

羊咳嗽声的特征参数提取与识别方法

宣传忠¹ 武佩¹ 张丽娜² 马彦华¹ 张永安³ 邬娟¹

(1. 内蒙古农业大学机电工程学院, 呼和浩特 010018; 2. 内蒙古师范大学物理与电子信息学院, 呼和浩特 010022;
3. 内蒙古农业大学计算机与信息工程学院, 呼和浩特 010018)

摘要: 为在设施圈养羊只产生呼吸道疾病的初期,通过监测其咳嗽声进行疾病预警和健康状况诊断,以内蒙古地区广泛推广的杜泊羊为例,对杜泊羊的咳嗽声信号进行自动采集和计算机识别,在不增加羊咳嗽声特征参数维数的前提下,提出一种改进的梅尔频率倒谱系数(MFCC),试验结果表明,该参数和短时能量、过零率组合的14维特征参数,经过羊咳嗽声隐马尔可夫模型(HMM)识别系统,其识别率、误识别率和总识别率分别达到了86.23%、7.17%和88.43%,该组合特征参数经主成分分析可降到9维,而通过BP神经网络改善的HMM咳嗽声识别系统,对咳嗽声的识别率、误识别率和总识别率分别达到了92.54%、5.37%和95.04%,满足了杜泊羊咳嗽声识别的要求。

关键词: 杜泊羊; 咳嗽声; 特征参数提取; 梅尔频率倒谱系数; 隐马尔可夫模型

中图分类号: S858.266.3 **文献标识码:** A **文章编号:** 1000-1298(2016)03-0342-07

Feature Parameters Extraction and Recognition Method of Sheep Cough Sound

Xuan Chuanzhong¹ Wu Pei¹ Zhang Li'na² Ma Yanhua¹ Zhang Yongan³ Wu Juan¹

(1. College of Machine and Electronics Engineering, Inner Mongolia Agricultural University, Huhhot 010018, China
2. College of Physics and Electronic Information Science, Inner Mongolia Normal University, Huhhot 010022, China
3. College of Computer and Information Engineering, Inner Mongolia Agricultural University, Huhhot 010018, China)

Abstract: In farming region of Inner Mongolia, animal husbandry is evolving from the traditional style to the modern style, which means the large-scale sheep breeding, intensive management and industrial development. However, the newly extensive stable breeding facilities are easily to make sheep suffer from respiratory disease. In the early stage, cough sound of sheep can be detected for early disease warning and health diagnosis. In this paper, taking Dorper sheep, which has been widely promoted in Inner Mongolia, for an example, cough sound signal of sheep was automatically collected and recognized by computer. Without increasing the dimension of sound signal feature parameters, an improved Mel frequency cepstrum coefficient (MFCC) was put forward. The experimental results demonstrated that the 14-dimensional parameters combined with improved MFCC, short-time energy and zero crossing rate were used in the hidden Markov model (HMM) cough sound recognition system, whose recognition rate, error recognition rate and total recognition rate reached 86.23%, 7.17% and 88.43% respectively. And the combination parameters can be reduced to nine dimensions using principal components analysis (PCA) method. Furthermore, the cough sound recognition system based on HMM was enhanced by a back-propagation (BP) neural network, and it's recognition rate, error recognition rate and total recognition

收稿日期: 2015-09-14 修回日期: 2015-11-16

基金项目: “十二五”国家科技支撑计划项目(2014BAD08B05)、国家自然科学基金项目(11364029)、内蒙古自然科学基金项目(2012MS0720)、内蒙古“草原英才”产业创新人才团队项目(内组通字[2014]27号)和内蒙古农业大学科技创新团队项目(NDTD2013-6)

作者简介: 宣传忠(1977—),男,副教授,博士生,主要从事设施养殖福利化测试与装备研究,E-mail: xcz@imau.edu.cn

通信作者: 武佩(1963—),男,教授,博士生导师,主要从事农业工程测试与控制研究,E-mail: jdwp@imau.edu.cn

rate reached 92.54% , 5.37% and 95.04% , respectively. Therefore, the recognition results meet the requirement of the Dorper sheep cough sound recognition.

Key words: Dorper sheep; cough sound; feature parameter extraction; Mel frequency cepstrum coefficient; hidden Markov model

引言

随着我国肉羊产业发展,内蒙古及周边的农牧区正进行着肉羊产业的结构调整,改变原来的个体散养、效益低下的养殖模式,实施养殖大户、联户经营和专业合作社等经营形式,实现羊只规模化饲养、集约化经营、标准化生产和产业化带动,促进传统畜牧业向现代畜牧业的转变。由于杜泊羔羊断奶时质量大、生长迅速且有早期采食的能力,在肉羊良种繁育推广中,杜泊羊被作为主要的杂交改良品种之一,内蒙古地区的杜泊羊规模化设施圈养改变了羊原来放养的生活习性,使羊的生活和生理过程发生了变化,但饲养人员往往忽视了这一变化。在冬季和早春枯草季节,羊只营养缺乏且机体抵抗力降低,加之养羊管理和环境调控不到位,易导致圈养杜泊羊产生一些呼吸道疾病,在发病初期,杜泊羊食欲和反刍减损,症状临床表现为咳嗽、呼噜和肠音,因此可以通过监测羊的咳嗽声进行早期疾病预警和研判杜泊羊的健康状况。

