一种新的关键词确认方法

戴海生[1],朱小燕[2],罗予频[1],杨士元[1]
(1. 清华大学自动化系,北京 100084;2. 清华大学计算机系,北京 100084)

摘 要: 本文提出了一种新的基于模型距离矩阵的关键词确认算法,并给出模型距离的定义及其训练方法,利用模型相对距离矩阵对语音识别结果进行确认. 对于关键词库较大的关键词检出系统,通过对关键词分段,得到扩展的模型距离矩阵确认算法,使得大词表确认问题得到很好的解决,并能够获得和小词表系统一样的确认效果. 为了对关键词库进行方便的操作,模型距离矩阵的更新算法使得用户可以很方便地修改关键词库内的关键词,而不必重新训练整个模型距离矩阵.

关键词: 隐马尔可夫模型; 最大似然准则; 语音确认; 关键词检出
中图分类号: TN912. 34 **文献标识码**: A **文章编号**: 0372-2112 (2005) 01-0101-05

A Novel Keyword Verification Algorithm

DAI Hai-sheng[1], ZHU Xiao-yan[2], LUO Yu-pin[1], YANG Shi-yuan[1]
(1. *Dept of Automation, Tsinghua University, Beijing* 100084, *China*;
2. *Dept of Computer Science & Technology, Tsinghua University, Beijing* 100084, *China*)

Abstract: This paper proposes a novel utterance verification (UV) algorithm based on model-distance matrix. It gives the definition of model-distance and training algorithm for model-distance matrix which is used to verify the recognition result. By using expanded model-distance matrix verification algorithm, it makes impossible to verify the result of a big keyword spotting system, and gets the same performance as on a small one. An updating algorithm to modify model-distance matrix is also introduced when vocabulary is modified.

Key words: hidden Markov model (HMM); maximum likelihood (ML); utterance verification (UV); keyword spotting

1 引言

近年来,随着自动语音识别技术的飞速发展及广泛的应用,语音确认技术越来越被重视并取得了相当不错的效果. 现有的语音确认方法可分为两类:一类是基于二次识别的语音确认方法. 一类是基于 Viterbi 解码识别结果的确认方法. 基于二次识别的语音确认方法就是要进行两次 Viterbi 解码过程才能得到最终的确认结果,包括 Chuang-Hsein 等人[1]提出的基于模糊搜索策略的关键词确认算法;以及 Benítez 等人[2]通过对可能的候选关键词对应的语音分别进行传统的 Viterbi 解码和加入关键词驻留信息的 Viterbi 解码的算法. 由于 Viterbi 解码的计算量非常大,使得基于二次识别的语音确认方法很难满足系统的实时性的要求.

基于 Viterbi 解码识别结果的确认方法又分为一阶段语音确认方法和两阶段语音确认方法. Lleida 等人[3]提出的基于似然比(Likelihood Ratio, LR)准则的一阶段确认方法对于语音识别和确认都取得了不错的效果,但是该方法存在两个很大的缺点:一是很大程度上依赖于语言模型;一是限制词表外的"非关键词"的范围且数目不能太多. 基于隐马尔可夫模型(HMM)最大似然(ML)准则的两阶段语音确认方法在关键词检出中得到了更为广泛的应用[4~12]. 关键词检出系统引入离线垃圾(Garbage)模型[7~8,10~11]来增强关键词检出和词库内外词的区分能力,离线垃圾模型的设计和训练难度较大,当关键词表改变时离线垃圾模型需要重新训练. 刘俊等[12]提出了动态评价方法与反关键词模型结合得到垃圾模型的似然得分,离线垃圾模型的再训练. 但是,该方法随着关键词表中的关键词数的增加,系统的性能有所下降.

