

# 管制指令语音识别在模拟飞行界面的实现<sup>①</sup>

陈亚青<sup>1</sup>, 蒋 豪<sup>2</sup>, 韩 丹<sup>2</sup>

<sup>1</sup>(中国民用航空飞行学院 民航飞行技术与飞行安全科研基地, 广汉 618307)

<sup>2</sup>(中国民用航空飞行学院 空中交通管理学院, 广汉 618307)

**摘 要:** 将语音识别技术应用到管制学员的培训中是提高管制模拟机应用效率的重要途径。在深入分析中文管制指令特征的基础上, 从语音识别基础技术出发, 设计一套适合于中英文管制指令语音特点的语音识别技术, 利用 MATLAB 完成相应程序的编写和模拟飞行界面的设计。通过形成独特的管制指令语音识别模板库, 最终在模拟飞行界面实现管制语音指令对航空器飞行的控制。从而为进一步实现自动化模拟机机长席位提供可能, 最终完全替代人工机长席位。

**关键词:** 管制指令; 语音识别; 模板库; MATLAB GUI; 管制模拟训练

## Realization of the Control Instruction of Voice Controlling Aircrafts' Flying in the Flight Simulation Interface

CHEN Ya-Qing<sup>1</sup>, JIANG Hao<sup>2</sup>, HAN Dan<sup>2</sup>

<sup>1</sup>(CAAC Academy of Flight Technology and Safety, Civil Aviation Flight University of China, Guanghan 618307, China)

<sup>2</sup>(College of Air Traffic Management, Civil Aviation Flight University of China, Guanghan 618307, China)

**Abstract:** It is an important way to improve the application efficiency of the control simulation machine by applying speech recognition technology to training of ATC students. On the basis of analyzing the characteristics of Chinese and English control instruction, fundamentals of speech recognition technology has been studied in this paper. And a set of speech recognition technology which is suitable for the phonetic characteristics of Chinese and English control commands has been designed. The corresponding procedures and flight simulation interface have been designed with MATLAB. Unique control instruction speech recognition template library is completed, and under the control of voice commands, the aircraft can fly very well in the flight simulation interface. The results provide reference for further realization of the automatic simulation of captain's seat.

**Key words:** control instruction; speech recognition; template library; MATLAB GUI; control simulation training

管制模拟机是培养空中交通管制人员的一个重要训练工具, 它根据管制场景采用计算机技术模拟各种管制情况, 管制学员通过与模拟机长进行联系, 实现对模拟界面飞机的管制。目前, 国内民航院校在使用管制模拟机时, 一般均由管制学员轮流担任飞行员席位人员, 这样大大减少了管制学员的练习训练时间, 降低了训练的效率<sup>[1,2]</sup>。

语音识别是涵盖信号处理、声学语音、模式识别、通信与信息原理、语言学、计算机科学、心理学等学科领域的综合技术<sup>[3]</sup>。它是实现人机智能交互的一种

方式, 通过对输入的语音识别出所代表的具体语义, 执行相应的操作动作, 满足人们的需要。

目前, 国外已成功地将语音识别技术应用于管制模拟机中, 研发出具体的产品, 成功地用于管制学员的培训<sup>[2][4]</sup>。相比之下, 国内语音识别技术起步较晚, 加之管制指令的特殊性, 使得我国的管制指令语音识别技术较国外发展缓慢。目前, 相关研究机构 and 高校主要研究工作集中在以下几个方面文献[5-7]: 分析了语音识别技术应用于管制模拟机中的设计规则、基于成熟的语音识别引擎结合中文指令特点给出了系统的设

① 基金项目: 2014 年中国民用航空飞行学院研究生创新基金项目(X2014-25); 2014 年国家级大学生创新创业训练计划(201410624039)

收稿时间: 2016-03-08; 收到修改稿时间: 2016-04-11 [doi: 10.15888/j.cnki.csa.005445]

计方案; 相关研究工作都缺乏系统性的设计, 也不能涵盖陆空通话标准用语中常用的指令; 此外, 现有的语音识别引擎的设计特点使得其在通用的办公环境中有了广泛的应用, 但是我国中文管制指令识别与管制环境的特殊性使得中文指令用语的实现变得困难。