如何快速诊断畜禽的健康状况以及评价畜禽福利化水平已成为畜牧科学领域中的研究热点。国内外学者对人的咳嗽声进行了相关的检测和分析^[1-4],而国内对家畜咳嗽声的数字化分析未见报道。文献[5-10]进行了猪咳嗽声的研究,指出咳嗽声可反馈动物的呼吸道疾病,猪发出的咳嗽声主要与猪舍的温度、通风等环境条件有关,同时借鉴人类语音识别技术的成果,对生猪的咳嗽声进行分类识别,系统地进行了猪咳嗽声信号的研究。FERRARI 等^[11]对奶牛的咳嗽声信号进行振幅、频率和持续时间的提取,开发了奶牛咳嗽声自动识别系统。对这些家畜咳嗽声信号的研究和监测,为构建家畜疾病早期的智能预警系统提供了理论依据。

通过监测杜泊羊咳嗽声来评估其生活环境适应性和健康状况,这种方法与传统生理生化参数指标检测方法相比,具有非侵入和无接触的优点。本文采集杜泊羊的咳嗽声信号,并进行信号的滤波、预加重、分帧加窗和端点检测等预处理,利用改进的梅尔频率倒谱系数(Mel frequency cepstrum coefficient, MFCC)提取杜泊羊的咳嗽声特征参数,并建立咳嗽声信号的隐马尔可夫模型(Hidden Markov model,

HMM),将改进的 MFCC 特征参数作为 HMM 识别系统的输入进行咳嗽声的训练和识别,并通过主成分分析(Principal component analysis,PCA)法降低特征参数的维数,以降低 HMM 系统的复杂度和提高系统的识别效率。

1 咳嗽声采集和预处理

杜泊羊咳嗽声试验数据采集在内蒙古赛诺草原羊业有限公司良种肉羊繁殖场进行,该繁殖场位于内蒙古自治区四子王旗,拥有杜泊种羊约 6 500 只,年可育肥杜泊杂交肉羊约 10 万只,是国内最大的黑头杜泊羊繁殖基地。

1.1 咳嗽声采集

杜泊羊咳嗽声信号的采集选用美国国家仪器有限公司生产的无线动态信号采集设备 NI WLS-9234(24 位分辨率、4 通道同步采集、102 dB 动态范围、同步采样率可达 51.24 kS/s),该设备将采集的声音数据通过 WiFi 无线的方式传输到监测计算机中,与采集设备相连接的声音传感器为北京声望声电技术有限公司生产的 MPA201(响应频率为 0.004~20 kHz,灵敏度为 40 mV/Pa)。杜泊羊咳嗽声信号进行单声道采集,采样精度为 16 位,采样频率为 32 kHz。

羊在患不同病时产生的咳嗽声均有所不同,如:当羊气管有异物,患胸膜炎、慢性支气管炎和肺结核等时,其咳嗽声干而短;当羊患咽喉炎、支气管炎和支气管肺炎等时,其咳嗽声湿而长;而当羊患急性喉炎、胸膜炎和喉水肿等时,其咳嗽声短而弱,因此羊的咳嗽声种类较多。试验在 3—4 月份羊病多发期进行,共采集 30 只 3~6 个月的杜泊羊 600 个咳嗽声样本和 300 个非咳嗽声样本,非咳嗽声包括 60 个饥饿声,60 个母羊寻羔声,60 个羊打斗声,60 个采食声和 60 个撞击声,其中有 15 只羊在发病初期,可直接采集其咳嗽声;剩下的杜泊羊采用诱发性咳嗽,使用致咳剂辣椒素或柠檬酸,刺激杜泊羊发出咳嗽声。使用致咳剂与羊患病所产生的咳嗽声都属于反射性咳嗽,试验表明,致咳剂与羊患病所产生的咳嗽声在时域和频域上具有相似性,故另采集 15 只羊的致咳剂诱发咳嗽声,用于增加咳嗽声的试验样本。咳嗽声的每次采样时间为 10~30 min 不等,采集的声信号以 wav 格式保存。采集 600 个咳嗽声样本用

