分析方法综合性能有着很大的提高<sup>[4]</sup>。AOK 时频分析方法应用广泛,已经应用于地震信号分析<sup>[5]</sup>,电路故障诊断分析<sup>[6]</sup>以及雷达信号分析<sup>[7]</sup>等领域中。

## 1 自适应最优核时频分析算法

### 1.1 最优核时频分布

最优核时频分布是一种主要的时频分析方法,它具有时频分辨率高的特点,但由于它是双线性表变换,在对多分量信号进行分析时会存在较严重的交叉项干扰而影响最后的分析结果。为了抑制交叉项的干扰,使信号的自主量能够分离出来,人们对最优核时频分布进行了改进,改进之后的时频分析统称为 Cohen 类双线性时频分布,其表达式为

$$C(t, \omega) = \frac{1}{2\pi} \iint A(\theta, \tau) \Phi(\theta, \tau) e^{-j\theta t - j\tau \omega} d\theta d\tau \quad (1)$$

式中:  $\Phi(\theta, \tau)$  为低通的核函数,  $A(\theta, \tau)$  为信号  $s(t)$  的模糊函数,  $t$  为时间,  $\omega$  为频率,  $\theta$  为频移,  $\tau$  为时移,它的定义为

$$A(\theta, \tau) = \int s(t + \frac{\tau}{2}) s^*(t - \frac{\tau}{2}) e^{j\theta t} dt \quad (2)$$

通过设计不同的核函数就能得到不同的分布特性,而为了在时频分布平面内不出现交叉项,则需要模糊函数域内有效去除互分量而仅保留自分量。将核函数定义为沿任意径向剖面都是 Gauss 型的二维函数可以达到上述目的。其表达式为

$$\Phi(\theta, \tau) = \exp(-\frac{\theta^2 + \tau^2}{2\sigma^2(\psi)}) \quad (3)$$

式中:  $\sigma(\psi)$  为控制径向高斯核函数在径向角  $\psi$  方向的扩展,  $\psi$  为径向与水平方向的夹角<sup>[8]</sup>。

### 1.2 自适应最优核时频分布

将核函数定义为沿任意径向剖面都是 Gauss 型的二维函数,虽然可以有效去除互分量而保留自分量的要求,但是其不适合于长时间非平稳信号的分析。为此,本文提出自适应最优核时频分布方法。定义短时模糊函数  $A(t, \theta, \tau)$ , 它是在窗函数截断后的一小段信号的模糊函数

$$A(t, \theta, \tau) = \int s^* \left(u - \frac{\tau}{2}\right) \omega^* \left(u - t - \frac{\tau}{2}\right) s \left(u + \frac{\tau}{2}\right) \omega \left(u - t + \frac{\tau}{2}\right) e^{j\theta u} du \quad (4)$$

式中:  $\omega(u)$  是对称的窗函数;  $t$  是  $\omega(u)$  的中心位置。令  $|u| > T$  ( $T$  是窗长) 时,  $\omega(u) = 0$ , 则在任一时刻  $t$ , 只有在  $[t - T, t + T]$  范围内的信号才可计算其模糊函数。对于信号的任意细节部分, 短时模糊函数都可准确地刻画出来, 此时可得与之相对应的自适应最优核函数  $\Phi_{\text{opt}}(t, \theta, \tau)$ 。则时间段  $[t - T, t + T]$  内信号的自适应最优核时频分布为

$$C_{\text{AOK}}(t, \omega) = \frac{1}{2\pi\alpha} \iint A(t, \theta, \tau) \Phi_{\text{opt}}(t, \theta, \tau) e^{-j\theta t - j\tau\omega} d\theta d\tau \quad (5)$$

## 2 鸟鸣声的 AOK 时频分析

本文进行分析和实验的鸟鸣声均取自鸟鸣声声音库(160 多种鸟), 该库鸟的种类齐全。声音库的鸟鸣声经音响外放(音响的频率响应范围在 60~20 kHz, 完全能够满足实验的要求), 再由 NI USB4432 搭配 NI IEPE 麦克风进行鸟鸣声数据的采集, NI USB-4432 提供了 108 dB 的动态范围, 能够在低频信号中精确地捕获高频率信息, 而且采集数据能够达到 102.4 k/s。