本文提出的基于模型距离矩阵的语音确认方法,是利用关键词表中关键词的声学模型的相似性对关键词进行确认的. 由于没有利用非关键词的信息,只需要训练模型距离矩阵,而不训练垃圾模型,这就使得本算法与关键词表相对独立,比较容易修改词表. 模型距离训练简单且训练量小,且对于词表外的非关键词和词库内的关键词以及词库大小不加以限制. 利用扩展的模型距离矩阵确认算法可以对大型关键词检出系统进行确认,很好地解决了关键词数增大系统性能下降的问题. 由于模型距离矩阵并不直接依赖于语言模型,所以

收稿日期:2002-10-12;修回日期:2004-11-10
基金项目:自然科学基金(No. 60272019)

本方法对于没有语言模型的孤立词系统同样可以得到很好的识别效果. 从参考文献[1～3,13,14]中可以看到,当前大型的关键词检出系统在错误拒绝率(False Rejection,FR)为 15 %左右时,错误接受率(False Acceptance,FA)大约也在 15 %左右,本文提出的基于模型距离矩阵的确认算法在错误拒绝率为 6.7 %时,错误接受率为 6.1 %,具有良好的确认性能.

本文的剩余部分是如下安排的:第二部分介绍基于模型距离矩阵的确认算法. 第三部分介绍关键词库修改时模型距离矩阵的更新算法;第四部分是实验和结果分析;第五部分是结论.

2 基于模型距离矩阵的语音确认算法

本部分中我们先给出模型距离和距离矩阵的定义和性质,以及距离矩阵的训练算法,然后给出基于模型距离矩阵的确认准则,以及用于大词表系统的扩展模型距离矩阵的语音确认算法.

2.1 模型距离的定义及性质

在关键词检出系统中,对于关键词模型 M_i 的一个训练语料 O_i,其相应于模型 M_j 的概率输出为 $P(O_i|M_j)$,$P(O_i|M_j)$ 越大,说明模型 M_j 和模型 M_i 越相似,或者称为模型之间的距离越小. 由于很难直接描述模型之间的这种距离特性,我们给出模型相对距离的定义,并把它简称为模型距离:对于关键词检出系统中的 N 个关键词模型 $M_1, M_2, \cdots\cdots, M_N$,其中模型 M_i 存在一个理想的训练语料 O_i,对任一模型 M_j 相应的概率输出值为 $P(O_i|M_j)$,则模型 M_j 对于模型 M_i 的相对距离就是

$$d(M_j, M_i) = \sum_{k=1}^{N} \mathrm{sign}(p(O_i|M_k), p(O_i|M_j)) \quad (1)$$

其中:

$$\mathrm{sign}(x, y) = \begin{cases} 1, & x > y \\ 0, & x \le y \end{cases}$$

从模型相对距离的定义可以看出:如果把关键词表中的关键词模型都和模型 M_i 的理想语料 O_i 进行基于似然准则的概率计算,并把计算结果按大小进行排序,其输出值越大的序越小,模型 M_j 在排序的中的序位正好等于 $d(M_j, M_i)$ 的大小. 很显然,模型距离具有以下几个性质:

(1) $d(M_j, M_i) \in \{0, 1, \cdots\cdots, N-1\}$;

(2) $d(M_i, M_i) = 0$;

(3) $\sum_{i=1}^{N} d(M_j, M_i) = \frac{N(N-1)}{2}$

2.2 模型距离矩阵

设关键词检出系统中的关键词数为 N,定义模型距离矩阵 $A_{NN} = [a_{ij}]_{NN}$,其中:$a_{ij} = E[d(M_j, M_i)]$($E[\cdot]$是求数学期望). 我们可以通过训练得到(a_{ij} $(1 \le i \le N, 1 \le j \le N)$ 的估计值 $\hat{a}_{ij}$. 对于关键词模型 M_i,共存在独立的 T 个相应的训练数据 $O_i(1), \cdots\cdots, O_i(T)$. 对于其中任一个训练数据 $O_i(t)$,由公式(1)中模型相对距离的计算公式可得:

$$\hat{d}_t(M_j, M_i) = \sum_{k=1}^{N} \mathrm{sign}(p(O_i(t)|M_k), p(O_i(t)|M_j)) \quad (2)$$