于特征参数提取,从各类声音样本中随机选取相应的 1/2 样本用于识别模型库的训练,剩余 1/2 数据作为识别系统的测试样本集。

采集的杜泊羊咳嗽声信号主要分为 2 类,第 1

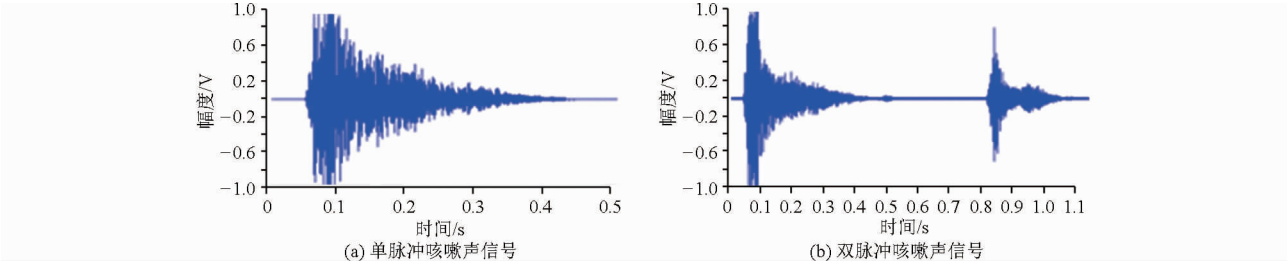


图 1 咳嗽声信号的时域波形

Fig. 1 Time-domain waveform of cough sound signal

1.2 咳嗽声预处理

由于杜泊羊咳嗽声信号的随机性,需要经过声信号的预处理,剔除与咳嗽声不相关以及干扰信息,预处理过程包括滤波与预加重、分帧与加窗和端点检测共 3 部分。

(1) 滤波与预加重

养殖场内采集的咳嗽声信号也包含着风机运行、羊走动等各种噪声信号,研究报告表明,咳嗽声的频率一般分布在 30 ~ 4 000 Hz 内,采用的数字滤波器的通带频率为 4 kHz,通带内衰减不大于 3 dB,阻带频率为 5 kHz,阻带内衰减大于 80 dB。通过预加重数字滤波器来提升信号的高频部分,使得信号从低频到高频的整个频带内频谱变平滑,用同样的信噪比去求频谱,利于咳嗽声信号特征参数的提取。

(2) 分帧与加窗

窗函数形状和长度是咳嗽声信号时域分析的 2 个重要参数,由于汉明窗能较好地消除旁瓣且主瓣相对较窄,使得信号帧与帧之间的起始与结束点处的连续性较好,故选择汉明窗。对于采样频率为 32 kHz 以下的信号,窗长一般在 100 ~ 300 之间,本文选取帧长为 256,帧移为 80。

(3) 端点检测

杜泊羊咳嗽声端点检测是从采集到的声信号中消除无声或部分清音,找出咳嗽信号的起止点,本文采用短时过零率与短时能量相结合的方法来检测信号的起止点。

2 咳嗽声特征参数提取

杜泊羊的咳嗽声信号能量随时间变化较大且时间短促,有着湍流性和突发性的特点,是一种典型的非平稳信号,而一般的语音信号相对比较平稳,因此咳嗽声信号没有明显的共振峰,对于那些与共振峰关联较大的特征参数提取方法,如线性预测倒谱系

数是单咳嗽脉冲,气流快速通过气管,咳嗽声信号以高频成分居多;第 2 类是双咳嗽脉冲,主要是由于气管痰较多被堵塞而形成,如图 1 所示,对于双咳嗽脉冲,在咳嗽声端点检测时将被分割成 2 个咳嗽声。

数(Linear predictive cepstrum coefficient,LPCC)就不能取得较好效果^[12-13]。基于咳嗽声信号的这一特征,本文选用基于人耳听觉特性的梅尔频率倒谱系数(MFCC)作为主要特征参数,并对该特征参数进行改进。

2.1 改进的 MFCC 参数

标准的倒谱参数 MFCC 只考虑到了杜泊羊咳嗽声的静态特性,而没有反映咳嗽声的动态特性,咳嗽声的动态特性可以用静态特征参数的差分来描述。目前学者常用的方法是组合 MFCC 与其一阶差分,或利用功率谱差来改进 MFCC 等,这些方法取得了较好的性能^[14],但增加了特征参数的维数,使得计算的复杂度增加,后续开发咳嗽声实时快速检测系统就变得困难。本文采用一种新的改进 MFCC 参数方法,在不增加特征参数维数的前提下,能够描述咳嗽声动态特性,设该特征参数为 MFCC_N,MFCC_N(用 β_N 表示)参数与 MFCC(用 β 表示)之间的关系式为

β_N = aβ + bΔβ + cΔ²β (1)

式中 a、b、c 代表权重,且 a + b + c = 1,Δβ 是 MFCC 的一阶差分,Δ²β 是 MFCC 的二阶差分。β 代表杜泊羊声道特性,Δβ 描述动态特性,表示特征参数变化的速度,Δ²β 作为平衡因子,表示特征参数变化的加速度。对于咳嗽声的 U 帧 V 维,标准 β 加一阶差分方法提取的特征参数矩阵的大小为 U × (2V),而改进的 β_N 特征参数矩阵的大小为 U × V,β_N 特征参数矩阵的维数没有增加。

