

限定性手写汉字识别的一种 可变形弹性匹配模型*

Constrained Handwritten Chinese Character Recognition Using a Deformable Elastic Matching Model

金连文 徐秉铮

(华南理工大学无线电所, 广州 510641)**

【提要】 本文提出了一种应用于手写体汉字识别的可变形弹性匹配模型. 在该模型中, 我们假定: 对于同一种汉字的不同的手写样本, 尽管由于书写风格的不同而使其形状各异, 但它们却具有相同的空间拓扑结构, 我们将用一组笔段位移矢量来描述不同的手写体之间的细节形状上的差异. 在位移矢量场的作用下, 一个可变形的模板汉字将逐步变形以趋近于输入的汉字, 从而在两汉字之中寻求一种最佳的匹配.

关键词: 可变形弹性匹配, 手写体汉字识别, 笔段位移矢量, 弹性模板

Abstract: This paper describes a deformable elastic matching approach to Handwritten Chinese Character Recognition (HCCR). We assume that different handwriting variations share the same topological structure, but may differ in shape details. The variations between different handwriting characters are modeled by a set of stroke displacement vectors (SDV). According to the SDV derived, a model character is deformed gradually, in an effort to transform itself much closer to an input character. Experiments show that the proposed elastic matching model can efficiently deal with local shape changes and variations between characters.

Key words: Deformable elastic matching, Handwritten Chinese character recognition, Stroke displacement vector, Flexible template

一、引 言

在计算机视觉及图象处理中, 已有许多基于刚体模型(Rigid Model)的目标识别及图象分析技术. 在这些模型中, 物体通常只进行线性变换和旋转变换. 然而, 客观世界中的物体不仅仅是作线性或旋转变换, 而可能会伸缩甚至变形, 在此情况下, 应用刚体模型就显得不太有效. 因此, 近年来, 将物体视为一种可变形弹性物体越来越多地引起了国内外学者的重视^[1~4, 6~7]. 许多可变形弹性模型已经建立起来并成功地应用于图象分析及识别, 如 Rubber Mask

Model^[1], Iterative Elastic Matching Technique^[2], Point Distribution Model^[3]及 Active Contour Model^[4]等等.

手写字符的识别多年以来一直是模式识别中一个极具挑战性的难题^[5]. 对于手写体汉字, 我们认为尽管不同书写风格写出来的汉字在形状上有很大的差异性, 但它们的基本拓扑结构是保持不变的, 因此我们可将手写汉字看成是一种可变形的物体, 不同的手写样本可视为从同一个能有效地表征该字基本拓扑结构的弹性模板经不同变形而来的版本.

在本文中, 我们将提出一种可变形弹性匹配模

* 1995年10月收到, 1996年2月修改定稿. 国家自然科学基金资助

** Jin Lianwen, Xu Bingzheng (Inst. of Radio Eng. and Auto., South China Univ. of Tech., Guangzhou 510641)

型并将之用于手写汉字体识别. 在该模型中, 每一种汉字可用一个或几个可变形的模板来代表. 当输入字符与模板进行比较时, 通过匹配这两个字符将导出一组位移矢量(Displacement Vectors), 在该位移矢量的作用下, 模板字符将逐步变形以寻求与输入字符之间的最佳匹配.

二、基于笔划的可变形弹性匹配模型

1. 手写体汉字的小笔划段特征表示

笔划是组成汉字的基本元素, 汉字的笔划特征受字体、字形大小影响较小, 是识别汉字的良好特征. 可惜笔划特征对实际的书写文本来说较难提取. 通常, 用小笔划段作为特征能较好地解决如上困难^[9].

小笔划段是汉字字符中的一段直线, 可用其两端点的坐标来表征. 假定是一给定汉字中的第 i 条小笔划段为 s_i , 则 s_i 可表示为:

$s_i = \{p_i^1, p_i^2\} = \{(x_i^1, y_i^1), (x_i^2, y_i^2)\}$, $p_i^1 = (x_i^1, y_i^1)$ 和 $p_i^2 = (x_i^2, y_i^2)$ 分别表示其两个端点.

若干小笔划首尾相连即构成了汉字笔划. 用小笔划段作为单元, 一方面易于提取, 另一方面又保留了汉字基本的笔划结构信息, 且在字体变化或噪声、干扰条件下, 仍能保持笔划结构的绝大部分信息. 值得注意的是, 各小笔划段之间的相互约束关系是不同的: 同一个笔划的各小笔划段之间的约束关系最强, 同一部件但不同笔划的小笔划段之间的约束关系次之, 不同部件的小笔划段之间的约束关系最弱.

