

基于隐马尔可夫模型和聚类的英语语音识别混合算法

朱 祥

(杨凌职业技术学院 文理分院, 陕西 杨凌 712100)

摘要: 与单词相比, 具有大量特征数据和复杂发音变化的英语语音在隐马尔可夫模型(HMM)中存在更多问题, 例如维特比算法的复杂度计算和高斯混合模型中的概率分布问题; 为了实现基于HMM和聚类的独立于说话人的英语语音识别系统, 提出了用于降低语音特征参数维数的分段均值算法、聚类交叉分组算法和HMM分组算法的组合形式; 实验结果表明, 与单个HMM模型相比, 该算法不仅提高了英语语音的识别率近3%, 而且提高系统的识别速度20.1%。

关键词: 英语语音识别; 隐马尔可夫模型; 聚类; 特征数据

Hybrid English Speech Recognition Algorithm Based on HMM and Cluster

Zhu Xiang

(Arts & Sciences Branch College, Yangling Vocational & Technical College, Yangling 712100, China)

Abstract: For English sounds with a large number of characteristic data and complex pronunciation variations, there are more problems in hidden Markov models (HMM) than words, such as complexity calculation of Wittby algorithm and probability distribution in Gaussian mixture model. In order to realize the speech-independent English speech recognition system based on HMM and clustering, a combination of segmented mean algorithm, clustering crossover grouping algorithm and HMM grouping algorithm is proposed to reduce the dimension of speech feature parameters. Experimental results show that compared with the single HMM model, the algorithm not only improves the recognition rate of English speech by 3%, but also improves the recognition speed of the system by 20.1%.

Keywords: English sentence recognition; hidden Markov model; clustering; feature data

0 引言

语言是人类用来交换信息的一种交流方式。语音处理是对语音信号^[1]及其处理方法的研究, 语音信号通常以数字形式处理, 语音识别是对语音信号进行转换的过程。每种语言都有自己的语法, 对于给定的一个英语语音, 如何利用已有的知识和技术, 认识和理解其内容是当前英语语音识别的难点。由于英语语音的语法与汉语不同, 所以语法分析^[2-3]给识别系统增加了额外的负担。

与英语单词相比, 更多的特征数据和更复杂的发音变化使英语语音的语音识别更加困难。首先, 英语语音的词汇量更大, 发音的单词之间没有明显的停顿。也就是说, 单词之间没有明确的边界。其次, 英语语音中每个单词的发音通常比较自然, 关联语的发音比孤立词的发音更随意, 因此协同发音的效果更为严重。此外, 受语境的影响, 在英语发音过程中, 英语语音在发音、节奏、语调、重音和语速等方面可能存在差异, 甚至即使是同一个说话人, 在不同的时间和环境下, 其韵律特征也是不同的。

隐马尔可夫模型 (Hidden Markov Model, HMM)^[4-7]

作为独立于大词汇量说话者的连续语音识别系统的主流技术, 已经取得了相当大的成功。通过分析语音信号的短期能量, 提取具有帧长特征的语音特征, 对整个语音进行隐马尔可夫建模。模型训练使用由许多说话者记录的训练集, 并使用统计理论来解决个体与整体之间的差异, 从而使说话者独立的单句隐马尔可夫建模具有鲁棒性。HMM使用概率模型来描述统计的发音。隐马尔可夫模型中的连续状态转换可以看作是短语音段的发声, 即一段相连的HMM, 它代表了语音段。HMM是近年来主要用于语音识别的方法。本文以HMM作为语音识别核心。在语音识别的过程中, 系统采用维特比算法^[8-10]进行解码, 从而找出正确的识别结果。在单个语音上使用隐马尔可夫建模, 可以描述每个语音中单词的相关性。在说话者进行充分训练的条件下, 可以较精确地实现说话者独立的简短的英语语音建模。然而, HMM需要语音信号的先验统计知识、分类决策能力较弱等问题, 包括维特比算法的复杂度计算和高斯混合模型中的概率分布问题, 这些缺点使得进一步提高单个HMM的识别性能变得困难。