图 2 为咳嗽声的 MFCC 特征参数图,X 轴、Y 轴和 Z 轴分别代表咳嗽声的帧数、维数和特征参数倒谱值,β 参数的 a、b、c 权重取值分别为 0.5、0.35、0.15。从图 2 可出,β 和 β_N 有相同幅度的变化趋势,但 β_N 参数不仅反映了声道特性,而且还能反映出咳嗽声的动态变化特性,具有更好的性能,本文采

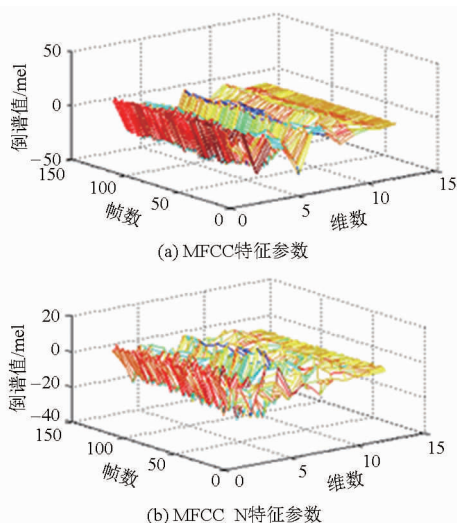


图2 咳嗽声的MFCC和MFCC_N特征参数

Fig. 2 MFCC and MFCC_N feature parameters of cough sound

用的 β 和 β_N 的维数均为12。

2.2 组合特征参数

单一的特征参数难以完整地描述咳嗽声信号的特征,为了提高咳嗽声的识别率,将MFCC_N特征参数与其他参数进行组合,当相关性不大的特征参数组合在一起时,能获得更优的识别性能,组合的特征参数有:

(1)短时能量(E):咳嗽声信号能量随时间变化较大,具有突发性和短暂性,短时能量度量语音信号的幅度变化,更能表现出咳嗽声的时域差别,而且维数少、计算简单。

(2)短时过零率(ZCR):表示在一帧语音中波形穿过横轴的次数。

(3)共振峰(F):采用线性预测法(LPC)来提取前3个共振峰参数。

(4)谱质心(SC):表示声信号频率成分的重心。

3 咳嗽声的识别

语音的识别方法有最小距离法、动态时间规整(Dynamic time warping, DTW)、人工神经网络法(Artificial neural network, ANN)、矢量量化(Vector quantization, VQ)技术和隐马尔可夫模型(HMM)等^[15-20]。人工神经网络只能解决静态模式分类且并不涉及时间序列处理,矢量量化技术不太合适语音时序变化,由于咳嗽声是一个时长分布较宽的时变过程,而基于统计方法HMM具有强大的时序建模能力,正适合于这种具有时变特征的对象识别。

3.1 咳嗽声的HMM模型

杜泊羊咳嗽声HMM模型 λ 用2个状态集合和3个概率矩阵表示为 $\lambda = \{S, O, A, B, \pi\}$,其中 S 表

示杜泊羊咳嗽声状态的集合,设 N 为咳嗽声状态的个数,则 $S = \{S_1, S_2, \dots, S_N\}$,咳嗽声一般都经历了开始、爆发、持续和结束4个阶段,可对应于咳嗽声的不同状态, q_t 表示 t 时刻的状态,状态是隐藏的, $Q = \{q_1, q_2, \dots, q_T\}$ 为状态序列, T 为咳嗽声的帧数; O 表示一组被观测值的序列, M 表示每个咳嗽声状态下输出的不同观测值个数, $O = \{V_1, V_2, \dots, V_M\}$;咳嗽声状态在 t 时刻从一个状态转移到下一个状态的概率为 $a_{ij} = P\{q_{t+1} = S_j | q_t = S_i\}$,状态转移概率矩阵 $A = \{a_{ij}\}_{N \times N}$;在状态为 j 时,其观测值的概率求解方式为 $b_j(v) = P\{O = v | q_t = S_j\}$,状态 j 时的观测概率密度分布 $B = \{b_j(v)\}$;咳嗽声初始时刻为某个状态的概率 $\pi_i = P\{q_1 = S_i\}$,初始状态为 $\pi = \{\pi_1, \pi_2, \dots, \pi_N\}$ 。描述杜泊羊咳嗽声的HMM声学模型如图3所示,为无跨越、由左至右连续模型。