### 2.1 鸟鸣声的预处理

#### 2.1.1 预加重

本文鸟鸣声的预处理过程主要包括预加重、滤波和归一化<sup>[9]</sup>。鸟鸣声信号与其他声音信号一样, 会随着频率的增长而呈现指数级衰减, 会使信号的高频成分越来越少, 低频成分越来越多, 从而不利于信号的分析<sup>[10]</sup>。信号的预加重是一种提高信号高频成分的方法, 而低频成分则保持原来的水平, 使信号的频谱变得平坦, 以便于进行频谱分析或声道参数的分析。预加重滤波器的传递函数为

$$H(z) = 1 - \partial z^{-1} \quad (6)$$

式中:  $\partial$  为预加重参数(预加重因子)。在声音识别系统中,  $\partial$  取 0.93 和 0.95 时识别率最高, 本文取 0.937 5。

#### 2.1.2 归一化

鸟鸣声信号的归一化处理就是将声音信号的每一个采样值除以该采样信号的幅度峰值

$$\tilde{x}(i) = x(i) / \max_{0 \leq i \leq n-1} x(i) \quad (7)$$

式中:  $x(i)$  为原始采集声音型号,  $\tilde{x}(i)$  为归一化后的信号,  $n$  为信号长度。

### 2.2 鸟鸣声的 AOK 时频分析

对采集的鸟鸣声信号进行 FFT 变换, 观察其频谱图, 可以看出鸟鸣声信号的频率一般都在 1 000 Hz 以内。根据奈奎斯特定理, 本文鸟鸣声信号的采样频率采用 2 000 Hz, 采集时长 5 s 或者 10 s。

将预处理过后的鸟鸣声进行 AOK 时频分析, 图 1 为公竹鸡鸣声信号的 AOK 时频分析结果。其中图 1(a) 是信号的功率谱图, 从图中可以看出信号包含的频率及其能量的大小, 但看不出时间轴的变化情况。图 1(b) 是预处理和归一化之后的鸣声信号。图 1(c) 是信号的边际谱图, 从图中也可以看出信号包含的频率及其能量的大小, 并且和功率谱图一一对应。图 1(d) 为信号的 AOK 分析时频谱图(横坐标是  $t/s$ , 纵坐标  $f/Hz$ ), 从图中可以清晰地看出信号包含的频率及其能量的大小以及时间轴的变化情况。从图中可以得到信号的能量主要集中在(100~200 Hz)且随着时间的变化呈现阶段性变化。图 2 为大天鹅鸣声信号的 AOK 时频分析结果, 从大天鹅鸣声信号的 AOK 时频图中可以看出信号的能量主要集中在 800~900 Hz。比较图 1 和图 2, 可以明显看出不同鸟类之间的鸣声差异, 为后期的鸟类识别提供了强有力的依据。图 3 和图 4 分别为四喜鸟和麻雀的鸣声信号的 AOK 时频谱图分析结果。从图中可以看出其与图 1, 图 2 的差别。从图 3 可以看出信号的能量分布主要集中在 400~500 Hz, 从图 4 可以