语音识别领域的大多数文献都对HMM中的聚类算法进行了改进, 并将其作为模式分类的方法, 进行优化模型参数的估计, 但对语音识别的效果并不理想。对于数据量大、发音变化较为复杂的英语语音, HMM的不足就更加明显, 其识别时间更长。为了有效提高识别效率, 本文在单个隐马尔可夫模型的基础上, 尝试将聚类算法与隐马尔

收稿日期: 2019-10-16; 修回日期: 2019-11-05。

基金项目: 杨凌职业技术学院人文社科研究基金项目 (GJ19100)。

作者简介: 朱祥(1982-), 男, 陕西凤翔人, 讲师, 硕士, 主要从事英语语言与教学方向的研究。

科夫模型相结合,并将其应用于英语语音识别中。根据英语语音的特点和语音之间的相似性,将英语语音的数据集分成若干组,每组由一些语音特征相似的语音组成。因此,在识别英语语音时,无需计算维特比解码中的所有语音,只需计算输入语音所属的选定组中的 HMM 参数即可。在选择合适的聚类组的情况下,系统将会节省大量的计算量,识别性能也会有很大的提高。这不仅为满足实时性的小型设备应用提供了一种新的语音识别参考方法,而且为新型英语语音评估系统奠定了语音识别的基础。

1 系统设计

聚类算法与隐马尔科夫模型相结合的算法流程如图 1 所示,首先对输入语音信号进行预处理,包括预加重,帧处理,窗口添加和端点检测。然后提取语音特征参数 MF-CC,通过分段平均算法降低 MFCC 的维数^[11-13],采用动态时间规整(DTW)算法^[14-15]确定语音特征聚类组 K,然后计算出 K 组内的 HMM 参数,最后输出识别结果并进行后处理。

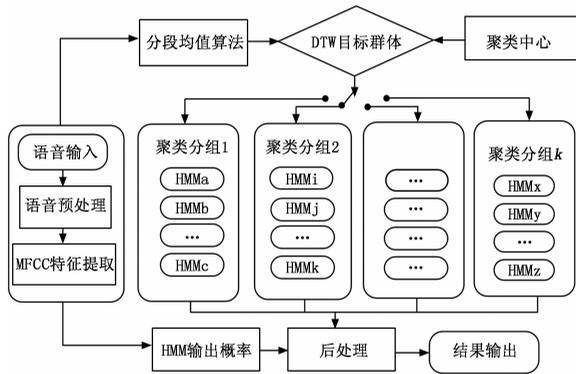


图 1 基于 HMM 和聚类的语音识别框架图

2 隐马尔科夫模型

隐马尔科夫模型(HMM)是由随机状态转移和符号输出组成的有限状态自动化。自动化模型概率生成过程的一系列符号是由从指定的开始状态,过渡到一个新的状态的过程。

以数学方式表达模型,如式(1)~(4)所示。

隐马尔科夫模型是一个五元数组 (S, A, Π, V, B) 。

$$H = (S, A, \Pi, V, B) \tag{1}$$

其中: S 为状态集。 $A = \{a_{i,j}\}$, 表示是从状态 i 到状态 j 的概率。

$$\sum_{j=1}^{N=j=1} a_{i,j} = 1, 1 \leq i \leq N \tag{2}$$

$\Pi = \{P[S_1(1)]\}$ 表示在 $t=1$ 时处于状态 S_i 的概率。

$$V = \{V_1, \dots, V_m\} \tag{3}$$

其中: m 为离散词汇表 V 中符号的个数。 $B = \{B_j(V_k)\}$ 表示符号 V_k 在状态 S_j 下的概率。

其中:

$$\sum_{k=1}^M b_j(V_k) = 1, 1 \leq j \leq N \tag{4}$$

3 核心算法

3.1 分段均值算法

由于 K-means 聚类算法^[6-17]具有随机选择采样点的迭代特性,再加上语音特征参数的维数较高,因此聚类结果的稳定性相对较差。为此,本文探讨了降维语音特征参数的分段均值算法,如图 2 所示。