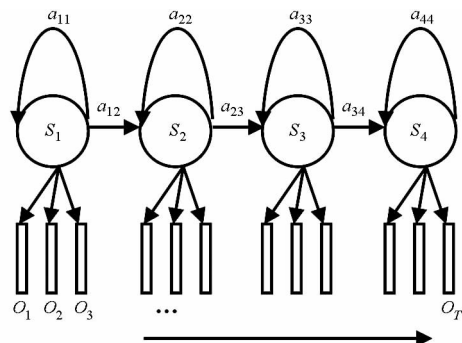


图3 咳嗽声的HMM声学模型

Fig. 3 HMM acoustic model of cough sound

状态数 N 越多,识别效果越好,但当状态数达到一定程度时,会因训练数据不足而效果变差,根据咳嗽声的特点,选 $N = 4$;每个状态输出的观测值 M 实际为线性叠加的概率密度函数的个数, M 值越大,对咳嗽声空间描述得越细,更加接近实际的咳嗽声空间,然而 M 值的增大需要增加训练的参数量,同时也需要更多的训练数据去获得可靠的结果,使得计算量大增,经试验比较,选择 $M = 3$ 。在杜泊羊的咳嗽声HMM模型中, π 和 A 的值可以随机选取,对识别率的影响不大,观测概率密度分布矩阵 B 初始化方法为:首先将咳嗽声信号的特征参数分成 N ($N = 4$)段,然后使用 K 均值法进行同一状态下的特征参数聚类,得到 M ($M = 3$)个子矢量集合并计算其方差和均值,经归一化处理,得出每个概率密度函数的权重,将这些权重值作为参数 B 的初值。

将采集到的600个咳嗽声样本和300个非咳嗽声样本的1/2用于HMM的训练,剩下的样本用于验证识别系统的性能。在训练阶段,使用Baum-Welch算法建立杜泊羊咳嗽声HMM参数,使观察序列的输出概率最大,每一个训练输入的咳嗽声都要

建立其 HMM 模型参数 λ_i , 所有的模型 λ_i 参数组成系统的 HMM 模型参数库。在识别阶段, 采用 Viterbi 算法计算测试样本与咳嗽声 HMM 参数库中每个参数 λ_i 匹配似然值, 选取输出概率 $P(\lambda_i)$ 最大值对应的模板作为咳嗽声的识别结果。

3.2 咳嗽声识别结果对比

用 3 个参数来描述咳嗽声识别性能, 分别为咳嗽声识别率(正确识别的咳嗽声占咳嗽声总数百分比)、咳嗽声误识别率(被误识别为咳嗽声的非咳嗽声占非咳嗽声总数百分比)和总识别率(正确识别出的咳嗽声和非咳嗽声总数占测试样本总数的百分比)。为了降低杜泊羊的咳嗽声识别系统的复杂度, 要求咳嗽声的特征参数维数较少。

将未参与 HMM 模型训练的 300 个咳嗽声和 150 个非咳嗽声样本进行识别试验, MFCC 参数及其组合参数作为 HMM 的输入, 识别结果如表 1 所示。

表 1 MFCC 及其组合参数的识别结果
Tab.1 Recognition result of MFCC combined

with other parameters				%
特征参数	维数	识别率	误识别率	总识别率
MFCC	12	73.33	17.55	76.37
MFCC + E	13	78.36	15.42	80.43
MFCC + SC	13	70.94	14.61	75.76
MFCC + ZCR	13	75.27	16.18	78.12
MFCC + F	15	71.53	15.02	76.01
MFCC + E + F	16	77.19	13.48	80.30
MFCC + E + ZCR	14	80.43	12.32	83.18
MFCC + F + SC	16	72.37	18.27	75.49
MFCC + E + SC	14	79.26	12.83	81.90
MFCC + F + ZCR	16	73.59	17.73	76.48

由表 1 可知, 杜泊羊咳嗽声的 MFCC 特征参数用于 HMM 识别时, 识别率、误识别率和总识别率分别为 73.33%、17.55% 和 76.37%, 当 MFCC 与短时能量 E 或短时过零率 ZCR 组合时, 识别性能均能提高, 而与谱质心 SC 或共振峰 F 组合时, 识别性能下降, 其中 MFCC + E + ZCR 组成的 14 维特征参数识别效果最佳, 其识别率、误识别率和总识别率分别为 80.43%、12.32% 和 83.18%, 这是因为杜泊羊咳嗽声的幅度变化超过了一般的语音, 非平稳性突出, 短时能量和过零率变化明显。使用改进的 MFCC_N 及其组合特征参数进行杜泊羊咳嗽声的识别时, 识别结果如表 2 所示。

从表 2 可看出, 改进的 MFCC_N 参数比 MFCC 参数识别性能有所提高, 其识别率和总识别率分别提高了 5.94% 和 5.70%, 误识别率降低了 5.23%, MFCC_N + E + ZCR 组合特征参数的识别率、误识别率和总识别率分别达到了 86.23%、7.17% 和

表 2 MFCC_N 及其组合参数的识别结果
Tab.2 Recognition result of MFCC_N combined
with other parameters %