2. 基于笔段的位移矢量场

给出两个汉字 M 和 T :

$$M = \{s_1, s_2, \dots, s_m\}, T = \{t_1, t_2, \dots, t_n\}$$

式中, s_i 是 M 中的第 i 条小笔划段, t_j 是 T 中的第 j 条小笔划段, m 和 n 分别为 M 、 T 的小笔划段数目, 匹配 M 与 T 实际上等价于匹配它们相对应的每条小笔划段. 我们将用一组笔划位移矢量来表示两个不贩手写汉

字相对应小笔划段之间的差异, 如图 1 所示.

字符 M 和 T 中的两个笔划段之间的位移矢量可以通过它们端点处的位移矢量来定义, 我们首先考虑笔 s_i 中一 endpoint p_i 到笔划段 t_j 的位移矢量, 如图 2 所示, 该位移矢量由下式给出:

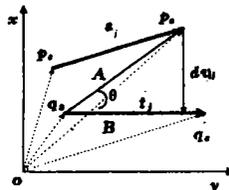


图 2 从笔划 s_i 上一点 p_i 到笔划段 t_j 的位移矢量

$$dv_{ij} = \begin{cases} \|A\| \sin(\theta) \frac{B \cdot t - B \cdot j}{\|B\|}, & 0 \leq \|A\| \cos(\theta) \leq \|B\| \\ q - p_i, & \text{其它} \end{cases} \quad (1)$$

式中: $A = p_i - q$, $B = q_i - q_j$, θ 为 A 和 B 之间的夹角,

$q = \begin{cases} q_i, & \|q_i - p_i\| \leq \|q_i - p_i\| \\ q_j, & \|q_j - p_i\| \leq \|q_i - p_i\| \end{cases}$, (i, j) 为二维坐标系中的基坐标; $\|x\|$ 是矢量 x 的模.

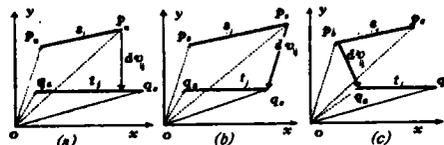


图 3 笔划 s_i 上点 p_i 到笔划段 t_j 的位移矢量的三种可能情况

图 3 给出了从 p_i 到 t_j 的位移矢量 dv_{ij} 的三种可能情况. 从笔划段 s_i 到 t_j 的笔划位移矢量可用其两端点的位移矢量 dv_{ij}^1 和 dv_{ij}^2 来表示. 从 t_j 到 s_i 的笔划位移矢量同理可定义, 我们将之表示为 dv_{ji}^1 、 dv_{ji}^2 , 如图 4 所示, 从物理学的观点来看, 这些笔划位移矢量可解释为一种“力场”, dv_{ij}^1 、 dv_{ij}^2 可解释为一种推力, 而 dv_{ji}^1 、 dv_{ji}^2 可解释为一种拉力. 在此“力场”的作用下, 笔划段 s_i 将逐渐向 t_j 靠近.

这样, 我们可以赋予字符 M 一定的可变形弹性

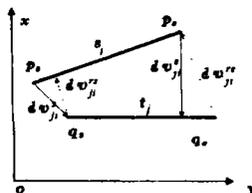


图 4 笔划段 s_i 和 t_j 之间的位移矢量场

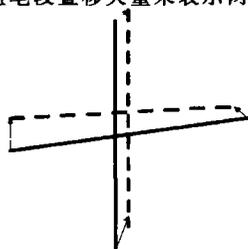


图 1 两个“十”字之间的笔划位移矢量

特性,不再将字符中的笔划段看成是固定不变的,而是在某种程度上响应这种位移矢量“力场”的作用,从而使该字符逐渐变化以寻求和字符 T 的结构上的最佳匹配。

3. 笔段位移矢量的平滑

由于汉字的字形结构复杂,因此两个汉字的笔段位移矢量场也可能会变得十分复杂,不能仅仅由式(1)来简单给出,而同时应考虑到各笔段之间的联系和相互关系。前面我们提到过同一笔划的各小笔段之间的约束关系最强,而不同部件的小笔段之间的约束关系最弱,因此在决定笔段位移矢量时,我们还应考虑到这种约束关系。首先,考虑一个小笔段 s_i 周围相邻的位移矢量场的影响。我们认为越靠近的笔段 S_j 的位移矢量对其影响越大,越远则越小。为此,我们使用了如下 Gaussian 平滑函数来提供这种效果:^[2]

