

文章编号:1006-1355(2018)02-0188-05

无人机识别的音频特征提取方法

丘恺彬, 李建良

(南京理工大学 理学院, 南京 210094)

摘要:针对无人机飞行时噪声产生的机理,分别选取基于经验模态分解(EMD)的能量比以及梅尔频率倒谱系数(MFCC)的特征提取算法实现无人机声信号的特征提取,并引用主成分分析(PCA)方法对特征集进行降维融合处理。最后选择矢量量化方法(VQ)作为分类器对不同类型的无人机目标进行分类与识别。实验结果表明特征融合后的分类性能要好于基于单一特征的分类性能,该方法较好地体现不同类型无人机之间的差异,分类结果准确率较高,具有良好的稳定性。

关键词:声学;经验模态分解;梅尔频率倒谱系数;特征提取;矢量量化方法;分类与识别

中图分类号:TN911.7

文献标志码:A

DOI编码:10.3969/j.issn.1006-1355.2018.02.035

Feature Extraction and Classification of Acoustic Signals of Drones

QIU Kaibin, LI Jianliang

(School of Science, Nanjing University of Science & Technology, Nanjing 210094, China)

Abstract : The mechanism of noise generation of drone flight is studied. The feature extraction method based on empirical mode decomposition (EMD) energy ratio and Mel-frequency reciprocal spectrum coefficient (MFRSC) is proposed. The acoustic signal extraction of the drone flight is realized. Then, the principal component analysis (PCA) method is used to combine the feature and reduce the dimension of the feature. Finally, the vector quantification (VQ) is chosen as a classifier to classify and identify the types of drone targets. The experimental results show that the classified performance of the combined feature is better than that of the single feature. The method can reflect the difference between the types of drones and have high accuracy and good stability for classification of features.

Keywords : acoustics; empirical mode decomposition (EMD); Mel-frequency reciprocal spectrum coefficient (MFRSC); feature extraction; vector quantification (VQ); classification and identification

无人机技术快速发展,在环境监测、通讯中继、大气取样及军事等领域得到广泛应用,但多地民用机场等净空保护区近期随之不断出现的“黑飞”事件给低空领域的管理、军民飞行安全等带来了极大麻烦。除出台各种无人机生产销售管控方法外,快速、有效侦测出无人机的黑飞过程具明确实用意义。目前无人机侦测常用的工具主要有雷达、调频无线电以及视频摄像等,国内外已经研发了一些专用装备,但由于无人机目标小且在低空慢飞行,这些装备可能受到诸如电子干扰、实物遮挡、盲区或能见度等影

响而具有局限性。

鉴此本文基于被动的音频采集处理原理,将声音感知的目标识别技术作为无人机侦测的补充手段。事实上已有文献报道国外企业正在开展基于此原理的反无人机商业开发。采用声音感知技术进行无人机侦测,本质上是利用无人机工作时不可避免出现的音频信号,通过多点布局的音频传感器采集、去噪,提取无人机自身独有的音频特征实现识别分类,进而可望基于音频传感器布局坐标进行定位跟踪。由于无人机的机动水平较强,其音频信号具有时间短、噪声强、非线性、不平稳、无先验信息的特点,传统音频处理的Fourier分析方法不能准确反映非平稳信号的时变特征,因此寻找适用于无人机的非平稳过程的音频信号处理方法将是出发点。本文采用基于EMD的能量比的特征提取方法,结合传统的MFCC特征提取方法,将两种特征用PCA降维处

收稿日期:2017-07-28

作者简介:丘恺彬(1992-)男,福建省龙岩市人,硕士研究生,研究方向为声音识别和定位。

E-mail: qkb0404@163.com

通信作者:李建良(1962-)男,研究生导师。

E-mail: lj16006@njjust.edu.cn

理后构成无人机声信号的综合特征向量,最后以VQ作为分类器实现不同型号的无人机声信号分类和识别。

1 无人机音频特征提取

1.1 基于EMD的IMF能量比特征提取

EMD的实质是依不同时间尺度特征将多分量信号分解成一系列单分量信号^[1],即本征模态函数(Intrinsic Mode Function, IMF),并提取出该复杂信号的趋势项,从而使得信号的瞬时频率具有物理意义,其分解步骤如下^[2-4]:

(1) 将信号 $x(t)$ 所有局部极大值点和局部极小值点分别用3次样条曲线连接起来构成 $x(t)$ 的上下包络线,两条曲线间包含所有的信号。计算其平均值曲线 $m_1(t)$,进而得

$$h_1(t) = x(t) - m_1(t) \quad (1)$$

将 $h_1(t)$ 作为待处理信号,重复上述求包络均值和差值的计算,得到

$$h_{1k}(t) = h_{1(k-1)}(t) - m_{1k}(t) \quad (2)$$

直到 $h_{1k}(t)$ 符合IMF分量的特征,从原信号中分解得到了第1个IMF,记为

$$c_1(t) = h_{1k}(t) \quad (3)$$

(2) 从原信号中减去 $c_1(t)$,得到第1阶剩余信号 $r_1(t)$

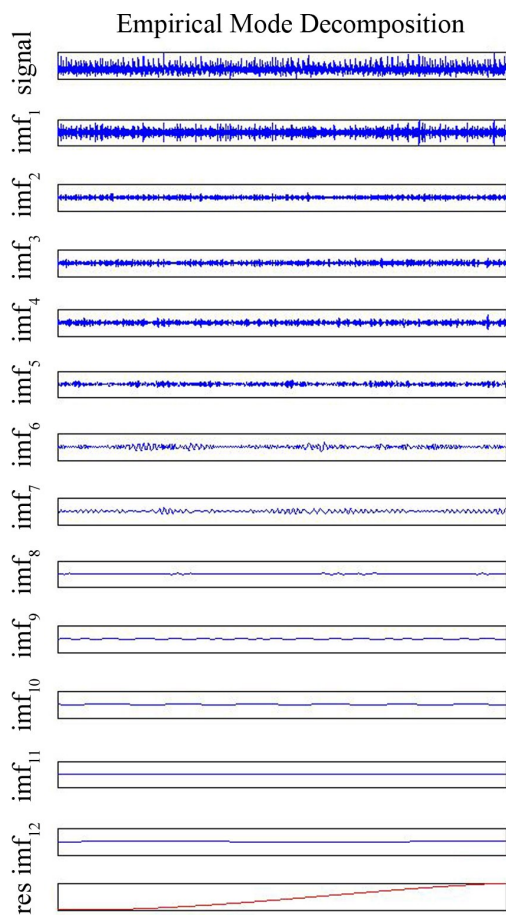
$$r_1(t) = x(t) - c_1(t) \quad (4)$$

把 $r_1(t)$ 作为新的信号重复步骤(1),将第2、第3直到第 $n-1$ 个IMF分量 $c_2(t)$ 、 $c_3(t)$ 、 \dots 、 $c_{n-1}(t)$,若第 N 阶IMF分量 $c_n(t)$ 或其剩余量 $r_n(t)$ 小于预先设定的值或 $r_n(t)$ 变成1个单调函数时,筛选结束。这样原始音频数据可表示为有限个IMF和1个余项的和