取其均值作为 $\hat{a}_{ij}$:$\hat{a}_{ij} = \frac{1}{T}\sum_{t=1}^{T} \hat{d}_t(M_j, M_i)$,也就是说:$a_{ij} = \frac{1}{T}\sum_{t=1}^{T} \hat{d}_t(M_j, M_i)$

由模型距离的性质式(3)知模型距离矩阵具有如下性质:

$$\sum_{j=1}^{N} a_{ij} = \frac{N(N-1)}{2}, (1 \le i \le N) \quad (3)$$

2.3 确认准则

关键词检出系统在进行关键词检出时,对于可能的关键词候选模型 $\hat{M}_i$(为了区别其与真正的关键词 M_i). $\hat{M}_i$ 在语音信号中相应的部分为 O_i,类似于模型相对距离的计算公式,模型 M_i 对应于 O_i 的似然输出值 $p_j = P[O_i|M_j]$,同样可以求出每个模型 M_j 与 $\hat{M}_i$ 的相对距离 $d(M_j, \hat{M}_i)$,我们可以通过比较每个 $d(M_j, \hat{M}_i)$ 与理想的 $d(M_j, M_i)$ 之间的相似程度(即选择一个合适的拒识量),来判断该候选词模型 $\hat{M}_i$ 是否为真正的关键词模型 M_i. 图 1 给出模型距离矩阵确认算法框图.

图 1 模型距离矩阵确认算法

为了达到对正确的关键词进行确认以及拒绝错误的语音输入,我们采用下式作为我们的拒识量:

$$\delta_i = \sum_{j=1}^{N} |a_{ij} - d(M_j, \hat{M}_i)| \quad (4)$$

当 δ_i 小于设定的阈值时,就接受识别结果;否则拒绝之. 由于 δ_i 会随着 N 的增大而相应的增大,这对于阈值的选择带来很大的困难. 因此,我们对其进行正规化处理:对于连续的 N 个整数序列$\{0, 1, 2, \cdots\cdots, N-1\}$,$d(M_j, \hat{M}_i)$ 是它的一个排列,a_{ij} 在理想的情况下,也是它的一个排列,此时,由附录可知:$\sum_{j=1}^{N} |a_{ij} - d((M_j, \hat{M}_i)| \le \frac{N^2}{2}$,由此,我们可以给出正规化处理以后的计算公式:

$$\delta_i = \frac{2}{N^2}\sum_{j=1}^{N} |a_{ij} - d(M_j, \hat{M}_i)| \quad (5)$$

显然有 $0 \le \delta_i \le 1$,从而达到正规化的目的. 图 4 中实验说明在没有正规化的时候,确认阈值的选取大致和关键词数的平方成正比,从而也说明我们选择 $\frac{2}{N^2}$ 作为系数进行正规化是合理的.

2.4 扩展的模型距离矩阵确认算法

模型距离矩阵确认算法对于小型关键词检出系统(关键词总数 < 50)来说具有比较不错的确认效果. 但是,对于一个大型的关键词检出系统来说,把所有的关键词作为一类计算模型相对距离不仅训练量大,而且训练的结果不够准确. 因此我们提出扩展的模型距离矩阵的确认算法.

扩展的模型距离矩阵的主要思想就是:把系统词库中总的关键词数分为 m 段,使得每段中的关键词数分别为 N_k $(1 \le k \le m)$ 个,每一段相当于一个独立的小型关键词检出系统,可以训练出一个独立的模型距离矩阵 $A(k) = [a_{ij}(k)]$. 对于

由关键词检出系统得到的一个在第 k 段中的候选关键词 $M_i(k)$,就可以利用模型距离矩阵 $A(k)$ 对 $M_i(k)$ 应用基本的模型距离矩阵确认算法进行确认.

扩展的模型距离矩阵的确认算法具体步骤如下:

(1) 对于一段输入语音信号,通过关键词检出系统得到候选关键词 $M_i(k)$ 及其对应的语音信号段 O_i,以及该候选关键词所在段相应的模型距离矩阵 $A(k)$,设该段中的关键词总数为 N_k.

