

变形变换及其在手写体汉字整形中的应用

金连文¹⁾ 黄建成²⁾ 尹俊勋¹⁾ 贺前华¹⁾

¹⁾(华南理工大学电子与通信工程系, 广州 510641)

²⁾(Motorola 电子(有限)公司中国研究中心, 上海 200002)

摘要 提出了一种新的一维变形变换方法并将之应用于手写体汉字的整形变换。尽管不同书写风格同类手写体汉字存在较大的随意性,但它们的拓扑结构是一致的,因此可将手写体汉字看成是一种可变形物体。通过选择适当的变形参数,该变形变换可以将一个汉字变形为24种不同的书写风格,利用一个变形系数可控制各种风格的变形程度,该变形变换可以矫正手写体汉字中偏左、偏右、偏上、偏下、偏胖、偏瘦等各种位置变形,从而达到对手写体样本进行非线性规范化处理的效果。实验结果表明,手写体汉字经变形变换后,识别率可提高大约1.5%左右,显示了本文所提算法的有效性。

关键词 变形变换 手写体汉字 整形处理

中图法分类号: TP391.43 **文献标识码:** A **文章编号:** 1006-8961(2002)02-0170-06

A Novel Deformation Transformation and Its Application to Handwritten Chinese Character Shape Correction

JIN Lian-wen¹⁾, HUANG Jian-cheng²⁾, YIN Jun-xun¹⁾, HE Qian-hua¹⁾

¹⁾(Department of Electronics and Communication, South China University of Technology, Guangzhou 510641)

²⁾(China Research Center, Motorola (China) Electronics LTD, Shanghai 200002)

Abstract In this paper, a novel 1D deformation transformation for handwritten Chinese character shape correction is proposed. From our point of view, handwritten Chinese character can be regarded as a kind of deformable object. Although there are large variations in the same category of character for different handwriting styles, the basic topological structures of them are the same. With a suitable selection of deformation parameters, the 1D deformation transformation could deform a given handwritten Chinese character into 24 different handwriting styles. A deforming parameter controls the deformation degree for each style. The proposed deformation transformation could be applied as a non-linear shape normalization method for Chinese character recognition to correct various position deformations such as left-slanting, right-slanting, top-slanting, bottom-slanting, too fat or too thin, etc in a handwritten Chinese character. Our experiment shows that the recognition rate can be improved by 1.5% or so by using the deformation transformation, demonstrating the effectiveness of the proposed transformation.

Keywords Deformation transformation, Handwritten chinese character, Shape correction

0 引言

基于模板的图形图象匹配是图象处理和模式识别中要解决的一项关键技术。传统的图形图象匹配

大多采用刚体模型(Rigid Model)匹配技术来进行,这种模型通过对模板进行旋转、平移、尺度变换、仿射变换等一些处理,在模板与待识别目标之间寻找最佳匹配。然而,客观世界中的物体不仅仅是作线性或仿射变换,还可能会伸缩,甚至做不规则的变形,

基金项目:国家自然科学基金(69802007);广东省自然科学基金(980602.970445);Motorola 研究基金

收稿日期:2000-07-20;改回日期:2001-03-08

在此情况下,应用刚体模型就显得不太有效,因此,最近几年来,在国内外有学者提出了所谓的可变形模板模型(Deformable Template Models,简称DTM),以解决这一难题.在可变形模板模型中,模板的几何/物理模型不是固定的,而是活性(Active)的,它可在一定范围内,根据各种约束条件进行非线性变形,以实现与待识别目标的最佳匹配.可变形模板模型理论及技术,由于其对图形图象的非线性变形的自适应处理能力,使其在许多实际应用场合比刚体模型匹配技术更为可靠有效,在模式识别(如人脸识别、手形识别)、目标检测、图象处理、手写体字符识别等领域中得到了广泛的应用^[1-7].

手写体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.手写体汉字识别的一个难点在于不同的书写者书写的汉字样本风格千差万别,笔迹不同、下笔轻重不同、肥瘦不均、字型极其不规则,如重心不稳定,有的字偏左、偏右、偏上、偏下等,发生了不同的变形,因此,不少学者提出许多线性或非线性的规范化及整形变换算法来对手写体汉字样本进行整形,并证明这些算法是行之有效的^[8,9].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,因此可将手写汉字看成是一种可变形的物体,而不同书写风格之间的差异可以通过适当的整形处理来加以矫正.