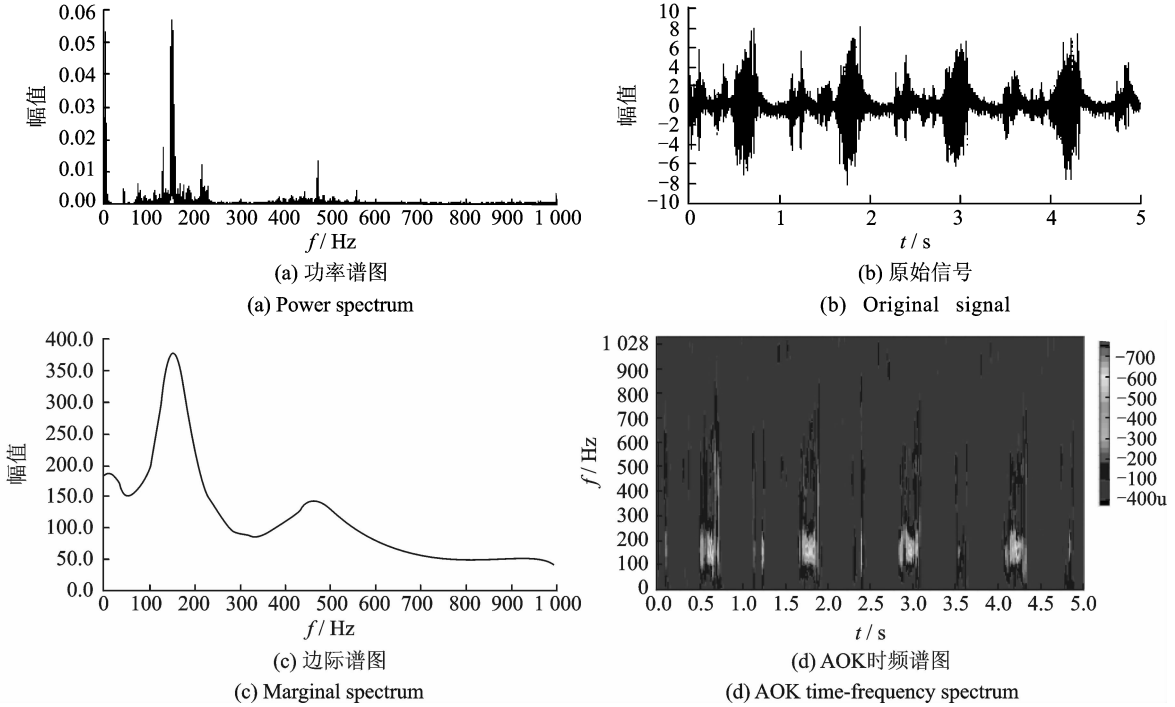


图1 公竹鸡鸣声信号的 AOK 时频分析结果

Fig. 1 AOK frequency analysis results of male bamboo crowing sound signal

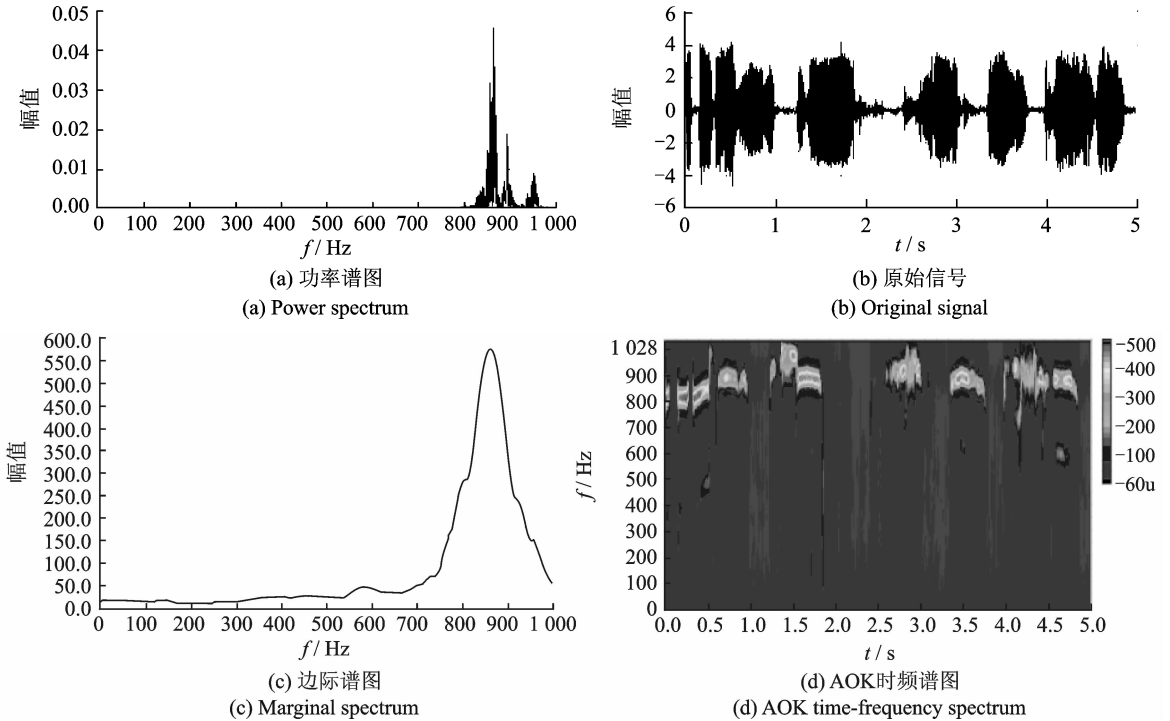


图2 大天鹅鸣声信号的 AOK 时频分析结果

Fig. 2 AOK frequency analysis results of swan sound signal

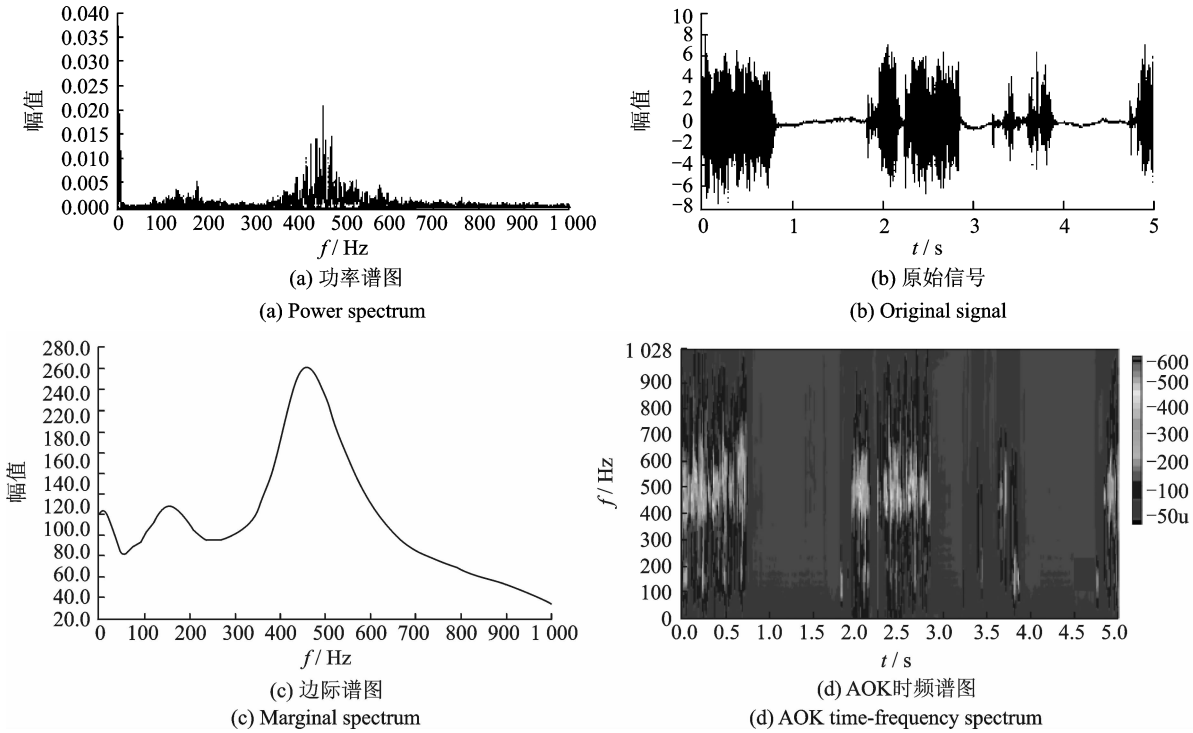


图 3 四喜鸟鸣声信号的 AOK 时频分析结果

Fig. 3 AOK frequency analysis results of Sixi bird sound signal

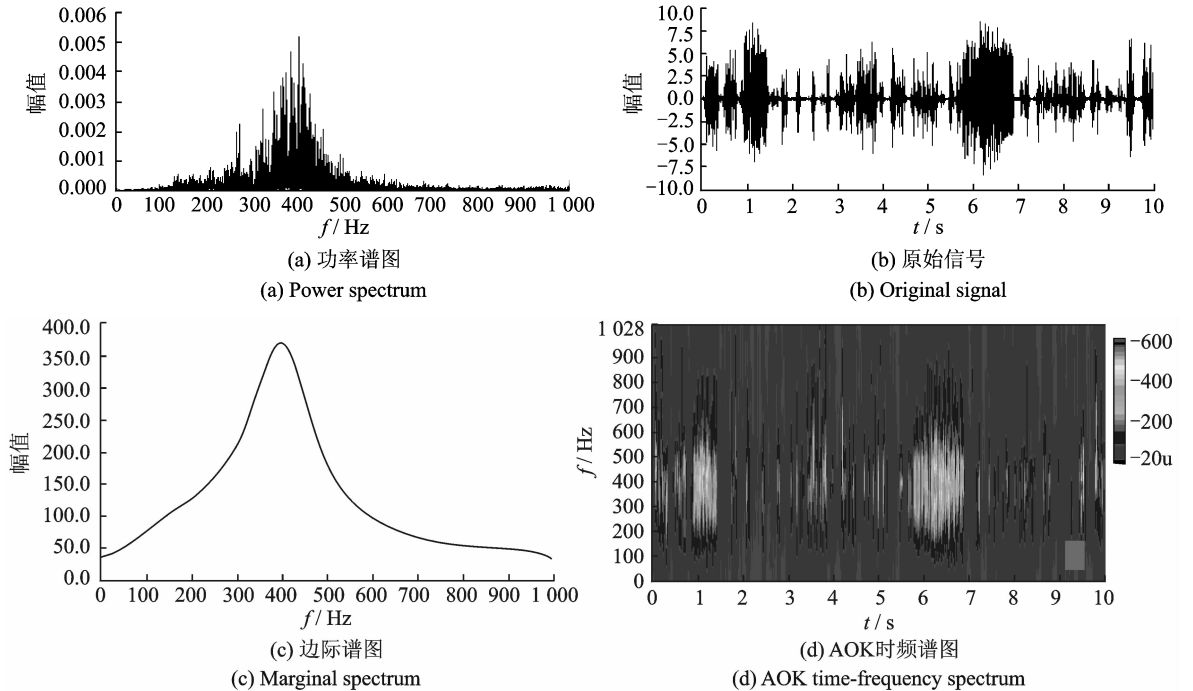


图 4 麻雀鸣声信号的 AOK 时频分析结果

Fig. 4 AOK frequency analysis results of sparrow sound signal

看出信号的能量分布集中在 350~450 Hz。比较图 3 和图 4, 可以明显看出两者既有相似性也有差异性。相似性表现在信号的能量分布有重叠部分, 差异性表现出两者的信号在时间轴上的变化。差异性也为后期相似鸟种之间的识别提供了依据。从以上分析可知, 从 AOK 时频谱图中可以清晰地看到不同鸟种的鸟类以及相似鸟种的鸟类鸣声信号的差异性。依此为依据, 可以根据鸟鸣声的 AOK 时频谱图的差异性来进行鸟类的识别。