(2) 利用关键词检出系统中的 Viterbi 解码过程得到 $p_j = P[O_i|M_j(k)]$,以及该段内每个模型 $M_j(k)$ 与 $M_i(k)$ 的相对距离 $d(M_j(k), M_i(k))$

(3) 利用模型距离矩阵 $A(k)$ 以及基本的模型距离矩阵算法对候选结果进行确认. 其中,对应的确认量计算公式为:

$$\delta_i(k) = \frac{2}{N_k^2}\sum_{j=1}^{N_K} |a_{ij}(k) - d(M_j(k), M_i(k))|$$

图 2 给出扩展的模型距离矩阵算法框图:

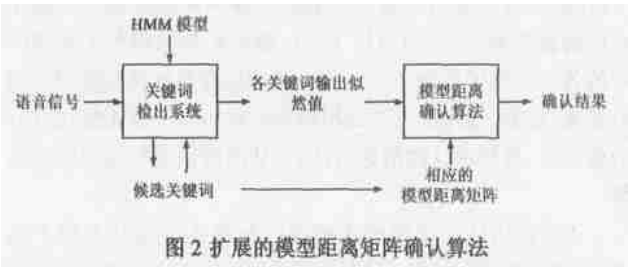

图 2 扩展的模型距离矩阵确认算法

3 模型距离矩阵的更新算法

一个实用的关键词检出系统应当允许用户自由地修改关键词库. 当关键词库发生变化时,我们可以利用模型距离矩阵的更新算法来得到新的模型距离矩阵,而不需要对整个模型距离矩阵重新训练,大大地节约了训练量. 由于扩展的模型距离矩阵确认算法可以视为由 m 个独立的模型距离矩阵组成的,因此,我们给出基于基本的模型距离矩阵的更新算法,扩展的模型距离矩阵的更新算法类似可得.

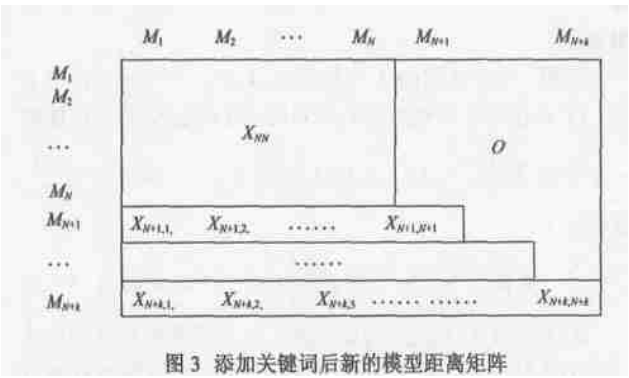

图 3 添加关键词后新的模型距离矩阵

3.1 添加关键词

设原距离矩阵为 $A_{NN} = [a_{ij}]_{NN}$,新的距离矩阵为 $\bar{A}_{N+1,N+1} = [\bar{a}_{ij}]_{N+1,N+1}$. 我们并不需要得到新的距离矩阵的所有数据,因为当识别结果为原来词表中的关键词时,我们可以忽略新添加的关键词的影响,并利用原距离矩阵进行判别;只有当候选关键词为新添加的关键词时,我们才需要新的距离矩阵进行判别,因此,我们只需要通过训练得到 $a_{N+1,k}(1 \le k \le N+1)$ 就可以了.

3.2 删除关键词

删除该关键词时,我们需要相应的更新模型距离矩阵. 设关键词表中的关键词数为 N,模型距离矩阵为 $A_{NN} = [a_{ij}]_{NN}$,我们要删除其中的关键词模型 M_j,得到新的距离矩阵为 $\bar{A}_{N-1,N-1} = [\bar{a}_{ij}]_{N-1,N-1}$,且:$\bar{a}_{tk} = \zeta_t y_{tk}$,其中 $\zeta_t = \frac{(N-1)(N-2)}{2\sum_{k=1}^{N-1} y_{tk}}$,$Y_{N-1,N-1} = [y_{ij}]_{N-1,N-1}$ 为引入的临时矩阵:

$$y_{tk} = \begin{cases} a_{tk} - sign(a_{tk}, a_{ij}) & , (t<j, k<j) \\ a_{t,k+1} - sign(a_{t,k+1} a_{tj} & , (t<j, k \ge j) \\ a_{t+1,k} - sign(a_{t+1,k}, a_{t+1,j}) & , t \ge j, k<j) \\ a_{t+1,k+1} - sign(a_{t+1,k+1}, a_{t+1,j}) & , t \ge j, k \ge j) \end{cases}$$