本文提出了一种用于手写体汉字变形整形的变形变换,该变换能对手写体汉字样本进行24种不同变形处理,以矫正偏左、偏右、偏上、偏下、偏胖、偏瘦等各种位置变形,达到对手写体样本进行规范化的处理效果,从而利于手写体汉字的识别处理.

1 变形变换的定义

手写体汉字经过数字化及二值化后,变成二维图象,该二值图象由若干黑象素点构成,因此,一个手写体汉字可表示为

$$C = \{p_1, p_2, \dots, p_n\} \quad (1)$$

式中, $p_i = (x_i, y_i)$ 表示第 i 个黑象素点的坐标, n 为该汉字的黑象素点总数目.

对汉字 C ,可按下式将其变换为新的字 C_{DT}

$$\begin{cases} C_{DT} = D(C) = \{D(p_1), D(p_2), \dots, D(p_n)\} \\ D(p_i) = (x_i, y_i) + (D_x, D_y) \\ \quad = (x_i, y_i) + (f_x(x_i, y_i), f_y(x_i, y_i)) \\ \quad = p_i + d_i \end{cases} \quad (2)$$

式中,称 $d_i = (f_x(x_i, y_i), f_y(x_i, y_i))$ 为位移矢量(displacement vector), $f_x(x_i, y_i), f_y(x_i, y_i)$ 为位移函数(displacement function).由此看到,对汉字 C 的变换相当于对汉字中各黑象素点进行位移变换(将其黑象素点位置进行变换).

位移函数 $f_x(x_i, y_i), f_y(x_i, y_i)$ 可以有多种不同的选择,从而得到的变换也是不同的,对于变形变换,其应具备如下特性:

(1) 变形变换是一种非线性变换,其可使图象形状发生非线性形变;

(2) 变形变换应保持原始图象的拓扑结构不变;

(3) 变形变换应满足一定的边界约束条件,例如对手写体汉字而言,变形变换应使汉字图象边界不变形,以保持汉字的大小及其方块形状;

(4) 变形变换应使图象中的曲线形状保持平滑性(smoothness)及连接性(connectivity).

满足上述条件的变换 D 我们称之为变形变换(Deformable Transform, DT).

在式(2)中,位移函数 $f_x(x_i, y_i), f_y(x_i, y_i)$ 是二元函数,这样得到的DT称之为2维变形变换.特别地,当 $f_x(x, y) = f(x)$ 及 $f_y(x, y) = f(y)$ 时,这样的变换 D 称之为—维可变形变换(1D DT).

限于篇幅,本文仅讨论1D可变形变换.

2 1D可变形变换

对手写体汉字 C 中的黑象素点的1D可变形变换定义为

$$D(p_i) = (x_i, y_i) + (f(x_i), f(y_i)) \quad (3)$$

即

$$\begin{cases} D(x_i) = x_i + f(x_i) \\ D(y_i) = y_i + f(y_i) \end{cases} \quad (4)$$

不失一般性,不妨将 x, y 的取值范围归一化为 $[0, 1]$.选择如下三角函数作为位移函数 $f(x)$

$$f(x) = \eta x [\sin(\pi\beta x + \alpha)\cos(\pi\beta x + \alpha) + \gamma] \quad (5)$$

式中, $\alpha, \beta, \gamma, \eta$ 为常数.特别地,当 $\alpha = 0, \beta = 1, \gamma = 0$ 时,有

$$f(x) = \eta x \sin \pi x \cos \pi x \quad (6)$$

图1给出了两种不同参数 η 的变形函数 $D(x) = x + f(x)$ 的曲线图.