### 3 图像纹理特征参数提取

#### 3.1 灰度共生矩阵

一幅图像的灰度共生矩阵能反映出图像灰度关于方向、相邻间隔以及变化幅度的综合信息。基于此, 灰度共生矩阵已经成为一种应用十分广泛的纹理分析方法。本文通过计算灰度共生矩阵的图像纹理特征参数提取 AOK 时频图像的特征。如图 5 所示, 灰度共生矩阵实质是从图像灰度为  $i$  的像元位置  $(x, y)$  出发, 统计与其距离为  $d$ , 灰度为  $j$  的像元位置  $(x + Dx, y + Dy)$  同时出现的频度  $P(i, j, d, \theta)$ 。本文的生成步长  $d = 2$ , 生成方向取  $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$  四个不同的方向。灰度共生矩阵的大小与灰度图的灰度级别有关, 本文的时频谱图灰度变换之后得到的灰度级别是 8, 最后得到的灰度共生矩阵是一个  $8 \times 8$  的矩阵<sup>[11-13]</sup>。

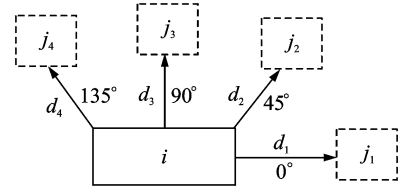


图 5 灰度共生矩阵的四个生成方向  
Fig. 5 Four generation directions of gray level co-occurrence matrix

#### 3.2 灰度共生矩阵的特征参数

灰度共生矩阵有 15 个特征参数<sup>[14]</sup>。本文选用具有代表性的对比度、相关、能量和逆差距(同质性) 4 个特征参数。

(1) 对比度 对比度反映图像的清晰度和纹理沟纹深浅的程度。图像纹理沟纹越深, 其对比度就越大, 图像就越清晰, 反之亦然。其表达式为