为了更好地实现将语音识别技术应用于我国管制模拟机中, 实现机长席位自动化, 代替现有的人工机长席位, 我们研究语音识别的基础技术, 结合管制指令特点, 设计一套中英文管制指令语音识别系统, 管制学员发出管制语音后, 管制指令语句经语音识别及分析后可自动转换为目标的执行动作, 从而提高培训效率和效果, 减轻培训人员的工作负担。

### 1 管制指令的特征分析

管制指令专业特征强, 要求发音高度清晰、指令公式化、简练性、词语的单一意旨性及结构祈使性等<sup>[8]</sup>。空中交通管制针对不同的飞行阶段可以分为塔台管制、进近管制和区域管制, 管制的手段有程序管制和雷达管制。这些管制模式的管制用语格式固定, 如航班呼号+指令动作+状态。同时, 管制指令词汇样本空间小、指令较少、重复率高且工作环境噪声低<sup>[1,2,5]</sup>。因此, 管制指令语音模板的训练可将动作短语作为单个词汇进行操作, 代替常用语音模板库建立需要逐字词进行训练不同的方式, 这样既能大大减少语音模板训练数量又能减轻语音合成难度, 同时也缩小了样本空间, 间接提高了语音匹配成功率的可能性。

### 2 系统设计

完整的计算机语音识别系统主要流程如图 1 所示, 包括语音的采集、预处理、特征参数提取、语音训练以及识别。

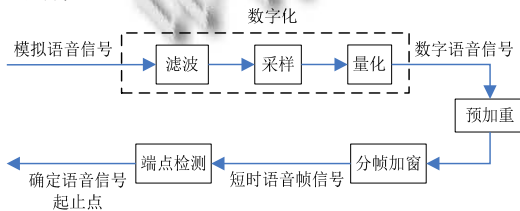


图 1 语音识别系统基本结构图

#### 2.1 语音信号预处理

如图 2 所示, 预处理包括滤波、采样、量化、预加重、分帧、加窗及端点检测等<sup>[2]</sup>。

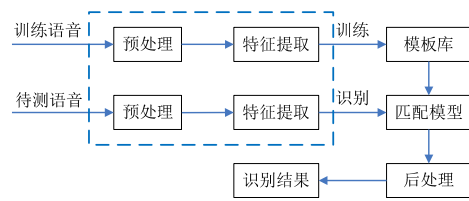


图 2 语音信号预处理结构图

基于文献[1][2], 利用式(1)进行预加重处理, 滤除低频干扰,  $\mu$  为预加重滤波器的系数, 取值范围一般在 0.93—0.97, 本系统取 0.9375; 选用汉明(Hamming)窗函数<sup>[9]</sup>, 其中  $N$  为窗长取 256,  $n$  为采样点, 如式(2); 端点检测方法采用基于短时能量和短时平均过零率的双门限检测方法对信号进行处理<sup>[10,11]</sup>, 分别如式(3)—(5), 其中  $x_i(m)$  为分帧加窗后的语音信号, 帧移为 80ms。

$$H(Z) = 1 - \mu z^{-1} \tag{1}$$

$$w(n) = \begin{cases} 0.54 - 0.46 \cos(2\pi n / (N-1)), & 0 \leq n \leq N-1 \\ 0, & \text{其他} \end{cases} \tag{2}$$

$$E_i = \sum_{n=1}^N x_i^2(m) \tag{3}$$

$$ZCR_i = \sum_{n=1}^N |sign[x_i(m)] - sign[x_i(m-1)]| \tag{4}$$

式中

$$sign[x_i(m)] = \begin{cases} 1 & x_i(m) \geq 0 \\ -1 & x_i(m) < 0 \end{cases} \tag{5}$$

#### 2.2 语音信号特征参数的提取

根据文献[2], 管制指令语音信号特征参数提取采用美尔频率倒谱系数(MFCC)方法, 在非特定人时, 该方法训练和识别速度快且识别率高<sup>[12]</sup>。