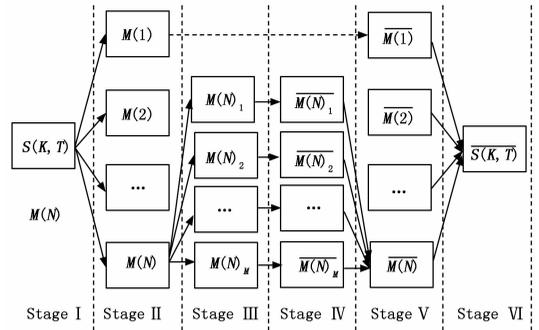


图 2 语音特征系数降维的分段平均算法

分段平均算法将语音特征参数分割成具有相同维数的片段,具体包括四个步骤:

1) 将语音特征参数定义为 $S(K, J)$, 其中 K 表示 MFCC 参数的阶数, J 表示分段帧数。假设 T 为分段之前的帧数,然后将语音特征参数分为 N 个片段,则为:

$$M(i) = S(K, J), J = \left[\frac{T}{N}(i-1) + 1 \right], \dots, \left[\frac{T}{N}i \right] \tag{5}$$

$M(i)$ 代表片段化语音特征参数的第 i 个片段。 N 的值设置为 HMM 的编号。

2) 将语音特征参数分成平均段后,我们将片段 M 继续分成 M 个平均片段 (M 的值设置为 HMM 的观察序列号)。子段的计算请参见上面的公式。

3) 每个子段的平均值由 $\overline{M(i)}_k$ ($k=1, 2, \dots, M$) 给出。

4) 将所有子段的均值合并到一个矩阵中,该矩阵表示降维后输出的语音特征参数,可以定义为 $\overline{S(K, T)}$, $\overline{S(K, T)}$ 的大小为 $K \times M \times N$ 。

表 1 中显示了图 2 中的参数总数。分段平均算法将特征参数矩阵的大小由 $T \times K$ 变为 $K \times M \times N$ 。也就是说,该算法成功地从矩阵中去除了帧长 T 。这意味着,经过分段平均计算后,矩阵(降维)保持相同的大小,确定了语音特征参数 K (语音特征参数的阶数)、 N (片段大小)和 M (子片段大小)的特征参数矩阵大小。这使得不同长度的语音可以被构造为相同大小的矩阵,从而极大地促进了语音特征聚类算法的实现。

3.2 聚类交叉分组算法

为了进一步提高语音特征聚类领域的性能,本文提出了一种新的二次训练方法:聚类交叉分组算法。如图 3 所示,聚类交叉分组算法包括 3 个步骤:

1) 使用 K-means 聚类算法对训练语音样本的特征进行聚类。

表 1 分段均值算法语音特征系数处理的参数表

Stage	I	II	III	IV	V	VI
矩阵大小	$T \times K$	$(\frac{T}{N}) \times K$	$(\frac{T}{NM}) \times K$	$(\frac{T}{NM} \times \frac{1}{T}) \times K$	$M \times K$	$(M \times N) \times K$
参数数量	$T \times K$	$T \times K$	$T \times K$	$K \times M \times N$	$K \times M \times N$	$K \times M \times N$

2) 使用动态时间规整 (DTW) 算法计算训练语音样本与聚类中心之间的距离。对于每个样本, 最小距离决定其目标组。

3) 检查目标组是否包含训练样本。如果包括, 则说明分类是正确的; 否则语音将被添加到目标组。

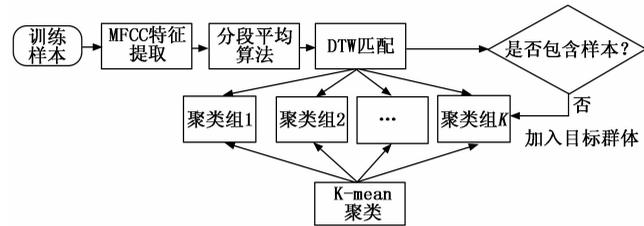


图 3 聚类交叉分组算法

3.3 HMM 分组算法

在基于单个 HMM 的识别系统中, 使用维特比算法进行解码操作时, 所有的模型参数都必须参与到计算中。假设系统词汇表的数量为 n , 那么 HMM 参数的数量为 n 。当识别一个语音时, 每个输出概率分别通过维特比算法在 n 个 HMMs 内计算。因为每个孤立的语音都有一个唯一的 HMM 参数, 我们能够将特征聚类组中的语音映射为相应的 HMM 参数。因此, 我们实现了如图 4 所示的 HMM 聚类分组模型。

