

基于纹理谱描述子的文档图像检索算法

马磊, 刘江

(山东山大鸥玛软件有限公司 数据研究中心, 济南 250100)

(lyjziyu@126.com)

摘要:新算法首先根据文档图像的特点分割图像文本区域,并将文档图像中字符的边缘信息使用纹理谱进行描述,计算纹理谱图像的直方图。相对于直接使用灰度直方图进行图像检索,该算法具有更好的区分度。实验结果表明,该方法具有很高的查准率,并对剪切、旋转操作表现出很好的稳定性,适合文档图像检索。

关键词:文档图像;图像分割;图像纹理;基于内容的图像检索;图像匹配

中图分类号: TP391.41 **文献标志码:** A

Document images retrieval based on texture spectrum descriptor

MA Lei, LIU Jiang

(Research Institute of Data Processing, Shandong Shanda Oumasoft Company Limited, Jinan Shandong 250100, China)

Abstract: The paper proposed a new method for document images retrieval based on texture spectrum descriptor. This algorithm firstly segmented the text area by document image characteristics, and then described the texture spectrum based on the text character edge, and computed the texture spectrum histogram. Compared with direct use of gray histogram image retrieval, this algorithm possesses high distinguish degree. The experimental results show that this method has high retrieval precision and resists cutting or rotating operations, and is suitable for document image retrieval.

Key words: document image; image segmentation; image texture; Content-based Image Retrieval (CBIR); image matching

0 引言

随着数字采集技术的发展,文档数字图像的规模也越来越大,图像的信息获取受到研究者的重视,基于内容的文档图像检索具有广泛的应用价值和前景。该检索方法针对文档图像的特点,使用纹理谱特征刻画文档图像中字符的特征,实现图像的相似度匹配。

基于内容的图像检索(Content-based Image Retrieval, CBIR)是利用图像本身的信息,通常以图像特征(颜色、纹理、形状与结构布局等)的相似性为检索依据,根据每幅图像都有的可比较特征进行检索。图像纹理反映的是图像的一种局部结构化特征,具体表现为像素点邻域内灰度级或者颜色的某种变化,而且这种变化是空间统计相关的,它由纹理基元和基元的排列两个要素构成,纹理分析方法有统计方法、结构方法和基于模型的方法^[1]。

图像的统计特征例如能量、方差、熵可以定义在其纹理特征上^[2],并可以实现感兴趣区域的内容检索。使用纹理直方图统计特性具有很高的效率和稳定性,对于图像的旋转、缩放,甚至剪切操作表现出良好的鲁棒性^[3]。纹理特征还可用于多目标人脸的检测^[4]。纹理特征与颜色特征相结合^[5]可达到较好的检索效果,相对于灰度直方图来说,基于局部不变特征的直方图在图像检索中是具有优势的^[6],另外纹理谱还可用于图像确认^[7]及图像分析^[8]。由于一般的文档图像字符边缘形态丰富,其纹理特征既具有一定的稳定性(相对于图像本身来说),同时又具有一定的区分度(不同文档图像之间),相比灰度直方图具有很大的优势,本文通过实验结果的分析也证明了这一点。

针对文档图像,研究纹理谱算子在图像检索中的应用,并

分析不同的文档图像类型其纹理谱的特点并作为全局统计特征。对于图文混合图像,使用行块分割定位文本区。文章给出了实验结果,并对实验结果进行了分析。

1 文档区域的定位

1.1 文档图像的特点

文档图像的检索,即使用纹理谱特征描述文档字符的特征。实际上,文档图像往往包含部分自然图像、表格、图形,因此从实用角度出发,检索算法先实现文本区的定位,用来提高检索算法的查准率。

文档图像的文本行的特点主要有以下几个:

- 1) 文本行中字符边界水平梯度较大;
- 2) 单个字符的各个部分可能不是一个完整的连通域,使用连通性算法复杂度相对较高;
- 3) 两个相邻的字符之间有一定的间隙;
- 4) 文本行之间有一定的间隙;
- 5) 文本行具有直线特点;
- 6) 字符在尺寸、语言类别、颜色、字体等方面复杂。

1.2 文本区计算方法

已经有很多算法用于在复杂图文混合图像中抽取文本信息^[9],本文使用文本块标记算法实现文档图像区域的定位,水平梯度差大的位置表示可能的文本行块,本文使用 $[-1, 0, 1]$ 对图像进行滤波。如图1所示,待计算梯度图像中包含了Lena图像和文本图像两部分:1) 图像中的黑线代表检测的行(第345行);2) 图像显示了检测行的梯度计算结果,即相邻像素灰度级的差值。水平方向上的梯度文本区是较大的,可以充分利用这一特性检测可能的文本区域。