特征参数	维数	识别率	误识别率	总识别率
MFCC_N	12	79.27	12.32	82.07
MFCC_N + E	13	83.54	10.16	85.64
MFCC_N + SC	13	76.39	10.55	80.74
MFCC_N + ZCR	13	81.46	11.85	83.70
MFCC_N + F	15	78.23	9.27	82.40
MFCC_N + E + F	16	83.76	8.66	86.29
MFCC_N + E + ZCR	14	86.23	7.17	88.43
MFCC_N + F + SC	16	79.59	12.63	82.18
MFCC_N + E + SC	14	85.14	8.18	87.37
MFCC_N + F + ZCR	16	80.33	12.19	82.82

88.43%, 因此改进的 MFCC_N 参数提高了识别系统的性能, 能够更好地表征杜泊羊咳嗽声的特性。

3.3 咳嗽声特征参数降维

MFCC_N + E + ZCR 的 14 维特征参数的总识别率最高, 为进一步提高识别系统的速度, 能够开发出实时监测杜泊羊咳嗽声的识别系统, 可通过主成分分析(PCA)舍去原特征参数中表征咳嗽声能力弱的特征分量, 选取表征特征最有效的部分, 这样既减少了数据的存储量, 又为语音识别减少了计算量, 同时还对咳嗽声特征参数进行了最优化, 提高了杜泊羊咳嗽声的识别效率^[21]。

设 m 为样本个数, n 为特征维数, $\mathbf{X}_{m \times n}$ 为原始数据, 其中 $m > n$, PCA 处理的步骤为:

(1) 计算原变量的协方差矩阵 $\mathbf{S}_{n \times n}$, 其中

$$S_{ij} = \frac{1}{m-1} \sum_{k=1}^m (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)$$

$(i, j = 1, 2, \cdots, n)$

(2)

式中 $\bar{x}_i, \bar{x}_j$ ——第 i, j 列的数学期望
 x_{ki}, x_{kj} ——原始数据 $\mathbf{X}_{m \times n}$ 中第 k 行的第 i, j 列的元素

S_{ij} ——协方差矩阵 $\mathbf{S}_{n \times n}$ 中的元素

(2) 求 $\mathbf{S}_{n \times n}$ 的特征值 λ_i 和与其正交化的特征参数, 将特征值按从大到小的顺序排列, 其前面 p 个特征参数对应为主成分, p 值的选取可通过特征值贡献率公式来确定, 公式为

$$L(p) = \frac{\sum_{i=1}^p \lambda_i}{\sum_{i=1}^n \lambda_i}$$

(3)

当选取的 p 值的贡献率大于 90% 时, 可认为这 p 个主成分的特征参数就能反映原变量的信息, 至此, n 维的特征参数降到 p 维。14 维的 MFCC_N + E + ZCR 特征参数与贡献率的关系如图 4 所示。

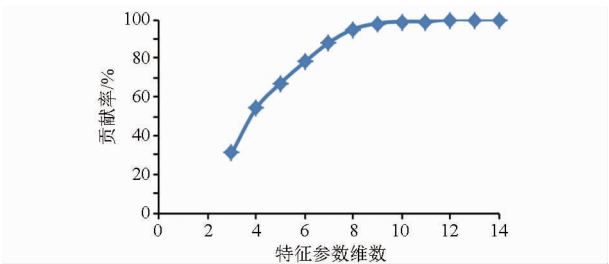


图 4 特征参数维数的贡献率

Fig.4 Contribution rate versus dimension of feature parameters

图 4 中,14 维的特征参数经过 PCA 处理后降到 9 维,累计贡献率为 98%,剔除了特征参数中冗余的特征,避免了多余特征的干扰。用 9 维的 MFCC_N + E + ZCR 参数进行 HMM 咳嗽声识别试验时,识别结果如表 3 所示,与表 2 中 14 维 MFCC_N + E + ZCR 参数比较,其识别率、总识别率分别下降了 1.72% 和 1.19%,误识别率提高了 0.14%,与降维之前的参数识别效果接近,因此优化的 9 维 MFCC_N + E + ZCR 特征参数可代替原 14 维的特征参数。

表 3 HMM 与 HMM-BP 模型识别结果对比

Tab.3 Comparison between HMM and HMM-BP

recognition results			%
模型类型	识别率	误识别率	总识别率
HMM	84.51	7.31	87.24
HMM-BP	92.54	5.37	95.04

3.4 咳嗽声识别性能的优化

HMM 是一种基于时序累积概率的动态信息处理方法,对动态时间序列有极强的建模能力,但其分类决策能力差,需要语音信号的先验统计知识,而且其抗噪性、自适应性也一般,这些不足使得单一的 HMM 的识别性能难以进一步提高。人工神经网络(ANN)通过对人脑组织的结构和功能模拟,有着与人类类似的学习记忆能力、概括推理能力以及强大的分类能力,缺点与 HMM 的优点刚好相反,对动态时间信号的描述能力差,而对信号分类和决策层面有着非常强的能力。结合 HMM 时序处理能力和 ANN 分类能力的优势,将 HMM 的输出作为 ANN 模型输入,ANN 模型作为 HMM 的后处理器,对杜泊羊咳嗽声信号进行分类,提高识别系统的性能,且 2 个模型相互独立。本文的 ANN 模型采用 BP(Back propagation)神经网络^[22]。