参数 ζ_t 的引入用来保证公式(3)得以满足:$\sum_{k=1}^{N-1}\bar{a}_{tk} = \sum_{k=1}^{N-1}(\zeta_t y_{tk}) = \zeta_t \sum_{k=1}^{N-1} y_{tk} = \frac{(N-1)(N-1-1)}{2}$ 从而避免了连续删除多个关键词时积累误差的引入.

4 实验及结果分析

本论文采用清华大学智能技术与系统国家重点实验室采制的 CIDS 语音数据库. 数据库是采用 11025Hz 采样率、16 位和单声道的语音信号. 全部数据取自自然发音,实验室环境,总共有 200 个关键词,说话者共 10 人(女音 2 人,男音 8 人),每个人的语音中都包括这 200 个关键词,而且每个关键词分别出现 10 次. 对上面的数据,以音节为基本识别单元,采用 $(1-0.95z^{-1})$ 预先加重,我们以帧长 256 点,帧移 128 点进行了 39 维的 Mel 谱计算,在连续语音识别的基础上,进行关键词检出与确认的实验. 通过错误拒绝率(FR)和错误接受率(FA)可以评价语音确认系统的性能,选择适当的确认阈值可以做出折衷的决定. 确认阈值由实验确定.

4.1 阈值确定

图 4 给出了采用未进行正规化处理的公式(4),对应于不同关键词数时,基于 FA + FR 最小的原则对应的确认阈值,从图中可以看到确认阈值大约和关键词数 N 的平方成正比($Threshold = N^2/10$). 所以在采用正规化后的公式(5)时,确认阈值为常数:

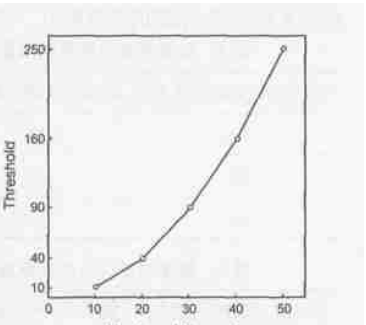

图 4 对应不同关键词数的确认阈值

$$Threshold = \left(\frac{N^2}{10}\right) \Big/ \left(\frac{N^2}{2}\right) = 0.2$$

从而也说明我们采用的正规化方法是比较合理的. 在下面的实验中,我们使用公式(5)作为确认准则,确认阈值就取 0.2.

4.2 模型距离矩阵确认实验

表 1 给出在模型距离矩阵确认算法中,不同关键词数对应的 FA,FR,FR + FA. 从中可以看到:关键词数在 40 左右时,可以得到相对较好的确认效果,FR + FA 约为 13 %左右.

表 1 不同关键词数相应的 FA,FR,FR + FA

N	10	15	20	25	30	35	40	45	50	100
FR(%)	5	6.7	6.7	8.0	8.3	7.1	5.4	7.4	9.0	7.2
FA(%)	14.7	10.0	9.5	9.4	8.1	8.0	7.5	7.2	6.6	10.1
FA + FR(%)	19.7	16.7	16.2	17.4	16.4	15.1	12.9	14.6	15.6	18.3

4.3 扩展的模型距离矩阵确认实验

表 2 给出在词表中的关键数较多时,应用基本的模型距离矩阵确认算法和扩展的模型距离矩阵确认算法的确认效果的比较,词表中总的关键词数为 200 个,在扩展的模型距离矩阵确认算法中共分成 5 段,每一段为 40 个关键词. 可以看到应用基本的确认算法时确认效果很差,而扩展的确认算法的确认效果和关键词数为 40 时确认效果相似,从而使得大词表的确认问题得到解决.