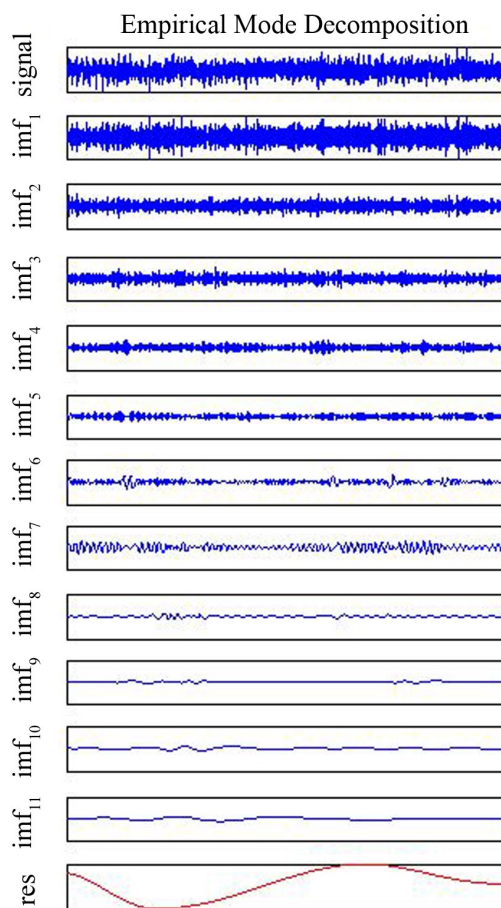
$$x(t) = \sum_{i=1}^n c_i(t) + r_n(t) \quad (5)$$

IMF反映了原信号中不同的频率成分,且 n 个IMF的频率从大到小排列,余项变为很低频率的脉动即趋势项。以大疆两种型号无人机精灵 Phantom 3和悟 Inspire工作时音频信号为例,进行EMD分解如图1所示,文中数值实验均是在Windows 10环境下用MATLAB 2011实现的。

从图1(a)和图1(b)中从上到下依次为两种不同型号无人机的声信号、 n 阶IMF分量和余项。可以看出精灵 Phantom 3声信号经由EMD分解后得到12个IMF分量,悟 Inspire声信号经由EMD分解后



(a) 精灵 Phantom 3 EMD 分解时域波形



(b) 悟 Inspire EMD 分解时域波形

图1 不同型号无人机声信号及EMD分解的时域波形

得到11个IMF分量。EMD分解可以突出信号的某些局部特征,且余项反映了原信号的趋势特征,有利于研究信号的变化趋势。

信号经EMD分解后得到有限个IMF分量,每个分量均代表机动目标内部某一部件的频率,因此各个分量的频域能量可能存在某种关系。据此提出基于EMD的IMF能量比的特征提取算法,过程如下^[5-6]。

对任一信号,定义第*i*阶IMF的能量为

$$E = \sum_{i=0}^{N-1} |c_i(t)|^2, \quad 1 \leq i \leq n+1 \quad (6)$$

以各个IMF分量的能量(E_i)相对于原信号能量(E)的能量比作为特征向量,即

$$k_i = \frac{E_i}{E} \quad (7)$$

精灵 Phantom 3 和悟 Inspire 两种无人机声信号提取IMF能量比特征如图2所示。

从图2中可以看出,两种类型的无人机信号分解得到的IMF分量的个数不同,在同一频段下,不同类的信号能量也有比较明显的区别,说明IMF可用于不同类目标的分类判别。

以基于EMD的IMF能量比作为特征对无人机类型进行识别,实验对象选取市场上常见的3种不同型号的无人机包括大疆精灵 Phantom3、大疆精灵 Phantom4 以及大疆悟 Inspire。由于 Phantom3 与 Phantom4 在形状、大小上较为相似容易混淆,可以更好地检验方法的有效性。

各取3型号无人机的35个声信号作为训练样本,采样频率为44.1 kHz,并将信噪比为5、10、20的3种不同程度的高斯白噪声加入原始声信号中得到带噪音频。提取EMD特征时以最大特征向量维数为基准,在其他向量尾部加0,使得3种类型无人机的特征向量维数相同。以3种类型的无人机包括所有训练样本在内的各60个声信号作为测试样本,使

用VQ作为分类器对EMD特征向量进行分类与识别实验,实验结果如表1所示。

表1 3种型号无人机的EMD特征识别率/(%)

方法	SNR/dB	无人机类型		
		精灵 Phantom 3	精灵 Phantom 4	悟 Inspire
EMD	5	85	83.3	88.3
	10	86.7	83.3	90
	20	90	85	93.3
	无噪声	90	88.3	95