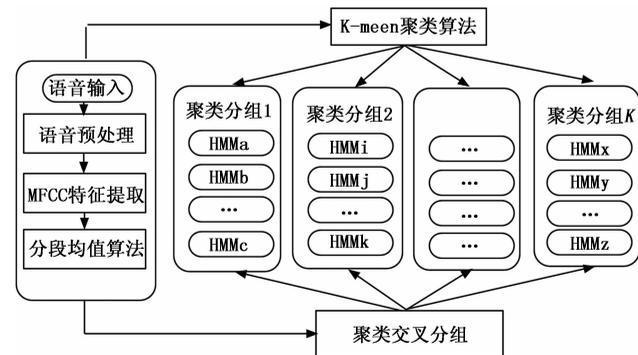


图 4 HMM 分组算法

由于聚类交叉分组算法具有良好的分组性能, 所以聚类组中的 HMM 参数的数量总是小于或等于系统词汇表的数量。此外, 改进的语音特征聚类模型可确保较高的分组准确率。因此, 本文提出将特征聚类模型与 HMM 相结合, 形成一个基于聚类与 HMM 的混合模型——英语语音识别系统。

在语音识别之前需要对语音信号进行预处理。语音预处理包含采样、量化、端点检测、预加重和加窗。在语音信号预处理之后, 识别特征特征参数以用于随后的识别计

算。在本文中, 线性预测系数 (LPC) 是最重要的特征参数。

线性预测的概念起源于在语音期间采样点的幅度与相邻采样点的幅度。如果语音信号的后采样序列是 $S(n)$, 则语音信号的当前样本, 即 $S(n)$, 是时间 n 的样本值。如果 $\hat{S}(n)$ 是 $S(n)$ 的预测值, 则由于在预测值和实际值之间必须存在误差, 因此预测误差可以表示为 $e(n)$, 如下所示:

$$e(n) = S(n) - \hat{S}(n) = S(n) - \sum_{j=1}^l a_j S(n-j) \quad (6)$$

其中: a 是线性预测编码, l 是线性预测的阶数。调整系数 a_j , 只要上式的平方误差值最小, 就可以得到最优的线性预测系数 a_j 。在求解线性预测系数之前, 先确定自相关系数, 然后利用所得到的自相关系数, 利用 Durbin 算法得到期望的线性预测系数。

在确定 LPC 之后, 从 LPC 推导出倒谱系数。倒谱系数^[18]将声道模型与激励信号分离, 可以更精确地计算声道参数, 从而控制语音频谱特性。倒谱系数 c_j 由线性预测系数 a_j 确定, 其中 l 是线性预测的阶数, 如下所示:

$$\dot{c}_j = a_1 \quad (7)$$

$$\dot{c}_j = a_j + \sum_{k=1}^{j-1} (1 - \frac{k}{j}) \cdot a_k \cdot c_{j-k}, (1 < j \leq l) \quad (8)$$

$$\dot{c}_j = \sum_{k=1}^p (1 - \frac{k}{j}) \cdot a_k \cdot c_{j-k}, (j > l) \quad (9)$$

根据预设的 HMM 状态编号, 将状态和帧与语音片段的音频部分平均分开, 并将帧中的特征向量用于计算平均值 M_j 和方差 Var_j , 如下式所示, 其中 A 是 HMM 的状态, i 是帧, j 是特征参数, T 是状态下的帧数, q 是倒谱特征向量的数量 (本文中设置 q 的值为 30)。

$$M_j = \frac{\sum_{i=1}^T A(i, j)}{T}, 1 \leq j \leq q \quad (10)$$

$$Var_j = \frac{\sum_{i=1}^T [A(i, j)]^2}{T} - (M_j)^2, 1 \leq j \leq q \quad (11)$$