表 2 FR,FA,FA + FR 在应用基本以及扩展的距离矩阵确认算法时的比较

	基本的模型距离矩阵确认算法	扩展的模型距离矩阵确认算法
FR(%)	13.7	6.1
FA(%)	15.6	6.7
FR + FA(%)	29.3	12.8

4.4 修改词表时使用更新算法对应的 FA + FR

表 3 和表 4 分别给出了利用更新算法得到模型距离矩阵相应的确认效果. 从表中可以看到,和直接训练的模型距离矩阵进行识别相比,添加关键词时的识别效果几乎没什么变化;删除关键词时,识别的效果略有下降,但是下降的不是很多.

表 3 添加关键词使用递推算法对应的 FA + FR

原词表中关键词数 N_1	添加后关键词数 N_2	FR + FA(%)
20	30	16.1
30	40	12.8
40	50	15.4

表 4 删除关键词使用递推算法对应的 FA + FR

原词表中关键词数 N_1	删除后关键词数 N_2	FR + FA(%)
30	20	17.8
40	30	15.7
50	40	13.2

通过以上的实验,我们可以看到:尽管模型距离确认算法对于关键词数在 30 以上时取得了相当不错的确认效果,但是当关键词数比较少时,确认的效果不是很好,此时可以采用添加反关键词模型以及一些常见的噪音模型使得关键词模型的总数达到 40 个左右. 还有,在添加关键词时,需要该关键词相应的 $T(T>3)$ 个训练语料;删除关键词时,模型距离矩阵的更新算法一定程度上损失了模型距离的精确性,此时确认效果会有所下降,尤其是当删除的关键词数超过原词库关键词总数的 1/3 时,这种现象尤其明显,而采用全部重新训练模型距离矩阵的方法使得训练量加大,给用户带来不便. 因此,我们设想将来可以采用基于模型参数的模型距离算法,使得模型距离的计算不依赖于词表中的模型,因此就不必要重新训练模型距离矩阵. 此外,由于没有对发音相近的词进行特殊的处理,当非关键词与关键词很相近时,确认的效果也会有所下降. 所以如何获得更好的确认效果,还需要进一步的研究.

5 结论

本文提出了一种新的基于模型距离矩阵的语音确认方法. 采用了两步骤策略,在关键词检出的基础上进行确认. 第一阶段,通过传统的关键词检出算法得到候选关键词以及相应的语音数据段,从而通过 Viterbi 解码的算法得到关键词库中的各个关键词模型的似然值以及该模型与候选关键词的相对距离;在第二阶段,通过调用模型距离矩阵计算正规化以后的确认量,并和确认阈值进行比较,从而确认或拒绝该识别结果.

本文提出的基于模型距离矩阵的确认方法具有如下优点:首先是直接利用基于 HMM 模型 ML 准则的语音识别器的输出结果中各个关键词模型的似然值大小,因而附加计算量小且确认过程简单;扩展的模型距离矩阵算法使得具有大型关键词库的系统可以转化为小型的关键词库相应的情况,并取得相同的确认效果. 其次,在关键词库被修改时,可以利用模型距离矩阵的更新算法得到新的模型相对距离,而不必全部重新训练. 再次,不对关键词库外的非关键词进行训练,因此非关键词的范围不受任何限制. 而由于确认过程中没有利用语言模型的信息,因此本确认算法不依赖于语言模型的好坏.

附录

证明 对于连续的 N 个整数 $\{k, k+1, \cdots\cdots, n\}\ (n = k + N - 1)$,存在两个 N 维数列 $\{a_i\}$ 和 $\{b_i\}$ 分别是这 N 个整数的一个排列,则有:

$$\sum_{i=1}^{N} |a_i - b_i| \leqslant \sum_{i=0}^{N-1} |(n-i) - (k+i)| \leqslant \frac{N^2}{2}$$

证明如下:

(一)先证明

$$\sum_{i=1}^{N} |a_i - b_i| \leqslant \sum_{i=0}^{N-1} |(n-i) - (k+i)| \quad (*1)$$

对于 $(*1)$ 的右边其实是对这 N 个连续整数的分别为从大到小排列和从小到大排列,然后两两作差,再取绝对值的和. 只要我们可以证明左式要达到最大,必然是 a_i 和 b_i 中的 n 和 k 分别作差,则余下的数列为 $N-2$ 个连续整数的排列,故可以利用第二数学归纳法证明之.