BP 神经网络又称多层前馈神经网络,它由输入层、隐含层以及输出层组成。隐含层可以有多个,本文采用 3 层 BP 神经网络,隐含层的神经元个数决定着网络的记忆容量、训练速度、泛化能力和输出响应的质量。隐含层节点数过少或过多,都会导致网络的识别能力和网络的容错性能降低。

(1)BP 神经网络的建模

利用 Viterbi 算法计算样本 HMM 语音模板的所有状态($N=4$)累积概率输出,状态累积概率进行时间规整后,将所有状态累积概率构成的向量 $\alpha = \{\alpha(1), \alpha(2), \alpha(3), \alpha(4)\}$ 作为 BP 神经网络的输入,因此 BP 网络的输入层含有 4 个神经元,反映了杜泊羊咳嗽声不同状态累积概率之间的相关性,输出层有 1 个神经元,表示声音信号与模板的相似概率,隐含层依靠经验选取 9 个神经元,BP 网络结构如图 5 所示。

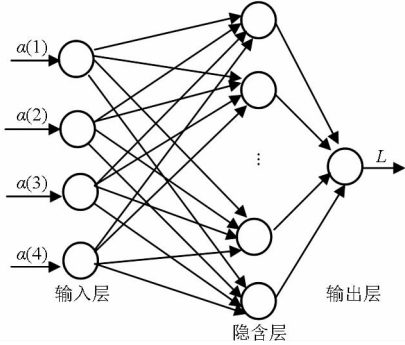


图 5 咳嗽声 HMM 外接的 BP 神经网络

Fig.5 HMM external connected BP neural networks of cough sound

(2)BP 神经网络的训练

将已训练好的 HMM 样本咳嗽声数据作为 BP 网络的训练样本,并对样本所有状态的累积概率向量 α 进行归一化,公式为

$$\alpha_v = \frac{\alpha - \alpha_{\min}}{\alpha_{\max} - \alpha_{\min}}$$

(4)

将归一化的向量 α_v 作为 BP 网络的输入,进行 BP 网络的学习和训练,得到 BP 网络各层之间的连接权值,输出值 L 为 1 表示是杜泊羊咳嗽声的概率为 100%,可预先设定咳嗽声一阈值,若 L 大于阈值则为咳嗽声,小于阈值则为非咳嗽声。阈值的选取对识别性能有着重要影响,通过试验当阈值为 0.68 时,得到的总识别率最高。

将降成 9 维的 MFCC_N + E + ZCR 特征参数代入 HMM 模型,并经过 BP 网络重新对所有测试样本进行识别,结果如表 3 所示。

由表 3 可知,HMM-BP 杜泊羊咳嗽声识别模型的识别性能得到了改善,其识别率和总识别率分别提高 8.03 个百分点和 7.80 个百分点,误识别率降低 1.94 个百分点,总识别率达到 95.04%,满足杜泊羊咳嗽声识别系统的要求。

4 结论

(1)利用声音传感器和无线数据采集设备采集杜泊羊的咳嗽声信号,并对咳嗽声信号进行预处理,

通过监测杜泊羊的咳嗽声信号可以为羊只的早期疾病预警和健康状况评估提供理论依据。

(2)提出了利用改进的 MFCC 参数提取杜泊羊的咳嗽声信号的方法,该方法在不增加特征参数的维数时,考虑到了咳嗽声信号的动态特性。改进的 MFCC 参数与短时能量 E 和短时过零率 ZCR 进行组合的 14 维特征参数,经过咳嗽声 HMM 识别系统,其识别率、误识别率和总识别率分别达到了 86.23%、7.17% 和 88.43%。将 14 维特征参数进行

PCA 分析后降到 9 维,可降低识别系统的复杂度和提高咳嗽声的识别效率。

(3)依据杜泊羊咳嗽声的时序特征,建立咳嗽声的 HMM 识别模型,并结合 HMM 时序处理能力和 BP 网络分类能力的优势,将 HMM 的输出作为 BP 网络输入,当 BP 网络的输出阈值为 0.68 时,咳嗽声的识别率、误识别率和总识别率分别达到 92.54%、5.37% 和 95.04%,提高了杜泊羊咳嗽声的识别精度。