#### 2.3 模板训练

语音识别是语音匹配过程, 语音模板库的好坏直接决定识别率的高低, 而识别的性能和模板训练的方式具有很大关系<sup>[13]</sup>。本文采用鲁棒性训练方法, 也就是得到的模版是在一致性较好的特征矢量序列对在沿 DTW 的路径上求平均的结果。

模板训练算法采用 HMM 参数训练算法 Baum-Welch。Baum-Welch 算法是极大似然(ML)准则的一个应用, 它采用了一种多次迭代的优化算法, 用 Lagrange 数乘法构造一个目标优化函数  $Q$ , 包含了所有 HMM 参数作为变量, 令  $Q$  对各变量求偏导推出  $Q$

达到极点时新的 HMM 参数相对于旧的模型参数之间的关系,从而得到 HMM 各参数的估计<sup>[13,14]</sup>.用新旧 HMM 模型参数之间的函数关系反复迭代运算,直到 HMM 模型参数不在发生明显的变化为止.

### 2.4 语音识别算法

利用 Viterbi 算法不仅可以找到一条足够好的状态转移路径,还可以得到该路径所对应的输出概率.采用对数形式的 Viterbi 算法,将避免进行大量的乘法计算,既能减少计算量又能保证有很高的动态范围,而不会由于过多的连乘而导致溢出问题.

对数形式的 Viterbi 算法如下:

1) 预处理

$$\tilde{\pi}_i = \log(\pi_i) \quad (6)$$

$$\tilde{b}_i(o_i) = \log[b_i(o_i)] \quad (7)$$

$$\tilde{a}_{ij} = \log(a_{ij}) \quad (8)$$

2) 初始化

$$\tilde{\delta}_1(i) = \log[\delta_1(i)] = \tilde{\pi}_i + \tilde{b}_i(o_i) \quad (9)$$

$$\psi_1(i) = 0, 1 \leq i \leq N \quad (10)$$

3) 递推

$$\tilde{\delta}_t(j) = \log[\delta_t(j)] = \max_{1 \leq i \leq N} [\tilde{\delta}_{t-1}(i) + \tilde{a}_{ij}] + \tilde{b}_j(o_t) \quad (11)$$

$$\psi_t(j) = \arg \max_{1 \leq i \leq N} [\tilde{\delta}_{t-1}(i) + \tilde{a}_{ij}], 2 \leq t \leq T, 1 \leq j \leq N \quad (12)$$

4) 终止

$$\tilde{p}^* = \max_{1 \leq i \leq N} [\tilde{\delta}_T(i)] \quad (13)$$

$$q_T^* = \arg \max_{1 \leq i \leq N} [\tilde{\delta}_T(i)] \quad (14)$$

5) 回溯最佳路径

$$q_t^* = \psi_{t+1}(q_{t+1}^*), 1 \leq t \leq T-1 \quad (15)$$

其中,  $\delta_t(i)$  为  $t$  时刻第  $i$  状态的累积输出概率,  $\psi_t(i)$  为  $t$  时刻第  $i$  状态的前序状态号,  $q_t^*$  为最优状态序列中  $t$  时刻所处的状态,  $p^*$  为最终输出概率.

## 3 系统实现

### 3.1 管制指令语音采集

设备: 麦克风和防噪网各一个, 安装好声卡和数字处理器的 PC 机一台.

采样频率 8KHz, 采样精度 16 位(bit), 单声道输出方式, 语音存储格式为\*.wav. 利用 MATLAB 编写语音采集程序, 注意: 在录音时, 应按需调整好 MIC 的音量, 以保证噪音电平较低且语音信号不会过强, 这样既避免语音信号的幅度过小, 又防止其幅度太高, 以致切顶造成失真.

### 3.2 系统设计

利用 Matlab 编程技术, 对采集到的语音经过处理后, 利用 Baum-Welch 训练算法建立管制指令语音的声学模型, 形成具有中英文特点的管制指令的模板库; 识别时通过 Viterbi 算法与模板语音比较的最佳匹配结果便是输出语音. 整个设计程序复杂, 实现时较为繁琐, 而 MATLAB GUI 技术能嵌入已有的仿真程序又能把仿真后的图形化结果以人机交互的动态方式直观呈现, 可以方便用户进行操作<sup>[15,16]</sup>. 为了使整个系统更为简便和人性化, 利用 MATLAB GUI 设计管制模拟训练界面, 如图 3 所示.