$$\text{con} = \sum_{i,j} |i - j|^2 p(i, j) \quad (8)$$

(2) 相关 相关是度量空间灰度共生矩阵元素在行或者列方向上的相似程度。当灰度共生矩阵的元素值均匀相等时, 相关值就越大, 反之, 相关值就越小。其表达式为

$$\text{cor} = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j) p(i, j)}{\sigma_i \sigma_j} \quad (9)$$

(3) 能量 能量是灰度共生矩阵元素的平方和, 也称二阶矩。反映了图像灰度分布均匀程度和纹理粗细程度。其表达式为

$$\text{ene} = \sum_{i,j} p(i, j)^2 \quad (10)$$

(4) 逆差距 逆差距又称同质性, 度量图像纹理局部变化的多少, 逆差距越大, 说明图像纹理的不同区域间缺少变化, 非常均匀, 反之亦然, 其表达式为

$$\text{hom} = \sum_{i,j} \frac{p(i, j)}{1 + |i - j|} \quad (11)$$

#### 3.3 基于时频谱图特征参数

本文的灰度图是时频谱图经灰度变化得到的, 从时频谱图上就可以看出灰度图在形状以及灰度值上存在着的差异。为了量化这种差异, 本文提取了图像 4 个方向 ( $\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ ,  $d = 2$ ) 灰度共生

矩阵的4个特征参数。由此可以得到每种鸟将共计提取了16个特征参数A1~A16;其中A1~A4代表对比度,A5~A8代表相关,A9~A12代表能量,A13~A16代表逆差距。