单个字符之间、两个相邻字符之间是有一定间隙,使用连

收稿日期:2010-03-31;修回日期:2010-07-27。

作者简介:马磊(1960-),男,山东济南人,研究员,博士,主要研究方向:数字图像识别、高速数据采集;刘江(1979-),男,山东济南人,工程师,硕士,CCF会员,主要研究方向:文档图像处理、海量图像检索算法。

通性检测算法并合并相邻字符获取文本块会增加算法的复杂性,本文使用梯度差获取可能的文本块,使用一个固定宽度为 n ,高度为 3 的模板实现字符之间的合并, n 的选取大于字符间距即可,高度为 3 时,可确保两个行块之间应至少大于两个像素的距离,而这种要求通常情况下可以满足。

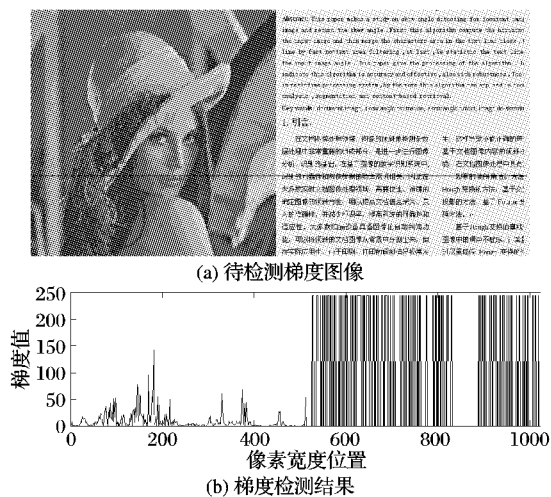


图 1 Lena 图像与文本图像的梯度值比较

如图 2 所示,中间为像素的梯度,当梯度大于一个给定的阈值,就进行目标扩展,从而实现文本行字符块的合并。

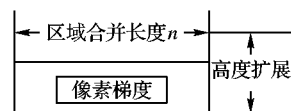


图 2 像素梯度扩展

1.3 纹理谱计算方法

文本行块用于正确地定位文档图像中的字符,避免噪声对该检索方法的影响,使用图像的二维模式刻画字符的基本形态,基本形态的频数统计得到纹理谱。

如图 4 所示,纹理谱考查一个目标中心点周围的像素形态,8 个位置可组合为 256 种基本形态,使用 0~255 表示目标中心的纹理值,图中标识了相应位置的权重系数,权重系数的选择不限于图 4 中的表示方法,只要能够区分出不同的纹理形态即可,在进行匹配算法时,两个图像纹理谱计算使用相同的权重系数即可,在这种情况下,相关性计算与权重系数选择无关。

记 8 个位置为 $p_i (i = 0, 1, \dots, 7)$, 则纹理谱的值计算为:

$$v = \sum_{i=0}^7 2^{p_i}; p_i \in \{0, 1, 2, 3, 4, 5, 6, 7\} \quad (1)$$

该纹理值反映了中心点周围的 8 个位置的像素值分布情况,即在较小的区域内像素关系(纹理形态),对该值的统计显示该文档图像的纹理形态的分布状况,可反映一般意义上的纹理特征,纹理特征的区分度为 256 种。

如图 5 所示,分别计算 Lena 图像、圆、矩形图像的纹理谱直方图。纹理谱图像如图(b)所示,计算图(b)的直方图得到图(c)。纹理谱直方图统计结果表明: Lena 图像的谱直方图包含有多种纹理形态,圆图像的谱直方图含有较少的纹理形态,矩形谱直方图纹理形态只有 4 种。为了说明图像之间的相似程度,可计算直方图之间的相关性,计算结果表明 Lena 图像与正方形和圆形的相关值接近于 0,表现为不相关,圆与正方形的相关值为 0.61,表现为部分相关,这恰好说明纹理

谱具有刻画图像内容的能力。如果直接使用灰度直方图进行相关值计算,正方形图像与圆形图像的相关值为 0.999,不能有效区别两个图像的内容,因此使用谱直方图进行内容检索是有效的、合适的方法。

