

FPGA 语音识别的四旋翼飞行器控制系统设计

靖 固, 张学松

(哈尔滨理工大学 计算机科学与技术学院, 黑龙江 哈尔滨 150080)

摘 要: 以 FPGA 语音识别命令作为四旋翼的航拍遥控命令, 将 FPGA 快速处理信息的能力与语音控制的简单、准确、灵活特点相结合。控制系统分为 2 部分, 四旋翼飞行器利用单片机, 采用模糊神经网络 PID 控制四旋翼螺旋桨的转速达到对机身的飞行高度和方向的控制; 语音遥控命令利用 FPGA, 采用优化的隐马尔可夫模型完成对机身的飞行姿态的命令控制。经仿真验证语音控制四旋翼控制系统具有很强的实时性, 能很好地控制四旋翼飞行器完成航拍任务。

关键词: FPGA; 四旋翼控制系统; 模糊神经网络 PID; 优化的隐马尔可夫模型

DOI: 10.15938/j.jhust.2017.06.018

中图分类号: TP391.41 **文献标志码:** A **文章编号:** 1007-2683(2017)06-0095-07

Design of FPGA Speech Recognition Four Rotor Aircraft Control System

JING Gu, ZHANG Xue-song

(School of Computer Science and Technology, Harbin University of Science and Technology, Harbin 150080, China)

Abstract: This paper uses FPGA speech recognition command as the four-rotor aerial remote control command. The FPGA can quickly process information combined voice control of the simple, accurate and flexible features. The control system is divided into two parts. The four-rotor craft uses the single-chip computer and it uses the fuzzy neural network PID to control the speed of the four-rotor propeller to achieve the control of the flying height and direction of the fuselage. The voice remote command uses FPGA to optimize the hidden markov model that completes the command control of the flight attitude of the fuselage. The simulation of the four-rotor control system has a strong real-time, and the system can control the four rotorcraft to complete the aerial mission better.

Keywords: FPGA; four-rotor control system; fuzzy neural network PID; optimized hidden Markov model

0 引 言

伴随着航拍等新型技术的崛起, 四旋翼飞行器成为近年来的热点研究主题^[1]。本文就是利用 FPGA 的快速处理信息的能力^[2], 将信息量较大的语音航拍遥控命令^[3]在 FPGA 上识别处理, 将 FPGA

快速处理信息的能力与语音控制的简单、准确、灵活性特点相结合, 使四旋翼飞行器控制系统更能满足特定系统的实时性、准确性和灵活性的要求^[4]。

1 系统整体设计

基于 FPGA 语音识别的四旋翼飞行器控制系统

收稿日期: 2016-04-06

基金项目: 黑龙江自然科学基金(F222805)。

作者简介: 靖 固(1961—), 女, 博士, 教授, E-mail: jinggu@hrbust.edu.cn;

张学松(1990—), 男, 硕士研究生。

$$F = R \sum_{i=1}^4 F_i, \quad (2)$$

$$R = \begin{bmatrix} \sin\alpha\cos\beta\cos\gamma + \sin\beta\sin\gamma \\ \sin\alpha\sin\beta\cos\gamma - \cos\beta\sin\gamma \\ \cos\alpha\cos\gamma \end{bmatrix}$$

其中: F 为总升力; R 为转换矩阵。

实现四旋翼飞行器飞行是要靠电机完成的,控制四旋翼电机的转速就能控制四旋翼的姿态以及方向。电机的数学模型为^[1]

$$G(s) = \frac{\omega(s)}{U(s)} = \frac{K_{sv}}{T_e s + 1} \quad (3)$$

其中: $\omega(s)$ 为电机输出转速; $U(s)$ 为电机的输入电控制量; K_{sv} 为常数; T_e 是电磁时间常数。

由此得到 4 个独立的电机控制量 U 与四旋翼飞行器转速 ω 之间的数学模型为

$$\begin{bmatrix} U_1 \\ U_2 \\ U_3 \\ U_4 \end{bmatrix} = \begin{bmatrix} F_1 + F_2 + F_3 + F_4 \\ F_4 - F_2 \\ F_3 - F_1 \\ F_2 + F_4 - (F_1 + F_3) \end{bmatrix} = \begin{bmatrix} K_l \sum_{i=1}^4 \omega_i^2 \\ K_l (\omega_4^2 - \omega_2^2) \\ K_l (\omega_3^2 - \omega_1^2) \\ K_l (\omega_4^2 + \omega_2^2 - (\omega_1^2 + \omega_3^2)) \end{bmatrix} \quad (4)$$