结合鸟鸣声的AOK时频分析中的4种鸟进行分析,将4种鸟的时频谱图进行灰度变换得到灰度图像,然后分别求取了灰度图像4个方向的灰度共生矩阵,最后提取了基于灰度共生矩阵的图像纹理特征参数如表1所示。

表1 图像纹理特征参数 A1-A16

Table 1 Image texture parameters A1-A16

鸟的种类	A1	A2	A3	A4	A5	A6	A7	A8
公竹鸡	0.774 8	0.860 7	0.099 6	0.860 6	0.244 2	0.095 1	0.907 3	0.096 4
大天鹅	0.418 5	0.529 8	0.193 9	0.525 1	0.615 5	0.478 2	0.820 3	0.486 7
四喜	0.450 8	0.536 8	0.102 7	0.533 9	0.349 0	0.126 6	0.850 9	0.135 1
地麻雀	0.677 5	0.769 0	0.105 4	0.767 6	0.219 2	0.028 9	0.878 8	0.031 9
鸟的种类	A9	A10	A11	A12	A13	A14	A15	A16
公竹鸡	0.414 5	0.412 5	0.573 2	0.412 8	0.841 0	0.834 4	0.985 4	0.834 7
大天鹅	0.636 6	0.632 3	0.686 7	0.632 9	0.925 0	0.915 5	0.961 0	0.915 8
四喜	0.582 0	0.581 1	0.708 8	0.581 3	0.896 1	0.889 9	0.985 1	0.890 0
地麻雀	0.438 2	0.436 9	0.620 1	0.437 1	0.840 9	0.834 5	0.985 3	0.834 7

纵观表1可以看出,对于任一种鸟类而言,基于不同方向的灰度共生矩阵提取的图像纹理特征参数存在着较大差异,这也为本文对每种鸟求取了4个不同方向的灰度共生矩阵来提取图像特征参数提供了依据。对比4种鸟,4个图像纹理特征参数在同一方向上也存在着差异,这验证了AOK时频谱图以及灰度变化得到的灰度图像存在着差异的说法,也从侧面说明了提取出来的特征值的可行性。

### 4 动态规整算法

动态规整(Dynamic time warping,DTW)算法是把时间规整和距离测度计算结合起来的一种非线性规整技术,是采用动态规划技术将一个复杂的全局最优问题化为很多局部最优问题,再逐步地进行决策的方法<sup>[15-16]</sup>。如图6所示是DTW算法的简易示意图,A,B分别为训练模板和测试模板的对应的矢量点,C则是两者之间的非线性映射关系。

本文使用DTW算法将待识别鸟的叫声生成的测试模板与已知种类的鸟的叫声生成的训练模板逐一进行匹配计算,从训练模板中选择最佳匹配模板来作为识别结果。图7就是系统的识别过程示意图。

