第30卷 第11期 计 算 机 工 程 2004年6月
Vol.30 №11 Computer Engineering June 2004
·人工智能及识别技术· 文章编号：1000—3428(2004)11—0128—02 文献标识码：A 中图分类号：TP391.4

手写体数字识别中一种新的倾斜校正的方法

王有伟，刘 捷
（山东大学计算机学院，济南 250061）

摘 要：介绍了在手写体数字识别预处理过程中一种新的倾斜校正的方法，该方法利用手写数字的图像的高与宽的比值是否最大来确定倾斜校正是否完毕。在校正过程中，手写体数字的图像高度逐渐变高，宽度逐渐变窄，所以当图像的高与宽的比值最大的时候倾斜角度最小。

关键词：模式识别；手写体数字识别；文件自动处理；倾斜校正

A New Method for Slant Correction in Recognition of Handwritten Numerals

WANG Youwei, LIU Jie
(Department of Computer, Shandong University, Jinan 250061)

【Abstract】 A new method for correcting askew handwritten numerals in the preprocess of recognition is presented in the paper. In order to know that the correcting process has ended, it should make sure that the ratio between the height and the width of the bitmap of the handwritten numerals is the greatest. In the process of slant correction, the height of the bitmap is increasing, while the width of the bitmap is decreasing, so the bitmap is least slanted when the ratio between the height and the width becomes the greatest.

【Key words】 Pattern recognition; Handwritten numerals recognition; Automated document processing; Slant correction

模式识别是用计算机来识别自然界各种模式(例如声音、指纹、虹膜、字符)的科学,是随着人工智能的发展而逐步被人们所重视的一门新的学问。手写数字、字符识别是模式识别的一个分支，在文件自动处理过程中，手写体字符的识别是非常重要的，大部分的文件自动处理过程基本上由以下几步来完成：(1)预处理，在这一阶段进行去噪、标准化、分割、倾斜校正以及提取骨架等操作；(2)特征提取，这一阶段提取每个字符所具有的不同特征；(3)设计分类器，这一阶段主要是根据上一阶段所提取的特征来设计分类器并训练分类器，使系统最终能够自动识别需要识别的字符。

预处理阶段如果进行得非常充分的话，就会提高所设计系统的识别率。而预处理中很重要的一个部分就是倾斜校正，如果这一部分的工作做得不够完善，那就会给后续工作造成麻烦甚至会影响到系统的识别率。

1 倾斜校正

需要进行倾斜校正的手写数字的图像主要有两种：一种是数字间完全没有限制的，可以连笔；另一种是数字间彼此孤立，没有任何联系。本文提出的倾斜校正方法是针对后一种情况的。从20世纪60年代以来，人们研究出了很多不同的对倾斜图像进行校正的方法。其中的很大一部分是通过矩的不变性特征来实现的，还有一部分方法是通过对待校正图像的一些特征点进行计算或是对灰度直方图进行分析从而估计出它倾斜的大致角度，根据得到的角度来把它校正至正常。

本文所采用的算法既没有用到矩的概念也不是用某种方法估算出图像的倾斜角度，而是用递归的方式来逐步完成校正工作的。该算法主要基于以下的观察：

(1)一般手写数字的倾斜度不会超过45°，即便是超过45°，也一定不会达到直角；

(2)当倾斜角度达到最小时，图像的高和宽的比值就达到最大。

这两个现象说明，只要在整个角度空间找到图像的高和宽的最大比值就找到了倾斜图像的校正位置。但是很显然，在整个角度空间来寻找这个最大值耗费的时间过长。因此如何得到图像高度和宽度的最大比值就是本文所要解决的主要问题。

在给出算法之前，先作这样的规定：图像中某点旋转方向为顺时针时，角度为正，为逆时针时，角度为负；旋转中心定在图像的几何中心，即点$(\frac{1}{2}h,\frac{1}{2}w)$（$h$，$w$分别指图像的高和宽）处。

对于任意给定的图像中的像素点(x_0,y_0)（该点不是背景点），绕旋转中心分别进行顺时针和逆时针旋转的方式如图1所示，图中(x_1,y_1)、(x_2,y_2)分别表示的是(x_0,y_0)顺、逆时针旋转角度θ后所到的点。