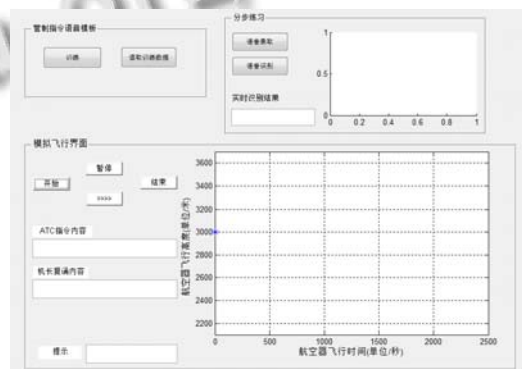


图 3 管制模拟训练界面

该界面主要分为三大模块: 管制指令语音训练(模板生成)、分步练习和模拟飞行界面.

制指令语音训练模块可以实现对采集到的语音进行训练. 点击该界面中“训练”按钮, 软件会自动执行内嵌的管制指令语音训练程序, 训练最终生成格式为.mat 的语音模板文件, 进行识别时只需提前读取模板即可, 训练进度可根据图 4 进度条观测.



图 4 训练进度界面

分步练习模块用于管制学员进行管制指令用语的模块化训练, 该模块能显示指令波形图, 同时也可直接查看识别结果, 促进管制学员管制用语的标准化、规范化.

模拟飞行界面模块主要用于显示航空器在指令下达后的运动轨迹, 以此判断指令的正确与否. 其中, 左半部分为操作按钮和指令显示栏; 右半部分采用二

维界面显示航空器当前高度, 纵坐标代表当前飞行的高度(单位/米), 横坐标表示当前航空器的飞行时间(单位/秒).

#### 4 实例应用

本文以“国航 101, 上高度 3600 并保持”为例, 指挥波音 747 进行模拟飞行, 设定爬升下降率都为 600 米每分钟.

首先, 在管制指令语音训练模块中, 分别将“国航”、“air china”、“保持”“maintain”、“3000”、“three thousand”、“上高度”、“climb”、“3600”、“下高度”、“descend”、“2400”、“two thousand and four hundred”等管制指令录入管制指令语音模板库中. 为了克服语音的偶然性, 每个语音指令分别由十人共采集 100 次.

在生成模板库之后, 如图 5 所示, 点击模拟飞行人机交互界面上的“开始”按钮, “提示框”内会出现“请输入您的语音指令”的字样, 这时输入管制语音指令“国航 101, 上高度 3600 并保持”, 左半部分操作区显示出管制语音指令的内容, 模拟机长会根据 ATC 指令作出答复并使用语音播报复诵内容. 右半部分的飞行模拟界面会出现一黑点和其坐标, 其中, 黑点表示当前给定指令的航空器, 航空器会根据管制指令内容做相应的模拟飞行. 图中表示第 13 秒, 航空器的高度为 3120 米, 并且在继续爬升.

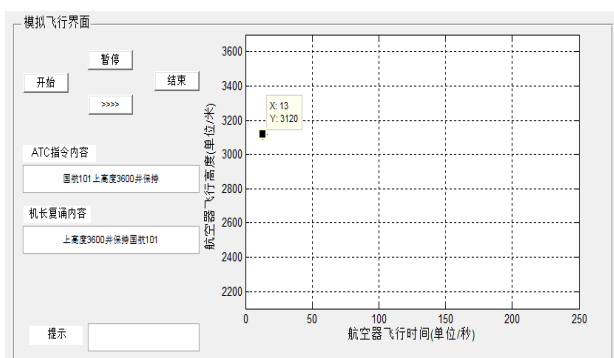


图 5 实例分析 I

如图 6 所示, 当 61 秒后模拟航空器到达 3600 米并保持这一高度. 点击交互界面按钮“>>>>”, 提示框会给出“请输入下一条指令”, 之后管制学员可与之前相同的操作方法, 继续输入拟操作管制语音指令实现对航空器的指挥. 如果暂停当前操作点击交互界面“暂停”按钮; 结束整个操作直接点击“结束”按钮.