其中: U_1 为垂直高度输入量; U_2 为俯仰角输入量; U_3 为偏航角输入量; U_4 为横滚角输入量。

2.3 模糊神经网络 PID 控制器

模糊神经网络 PID 控制器总体结构框图如图 6 所示。

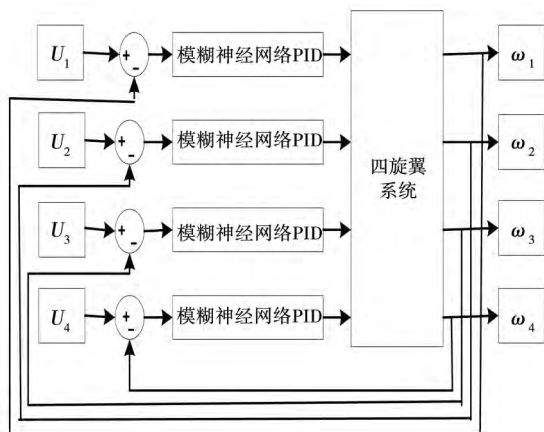
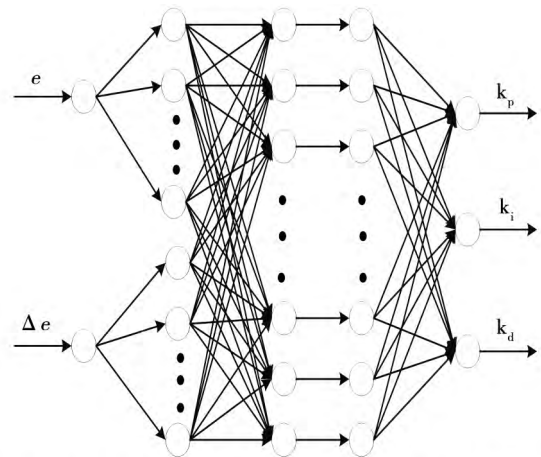


图6 模糊神经网络控制器总体结构框图

偏差和偏差变化率作为模糊神经网络 PID 控制器的输入,经过学习和训练的模糊神经网络输出 P、I、D,再经过比例因子调整后得到最佳的参数,输入给 PID 控制器。模糊神经网络^[1]控制器有输入层、隶属函数层、模糊推理层、归一化层、输出层组成。其结构如图 7 所示^[1]。