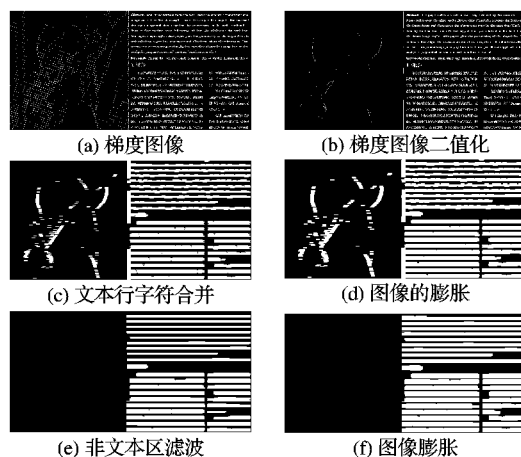


图 3 文本区域的定位过程

7	6	5
4	*	3
2	1	0

图 4 纹理谱计算方法

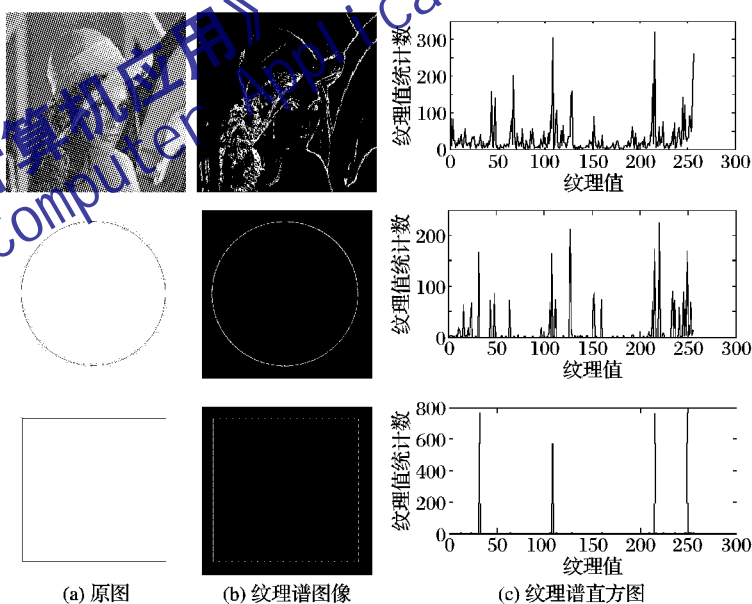


图 5 纹理谱图像与直方图

2 纹理谱匹配算法

为提高检索算法的查准率,提出了使用版面分析的方法定位文本行块。

假设文本图像的纹理直方图为 $h_1[k]$ 、 $h_2[k]$, 则直方图相关性表示为:

$$\text{corr}(h_1[k], h_2[k]) = \frac{\sum_{k=0}^{255} (h_1[k] - \bar{h}_1) \times (h_2[k] - \bar{h}_2)}{\sqrt{\sum_{k=0}^{255} (h_1[k] - \bar{h}_1)^2 \times \sum_{k=0}^{255} (h_2[k] - \bar{h}_2)^2}} \quad (2)$$

相关值用于度量文本区基本形态的相似性,值的取值范围为 $[-1, 1]$ 。当计算值为 1 时,表示两个图像的基本形态完全相同。

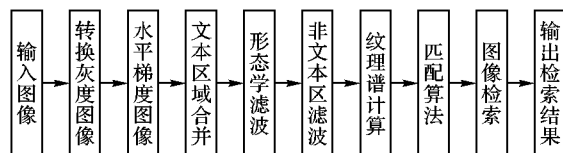


图6 算法的流程

3 实验结果与分析

为了测试该算法可靠性,进行相关的实验,一组实验检测同一个图像抗图像旋转的能力,如表1所示。

表1 偏转角度与相关值的关系

偏转角/(°)	相关值	偏转角/(°)	相关值
1	0.9450	16	0.9206
5	0.9457	21	0.8923
11	0.9385	26	0.8688

如表1所示,偏转角度为1°时,相关值计算为0.945;偏转角度为26°时,相关值具有较好的保持特性,计算为0.8688。

为了测试该算法检索相似图像的能力和区分不同内容图像的能力,使用99个不同人的手写体英语(图像中字符数相当,数据来源于高三英语答卷扫描图像,扫描分辨率100 DPI,256级灰度图像),两两进行相关值计算,共计4950个不同的相关值,便于比较检索结果与主观评价是否一致。

如图7所示,通过计算不同图像的纹理谱直方图的相关性,得到检索算法给出的高相似性和低相似性对应的图像,统计表明,相关值大于0.6的占总数的92.1%,高相关值(大于0.9)占总数的34%,相关性最低为0.11,对比原始图像,检索结果与主观感受是一致的(如图8所示)。

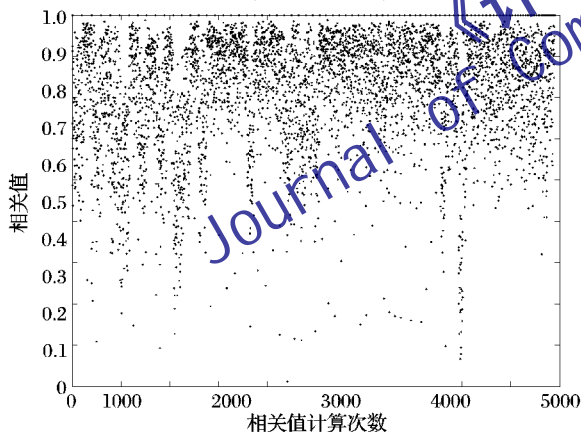


图7 相关值计算结果(英语手写体)

