



关于 HMM 相对可靠性量度

谢锦辉

(华中理工大学电子与信息工程系, 武汉 430074)

高雨青

(中科院自动化所, 北京 100080)

关键词: 语音识别, HMM, 相对可靠性量度.

一、引言

隐 Markov 模型 (Hidden Markov Models, HMM) 是近年来在语音处理领域中十分活跃的方法。但是, 由于 HMM 参数很多, 一般难以提供足够多的训练数据, 以得到良好的模型。在大词汇量语音识别时, 尤其如此。因此, 很有必要对两个或多个针对同一事件的表示不同程度的细节和鲁棒性的模型进行合并。这样, 确定每个 HMM 的相对可靠性, 以便给出合并时每个 HMM 相应的权值, 就是一个十分重要的问题。本文首先基于作者已做的语音识别中 HMM 训练算法工作^[1], 从著名的 HMM 训练算法——Baum-Welch 算法^[2]着手, 导出了一种 HMM 相对可靠性量度公式, 然后, 介绍了作者在法国 LIMSI-CNRS 的一个大词汇量连续语音识别系统上给出的相应实验验证。

二、HMM 相对可靠性量度公式

HMM 训练的 Baum-Welch 算法, 是利用重估公式, 逐步改进模型参数, 最后获得一个(局部)最佳值。以离散 HMM 为例, $A = \{a_{ij}\}$ 为状态转移概率矩阵, $B = \{b_{ik}\}$ 为观察值输出概率矩阵, 考虑有 L 个观察值序列用于训练, 用类似于文献[1]的方法, 重估公式可以写成

$$\bar{a}_{ij} = \frac{\text{从状态 } i \text{ 到 } j \text{ 的转移次数的期望值}}{\text{从状态 } i \text{ 转移出去次数的期望值}}$$
$$= \frac{\sum_{l=1}^L \text{第 } l \text{ 个训练序列从状态 } i \text{ 到 } j \text{ 的转移次数}}{\sum_{l=1}^L \text{第 } l \text{ 个训练序列状态 } i \text{ 的状态数}}$$

$$\bar{a}_{ij} \triangleq \frac{\sum_{l=1}^L \text{trans-counts}(i, j, l)}{\sum_{l=1}^L \text{state-counts}(i, l)}, \quad (1)$$

\bar{b}_{ik} = $\frac{\text{位于状态 } i \text{ 且观察到输出矢量 } k \text{ 的次数的期望值}}{\text{位于状态 } i \text{ 的次数的期望值}}$

$$= \frac{\sum_{l=1}^L \text{第 } l \text{ 个训练序列位于状态 } i \text{ 输出矢量为 } k \text{ 的矢量数}}{\sum_{l=1}^L \text{第 } l \text{ 个训练序列位于状态 } i \text{ 的状态数}}$$

$$\triangleq \frac{\sum_{l=1}^L \text{vect-counts}(k, j, l)}{\sum_{l=1}^L \text{state-counts}(j, l)}. \quad (2)$$

注意到

$$\text{state-counts}(j, l)|_{j=i} = \text{state-counts}(i, l), \quad (3)$$

那么,重估公式(1)和(2)可写成

$$\bar{a}_{ij} = \sum_{l=1}^L R_{il} \cdot \frac{\text{trans-counts}(i, j, l)}{\text{state-counts}(i, l)} = \sum_{l=1}^L R_{il} \cdot \bar{a}_{ijl}, \quad (4)$$

$$\bar{b}_{ik} = \sum_{l=1}^L R_{jl} \cdot \frac{\text{vect-counts}(k, j, l)}{\text{state-counts}(j, l)} = \sum_{l=1}^L R_{jl} \cdot \bar{b}_{ikl}, \quad (5)$$

这里,

$$R_{jl} = \frac{\text{state-counts}(j, l)}{\sum_{l'=1}^L \text{state-counts}(j, l')}. \quad (6)$$

分析(4)和(5)式,可以得到结论:当用 L 个训练序列获取 HMM 参数时,在每次迭代时,可以分别用每个训练序列获取 HMM 参数,再加以合并,而且,合并的权值仅仅取决于状态数。这样,可以认为,正是状态数描述了 HMM 相对可靠程度。因此,当需要合并 L 个 HMM 时,对于任一状态 j ,合并时的权值可由(6)式确定。这种 HMM 相对可靠性量度公式,具有两个特点:1)它是由 Baum-Welch 算法中重估公式导出,因而在最大似然意义上是(局部)最佳的。2)它不同于一般权值选取方法,不是合并时每个 HMM 选取一个权值,而是在一个 HMM 中,每个状态对应一个权值,因此具有更好的性能。当然,训练 HMM 时必须保存每个状态的状态数,以备合并时使用。此外,这样选取的权值,必须要求训练数据足够多,使权值在大数定理的基础上反映 HMM 真实的可靠程度。