参 考 文 献

- 1 陈冲,尤鸣宇,刘家铭,等. 基于高频子带特征的咳嗽检测方法[J]. 南京大学学报:自然科学版,2015,51(1):157-164.
CHEN Chong, YOU Mingyu, LIU Jiaming, et al. Cough detection based on high-frequency subband features[J]. Journal of Nanjing University: Natural Sciences, 2015, 51(1):157-164. (in Chinese)
- 2 YUSUF A Amrulloh, Udantha R Abeyratne, Vinayak Swarnkar, et al. Automatic cough segmentation from non-contact sound recordings in pediatric wards[J]. Biomedical Signal Processing and Control, 2015, 21(8):126-136.
- 3 THOMAS Drugman. Using mutual information in supervised temporal event detection: application to cough detection[J]. Biomedical Signal Processing and Control, 2014, 10(3):50-57.
- 4 Arzu Tuzuner, Sule Demirci, Gulden Bilgin, et al. Voice assessment after treatment of subacute and chronic cough with inhaled steroids[J]. Journal of Voice, 2015, 29(4):484-489.
- 5 MITCHELL Silva, VASILEIOS Exadaktylos, SARA Ferrari, et al. The influence of respiratory disease on the energy envelope dynamics of pig cough sounds[J]. Computers and Electronics in Agriculture, 2009, 69(1):80-85.
- 6 SARA Ferrari, MITCHELL Silva, MARCELLA Guarino, et al. Cough sound analysis to identify respiratory infection in pigs[J]. Computers and Electronics in Agriculture, 2008, 64(2):318-325.
- 7 EXADAKTYLOS V, SILVA M, AERTS J M, et al. Real-time recognition of sick pig cough sounds[J]. Computers and Electronics in Agriculture, 2008, 63(2):207-214.
- 8 SILVA M, FERRARI S, COSTA A, et al. Cough localization for the detection of respiratory diseases in pig houses[J]. Computers and Electronics in Agriculture, 2008, 64(2):286-292.
- 9 VAN HIRTUM A, BERCKMANS D. Considering the influence of artificial environmental noise to study cough time-frequency features[J]. Journal of Sound and Vibration, 2003, 266(3):667-675.
- 10 VAN HIRTUM A, BERCKMANS D. Fuzzy approach for improved recognition of citric acid induced piglet coughing from continuous registration[J]. Journal of Sound and Vibration, 2003, 266(3):677-686.
- 11 FERRARI S, PICCININI R, SILVA M, et al. Cough sound description in relation to respiratory diseases in dairy calves[J]. Preventive Veterinary Medicine, 2010, 96(3):276-280.
- 12 VAN HIRTUM Annemie, DANIEL Berckmans. Automated recognition of spontaneous versus voluntary cough[J]. Medical Engineering & Physics, 2002, 24(7):541-545.
- 13 AERTS J M, JANS P, HALLOY D, et al. Labeling of cough data from pigs for on-line disease monitoring by sound analysis[J]. Transactions of the ASAE, 2004, 48(1):351-354.
- 14 陈迪,龚卫国,李波. 噪声鲁棒性说话人识别语音高频加权 MFCC 提取[J]. 仪器仪表学报, 2008, 29(3):668-672.
CHEN Di, GONG Weiguo, LI Bo. High frequency weighted MFCC extraction for noise robust speaker verification[J]. Chinese Journal of Scientific Instrument, 2008, 29(3):668-672. (in Chinese)
- 15 李志忠,滕光辉. 子带特征参数在家禽应激发声识别中的应用[J]. 农业机械学报, 2009, 40(3):143-146.
LI Zhizhong, TENG Guanghui. Application of subband feature of poultry vocalization for identifying their status[J]. Transactions of the Chinese Society for Agricultural Machinery, 2009, 40(3):143-146. (in Chinese)
- 16 CHEDAD A, MOSHOU D, AERTS J M, et al. AP - animal production technology: recognition system for pig cough based on probabilistic neural networks[J]. Journal of Agricultural Engineering Research, 2001, 79(4):449-457.
- 17 ANDREY Temko, CLIMENT Nadeu. Classification of acoustic events using SVM-based clustering schemes[J]. Pattern Recognition, 2006, 39(4):682-694.
- 18 MOSHOU D, CHEDAD A, VAN HIRTUM A, et al. Neural recognition system for swine cough[J]. Mathematics and Computers in Simulation, 2001, 56(4):475-487.
- 19 MARKO Kos, ZDRAVKO Kačić, DAMJAN Vlaj. Acoustic classification and segmentation using modified spectral roll-off and variance-based features[J]. Digital Signal Processing, 2013, 23(2):659-674.
- 20 GUARINO M, JANS P, COSTA A, et al. Field test of algorithm for automatic cough detection in pig houses[J]. Computer and Electronics in Agriculture, 2008, 62(1):22-28.
- 21 TRENDAFILOVA I, CARTMELL M P, OSTACHOWICZ W. Vibration-based damage detection in an aircraft wing scaled model using principal component analysis and pattern recognition[J]. Journal of Sound and Vibration, 2008, 313(3):560-566.
- 22 ANUJ Mohamed, RAMACHANDRAN NAIR K N. HMM/ANN hybrid model for continuous Malayalam speech recognition[J]. Procedia Engineering, 2012, 30:616-622.