3.4 维特比算法

维特比算法是一种动态规划算法, 用于寻找最有可能产生观测事件序列的维特比路径—隐含状态序列, 特别是在马尔可夫信息源上下文和隐马尔可夫模型中。经过转换、转移概率、观测/发射概率等过程, 形成训练有素的隐马尔可夫模型。该模型由转换和可观察/发射概率矩阵的值以及从测试数据集得出的符号序列组成, 此时采用维特比算法, 由此得到最可能产生该符号的状态序列。高斯概率函数可以

表 2 不同识别方法的语音 1(15 个样本)的识别时间表

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
单一模型	1.44	1.52	1.86	1.64	1.44	1.45	1.41	1.37	1.94	1.79	1.62	1.64	1.40	1.59	1.52
混合模型	1.23	1.25	1.30	1.25	1.26	1.24	1.26	1.26	1.31	1.29	1.25	1.24	1.22	1.29	1.23

用来确定状态和帧的相似性概率值。较高的概率值表示相应帧与状态之间的相似度较高，如式 (12) 所示：

$$G_i(x_T) = \frac{1}{\sqrt{(2\pi)^d |R_i|}} \exp\left\{-\frac{1}{2}(x_T - \tau_{R_i})^T R_i^{-1} (x_T - \tau_{R_i})\right\} \quad (12)$$

其中： $G_i(x_T)$ 为每个状态对应于其帧的概率值， d 为特征向量维， x_T 为特征向量， τ_{R_i} 为状态的平均值， R_i 为密度函数的协方差矩阵， $G_i(x_T)$ 为特征向量 x_T 和状态 i 之间相似度的概率值。

通过高斯概率密度函数可以确定帧与状态之间的概率值。HMM 具有许多用于状态转换的可选路径，并且需要找到所有可能路径中具有最大总概率值的路径。本文使用的维特比算法步骤如下，其中 $\theta_i(i)$ 为在时间 t 停留在状态 i 的概率。 $\phi_i(i)$ 为在时间 t 到达状态 i 的概率， p 为维特比算法的最终概率值， S_T 为最佳状态序列。

步骤一：初始化。

$$\theta_i(i) = \pi b_i(o_1), 1 \leq i \leq N \quad (13)$$

$$\phi_i(i) = 0$$

步骤二：递归运算。

$$\theta_{t+1}(j) = \max_{1 \leq i \leq N} [\delta_t(i) \cdot a_{ij}] \cdot b_j(o_{t+1}) \quad (14)$$

$$\phi_{t+1}(j) = \arg \max_{1 \leq i \leq N} [\delta_t(i) \cdot a_{ij}] \quad (15)$$

$$1 \leq t \leq T-1, 1 \leq j \leq N$$

步骤三：确定概率最大路径。

$$p = \max_{1 \leq i \leq N} [\delta_t(i)] \quad (16)$$

$$S_T = \arg \max_{1 \leq i \leq N} [\delta_t(i)] \quad (17)$$

步骤四：路径回溯。

$$S_t = \Psi_{t+1}(S_{t+1}), t = T-1, T-2, T-3, 1 \quad (18)$$

所需的命令被训练成模型，这些模型可用作语音识别的参考数据库。在识别期间，根据先前的过程确定特征参数。使用维特比算法比较数据库的参考模型，以确定每个模型的概率值并找到最佳状态序列。当语音信号的帧序列对应于状态序列时，自动解决语音信号的时间扭曲问题。语音训练过程的关键是识别帧与状态之间的相关性。通过维特比的连续路径回溯来更新帧与状态之间的关系，直到确定总概率最大的路径为止。识别过程中最重要的一步是比较训练的参考模型，获得参考模型中最大的总概率值。

4 结果验证

为了验证所提模型的有效性，在非特定人英语语音识别系统中，比较了单 HMM 和基于 HMM 和聚类的混合模型的识别率和时间。系统参数的数量为 30。本实验选取 30 个不同的英语语音作为标准句，30 个个体记录 900 个英语语音作为训练样本，15 个个体记录 450 个英语语音作为测