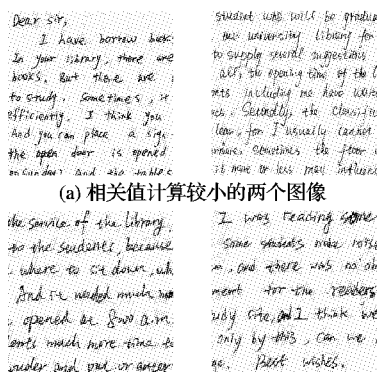


图8 检索结果的图像示例

如图8所示:图(a)计算结果显示出两幅图像的纹理具有很强的不相关性,相关值计算为0.11,主观上可以判断出

这是两个不同人答卷;图(b)两幅图像的相关值计算较高,计算为0.99,主观上不容易断定这是两个人的答卷。这说明该算法对文档图像的内容具有较好的区分能力,如果直接使用灰度直方图计算这两组数据的相关性,其相关值均在0.9以上,不能进行有效地区分。

如图9所示,99幅不同人的手写体英语图像,使用图像一半区域与整图的纹理谱计算相关值,子图像与原图表现出高度相关性,最小值为0.9853,该结果显示出该算法对于子图检索具有很好的稳定性,结果与实际的主观感受是一致的。

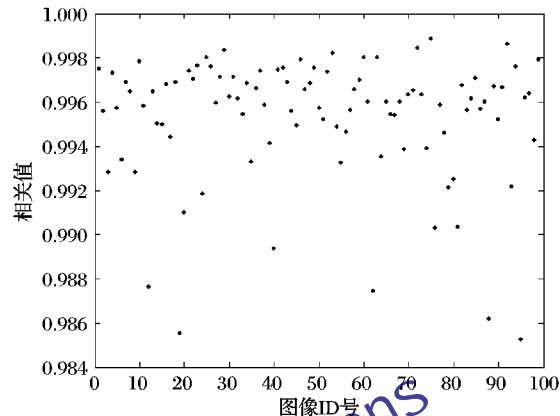


图9 图像剪切相关值计算结果

4 结语

实验结果表明,该方法具有很强的抗噪声能力,适合图文混合文档图像检索,纹理谱方法可有效区分文本基本形态,对于图像剪切和旋转操作表现出很好的稳定性,相对于灰度直方图检索方法,具有更好的区分能力,在文档检索领域具备实用性。另外,该方法可用于不同语种的文档图像分类、笔迹鉴别等相关领域。

参考文献:

- [1] 施智平, 胡宏, 李清勇, 等. 基于纹理谱描述子的图像检索[J]. 软件学报, 2005, 16(3): 1039-1045.
- [2] 沈项军, 崔志明, 龚声蓉, 等. 一种基于感兴趣内容的图像检索方法[J]. 计算机工程, 2008, 34(1): 224-226.
- [3] FAN ZHIGANG, LI JILIN, WU BO, et al. Local patterns constrained image histograms for image retrieval [C]// 15th IEEE International Conference on Image Processing. Washington, DC: IEEE, 2008: 941-944.
- [4] AN K H, PARK S H, CHUNG YUNSU, et al. Learning discriminative multi-scale and multi-position LBP features for face detection based on Ada-LDA [C]// 2009 IEEE International Conference on Robotics and Biomimetics. Washington, DC: IEEE, 2009: 1117-1122.
- [5] VIET C N, MARSHALL I. An efficient obstacle detection algorithm using color and texture [EB/OL]. [2010-02-10]. <http://www.waset.org/journals/waset/v60/v60-22.pdf>.
- [6] HALAWANI A, BURKHARDT H. On using histograms of local invariant features for image retrieval [EB/OL]. [2010-02-10]. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.116.6736&rep=rep1&type=pdf>.
- [7] HEUSCH G, RODRIGUEZ Y, MARCEL S. Local binary patterns as an image preprocessing for face authentication [C]// Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition. Washington, DC: IEEE, 2006: 9-14.
- [8] PIETIKÄINEN M. Image analysis with local binary patterns [EB/OL]. [2009-12-10]. <http://www.springerlink.com/content/ce5kvfyqrbtxym3/fulltext.pdf>.
- [9] GOPALAN C, MANJULA. Text region segmentation from heterogeneous images [J]. IJCSNS International Journal of Computer Science and Network Security, 2008, 8(10): 108-113.