注:训练样本个数为35,测试样本个数为60

从表1可以看出,对于不同的信噪比,精灵 Phantom 4 的正确识别率不如其它两种无人机,由于精灵 Phantom4 与精灵 Phantom3 外形相似,导致精灵 Phantom4 的识别率较低。在没有噪声干扰的情况下,3种无人机的识别结果较好,但是在噪声环境下,整体的识别率较低,因为产品要最终应用到实际中,所以考虑一种广泛用于语音识别的特征提取算法—MFCC。

1.2 基于MFCC的特征提取

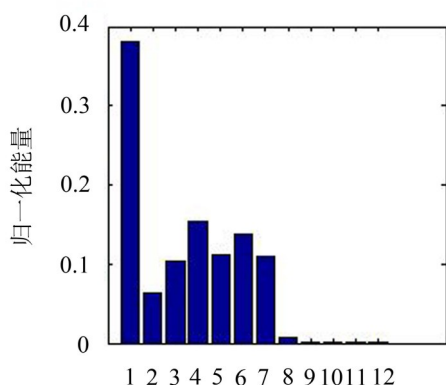
MFCC是一种传统且非常重要的特征提取算法,它模拟了人耳的听觉感知特性来分析音频的频谱。根据实验结果,人的主观感知频域的划定并不是线性的,有下面的公式

$$F_{\text{mel}} = 1125 \log \left(1 + \frac{f}{700} \right) \quad (8)$$

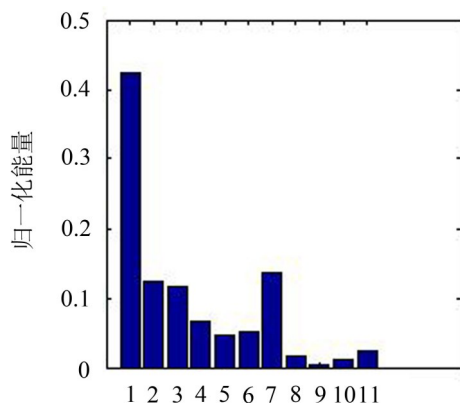
其中: F_{mel} 是以美尔(Mel)为单位的感知频率; f 是以Hz为单位的实际频率。

MFCC特征提取的实质是将时域内的频谱通过非线性的频谱进行转化,通过离散余弦变换得到MFCC倒谱。其计算流程如图3所示。

将 Phantom 3 分别和 Phantom 4、Inspire 之间的MFCC参数进行比较,并选择其中几帧,如图4所



(a) 精灵 Phantom 3 声信号 IMF 能量比特征



(b) 悟 Inspire 声信号 IMF 能量比特征

图2 不同型号无人机的IMF能量比特征

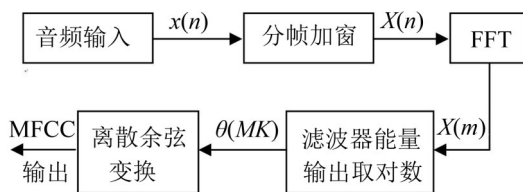


图3 MFCC计算流程图

示。横坐标是 Phantom 3 的 MFCC 参数,纵坐标是 Phantom 4 或 Inspire 的 MFCC 参数,如果某 1 阶的特征相似,则系数分布应在虚线附近。

对 3 种无人机的声信号提取 16 阶 MFCC 参数,同样在原信号中加入信噪比为 5、10、20 3 种不同程度的高斯白噪声,识别结果如表 2 所示。

表 2 3 种型号无人机的 MFCC 特征识别率/(%)

方法	SNR/dB	无人机类型		
		精灵 Phantom3	精灵 Phantom4	悟 Inspire
MFCC	5	85	85	88.3
	10	85	88.3	91.6
	20	88.3	90	95
	无噪声	90	90	95