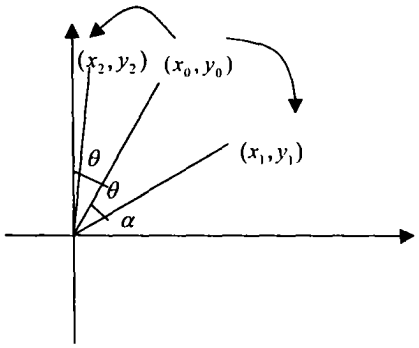


图1 任意点的顺、逆时针旋转示意图

作者简介：王有伟(1979－)，男，硕士生，主要研究模式识别和计算机网络；刘 捷，教授

收稿日期：2003-05-15 E-mail：alittlekid@163.com

旋转前：$x_0 = r\cos(\alpha)$，$y_0 = r\sin(\alpha)$，r 是坐标点离原点的距离，那么顺、逆时针旋转后的点 (x_1, y_1)、(x_2, y_2) 分别为：

顺时针旋转角度 θ 后，点 (x_1, y_1)：

$$x_1 = r\cos(\alpha-\theta) = r\cos(\alpha)\cos(\theta) + r\sin(\alpha)\sin(\theta) = x_0\cos(\theta) + y_0\sin(\theta)$$

$$y_1 = r\cos(\alpha-\theta) = r\sin(\alpha)\cos(\theta) - r\cos(\alpha)\sin(\theta) = -x_0\sin(\theta) + y_0\cos(\theta)$$

逆时针旋转角度 θ 后，点 (x_2, y_2)：

$$x_2 = r\cos(\alpha+\theta) = r\cos(\alpha)\cos(\theta) - r\sin(\alpha)\sin(\theta) = x_0\cos(\theta) - y_0\sin(\theta)$$

$$y_2 = r\cos(\alpha+\theta) = r\sin(\alpha)\cos(\theta) + r\cos(\alpha)\sin(\theta) = x_0\sin(\theta) + y_0\cos(\theta)$$

以上的旋转方法就是该算法对手写数字的位图图像进行倾斜校正时所要用到的，下面给出倾斜校正的具体算法：

(1) 先假定初始旋转角度 θ 为45°，初始位图图像 B 为活动位图。

(2) 如果旋转角度 θ >1°，转到(3)。

(3) 设定活动位图为 B_0，求出 B_0 的高度 h_0、宽度 w_0 并求出两者的比值 $\mu_0 = \frac{h_0}{w_0}$。将 B_0 的所有像素点利用上面的方法分别进行顺、逆时针旋转角度 θ，得到的图像赋值为 B_1、B_2，并求出 B_1 的高度 h_1、宽度 w_1、比值 $\mu_1 = \frac{h_1}{w_1}$ 和 B_2 的高度 h_2、宽度 w_2、比值 $\mu_2 = \frac{h_2}{w_2}$。

(4) 求出 μ_0、μ_1、μ_2 中最大的一个，将它所对应的位图图像赋值为活动位图 B_0。并把旋转角度作改变：$\theta \leftarrow \frac{\theta}{2}$。然后回到(2)重新开始。

这个算法是个逐步接近以取得最佳值的算法，下面解释一下该算法是如何通过多次循环来达到校正倾斜的手写数字的目的的。

如果原图像的倾斜角度大于初始角度 θ，那么经过第一轮循环以后，得到的图像还跟原图像在同一个倾斜方向；如果原图像的倾斜角度小于初始角度 θ，那么经过第一轮循环后，可能得到一个新的倾斜的图像，它的倾斜角度可能为正也可能为负。

如果经过某轮循环后，倾斜角度有改变，新的倾斜角为 $+\alpha$ 或 $-\beta$，那么会有两种情况：α（或 β）$<\theta/2$ 或者 $\theta/2 \le \alpha$（或 β）$\le\theta$。在第一种情况下，下一轮循环旋转后得到的倾斜角度 θ' 大于 $\theta/2$，则循环后保持原来的倾斜角度 α（或 β）不变；在第二种情况下，下一轮循环旋转后得到的倾斜角度 θ' 小于等于 $\theta/2$，则将 θ' 赋值给 α（或 β）。如果这两种情况一直出现，那么循环一直到 $\theta \le 1°$ 为止。如果连续几轮循环后的高度与宽度的比值均保持不变，说明图像倾斜校正已经完毕或者是原来的图像根本就不倾斜。

校正后的倾斜角度不一定要是0，因为当倾斜角度小于一个极限值时，对识别结果几乎就没有影响了。假设这个极限值为 ε，若在第n次校正角度小于等于 ε，那么可以计算该算法的循环次数：

$$\frac{\theta}{2^{n-1}} \le \varepsilon \Rightarrow \frac{\theta}{\varepsilon} \le 2^{n-1}$$

$$n \ge \log_2\frac{\theta}{\varepsilon} + 1$$

因为倾斜角度的最初值为45°，最小的极限值是1°，所以由上面的公式可以算出这个算法所需要的循环次数，n为8。也就是说，只要倾斜角度在90°以内，总可以用小于8次的循环来完成校正工作。