三、实验与讨论

实验是在法国 LIMSI-CNRS 进行的。实验系统是一个基于离散 HMM, 使用美国 DARPA 语音数据库的连续语音识别系统。限于篇幅, 这里只列出一组实验数据, 更详尽的实验报告见 LIMSI-CNRS 内部资料^[3]。实验数据如表 1 所示。在表 1 中, 测试数据为 DARPA-RM1 的前 40 个测试句子, 训练数据为 DARPA-RM1 的全部 600 个训练句子。系统 A 为用 Baum-Welch 算法训练的原始实验系统(使用 600 个训练句子); 系统 B 为用前 300 句子和后 300 个句子分别训练两个原始实验系统, 再用本文导出的公式确定权值进行合并而得到的系统; 系统 C 为对同上两个原始实验系统, 人为选取权值为 0.5 进行合并而得到的系统。

表 1 连续语音识别字识别性能(百分比值)

	系统	正确	替代错误	消去错误	插入错误
利用词对 (word-pair) 句法信息	A	82.1	15.7	2.2	6.6
	B	82.1	15.4	2.5	7.2
	C	81.8	15.7	2.5	7.2
不利用任 何句法信息	A	46.7	50.8	2.5	24.1
	B	48.0	48.9	3.1	19.7
	C	46.1	51.1	2.8	20.1

实验表明, 用本文导出的相对可靠性量度公式确定的权值去合并两个 HMM 明显优于人为选取权值去合并 HMM。但由于训练数据量的限制, 受大数定理影响, 并不是所有实验结果与前面理论分析一致。

在语音处理中, 尤其是大词汇量(连续)语音识别时, 常常会面临因训练数据不足而必须合并两个或多个 HMM 的情形^[4]。例如, 通常一些出现次数很少的输出矢量没有包含在整个训练数据中, 这样训练出的 HMM 就会有不少为零的概率参数值。而事实上, 在实际语音识别测试时, 这些矢量又可能出现, 因而需要对训练好的 HMM 进行平滑处理。而一般为了避免过分平滑, 总是要合并平滑前后的两个 HMM, 以获取较好的结果模型。因此, 本文导出的 HMM 相对可靠性量度公式是很有实用价值的。

本文工作虽是针对离散 HMM 的, 但结论可推广到任何其它类型的 HMM。

致谢. 感谢法国 LIMSI-CNRS 的 Groupé Communication Parlée 的同事们对作者提供的帮助, 尤其是 Martine ADDA 博士对本文实验工作给予的支持。

参 考 文 献

- [1] Gao Yu-Qing and Xie Jin-Hui, A Model Block-Training Method for HMM Based Speech Recognition systems, Proc. ICASSP'90, 541—544.
- [2] Rabiner L.R., A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition, Proc. of IEEE, 77(1989), (2), 257—285.

- [3] Xie Jin-Hui, Etude d'une Méthode de Lissage par RRM (Relative Reliability Measure) de Modèles Markoviens Cachés et Son Application à la Parole Continue, Notes et Documents LIMSI, No. 91-16, Dec. 1991.
- [4] Lee K. F. and Hon H. W., Speaker-Independent Phone Recognition Using Hidden Markov Models, *IEEE Trans. ASSP-37* (1989), (11), 1641-1648.

ON AN HMM RELATIVE RELIABILITY MEASURE

XIE JINHUI

(Huazhong University of Science and Technology, Wuhan 430074)

GAO YUQING

(Institute of Automation, Academia Sinica, Beijing 100080)

Key words: Hidden Markov Models; Speech Recognition; Relative Reliability Measure.

(上接封三)

- On the Reasoning Method of Syntactic Knowledge System
 Yang Guangzheng (628)
- A Study of H^∞ State Feedback Control and Its Optimal Solution
 Yang Fuwen, Tu Jian (633)

Research Notes

- Neuron Intelligen Control For Electroslag Remelting Processes
 Wang Ning, Tu Jian, Chen Jinjiang (636)
- On an HMM Relative Reliability Measure Xie Jinhui, Gao Yuqing (640)

本期责任编辑 武海波