注:训练样本个数为 35,测试样本个数为 60

对比表 1 和表 2 发现,对于 Phantom 4,MFCC 的识别效果要好于 EMD,对于 Phantom 3 的识别,MFCC 的效果又稍差于 EMD,而对于 Inspire,两种方法的识别效果相近。所以考虑将以上两种特征融合为一个新的特征。

2 EMD 与 MFCC 的特征融合

2.1 基于主成分分析的特征降维

由上述实验可知两个特征表现互有好坏,所以考虑将两个特征融合得到一个新的特征。但如果只

是简单地将两个特征结合,则会得到一个高维特征,也意味着计算复杂度的增加。本文利用主成分分析的最优降维性质,在特征降维的同时,还能保证系统的识别性能。主成分分析(Primary Component Analysis, PCA)是一种在均方误差最小意义上最优的线性降维方法,通过 PCA 变换,特征向量中的主要成分能够被保留下来^[7]。

设 X 为 p 维向量,且 X 的协方差存在,令 λ_i 为 X 的特征值(这里的 i 代表从大到小的第 i 个位置), ϕ_i 为 λ_i 对应的特征向量, X 向量的第 i 个主成分因子则定义成

$$\gamma_i = \phi_i^T X \quad (9)$$

这里所有 γ_i 构成了特征空间一个正交基,它们之间不存在冗余。则 γ_i 的方差贡献率为

$$vcr_i = \frac{\lambda_i}{\sum_{k=1}^p \lambda_k} \quad (10)$$

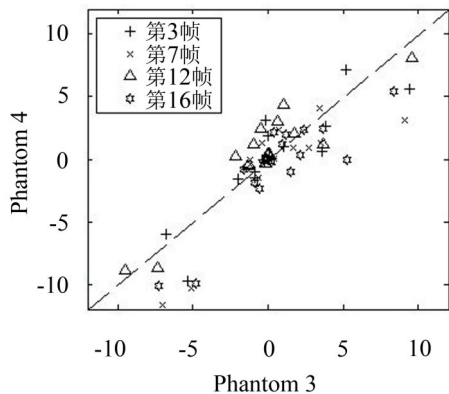
根据式(10)的结果进行大小排序,计算 $\lambda_i / \sum_{k=1}^p \lambda_k$ 值,定义其为主成分因子 $\gamma_1, \dots, \gamma_m$ 所对应累计方差贡献率,根据贡献率的大小即可将特征降到 m 维。

2.2 实验分析

将训练集、测试集合分类器设置成与上一组实验相同,然后应用 PCA 对两种特征进行降维处理,得到 16 维的新的特征向量,其中包括 7 阶 EMD 分量和 9 阶 MFCC 分量,其他实验变量的设置与前两个实验相同,分类结果如表 3 所示。

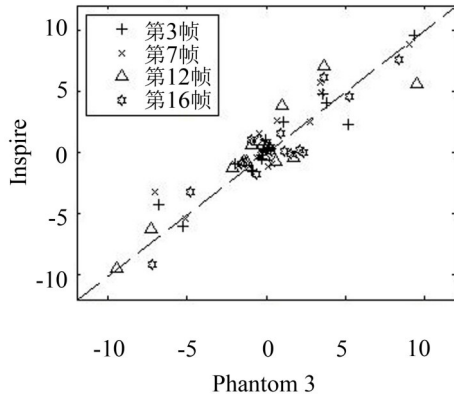
对比表 1、表 2 和表 3 的分类结果可以看出,融合后的特征比使用单一特征具有更高的识别率,所以对于这种复杂的机动目标来说,选择多种特征能更全面地表现不同类型无人机声信号的特征。

Phantom3和Phantom4之间的MFCC参数匹配比较



(a) Phantom 3 与 Phantom 4 的 MFCC 参数比较

Phantom3和Inspire之间的MFCC参数匹配比较



(b) Phantom 3 与 Inspire 的 MFCC 参数比较

图 4 不同型号无人机的 MFCC 参数比较

表3 3种型号无人机的EMD+MFCC特征识别率/(%)