从上面所描述的算法过程可以看出，该算法在每次循环中所能校正的倾斜角度分别是：θ，$\theta/2$，$\theta/4$，…，由此可知该算法能校正的最大倾斜角度就是将每一次循环所校正的角度进行总和：$\theta + \theta/2 + \theta/4 + \ldots + \theta/2^{n-1} + \cdots = \frac{\theta}{1-1/2} = 2\theta$ =90°。也就是说，本文所提出的算法能对倾斜角度在90°以内的任何手写数字的图像进行校正。当然，也可以通过改变 θ 值的大小来改变所能校正的倾斜角度的范围。

此外，在图2中给出了几个校正前后的手写体数字的图像的对比（上面一行是校正以前的图像，下面一行是校正以后的图像）：

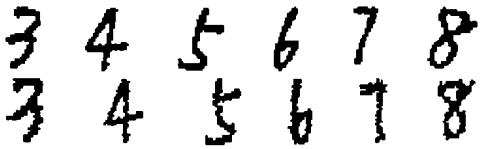


图2 几个手写体数字的校正前后对比

从图2中可以看出，经过倾斜校正后，手写数字的形状并没有发生任何变化，只是多了一些像毛刺一样的东西，但这些并不会影响分类器的识别结果。因为在校正过程后还要进行细化（即提取骨架）的工作，在细化过程中这些所谓的毛刺也都被处理掉了。由于细化的工作不是本文讨论的内容，因此不详细解释。

2 与参考文献[1]中所用算法的比较

参考文献[1]中所用的算法大体思想是利用了这样的一个现象：当手写体数字的图像倾斜度最小时，它的宽度达到最小。假设图像中原来的像素坐标点为 (x, y)，旋转后的像素点坐标为 (x', y')，该算法对图像中像素点的旋转是这样的：$x' = x - y\times\tan(\theta)$，$y' = y$。它的旋转中心是图像的左下角；同本文的其他假定一致：旋转方向也是顺时针为正逆时针为负；旋转角度 θ 最初的大小也是45°，递减的规律也是 $\theta \leftarrow \frac{\theta}{2}$，角度的最小极限也是1°；判定倾斜与否的标准是图像的宽度是否达到了最小（连续几次循环宽度值都不变）或者旋转角度是否达到了最小极限（$\theta \le 1°$）。

(a)

原图　　(b) 用本文方法所作倾斜校正　　(c) 用文献[1]中的方法所作倾斜校正

图3 本文所用方法与参考文献[1]中所用方法的比较

由图3可见，本文中方法所做的倾斜校正并不改变图像

（下转第137页）

自适应视频编码是根据网络带宽调整视频编码的速率控制技术，目的是在给定码率情况下获得最佳的编码质量。本文采用MPEG4和H.263+ 压缩标准，控制视频压缩码率。

发送速率控制是根据网络带宽将视频传输速率规约到目标码率范围之内的速率控制技术。经过速率调整，视频流的输出速率可以和网络带宽相匹配。在本系统中，发送端控制视频数据的传输速率，直到与网络带宽相匹配。

5 实验分析

本文对网络自适应传输控制算法在局域网内进行了模拟实验。实验采用的硬件配置是两台P4计算机，分别用作发送端和接收端。操作系统是Win2000，网络为10M/100Mbps自适应以太网，采用Foreman（QCIF）视频序列，初始发送速率为512kb。

试验的数据统计和带宽模拟是采用Sniffer Pro软件完成的，该软件用于网络监控和状态分析。为了在局域网内模拟带宽变动的情况,这里使用Sniffer Pro中的网络群发功能，定时向网络中发送一些与测试无关的数据的方式来占用一部分带宽，通过这种方式造成局域网内的带宽波动。然后分别测试在这种带宽变化的条件下，使用和不使用自适应策略算法对接收端的丢包率的影响。以下是模拟环境中的试验结果数据曲线图。

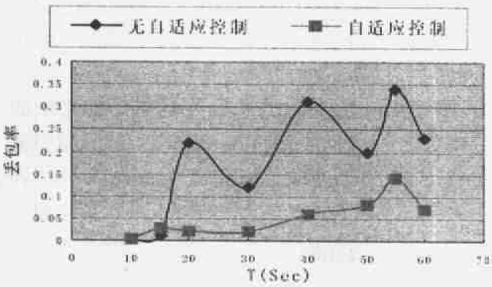
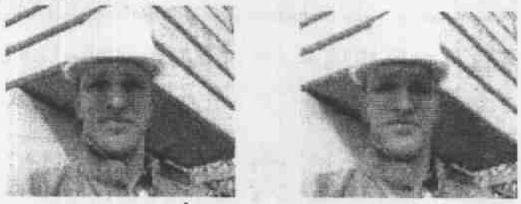