从图1可以看到,选择如式(5)所示的三角函数构成的变形函数 $D(x)$ 是一非线性函数,它将 x 域中的均匀区间,变换为不均匀区间,而不同的参数 η

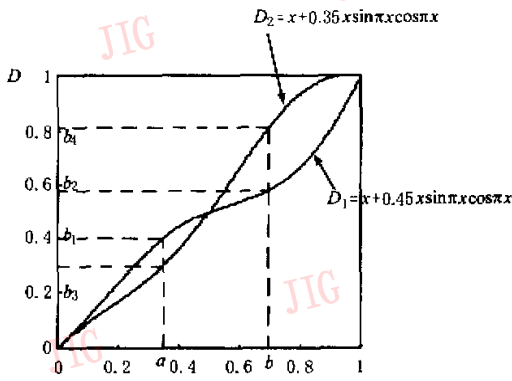


图 1 两条不同的变形曲线

及不同的区间控制着不同的变换效果,称 η 为变形系数.以图 1 为例,对 $[a, b]$ 区域,变换 D_1 将该区间压缩为区间 $[b_1, b_2]$,而变换 D_2 则将该 $[a, b]$ 放大为区间 $[b_3, b_4]$,对于区间 $[0, a]$,变换 D_1 将 $[0, a]$ 中心点往原点方向压缩,而 D_2 则将 $[0, a]$ 向远离原点方向扩展.三角函数这一特殊的非线性在变形变换中非常有用,即可以通过选取不同的区间,而得到不同的非线性变换效果,同时又能保证在边界处(即 x 取值为 0 和 1 时)函数取值不变.

在式(5)中,考虑 $[0, 1]$ 区间的子区间 $[a, b]$,令:

$$\pi\beta x + \alpha \Big|_{x=0} = a$$

$$\pi\beta x + \alpha \Big|_{x=1} = b$$

得到:

$$\alpha = a, \beta = \frac{b-a}{\pi} \tag{7}$$

根据变形变换的第 3 条约束特性,置移函数 $f(x)$ 应满足如下边界条件(对汉字图象而言,此条件保证了其变形后边界不发生变化,即保证汉字的方块形状不变):

$$f(0) = 0, f(1) = 0$$

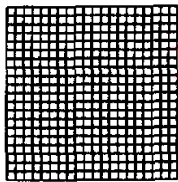
这样就得到

$$y = -\sin b \cos b \tag{8}$$

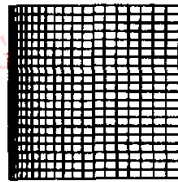
将式(5)、式(7)、式(8)代入式(4),考虑到 x 及 y 可进行独立变换,这样就得到对汉字 C 的可变形变换定义为

$$\begin{cases} D(x_i) = x_i + \eta_1 x_i [\sin[(b_1 - a_1)x_i + a_1] \times \\ \cos[(b_1 - a_1)x_i + a_1] - \sin b_1 \cos b_1] \\ D(y_i) = y_i + \eta_2 y_i [\sin[(b_2 - a_2)y_i - a_2] \times \\ \cos[(b_2 - a_2)y_i - a_2] - \sin b_2 \cos b_2] \end{cases} \tag{9}$$

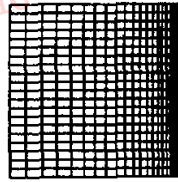
式中, $0 \leq a_1 < b_1 \leq 1, 0 \leq a_2 < b_2 \leq 1, \eta_1, \eta_2$ 为常数.值得注意的是,可以通过选择不同的变换区间 $[a_1, b_1], [a_2, b_2]$ 及不同的变形系数 η_1, η_2 而得到不同的变形效果(如图 2 所示)(参数 $a_1, a_2, b_1, b_2, \eta_1, \eta_2$ 的选择可依据图 1 所示的非线性曲线的分段非线性



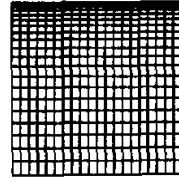
(a) 原始网格图



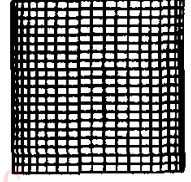
(b) 水平左拉效果
($a_1=0, b_1=0.5, \eta_1>0$,
垂直方向不变形)



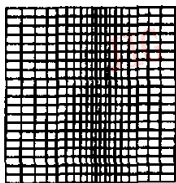
(c) 水平右拉效果
($a_1=0.5, b_1=1, \eta_1>0$,
垂直方向不变形)