实现模糊神经网络 PID 控制器的算法设计流程如图 8 所示。



输入层 隶属函数层 模糊推理层 归一化层 输出层

图7 模糊神经网络结构框图



图8 模糊神经网络PID控制器的算法设计流程图

经 MATLAB 仿真得到的阶跃响应仿真图如图 9 所示。

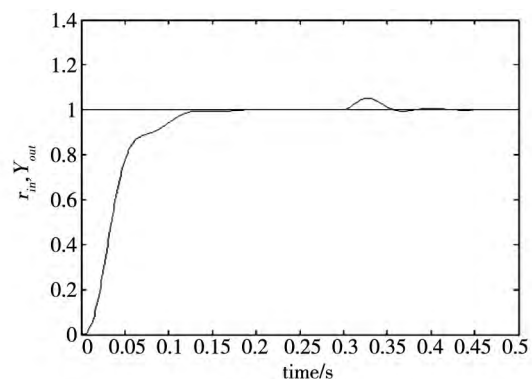


图9 模糊神经网络PID控制器的阶跃响应仿真图

仿真结果表明模糊神经网络控制器能够很好地控制四旋翼飞行器平稳飞行,当有干扰的时候也能很快的恢复到平稳飞行状态。

3 语音控制命令识别系统

3.1 优化 MFCC 特征

中国有 56 个民族,每个民族都有自己的方言且各有特点,对于语音识别加大了识别的难度。在研究中发现,梅尔倒谱系数能很好地区分这些特征,并将语音识别出来。

梅尔倒谱系数(MFCC)^[3]与频率的关系为

$$\text{Mel}(f) = 2595 \lg\left(1 + \frac{f}{700}\right) \quad (5)$$

其中 f 表示频率。实现流程如图 10 所示。

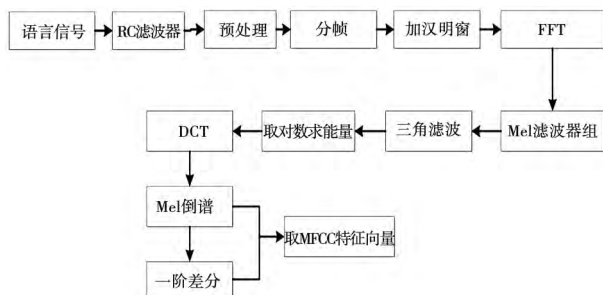


图 10 MFCC 特征提取过程

首先对采集的语音信号进行 RC 滤波和分帧,通过 RC 滤波器滤除语音命令输入的干扰信号,本系统帧长取 256,帧移 60;其次,设置预加重滤波器,能消除部分噪声,以提升清音部分的能量;再其次,加汉明窗,端点检测,在第 n 帧的短时能量大于噪声的短时能量,即第 n 帧短时能量 E_n ,并用短时过零率来区分无声、浊音以及清音。再进行 FFT 变换是为了得到语音信号的功率谱,而语音信号是在时域下的信号,必须变到频域下才能进行处理从而得到想要的数。

三角滤波器的作用有两方面:第一,平滑频谱,消除谐波,凸显原语音的共振峰;第二,降低运算量。Mel 滤波器组有 M 个,中心频率为 $f(m)$,本系统取 $m = 24$ 。计算滤波器组输出的对数能量,经 DCT 变换再加一阶动态差分,构成 MFCC 系数。

由于 MFCC 提取的特征维数较大,对之后的 HMM (隐马尔科夫模型) 语音识别效率带来很大的影响,所以采用如下优化方法:用主成分分析 PCA 算法对 MFCC 提取的 24 维参数进行优化,一方面减

少 HMM 匹配时的计算量,另一方面使四旋翼飞行的实时性达到很好的效果。主成分分析 PCA 算法如图 11 所示。

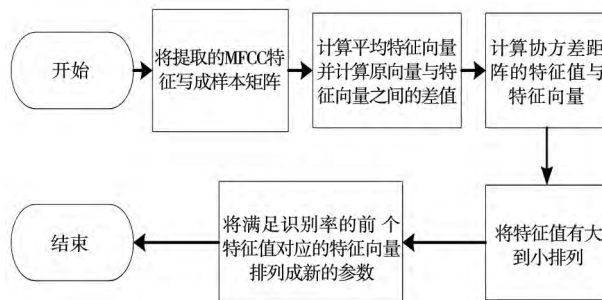


图 11 优化的 PCA 算法流程图

其中: $r <$ 总样本个数 n 。实现降维的目的,且保证了一定的识别率。

3.2 优化的隐马尔科夫算法

隐马尔科夫模型^[4]是一种双重随机过程:①Markov链是状态与转移矩阵具有一一对应的关系,是 HMM 的基本随机过程;②第二个随机过程是描述状态与观察值之间的概率关系。在使用者角度来看,HMM 内的随机过程是透明的,状态转移是看不到的,因而叫做隐马尔科夫模型。

隐马尔科夫模型需要解决的 3 个问题:

1) 估计问题:给定观察序列 $O = \{O_1, O_2, O_3, \dots, O_T\}$ 和参数模型 $\lambda = (\pi, A, B)$,如何确定观察序列与模型之间达到最佳匹配,其中 T 为语音信号帧数, π 为初始状态概率矢量, A, B 分别为状态转移概率矩阵和 HMM 模型的观察值概率矩阵;

2) 解码问题:已知观察序列 $O = \{O_1, O_2, O_3, \dots, O_T\}$ 和参数模型 $\lambda = (\pi, A, B)$,怎样选择最佳状态序列;

3) 模型训练问题:怎样修正观察序列的参数模型 $\lambda = (\pi, A, B)$,使 $P(O|\lambda)$ 概率最大,达到最佳识别效果。

针对以上问题,本文分别采用前向后向算法、Viterbi 算法、Baum—Welsh 算法。由于 Baum—Welsh 算法在进行训练时对初始参数要求较高,此外四旋翼飞行器要求实时性较高,所以对 Baum—Welsh 算法进行优化如下:

对于没有约束的随机过程来说, B 矩阵可以是任意值,但对于本文的约束,让 B 矩阵的值全部为 1,利用 Viterbi 算法划分状态序列,再利用 K —均值算法对 B 矩阵进行重新估计,得到新的 B_1 矩阵和新的 λ_1 ,用 λ_1 代入 Baum—Welsh 算法中重新计算,