对端点 p_i^j :

$$DV_i^j = \gamma^{-1} \left[\frac{\sum_j dv_{ij}^j w_{ij}^j}{\sum_j w_{ij}^j} - \frac{\sum_j dv_{ij}^i w_{ij}^i}{\sum_j w_{ij}^i} \right] \quad (2)$$

对端点 p_i^i :

$$DV_i^i = \gamma^{-1} \left[\frac{\sum_j dv_{ij}^j w_{ij}^j}{\sum_j w_{ij}^j} - \frac{\sum_j dv_{ij}^i w_{ij}^i}{\sum_j w_{ij}^i} \right] \quad (3)$$

式中, $w_{ij}^i = \exp\{-\|dv_{ij}^i\|/\sigma_k\}$; $w_{ij}^i, w_{ij}^j, w_{ij}^i$ 可同理定义。

平滑实际上是对笔段 s_i 周围的位移矢量场中的“拉力场”和“推力场”的加权和, γ 可视为一个平滑控制因素, σ_k 是在第 k 个迭代过程中的松弛因子。经验表明: σ_k 取如下的值时效果会较理想:

$$\sigma_k = \sigma_0 / f^k \quad (4)$$

式中 f 是一个取值于 $[1, 2]$ 之间的常量。

上面我们讨论了两个字符之中的位移矢量场中的平滑问题,另一方面,考虑到同一个字符中相邻笔段之间的约束关系,我们将对由式(2), (3)中求得的位移矢量按下式作进一步的加权平滑,以求得最终的笔段平滑位移矢量(SDV)。

对端点 p_i^j :

$$SDV_{ij} = \sum_j DV_j^j G_j^j N_{ij} - \sum_j G_j^i \quad (5)$$

对端点 p_i^i :

$$SDV_{ij} = \sum_j DV_j^j G_j^j N_{ij} - \sum_j G_j^i \quad (6)$$

式中, $G_j^j = \exp\left\{-\frac{\|DV_j^j\|}{\tau}\right\}$, $G_j^i = \exp\left\{-\frac{\|DV_j^i\|}{\tau}\right\}$

$$N_{ij} = \begin{cases} 1, & s_i \text{ 与 } s_j \text{ 相邻} \\ 0, & \text{其他} \end{cases}$$

我们发现:式(5)、(6)在变形匹配过程中,对于保持一个字符的拓扑结构起着十分重要的作用。

4. 弹性迭代匹配过程

给定输入字符 $T = \{s_1, s_2, \dots, s_m\}$ 和模板字符 $M = \{t_1, t_2, \dots, t_m\}$, 在进行可变形弹性匹配时,模板字符 M 根据与 T 进行笔段匹配所在的笔段位移矢量(SDV^i, SDV^j), 逐次变形以寻求与字符 T 的最佳匹配。变形字符 M 的过程是通过变形其各小笔段而进行:

$$\{s_i^{t+1}\} = \{s_i^t\} + \{SDV_i^t, SDV_j^t\} \quad (7)$$

式中 t 是迭代次数, 当下述条件满足时, 迭代过程即告停止。

$$\frac{1}{m} \sum_{i=1}^m [\|SDV_i^t\|] + [\|SDV_j^t\|] \leq \epsilon \quad (8)$$

这里 ϵ 是一个很小的常数。

字符 M 与 T 的匹配相似度可用下式所定义的差异度来描述:

$$\text{Dissimilarity}(M, T) = \frac{1}{m} \sum_{i=1}^m [\|dv_i^* \| + \|dv_j^* \|] + \frac{1}{n} \sum_{j=1}^n [\|dv_j^* \| + \|dv_i^* \|] \quad (9)$$

这里, $dv_i^* = \min\{dv_{ij}\}$ 是字符 M 中从笔段 s_i 的起点所得到的最小位移矢量, $dv_j^* = \min\{dv_{ji}\}$ 是字符 T 中从笔段 t_j 的起点所得到的最小位移矢量; dv_i^*, dv_j^* 同法可定义。