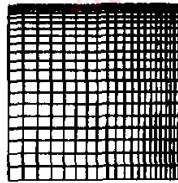
(d) 垂直上拉效果
($a_2=0, b_2=0.5, \eta_2>0$,
水平方向不变形)



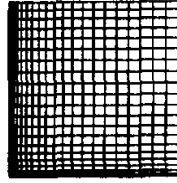
(e) 水平左右扩展
($a_1=0, b_1=1, \eta_1>0$,
垂直方向不变形)



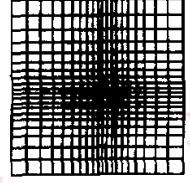
(f) 水平压缩效果
($a_1=0, b_1=1, \eta_1<0$,
垂直方向不变形)



(g) 右上压缩效果
($a_1=0.5, b_1=1, \eta_1>0$,
 $a_2=0, b_2=0.5, \eta_2>0$)



(h) 左下压缩效果
($a_1=0, b_1=0.5, \eta_1>0$,
 $a_2=0.5, b_2=1, \eta_2>0$)



(i) 中心压缩效果
($a_1=0, b_1=1, \eta_1<0$,
 $a_2=0, b_2=1, \eta_2<0$)

图 2 对一均匀网格图进行可变形变换的效果

性视具体情况而定)。

3 对手写体汉字的 24 种整形处理

一般地,根据图 1 所示的非线性变形函数不同的非线性特性,通过选择不同的 $a_1, a_2, b_1, b_2, \eta_1, \eta_2$ 值,得到了对汉字进行如图 3 所示的 24 种整形效果。