(1) $N = 1$ 时,显然有 $\sum_{i=1}^{N} |a_i - b_i| = 0 = \sum_{i=0}^{N-1} |(n-i) - (k+i)|$,故:$(*1)$ 式成立.

(2) 设对于任意的 $N < M$, (*1) 式都成立.

(3) 对于连续的 M 个整数序列,$\{k, k+1, \cdots\cdots, n\}$,且 M 维数组 a_i 和 b_i 分别是这 M 个整数的一个排列,不妨设:$a_{i1} = k$, $a_{i2} = n$, $b_{j1} = k$, $b_{j2} = n$,则:

$$\begin{aligned}
&|a_{i1} - b_{i1}| + |a_{i2} - b_{i2}| + |a_{j1} - b_{j1}| + |a_{j2} - b_{j2}| \\
&= (b_{i1} - k) + (n - b_{i2}) + (a_{j1} - k) + (n - a_{j2}) \\
&= (n - k) + (n - k) + (a_{j1} - b_{i2}) + (b_{i1} - a_{j2}) \\
&\leq |n - k| + |k - n| + |a_{j1} - b_{i2}| + |a_{j2} - b_{i1}| + \sum_{i=1}^{N} |a_i - b_i| \\
&\leq \sum_{1 \leq i \leq N}^{i \neq i1, i2, j1, j2} |a_i - b_i| + |n - k| + |k - n| + |a_{j1} - b_{j2}| \\
&\quad + |a_{j2} - b_{i1}| \\
&= \left\{ \sum_{1 \leq i \leq N}^{i \neq i1, i2, j1, j2} |a_i - b_i| + |a_{j1} - b_{i2}| + |a_{j2} - b_{i1}| \right\} \\
&\quad + |n - k| + |k - n|
\end{aligned}$$

显然有:$\left\{ \sum_{1 \leq i \leq N}^{i \neq i1, i2, j1, j2} |a_i - b_i| + |a_{j1} - b_{i2}| + |a_{j2} - b_{i1}| \right\}$ 分别为 a_i 和 b_i 中抽出 k 和 n 后(也就是新的 $N-2$ 个连续整数序列$\{k+1, \cdots\cdots, n-1\}$的两个排列)两两作差,然后取绝对值的和.所以它应当满足归纳假设:

$$\left\{ \sum_{1 \leq i \leq N}^{i \neq i1, i2, j1, j2} |a_i - b_i| + |a_{j1} - b_{i2}| + |a_{j2} - b_{i1}| \right\} \leq \sum_{i=0}^{N-3} |(n - 1 - i) - (k + 1 + i)|$$

故:$\sum_{i=1}^{N} |a_i - b_i| \leq \sum_{i=0}^{N-3} |(n - 1 - i) - (k + 1 + i)| + |n - k| + |k - n| = \sum_{i=0}^{N-1} |(n - i) - (k + i)|$

综上可知:(*1) 成立,证明完毕.

(二) 再证明 $\sum_{i=0}^{N-1} |(n - i) - (k + i)| \leq \frac{N^2}{2}$

(1) N 为偶数时 ($N > 0$),

$$\sum_{i=0}^{N-1} |(n - i) - (k + i)| = 2 \sum_{i=0}^{\frac{N}{2}-1} (2i + 1) = \frac{N^2}{2}$$

(2) N 为奇数时 ($N > 0$),

$$\sum_{i=0}^{N-1} |(n - i) - (k + i)| = 2 \sum_{i=0}^{\frac{N-1}{2}} (2i) = \frac{N^2 - 1}{2}$$

故:任意的整数 $N > 0$,有 $\sum_{i=0}^{N-1} |(n - i) - (k + i)| \leq \frac{N^2}{2}$,当 N 为偶数时,等号可达到;当 N 为奇数时,但 N 较大时,$\frac{N^2 - 1}{2}$ 趋近于 $\frac{N^2}{2}$.