得到 λ_2 , 比较 λ_1 和 λ_2 的差值与设定的阈值的大小, 得最佳值。

具体优化的 HMM 流程图如图 12 所示。

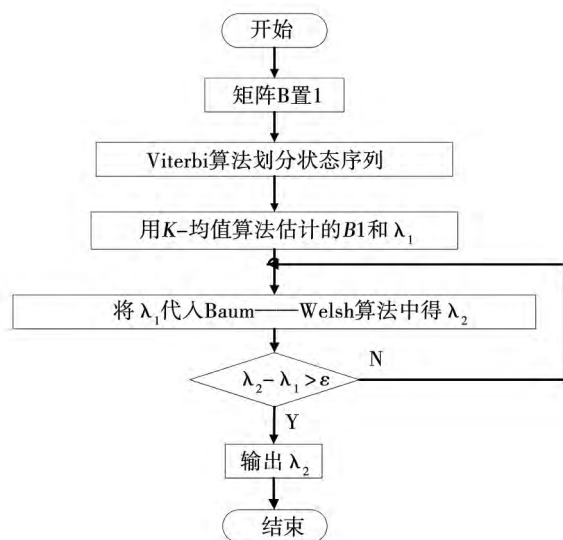


图 12 优化 HMM 参考模板程序流程图

整体语音识别程序流程图如图 13 所示。

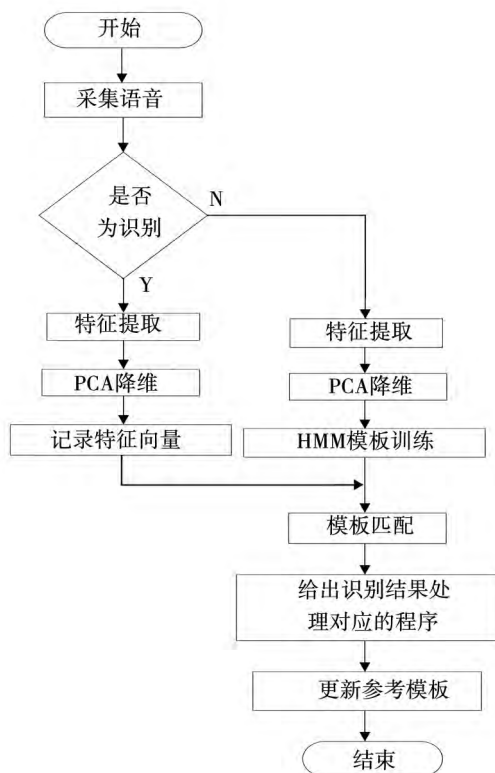


图 13 优化的隐马尔科夫算法流程图

对 HMM 孤立词识别的结果, 发送到四旋翼上进行控制。由摄像头当前的位置以及控制者想要的位置对四旋翼进行命令编码。当前的指令系统由

14 条指令组成, 分别为起飞、降落、前进、后退、左行、右行、上升、下降、停止、左旋、右旋、悬停、拍照和摄像。对 14 条指令进行编码, 用 5 位二进制位表示, 第一位为预留位, 方便以后添加其他指令, 后四位为指令位。且增加指令的数目是以指数形式增长, 成本大大降低。开发者的命令编写也比较简单, 由于系统已经成型, 算法也没有多大改动, 只是增加几条判断语句。控制系统指令编码表如表 1。由于时间有限本系统只设计了当前的 14 条指令, 以后还可以加角度控制命令, 如左转 5 度和右转 10 度等一些命令。使四旋翼飞行器的指令系统更加完整, 控制更加的准确和简便。

表 1 四旋翼飞行器动作对应指令编码表

动作	编码	动作	编码
起飞	00001	降落	00010
前进	00011	后退	00100
左行	00101	右行	00110
上升	00111	下降	01000
左旋	01001	右旋	01010
停止	01011	悬停	01100
拍照	01101	摄像	01110