试样本。以英语语音 “Can I have breakfast served in my room?” 为例说明了不同识别方法的识别率和识别时间。例如，对比学生 1 中的语音，系统给出了如图 5 所示的识别结果。

语音识别结果			
句子 “Can I have breakfast served in my room?” 的识别结果			单一模型
匹配时间	1.7834 s	总识别时间	1.8527 s
句子 “Can I have breakfast served in my room?” 的识别结果			混合模型
匹配时间	1.344s	总识别时间	1.4133 s

图 5 学生 1 的语音识别结果

无论是单一模型还是混合模型，识别结果都是正确的，但前者的总识别时间为 1.852 7 秒，后者的总识别时间为 1.413 3 s，仅为前者的 76.22%。也就是说，识别时间减少，并且系统效率提高。比较所有学生（学生 1 至 15）的语音，结果如表 2 所示。实验表明，单个 HMM 和所提模型的识别率均为 100%。但前者的平均识别时间为 1.575 3 s，后者的平均识别时间为 1.258 7 s，仅为前者的 79.90%，从而提高了识别效率。

表 3 是不同识别方法的总体识别性能比较。实验结果表明，与基于单个 HMM 的英语语音识别系统相比，基于 HMM 和聚类（提出的模型）的英语语音识别系统的平均识别率提高了 2.89%，而平均识别时间占 仅前者的 69.25%，提高了系统效率。

表 3 采用不同方法的总体识别性能表

	单一模型	混合模型
平均识别率	96.89%	99.78%
平均识别时间/s	1.7684	1.2248

在使用维特比算法获得状态与框架之间的新关系之后，更新旧状态下的平均值和方差，并使用高斯密度函数再次确定状态与框架之间的更新概率。使用维特比算法获得新的总概率值，更新将持续到最大总概率值收敛为止。

下面以常用的循环神经网络为例。例如语音 “Population growth has been slow in recent years”，系统识别过程如下所示。

1) 首先第一个词 “Population” 作为语音输入，此时产生第一个隐含状态 h_1, h_1 包含了该词的信息。

2) 然后第二个词 “growth” 作为语音输入，并和 h_1 进行融合，此时产生第二个隐含状态 h_2, h_2 包含了前两个词 “Population growth” 的信息。

3) 以此类推，将例句中所有的词输入系统，每输入一

个单词都会同前一时刻的隐含状态进行融合产生一个包含当前词信息和前面所有词信息的新的隐含状态。

当把整个语音所有的词输入进去之后，最后的隐含状态理论上包含了所有词的信息，便可以作为整个语音的语义向量表示，该语义向量称为源语言语音的上下文向量。

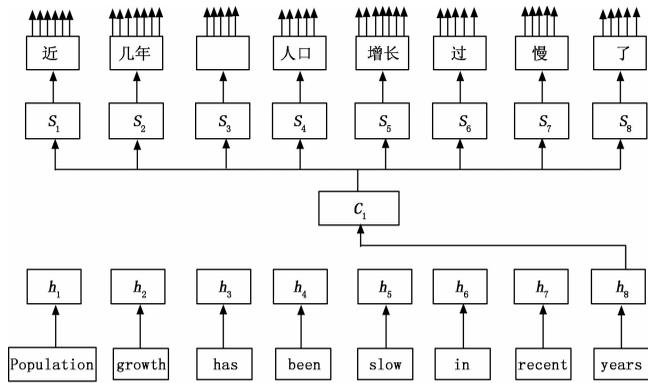


图 6 基于的循环神经网络翻译过程

给定源语言的上下文向量，HMM 与聚类的识别系统对该语音片段进行识别，过程如下：

- 1) 产生第一个隐含状态 S_1 ，并基于该隐含状态预测第一个目标语言词“近”。
- 2) 将第一个目标语言词“近”、隐含状态 S_1 和上下文向量 C_1 作为输入，此时产生第二个隐含状态 S_2 。隐含状态 S_2 包含了第一个词“近”的信息和源语言语音的信息，用来预测语言语音第二个词“几年”。
- 3) 以此循环类推，直到预测到一个语音的结束符为止。

5 结论

在英语语音识别方法和传统的 HMM 语音识别技术的基础上，针对英语语音识别系统的实现，提出了一种基于 HMM 和聚类的改进算法。实验结果表明，按照本文方法改进的系统，不仅提高了系统的识别率，而且减少了系统的计算量（即识别时间），从而达到提高系统性能的目的。但是如何确定聚类群，进一步提高识别效率，并应用于更大规模的英语语音识别中，还需要进一步的研究。

参考文献：

[1] 张震, 王化清. 语音信号特征提取中 Mel 倒谱系 MFCC 的改进算法 [J]. 计算机工程与应用, 2008, 44 (22): 54-55.