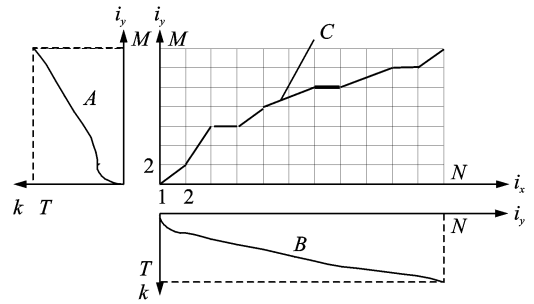


图6 DTW示意图

Fig. 6 DWT schematic diagram

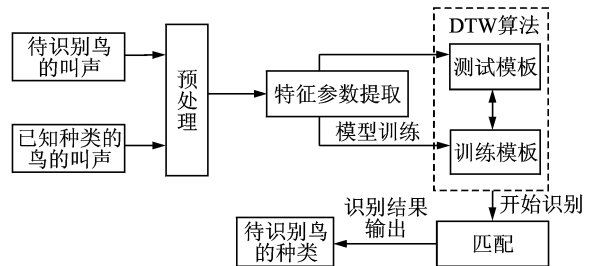


图7 系统的识别过程示意图

Fig. 7 Schematic diagram of system identification process

### 5 实验结果分析

本文选取了杭州地区常见的 40 种鸟作为实验对象<sup>[17-18]</sup>。第 1 次实验首先对 40 种鸟 1 次进行 3 次鸟鸣声数据的采集,将采集的数据进行模型训练生成训练模板。然后再对每种鸟依次进行采集和识别,每种鸟识别 10 次。实验结果取得了很好的识别率,大多数的鸟类都能识别出来,但是仍有少部分的鸟类出现了误识别。因此,进行了第 2 次实验,实验在第 1 次实验的基础之上对每种鸟都增加了 1 个训练模板,将每种鸟的训练模板数增加到 4 个,然后同样对每种鸟进行采集和识别,同样识别 10 次。实验结果表明,增加训练模板数,鸟类的识别率会提高,主要体现在实验 1 出现的误识别数减少了。为了解决鸟类识别率是否会随着训练模板数的增加而提高这个问题,紧接着做了 3 次实验,实验在实验 2 的基础上依次增加训练模板的数量,实验结果表明鸟类的识别率随着训练模板数的增加不是一直提高的,会有个临界值。表 2 给出的以上实验的鸟类训练模板数以及对应的识别率。表 3 给出的是训练模板数为 6 时每种鸟对应的识别情况。

表 2 训练模板数和识别率

Table 2 Training module number and recognition rate

训练模板数	鸟类识别率/%
3	90.5
4	93.75
5	95.5
6	96
7	95.25

表 3 鸟类的识别情况

Table 3 Identification of birds

鸟的种类	白面水鸡	公竹鸡	白胸翡翠	夜游	大天鹅	禾花雀	灰喜鹊	五色鸟	麻雀	鹌鹑	八哥	白颊噪鹛	相思鸟	
识别总数	10	10	10	10	10	10	10	10	10	10	10	10	10	
误识别数	0	0	1	0	0	2	0	0	2	0	0	0	1	
鸟的种类	伯劳	画眉	斑鸠	白纹腰鸟	斑胸田鸡	大山雀	戴胜	董鸡	杜鹃	丝光掠鸟	四喜鸟	鹳鸟	乌鸫	
识别总数	10	10	10	10	10	10	10	10	10	10	10	10	10	
误识别数	0	1	0	0	0	1	0	0	2	1	1	0	0	
鸟的种类	黑田鸡	黑尾蜡嘴	红耳鹎	松鸦	红胸秧鸡	红毛鸡	长尾蓝鹊	黄雀	黄腾鸟	金翅雀	金丝雀	麻鸡	翠鸟	山胡
识别总数	10	10	10	10	10	10	10	10	10	10	10	10	10	10
误识别数	0	1	0	0	1	0	0	0	0	0	1	0	0	1

### 6 结束语

本文通过对鸟鸣声信号的研究,设计了一套基于 AOK 时频分析的鸟类识别系统。本系统可以尽可能地降低外界噪音以及不同时间段鸟鸣声差异性对实验结果的影响,通过 AOK(自适应最优核时频分布)时频谱图可以明显看出不同鸟类的时频差异,基于时频谱图获取的特征值数据量小且具有代表性。识别过程只需要计算测试模板和训练模板的距离差即可以得到最后的识别结果,具有计算量小、识别效率高等优点。

#### 参考文献:

[1] 薛宇志. 依照鸟类鸣叫与鸣唱声识别其种类[D]. 中国台北:国立台北科技大学,1998.  
Xue Yuzhi. According to the birds chirping and singing sound to identify the type [D]. China, Taipei: National Taipei University of Science and Technology, 1998.

[2] 陈雨红, 杨长春, 曹齐放, 等. 几种时频分析方法比较[J]. 地球物理学进展, 2006, 21(4): 1181-1185.  
Chen Yuhong, Yang Changchun, Cao Qifang, et al. Several time-frequency analysis methods comparison [J]. Journal of Ge-



ophysics, 2006, 21 (4): 1181-1185.