语音识别的四旋翼飞行器控制系统整体流程图如图 14 所示。

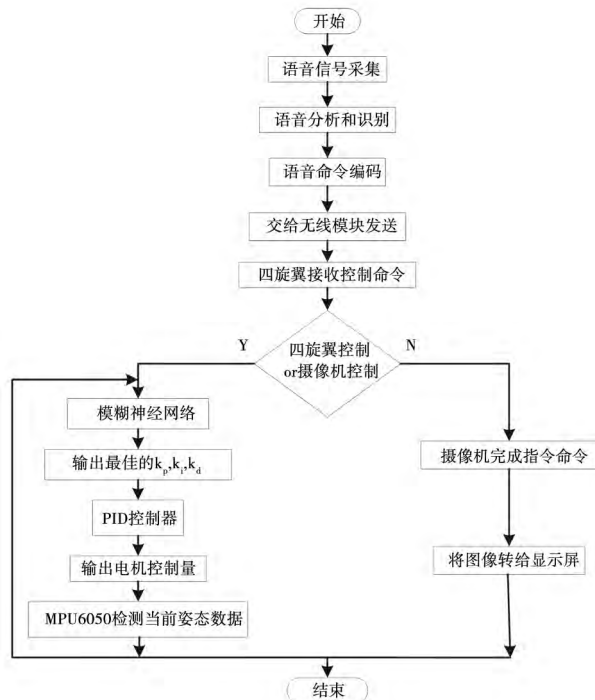


图 14 语音控制命令识别系统整体流程图

将式(6)和图12算法在MATLAB做仿真实验,实验结果如表2所示。

表2 两种算法识别率比较

系统	算法	识别率/%	时间/s
原传统	HMM	94.8	0.98
优化后系统	优化 HMM	98	0.862

最终用MATLAB来实现的语音识别命令的仿真图如图15所示。

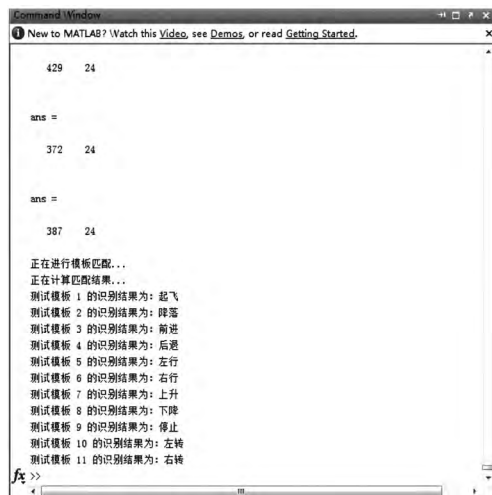


图15 语音识别命令仿真结果图

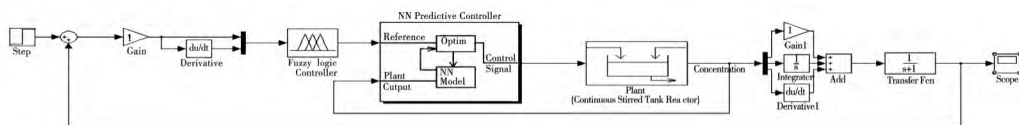


图16 模糊神经网络PID控制规则的四旋翼飞行器控制系统的SIMULINK仿真模型图

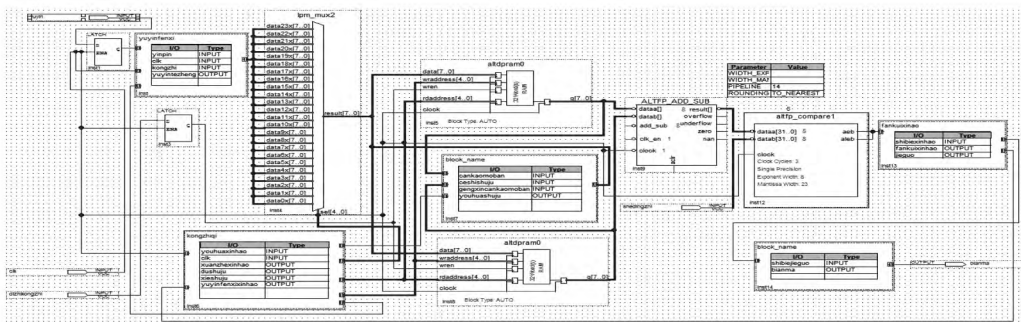


图17 FPGA控制命令识别系统的顶层框架图

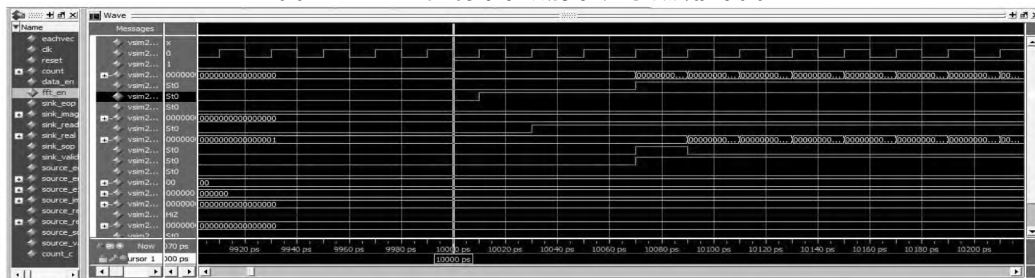


图18 FPGA语音识别顶层时序仿真图

4 FPGA实现与系统仿真

利用MATLAB对四旋翼控制系统进行算法仿真。基于模糊神经网络PID控制规则的四旋翼飞行器控制系统的SIMULINK仿真模型如图16所示。四旋翼飞行器控制系统由一个计算偏差和偏差变化率的结构、模糊控制器、神经网络训练、PID控制器和系统传递函数构成。再系统前加一个阶跃响应发生器,系统最后接一个图像接收器,目的是显示本系统的调节状态。

FPGA控制命令识别系统的顶层框架如图17所示,顶层框架采用自下而上的设计方法,首先分别设计MFCC特征提取、优化的HMM语音识别、命令存储及控制等模块,仿真验证后封装,封装后通过各模块的硬联逻辑形成顶层结构。顶层时序仿真如图18所示,从仿真图中可以看出语音识别的时间为10 ns,且在每个时钟周期上顶层数据的传输稳定正常,说明了顶层数据的传输能够快速、准确地识别,反过来证明了优化的动态时间规整算法的正确性。最后,将已编好的语音识别识别模块下载到FPGA封装后与Quartus II板上的麦克和显示器相连,因控制命令简单可达100%。

5 结 语

通过 SIMULINK 的仿真与 Quartus II 的仿真实验验证了四旋翼控制器系统和语音控制命令识别系统具有很强的实时性,能很好地实现四旋翼飞行器的稳定工作,且数据能够很好地交互,完成语音控制四旋翼飞行器进行航拍的任务。

参 考 文 献:

- [1] 刘志华,郭付才,彭新伟,等. 基于 CY7C68013A 的 FPGA 配置和通信接口设计 [J]. 电子技术应用,2013,39(2):46-54.