这 24 种整形处理的实际应用意义在于:如果某个手写体汉字发生了位置变形(例如偏左),则可以进行相应的整形处理(右拉)来对其进行矫正,以得到规范化的汉字样本。

此外,对某种整形效果,还可以通过控制不同变形系数的取值而得到变形程度不同的处理效果,如图 4 所示。

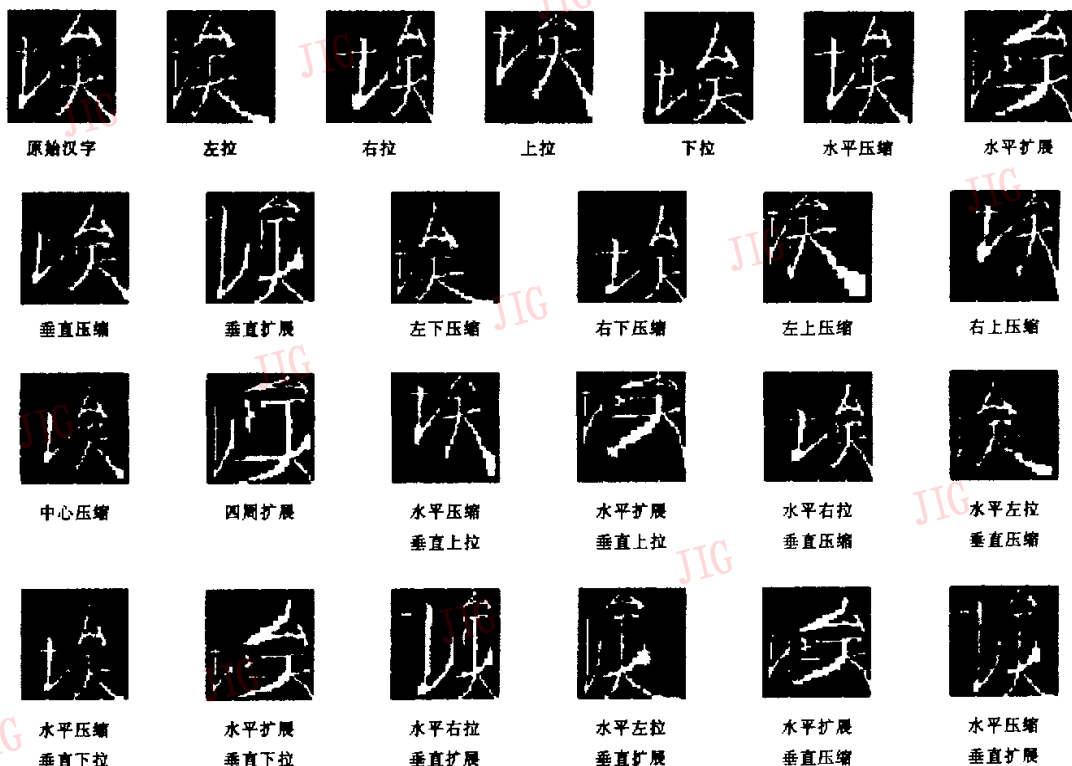


图 3 24 种不同整形效果



(a) 原始图像

(b) η 取值为 0.8~2 时不同的左压缩效果

(c) η 取值为 0.8~2 时不同的右压缩效果

图 4 不同的 η 得到变形程度不同的效果

4 与余弦变形变换的比较

余弦变形变换最早是 Guo 等提出的一种非线性变形变换方法^[8,10],其特点是:将余弦函数的非线性区间和准线性区间并用,以矫正文字的变形部分,保持非变形部分.这种变换具有矫正文字偏向上下左右某一方向的变形和将信息量最为集中的中心部

分扩展的功能.余弦变形变换的基本思想是:利用余弦函数的非线性区间和准线性区间,对输入文字矫正变形部分,保持其非变形部分.余弦变形变换的定义为

$$X^{r,\theta}(x) = m - n\cos(\varphi x + \theta)$$

这里, φ 和 θ 分别代表变换域的宽度和初始相位, $\theta \geq 0, \varphi > 0, 0 \leq \varphi + \theta \leq \pi, 0 \leq x \leq 1$.根据不同的初始相位及变换区间,可得到如图5所示的6种整形效果.



图5 余弦变形变换的6种整形效果

比较图3及图5可知,用本文提出的变形变换得到的24种整形效果实际上已完全包含了余弦变形变换的6种整形效果,余弦变形变换可视为本文提出方法的一个子集.

5 变形变换在汉字识别中的初步应用

手写体汉字识别的一个难点在于对不同的书写样本有较大的随意性及书写变形,因此,不少学者提出了许多线性或非线性的规范化及变形变换算法用以对手写体汉字样本进行整形,并已证明这些算法是行之有效的^[8~10].为将本文提出的变形变换应用于汉字识别,现将一个未知待识别汉字进行如图3所示的24种整形变换,每种变换取3种不同的 η 参数值,这样每个汉字将产生72个整形模板.假设:如果某个汉字发生了如图3所示的位置变形中的一种,那么通过变形变换,其中有一个样本一定是较规范并且容易辨认的.类似于文献[10]的方法,实验中将72个整形模板及原始未整形的样本与标准模板进行比较,选取距离最近的模板作为最后的识别输出.对国标一级字库16~20区共470字的实验结果如表1所示,实验数据为40套训练样本,20套测试训练,使用的特征为弹性网格方向特征^[11].

从表1可以看到,通过变形变换能有效地提高识别率,但并不是对所有的样本都有提高(如样本3、10、14).这是由于这些样本都比较规范(识别率已经很高),没有整形的必要,而整形后反而带来一定的误差,使识别率略有下降.通过实验还发现,对于书写不规范的样本,变形变换特别有用.

表1 变形变换识别性能

样本	单位: %	
	整形前识别率	整形后识别率
1	94.04	95.75
2	93.83	95.75
3	97.66	97.45
4	93.62	93.83
5	92.77	94.68
6	88.72	91.28
7	92.77	94.25
8	93.83	96.17
9	83.83	85.96
10	97.87	96.60
11	91.06	92.55
12	95.74	96.81
13	95.96	96.38
14	99.15	98.72
15	89.15	92.55
16	83.83	85.96
17	90.64	91.91
18	94.04	95.53
19	91.91	94.68
20	83.83	85.10
总识别率	92.18	93.60

6 结语

本文提出一种用于手写体汉字的变形变换,通过控制所提出的置移函数的参数及变形系数,实现对手写体汉字的24种不同整形效果的变形处理,初步实验表明,该方法在手写体汉字识别的整形处理中有较大的应用潜力.

变形变换在实际应用中,仍然存在一些具体问题需要解决,主要表现为 3 个方面:

(1) 参数的选择问题 本文提出的变形变换用 a, b, η 3 个参数决定变形的效果,即决定变形的类型及变形的程度.在实际应用中如何选择合适的变形类型及合理的变形程度是一个十分关键的问题,目前还只能靠实验及经验来确定.能否设计一些自动搜索最佳参数的算法,将是一个值得研究的问题.如果参数选择恰当,相信识别率还可以进一步得到提高.