三、实 验

我们选用了 200 种常用汉字作为实验数据, 每一种汉字含有 30 个不同的书写样本。每个字符被预先扫描并归一化为 64×64 的图象数据, 然后, 我们应用了一种基于细化及特征点的笔划提取方法将一个汉字中的所有笔划段提取出来^[8]。

每一种汉字我们选用两个较规范的样本作为模板汉字, 其余的样本用于测试, 当输入一个未知的测试数据时, 每一个模板汉字将根据上文所述的弹性匹配算法进行变形以匹配该输入汉字。图 5 显示了匹配两个“不”字的动态过程。由图中可见, 经过六次迭代过程之后, 模板汉字已经变得与输入汉字十分接近了。图 6 给出了另外两个例子。

当所有的模板汉字与输入的未知汉字进行过匹



图 5 对两个“不”字的弹性匹配过程。图(a)为输入字符,图(b)为可变形弹性模板,由图(c)可见,经过几次迭代匹配过程之后,模板已变得与输入字符十分接近。

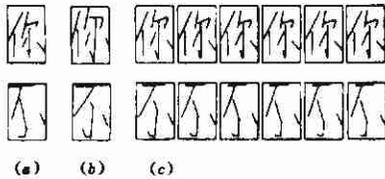


图 6 分别对两个“你”字和“不”字的弹性匹配过程。图(a)为输入字符,图(b)为可变形弹性模板,图(c)为匹配过程。

配比较之后,最终给出最好匹配度(即具有最小的差异度)的模板将作为识别结果。对 200 种汉字共 5600 个手写样本进行测试我们得到了 94.6% 的识别率。

四、结束语

本文提出了一种可变形弹性匹配模型并将之用于手写体汉字识别,实验结果是令人鼓舞的。

该模型的不足之处在于匹配算法的运算量较大。当实际用于大词汇量的汉字识别时,希望能有更好的预分类算法来作粗分类,将大量汉字粗分为一系列的小字量集,然后再用此方法作为最终的匹配识别。此外,对于同类汉字如果所匹配的手写样本的噪声较大或整体的匹配程度太差的情况下,有可能会较难决定正确的位移矢量场。在此情况下,利用一些约束条件(如手写汉字的一些结构特征)来协助决定位移矢量场,也许值得我们更进一步的研究。

本文的工作得到了香港大学陈国平博士的协助与指导,谨此表示感谢。

参 考 文 献

- 1 B. Widrow. The Rubber Mask Technique, Part one and Part two. Pattern Recognition, 1973, 5: 175~211
- 2 D. J. Burr. Elastic Matching of Line Drawings. IEEE Trans on Pattern Anal. Machine Intell., 1981, PAMI-3(6): 709~713
- 3 A. Hill, T. F. Cootes and C. J. Taylor. A Generic System for Image Interpretation Using Flexible Templates. Proc. of British Machine Vision Conf., 1992: 276~285
- 4 M. Kass, A. Witkin and D. Terzopoulos. SNAKE Active Contour Model. Proc. of Intel. Conf. on Computer. Vision. 1987: 259~268
- 5 Shunji Mori, C. Y. Suen and Kazuhiko Yamamoto. Historical Review of OCR Research and Development. Proc. of the IEEE, 1992, 80(7): 1029~1057
- 6 Jean-Michel Bertille. An Elastic Matching Approach Applied to Digit Recognition. in Proc. of the Second Intel. Conf. on Document Anal. and Proc., 1993: 82~85
- 7 Toru Wakabayashi. Shape Matching using LAT and its Application to Handwritten Numeral Recognition. IEEE Trans, 1994, PAMI-16(6): 618~629
- 8 Fang-Hsuan Cheng and Wen-Hsing Hsu. Three Stroke Extraction Methods for Recognition of Handwritten Chinese Characters. Proc. of the 1986 Intel. Conf. on Chinese Computing, 1986: 191~195
- 9 张忻中. 汉字识别技术. 北京清华大学出版社, 1992



金连文 1968年生,1991年毕业于中国科技大学无线电系,获学士学位。现正在华南理工大学攻读博士学位。研究兴趣包括:模式识别,中文信息处理,神经网络等。



徐杰琼 华南理工大学教授,博士生导师,IEEE高级会员。研究兴趣包括:通信理论,模式识别,中文信息处理,神经网络,模糊信息处理等。