方法	SNR/dB	无人机类型		
		精灵 Phantom 3	精灵 Phantom 4	悟 Inspire
EMD + MFCC	5	93.3	90	95
	10	95	91.6	96.6
	20	96.6	95	100
	无噪声	96.6	96.6	100

注:训练样本个数为35,测试样本个数为60

3 结 语

本文从无人机噪声的产生机理出发,使用PCA综合了两种特征提取技术对不同类型的无人机声信号进行特征提取:MFCC以及基于EMD和能量比的特征提取算法。最后使用VQ作为分类器对两种单一的特征和结合后的特征进行分类与识别实验,可以看出,在单一特征的表现上,MFCC的识别率要略高于EMD,而将两种特征融合后比起单一的特征分类效果更显著,识别率更高。实验证明了利用声音感知技术来识别与监管无人机的可行性。

参考文献:

- [1] NORDEN E HUANG, ZHENG SHEN, STEVEN R LONG. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis[C]. Proceedings of the Royal Society London, 1998 (454): 903-995.
- [2] 吕艳新. 基于EMD和能量比的战场声目标分类与识别[J]. 振动与冲击, 2008, 27(11): 51-55.
- [3] BOUDRAA O, CEXUS J C, SAIDI Z. EMD-based signal filtering[J]. IEEE Transactions on Instrumentation and Measurement, 2007, 56(6): 2196-2202.
- [4] 赵知劲, 黄艳波. 基于经验模态分解的单通道盲源分离算法[J]. 计算机应用研究, 2017, 34(10): 3010-3012.
- [5] 赵天青, 梁旭斌, 许学忠, 等. EMD在目标声信号特征提取中的应用研究[J]. 南京大学学报(自然科学版), 2015 (S1): 102-106.
- [6] 徐可君, 秦海勤, 江龙平. 基于EMD和HHT的航空发动机转子-机匣振动信号分析[J]. 振动与冲击, 2011, 30(7): 237-240.
- [7] ASTON JOHN A D, CHIOU JENG-MIN, EVANS JONATHAN P, et al. Linguistic pitch analysis using functional principal component mixed effect models[J]. Journal of the Royal Statistical Society Series C-Applied Statistics, 2010, 59(2): 297-317.

(上接第178页)

- [6] AMIN S L, JALIL S. Pitch detection of speech signal using wavelet transform[J]. International Journal of Scientific Engineering and Technology, 2015, 5(4): 322-324.
- [7] 杨龙. 基于小波变换的自相关基音周期检测方法[J]. 信息通信, 2015, 1(9): 72-73.
- [8] 陈小利, 徐金甫. 利用小波变换加权自相关的基音周期检测法[J]. 数据采集与处理, 2007, 22(4): 463-467.
- [9] 潘峥嵘, 戴芮, 张宁. 改进的基音周期检测算法研究[J]. 计算机工程与应用, 2015, 51(10): 223-226.
- [10] 宋之用. MATLAB在语音信号分析与合成中的应用[M]. 北京: 北京航空航天大学出版社, 2013: 251-258.
- [11] 宗源, 李平, 曾毓敏, 等. 基于EMD的ACF基音周期检测改进算法[J]. 南京师大学报(自然科学版), 2013, 36(3): 42-47.

- [12] 王建军, 侯盼卫, 苏新彦, 等. EMD去噪法在FMCW雷达信号处理中的应用[J]. 科学技术与工程, 2014, 14(27): 66-70.
- [13] 李曼曼, 杨鸿武, 洪宁, 等. 基于EMD的带噪语音端点检测[C]. 第十一届全国人机语音通讯学术会议, 西安, 2011: 60-64.
- [14] 张雪英. 数字语音处理及MATLAB仿真[M]. 北京: 电子工业出版社, 2010: 220-224.
- [15] 李辉, 戴李辉, 戴蓓蓓, 等. 基于前置滤波和小波变换的带噪语音基音周期检测方法[J]. 数据采集与处理, 2005, 20(1): 100-104.