(2) 自动检测汉字的变形问题 在实验中发现,写得比较规范的样本并没有进行整形的必要,如何判断汉字的书写变形,以决定是否对其进行变形变换,是一个亟待解决的难题.目前一个可行的思路是在整形前先进行识别处理,并判断识别结果的可信度,对于较规范的样本,一般其识别可信度较高,可不必进行变形处理,而对不规范的样本,识别可信度较低,此时再对其进行变形变换,矫正其书写变形,从而提高识别精度.

(3) 速度问题 变形变换要通过产生多个变形模板来进行匹配识别,因此识别处理速度较慢.解决这一问题的一个办法是先进行预分类,然后再进行变形变换处理及识别;另外,如果能成功解决前两个问题,变形模板数也将大大降低,识别速度则会相应提高.

变形变换方法也可推广应用到目标识别、图象图形匹配等其他领域,这也将值得进行进一步的研究.

参 考 文 献

- Jain A K, Zhong Y, Lakshmanan S. Object matching using deformable templates. [J] IEEE Trans. Pattern Anal. And Machine Intell., 1996, 18(3): 267~278.
- 金连文, 徐秉坤. 限定性手写体汉字识别的一种可变形弹性匹配模型[J]. 电子学报, 1997, 25(5): 35~38.
- Yuille P W, Hallinan, Cohen D S. Feature extraction from faces using deformable templates [J]. Int. Journal on Computer Vision, 1992, 12(2): 133~144.
- Terzopolous J, Platt, Barr A et al. Elastically deformable models [J]. Computer Graphics, 1987, 18(3): 205~214.
- Christensen R D, Rabbitt, Miller M I. Deformable templates using large deformation kinematics [J]. IEEE Trans. Image Processing, 1996, 5(10): 1435~1447.
- Berger M, Danuser G. Deformable multi template matching with application to portal images [A]. In: Proceedings IEEE International Conference on Computer Vision and Pattern Recognition [C], San Juan, USA, 1997: 374~379.
- Tsang C K Y, Chung Fu-Lai. Development of a structural deformable model for handwriting recognition [A]. In: Proceedings of Fourteenth International Conference on Pattern Recognition [C], USA, 1998, 2: 1130~1133.
- Guo J, Sun N, Nemoto Y et al. Recognition of handwritten characters using pattern transformation method with cosine function [J]. IEICE of Japan Trans. 1993, J76-D-II(4): 835~842.
- GOU Da-yin, DING Xiao-qing, WU You-shou. A handwritten chinese character recognition method based on image shape correction [A]. In: Proc. of 1st National on Multimedia and Information Networks [C], Beijing, 1995: 254~259.
- 张洪刚. HCL2000 汉字数据库与高精度识别算法的研究 [硕士论文]. 北京邮电大学, 1999 年.
- JIN Lian-wen, GANG Wei. Handwritten chinese character recognition with directional decomposition cellular features [J]. Journal of Circuit, System and Computer, 1998, 8(4): 517~524.



金连文 1996 年获华南理工大学电子与通信工程系博士学位,现为华南理工大学电子信息学院副教授.目前主要研究方向为模式识别、图象处理.在国内外学术刊物上已发表论文 20 多篇,主持多项自然科学基金项目的研究.

黄建成 博士, Motorola 公司中国研究中心主任研究员,长期从事信号处理、通信等领域的研究工作,在国际著名学术刊物上发表论文数十篇,获专利多项.



尹俊勋 华南理工大学电子信息学院教授、博士生导师.主要研究方向为信号处理、模式识别、通信编码等.在信号处理领域已发表学术论文 30 余篇,目前是多项广东省自然科学基金及国家自然科学基金资助项目的负责人.



黄建华 博士,华南理工大学电子信息学院副教授,1993 年获华南理工大学无线电博士学位.主要研究领域为遗传算法、语音信号处理、语音识别、语音合成、神经网络等.在模式识别及信号处理领域已发表论文近 30 篇.