[2] 张智义. 英汉语情态动词实句的生成语法分析 [J]. 外语教学与研究, 2018, 50 (6): 65-74.

[3] 莫启扬, 文旭. 英语分词倒装构式的认知语法分析 [J]. 外语研究, 2018, 168 (2): 13-18.

[4] 贾晓立, 江晓波, 蒋三新, 等. 利用结构特征的语音压缩感知重建算法 [J]. 上海交通大学学报, 2017, 51 (9): 1111-1116.

[5] 李谈笑. 基于隐马尔可夫模型的英语单词词性智能分析判断研究 [J]. 通讯世界, 2019, 26 (4): 290-291.

[6] Zeinali H, Sameti H, Burget L. HMM-Based Phrase-Independent i-vector Extractor for Text-Dependent Speaker Verification [J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2017: 1-1.

[7] Mohammadiha N, Leijon A. Nonnegative HMM for Babble Noise Derived From Speech HMM: Application to Speech Enhancement [J]. IEEE Transactions on Audio, Speech, and Language Processing, 2013, 21 (5): 998-1011.

[8] 袁金仕. Viterbi 算法：分析与实现 [D]. 长沙：国防科学技术大学, 2004.

[9] 赵力, 邹采荣, 吴镇扬, 等. 基于 3 维空间 Viterbi 算法的汉语连续语音识别方法 [J]. 电子学报, 2000, 28 (7): 84-87.

[10] 袁俊. HMM 连续语音识别中 Viterbi 算法的优化及应用 [J]. 电子技术, 2001, 28 (2): 46-49.

[11] 张策, 韦鹏程, 石熙. 小语种库重庆话语音识别的研究 [J]. 计算机测量与控制, 2018, 26 (11): 252-255.

[12] Ai O C, Hariharan M, Yaacob S, et al. Classification of speech dysfluencies with MFCC and LPCC features [J]. Expert Systems with Applications, 2012, 39 (2): 2157-2165.

[13] 陈昌浩, 范太华. 改进的 HMM 模型在特征抽取上的应用 [J]. 计算机测量与控制, 2018, 26 (4): 217-220.

[14] 陈锦飞, 徐欣, CHENJinfei, 等. 基于梅尔频率倒谱系数与动态时间规整的安卓声纹解锁系统 [J]. 计算机工程, 2017, 43 (2): 201-205.

[15] 李凤. 非平衡时序数据的动态时间规整过采样方法研究 [D]. 广州：暨南大学, 2018.

[16] 沈泓, 刘顺. 基于 K-means 聚类算法的数据分析模型应用研究 [J]. 软件导刊, 2017, 16 (3): 103-107.

[17] 杨华晖, 孟晨, 王成, 等. 基于目标特征选择和去除的改进 K-means 聚类算法 [J]. 控制与决策, 2019, 34 (6): 101-108.

[18] 茅正冲, 王俊俊. 基于耳蜗倒谱系数和 Teager 能量算子相位融合的说话人识别系统 [J]. 南京理工大学学报, 2018, 42 (1): 82-88.

[5] 帅定新, 谢运祥, 王晓刚. 基于状态反馈精确线性化 Buck 变换器的最优控制 [J]. 中国电机工程学报, 2008 (33): 1-5.

[6] 刘彦呈, 古龙瑞, 张勤进. DC-DC 变换器新型趋近律滑模控制 [J]. 电力系统及其自动化学报, 2018, 30 (1): 64-68.