- [2] 靖固,宋振伟,王铮. 车道线识别系统算法设计及其 FPGA 实现 [J]. 哈尔滨理工大学学报,2013,18(6):75-77.
- [3] 王强. 小型无人直升机航拍系统设计与实现 [D]. 上海: 上海交通大学. 2013:36-48.
- [4] ASHUTHA K, ANKITHA K. Smart Shopping Cart Using Embedded System and Wireless Modules [J]. Recent Patents on Computer Science,2016,9(3):2213-2759.
- [5] 张海涛,方明,付飞. 嵌入式轮式机器人实验平台设计与实现 [J]. 吉林大学学报(信息科学版),2016(5):670-675.
- [6] HECTOR Perez Tijero, MARIO Aldea Rivas, DANIEL Medina Ortega. Multiprocessor Platform for Partitioned Real-time Systems [J]. Softw. Pract. Exper.,2017,47(1):1002-1014.
- [7] 李德仁,李明. 无人机遥感系统的研究进展与应用前景 [J]. 武汉大学学报(信息科学版),2014,39(5):505-513,540.
- [8] 汪绍华,杨莹. 基于卡尔曼滤波的四旋翼飞行器姿态估计和控制算法研究 [J]. 控制理论与应用(英文版),2013,30(9):1109-1115.
- [9] 昂海松. 微型飞行器的设计原则和策略 [J]. 航空学报,2016,37(1):69-80.
- [10] 邹继斌,赵博,梁维燕,等. 多单元永磁同步电机数学模型与转矩波动抑制 [J]. 电机与控制学报,2012,16(5):46-51.
- [11] 刘长征,张磊. 语音识别中卷积神经网络优化算法 [J]. 哈尔滨理工大学学报,2016(3):34-38.
- [12] 刘志军,张杰. 模糊数据流的进化粒度神经网络分类算法 [J]. 哈尔滨工程大学学报,2016,37(3):3-7.
- [13] 张戈. 基于 FPGA 实时语音识别系统的实现 [D]. 哈尔滨: 哈尔滨理工大学学报,2012:29-40.
- [14] VEISI H, SAMETI H. Hidden-markov-model-based Voice Activity Detector with High Speech Detection Rate for Speech Enhancement [J]. IET Signal Processing, 2012(6):54-63.

(编辑: 温泽宇)