- [3] Care M. Time-frequency analysis[J]. Eos, Transactions American Geophysical Union, 1982, 63(51):13-20.
- [4] 李亚安, 王军, 雷粉霞. 自适应时频分布在抑制交叉项中的应用[J]. 系统工程与电子技术, 2004, 26(11):1567-1569.  
Li Yaan, Wang Jun, Lei Fenxia. Adaptive time-frequency distribution in the cross terms of application [J]. Journal of Systems Engineering and Electronics, 2004, 26 (11): 1567-1569.
- [5] 马会星, 王立振, 杨军. 基于 AOK 时频分布的窄带雷达目标架次分辨方法[J]. 空军雷达学院学报, 2012, 26(2):98-100.  
Ma Huixing, Wang Lizhen, Yang jun. Based on the AOK time-frequency distribution of narrowband radar target vehicles discrimination method [J]. Journal of Air Force Radar College Journal, 2012, 26 (2): 98-100.
- [6] 刘小龙, 王华, 赵淑娥, 等. 自适应最优核时频分布在地震储层预测中的应用[J]. 中南大学学报:自然科学版, 2012, 43(8): 3114-3120  
Liu Xiaolong, Wang Hua, Zhao Shue, et al. The adaptive optimal kernel time-frequency distribution in the application of seismic reservoir prediction [J]. Journal of Central South University: Natural Science Edition, 2012(8): 3114-3120.
- [7] 金国彬, 李玲, 黄邵平, 等. 基于 AOK TFR 理论的电力系统暂态信号分析新方法[J]. 电力自动化设备, 2008, 28(5):63-67.  
Jin Guobin, Li ling, Huang Shaoping, et al. Based on the AOK TFR theory, a new method of power system transient signal analysis [J]. Electric Power Automation Equipment, 2008, 28 (5): 63-67.
- [8] 胡广书. 现代信号处理教程[M]. 北京:清华大学出版社, 2004.  
Hu Guangshu. Modern signal processing tutorial [M]. Beijing: Tsinghua University Press, 2004.
- [9] 白雅, 尚志远. 珍稀鸟类朱鹮鸣声特征与行为关系的研究[D]. 西安:陕西师范大学, 2005.  
Bai Ya, Shang Zhiyuan. Rare birds chirp crested ibis characteristics and behavior relation research [D]. Xian: Shanxi Normal University, 2005.
- [10] 张万里, 刘桥. Mel 频率倒谱系数提取及其在声纹识别中的作用[J]. 贵州大学学报, 2005, 22(2):5.  
Zhang Wanli, Liu qiao. Mel frequency cepstrum coefficient of extraction and its application in the role of voiceprint recognition [J]. Journal of Guizhou University, 2005, 22(2):5.
- [11] 庄军, 李弼程. 一种基于灰度共生矩阵的文本图像识别方法[J]. 计算机工程, 2006, 32(3):214-216.  
Zhuang Jun, Li Bicheng. A text image recognition method based on gray level co-occurrence matrix [J]. Computer Engineering, 2006, 32 (3) : 214-216.
- [12] 李慧娜, 郭超峰, 平原. 灰度共生矩阵在指纹图像分割中的应用[J]. 数据采集与处理, 2012, 27(1):63-67.  
Li Huina, Guo Fengchao, Ping Yuan. Application of gray level co-occurrence matrix for fingerprint segmentation [J]. Journal of Data Acquisition and Processing, 2012, 27(1):63-67.
- [13] Haralick R M. Statistical and structural approaches to texture[J]. Proceeding of IEEE, 1975, 67(5):786-504.
- [14] Baraedi A, Parmiggian F. An investigation of texture characteristics associated with gray level co-occurrence matrix statistical parameters [J]. IEEE Trans On Geo-Science and Remote Sensing, 1995, 33(2):293-303.
- [15] 汪力, 叶桦, 夏良正. 一种结合形状序列和 DTW 的动作识别方法[J]. 数据采集与处理, 2009, 24(5):615-619.  
Wang Li, Ye Hua, Xia Liangzheng. Activity recognition method for combining shape sequence and DTW [J]. Journal of Data Acquisition and Processing, 2009, 24(5):615-619.
- [16] Michael K. Lawrence B, Rabiner R. An adaptive, ordered, graph search technique for dynamic time warping for isolated word recognition[J]. IEEE Transactions on Acoustics, Speech, and Signal Processing, 1982, 1(4):535-544.
- [17] 朱曦, 陈勤娟, 詹伟君, 等. 杭州市鸟类区系研究[J]. 浙江林学院学报, 2002, 19(1):36-47.  
Zhu Xi, Chen Qinjuan, Zhan Weijun, et al. Hangzhou birds fauna study [J]. Journal of Zhejiang Forestry College, 2002, 12 (1): 36-47.
- [18] 于炜, 吕敏, 刘锦. 杭州植物园野生鸟类多样性调查[J]. 野生动物, 2011, 32(5):277-281.  
Yu Hui, Lü Min, Liu Jin. Wild birds diversity investigation in Hangzhou botanical garden [J]. Journal of Wildlife, 2011, 32 (5): 277-281.

#### 作者简介:



**孙斌** (1972-), 男, 教授, 研究方向: 过程参数检测与远程测控系统设计、计量及标准装置自动化、远程实验及校准系统设计, E-mail: bsun@555. cjlu. edu. cn.

**万鹏威** (1989-), 男, 研究生, 研究方向: 语音识别及信号处理。

**赵玉晓** (1972-), 女, 副教授, 研究方向: 计量测试技术。

**陶达** (1992-), 男, 本科, 研究方向: 语音识别及处理。