[7] 衣丽葵, 朱红兰, 李光彦. 一类不确定的 Buck 型直流斩波器的自适应逆推设计 [J]. 沈阳工程学院学报 (自然科学版), 2006 (2): 127-129.

[8] 冯均永, 史旺旺. CCM 模式下 Boost 电路的反步法非线性控

[2] 张智义. 英汉语情态动词实句的生成语法分析 [J]. 外语教学与研究, 2018, 50 (6): 65-74.

[3] 莫启扬, 文旭. 英语分词倒装构式的认知语法分析 [J]. 外语研究, 2018, 168 (2): 13-18.

[4] 贾晓立, 江晓波, 蒋三新, 等. 利用结构特征的语音压缩感知重建算法 [J]. 上海交通大学学报, 2017, 51 (9): 1111-1116.

[5] 李谈笑. 基于隐马尔可夫模型的英语单词词性智能分析判断研究 [J]. 通讯世界, 2019, 26 (4): 290-291.

[6] Zeinali H, Sameti H, Burget L. HMM-Based Phrase-Independent i-vector Extractor for Text-Dependent Speaker Verification [J]. IEEE/ACM Transactions on Audio, Speech, and Language Processing, 2017: 1-1.

[7] Mohammadiha N, Leijon A. Nonnegative HMM for Babble Noise Derived From Speech HMM: Application to Speech Enhancement [J]. IEEE Transactions on Audio, Speech, and Language Processing, 2013, 21 (5): 998-1011.

[8] 袁金仕. Viterbi 算法：分析与实现 [D]. 长沙：国防科学技术大学, 2004.

[9] 赵力, 邹采荣, 吴镇扬, 等. 基于 3 维空间 Viterbi 算法的汉语连续语音识别方法 [J]. 电子学报, 2000, 28 (7): 84-87.

[10] 袁俊. HMM 连续语音识别中 Viterbi 算法的优化及应用 [J]. 电子技术, 2001, 28 (2): 46-49.

[11] 张策, 韦鹏程, 石熙. 小语种库重庆话语音识别的研究 [J]. 计算机测量与控制, 2018, 26 (11): 252-255.

[12] Ai O C, Hariharan M, Yaacob S, et al. Classification of speech dysfluencies with MFCC and LPCC features [J]. Expert Systems with Applications, 2012, 39 (2): 2157-2165.

[13] 陈昌浩, 范太华. 改进的 HMM 模型在特征抽取上的应用 [J]. 计算机测量与控制, 2018, 26 (4): 217-220.

[14] 陈锦飞, 徐欣, CHENJinfei, 等. 基于梅尔频率倒谱系数与动态时间规整的安卓声纹解锁系统 [J]. 计算机工程, 2017, 43 (2): 201-205.

[15] 李凤. 非平衡时序数据的动态时间规整过采样方法研究 [D]. 广州：暨南大学, 2018.

[16] 沈泓, 刘顺. 基于 K-means 聚类算法的数据分析模型应用研究 [J]. 软件导刊, 2017, 16 (3): 103-107.

[17] 杨华晖, 孟晨, 王成, 等. 基于目标特征选择和去除的改进 K-means 聚类算法 [J]. 控制与决策, 2019, 34 (6): 101-108.

[18] 茅正冲, 王俊俊. 基于耳蜗倒谱系数和 Teager 能量算子相位融合的说话人识别系统 [J]. 南京理工大学学报, 2018, 42 (1): 82-88.

[9] 王攘攘, 张广明, 梅磊. 基于状态观测器的 DC-DC 升压变换器滑模控制 [J]. 电气传动, 2019, 49 (5): 55-58.

[10] 曹文思, 杨育霞. 基于状态空间平均法的 BOOST 变换器仿真分析 [J]. 系统仿真学报, 2007 (6): 1329-1330.

[11] 刘金琨. 机器人控制系统的设计与 MATLAB 仿真 [M]. 北京：清华大学出版社, 2016.

[12] 贾志东, 姜久春, 程龙, 等. 适用于 Boost 变换器的自适应模型预测控制算法 [J]. 中国电机工程学报, 2018, 38 (19): 5838-5845.

(上接第 102 页)