图4 自适应实验效果对比图

比较图4中的两条曲线，容易看到在使用网络自适应传输控制技术之后，接收端的丢包率明显下降。这是因为使用网络自适应传输控制后，当服务器监测到网络出现拥塞，带宽下降后，服务器会自动调整发送策略，即从占用较多带宽，质量较高的视频流向占用带宽较少，质量稍微差一点的视频流上转移。因此如果出现持续的低带宽情况，网络自适应传输控制算法的表现结果就是在最初阶段丢包率会较大，随后就会逐渐适应低带宽的环境，丢包率会呈下降趋势。

图5显示了视频传输系统在采用网络自适应传输控制技术后的图像效果对比。



(a) 有自适应控制　　(b)无自适应控制

图5 有无网络自适应传输控制对比

根据视频压缩原理，在多数情况下，视频流系统只传送图像的变化部分。当视频数据包在传送过程中丢失后，解码器进行帧间预测就变得较为困难，这时就会在图像中出现一些粗大的色块。这种色块越多就表明接收到的图像帧越不完整。上述图像效果的对比，验证了网络自适应传输控制技术的合理性。

6 结论

从实验结果可以看出，网络自适应传输控制技术可对网络带宽的变动作出及时的响应，并能有效地避免网络拥塞，保证了视频的传输质量。同时，该技术对视频编码器的依赖小，不需要复杂的视频自适应编码算法，也不占用冗余的存储空间，因此在对视频流的实时性要求高的应用上具有较突出的优势。该技术为实时视频传输提供了有效的拥塞控制机制。随着网络技术和视频通信技术的发展，还需要对该技术进行深入的探讨和研究。

参考文献

1 Wu D, Hou Y T, Zhu W, et al. Streaming Video over the Internet : Approaches and Directions. IEEE trans on CSVT, 2001,11(3):282-300

2 Wu D, Hou Y T, Zhu W, et al. On End to End Architecture for Transporting Mpeg-4 Video over the Internet. IEEE Trans.on CSVT,2000, 10:923-941

3 Jacobs S, Eleftheriadis A. Streaming Video Using TCP flow Control and Dynamic Rate Shaping. Journal of Visual Commun. and Image Representation, 1998-09

4 Floyd S, Fall K. Promoting the Use of End-to-end Congestion Control in the Internet. IEEE Trans. On Network, 1999,7: 458-472

5 Handley M, Pahye J, Floyd S. TCP Friendly Rate Control(TFRC): Protocol Specification. IETF Internet Draft, Draft-ietf-tsvwg-tfrc-02, 2001-05

☆☆

（上接第129页）

原来的大致形状，保持了原有图像的所有性质。而文献[1]中所用方法由于仅改变横坐标而不改变纵坐标，因此会改变图像的形状，这是在识别系统的预处理阶段不希望得到的结果，因为图像形状的改变可能会对识别率造成一定的影响。

3 总结

本文提出了一种对于孤立的手写体数字的图像进行倾斜校正的新方法。该方法是通过计算二值图像的高度和宽度的比值来估计图像的倾斜角度，该比值越小，则倾斜角度越大，反之，倾斜角度就越小。因为通过观察，我们注意到，当某一手写体数字的二值图像的高度和宽度的比值达到最大时，该图像就最接近于直立。用此方法在对20多人写的300多个数字的倾斜校正后，得到了令人满意的效果。

参考文献

1 Nagendraprasad NV, Gupta A, Feliberti V. Correction of Slant in Handwritten Numerals for Automated Document Processing [J].Engng. Applic. Intell., 1995, 8(4): 469-472

2 Khotanzad A, Yaw Hua Hong. Invariant Image Recognition by Zernike Moments [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1990, 12(5):489-497

3 Reiss T H. The Revised Fundamental Theorem of Moment Invariants [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 1991, 13(8): 830-834

4 Trier O D, Jain A K, Taxt T. Feature Extraction Methods for Character Recognition——A survey [J]. Pattern Recognition, 1996, 29(4): 641

5 Slavik P, Govindaraju V. Equivalence of Different Methods for Slant and Skew Corrections in Word Recognition Applications [J]. IEEE Transaction on Pattern Analysis and Machine Intelligence, 2001, 23(3): 323-326

6 何　斌, 马天予, 王运坚等. Visual C++数字图像处理(第二版)[M]. 北京：人民邮电出版社，2002