**《第四章 基于统计模型的语音识别技术》知识点汇编**

1. 语音识别任务的分类。（见课件）
2. 掌握DTW算法，（见课件），它更适用于特定人孤立（命令）词连续语音识别任务。
3. 对非特定人的语音识别任务，不同人群的命令词的声学表示序列有一个较大的动态变化范围，因而人们想到用统计模型来刻画这一动态变化。由于是对序列进行统计建模，所以采用了随机过程。
4. 马尔科夫链这种随机过程只能描述离散状态在时序上的随机性，而语音是声学特征矢量的序列，从而我们采用了两种统计模型的耦合：马尔科夫链和观察概率分布。这被称为隐马尔科夫模型。观察概率分布可以是高斯分布或混合高斯分布。
5. 由于是两个随机过程的耦合，训练和应用这个模型都比较困难，我们将其称之为HMM的三个基本问题。
6. 掌握计算问题的方法，掌握前向算法，能证明前向函数的递推公式。（见课件）
7. 掌握估计最佳状态链的有效方法，掌握viterbi算法。（见课件）
8. 掌握Baum-Welch算法。（见课件）
9. 掌握利用HMM解决孤立词识别问题的有效方法。
10. 掌握利用HMM模型解决数字串识别问题的有效方法。将数字HMM连接成识别网络，形成一个新的大HMM。通过viterbi算法在识别网络HMM模型上找到最佳状态链，作为识别结果。
11. 对于连续语音识别问题，由于词汇太多，没有足够的语料来训练每个词的HMM，因而选择在子词上建立HMM，然后用子词HMM拼接成词的HMM。由于相近词比较多，识别时特别容易混淆，所以引入了语言模型来约束识别结果，让其在句子层级更具合理性。这就是HMM的三层框架。（见课件）。
12. 由于Baum-Welch算法比较耗时，也容易收敛于局部极值点，因而需要用强制对齐方法给HMM选择一个较好的初值。
13. 强制对齐是依赖viterbi算法实现的。