参考文献:

[1] Chung-Hsien Wu, Yeou-Jiunn Chen. Multi-keyword spotting of telephone speech using a fuzzy search and keyword-driven two-level CBSM [J]. Speech Communication, 2001, 33: 197 - 212.

[2] M C Benítez, A Rubio, P García, A de la Torre. Different confidence measures for word verification in speech recognition [J]. Speech Communication, 2000, 32: 79 - 94.

[3] E Lleida, R C Rose. Utterance verification in continuous speech recognition: decoding and training procedures [J]. IEEE Trans on Speech and Audio Processing, 2000, Vo8(2): 126 - 139.

[4] J Rohlicek, W Russel, S Roukos, H Gish. Continuous hidden Markov modeling for speaker independent word spotting [A]. in Proc Int Conf Acoustics, Speech and Signal Processing [C]. 1989. 627 - 630.

[5] W Chou, B-H Juang, C-H Lee. Segmental GPD training of HMM based speech recognizer [A]. in Proc Int Conf On Acoustics, Speech, and Signal Processing [C]. 1992. 473 - 476.

[6] R Sukkar. Rejection for connected digit recognition based on GPD semental discrimination [A]. in Proc Int Conf Acoust, Speech and Signal Processing [C]. 1994. 393 - 396.

[7] Mazin G Rahim, Chin-Hui Lee and Biing-Hwang Juang. Dicriminative utterance verification for connected digits recognition [J]. IEEE trans on speech and audio processing [J]. 1997, 5(3): 266 - 277.

[8] J Wilpon, L Rabiner, C-H Lee, E Goldman. Automatic recognition of keywords in unconstrained speech using hidden Markov models [J]. IEEE Trans. Acoustics, Speech, Signal Processing [J]. 38(11): 1870 - 1990.

[9] R C Rose, B H Juang, C H Lee. A training procedure for verifying string hypothesis in continuous speech recognition [A]. ICASSP-95 Detroit [C]. 1995. 281 - 284.

[10] C de la Torre, L Hernández-Gómez, F J Caminero-Gil, C Martín del Álamo. On-line Garbage Modeling for Word and Utterance Verification in Natural Numbers Recognition [A]. ICASSP-96, Atlanta [C]. 1996. 845 - 848.

[11] H Bourlard, B D'hoore, J-M Boite. Optimizing recognition and rejection performance in wordspotting systems [A]. ICASSP-1994 [C]. 1994. 373 - 376.

[12] 刘俊,朱小燕. 基于动态垃圾评价的语音确认方法 [J]. 计算机学报, 2001, 24(5): 480 - 486.

[13] R Paredes, A Sanchis, E Vidal, A Juan. Utterance verification using an optimized k-nearest neighbour classifier [A]. EUROSPEECH-2003, Geneva [C]. 2003.

[14] A Sanchis, A Juan, E Vidal. Improving utterance verification using a smoothed naive bayes model [A]. ICASSP-2003 [C]. 2003.

作者简介:

戴海生　男,1976 年 2 月出生于江苏省邳州市,1994 年进入清华大学自动化系自动化专业学习,1999 年获工学学士学位并升入清华大学自动化系攻读工学博士学位,现为清华大学自动化系检测技术与自动化装置研究所博士研究生,研究兴趣为语音识别、关键词检出、语音确认、噪声去除等. E-mail: dai@mail.au.Tsinghua.edu.cn.

朱小燕　女,1982 年获北京科技大学学士学位,1987 年获日本神户大学硕士学位,1990 年获日本名古屋工业大学博士学位,1993 年到清华大学任教,现为清华计算机系教授,博士生导师. 主要研究领域为智能信息处理,其中包括:模式识别、神经元网络、机器学习、自然语言处理、信息提取等.