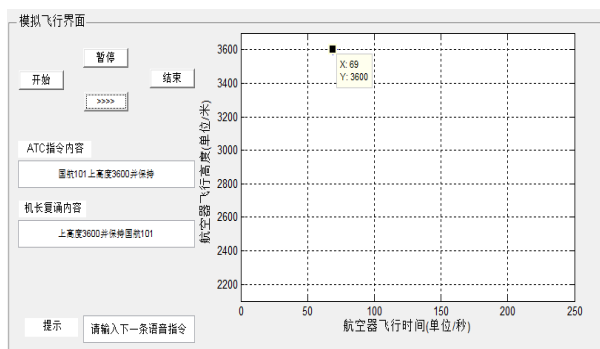


图 6 实例分析 II

#### 5 总结

论文对中文管制指令在语音识别方面的特征进行了深入分析, 将动作短语作为单个词汇进行处理, 提出了一套适合于管制学员自主进行训练的中英文管制培训的系统, 并完成了系统设计主要工作. 基于 MATLABGUI 技术, 设计了管制指令语音训练界面、分步训练界面和模拟飞行界面. 利用模拟飞行界面, 实现了用识别后的管制指令语音控制航空器的飞行, 同时界面会根据语音管制指令语音复诵相应指令并文字显示, 练习者可以清晰明了地获得自己的操作结果, 以便检查, 从而不断提升管制用语的标准化, 促进管制用语规范化.

在本文研究的基础上, 下一步需充分完善管制指令语音模板库, 完成对所有管制指令的模板设计, 最终将该系统嵌入到现有管制模拟培训机中, 完全实现机长席位的自动化功能. 从而促进语音识别技术广泛应用于管制模拟机教学, 提高管制培养质量.

#### 参考文献

- 何昕, 韩丹. 管制指令语音识别模板库的关键技术研究. 中国西部科技, 2015, (3): 1-3.
- 何昕, 蒋豪, 韩丹. 管制指令特征参数提取研究. 科学技术与工程, 2015, 15(20): 89-94.
- 韩纪庆, 张磊, 郑铁然. 语音信号处理. 北京: 清华大学出版社, 2013.
- Pardo JM, Ferreiros J, Fernandez F, et al. Automatic understanding of ATC speech: Study of prospectives and field experiments for several controller positions. IEEE Trans. on Aerospace & Electronic Systems, 2011, 47(4): 2709-2730.
- 袁伟伟, 胡军, 刘万凤. 空管模拟训练中指令的语音识别与合

- 成技术研究.电子设计工程,2013,(6):8-11.
- 6 由扬.空管模拟训练机的语音应用系统设计研究[硕士学位论文].南京:南京航空航天大学,2002.
- 7 李锐.语音技术在塔台模拟机上的应用[硕士学位论文].成都:四川大学,2004.
- 8 MH/T 4014-2003,空中交通无线电通话用语标准.
- 9 宋知用.MATLAB 在语音信号分析与合成中的应用.北京:北京航空航天大学出版社,2013.
- 10 Junqua JC. Robustness and cooperative multimodel man-machine communication applications. Proc. Second Venaco Workshop and ESCA ETRW. 1991, 9. 101-112.
- 11 王佳乾.基于 HMM 语音关键词识别技术研究[硕士学位论文].南京:南京理工大学,2013.
- 12 Bhasher PV, Rao SRM. A computer aided MFCC-HMM based speech controlled automation system. International Journal of Electronic Communications Engineering Advanced Research, 2014, 2(2): 148-155.
- 13 王秀景.语音识别技术研究是实现[硕士学位论文].青岛:山东科技大学,2011.
- 14 何英,何强.MATLAB 扩展编程.北京:清华大学出版社,2002.
- 15 王光艳,赵晓群,王霞.基于 MATLAB GUI 的语音信号特征提取系统设计.河北工业大学学报,2010,4:14-18.
- 16 罗华飞.MATLAB GUI 设计学习手记.北京:北京航空航天大学出版社,2014.