

基于形状不变矩的图像检索算法的研究

黄 勇, 王崇骏, 王 亮, 杭 燕, 陈兆乾

(南京大学 计算机软件新技术国家重点实验室, 江苏 南京 210093)

摘要: 描述了一种基于 Hu 形状不变矩的图像全局形状特征提取方法和算法 IMS。实验结果表明, 使用 IMS。算法提取的形状特征向量具有对平移、旋转和尺度变化的不变性, 适合于进行图像形状的检索。

关键词: 图像检索; 图像形状; 形状特征提取; 形状不变矩

中图法分类号: TP393 文献标识码: A 文章编号: 1001-3695(2004)07-0256-02

Research on Image Retrieval Algorithm Based on Moment Invariant

HUANG Yong, WANG Chong-jun, WANG Liang, HANG Yan, CHEN Zhao-qian

(State Key Laboratory for Novel Software Technology, Nanjing University, Nanjing Jiangsu 210093, China)

Abstract: Describes an extraction method of global shape feature of image based on Hu moment invariant and the correlative algorithm IMS. The experimental results illuminate that the shape features extracted by IMS approach are invariant to translation, rotation and scaling to be suitable for shape-based image retrieval.

Key words: Image Retrieval; Image Shape; Shape Feature Extraction; Moment Invariant

形状特征索引技术的研究目前已在基于内容的图像检索中得到充分重视。Jain 等人^[1]使用封闭直线段描述形状, 并计算形状中每条直线段的斜率, 然后按固定的角度间距统计这些直线段的斜率直方图, 并计算形状的不变矩, 通过斜率直方图和不不变矩的加权距离得出图像形状的相似程度。斜率直方图的概念来自于 Jain 等人曾经提出的图像边缘点直方图, 并分析了该方法的优缺点, 试图通过对直方图的平滑解决图像有细微旋转的问题。对 400 多幅商标图像进行了实验, 发现将形状特征与颜色特征结合进行检索, 取得了更好的效果。1972 年, Zahn C T 等人^[2]就使用傅里叶描述子(Normalized Fourier Descriptors)描述形状的边界信息, 并发现对这些系数反变换后仍可恢复其边界信息, 且傅里叶描述子具有良好的平移、旋转和伸缩不变性, 非常适合构造形状的边界信息。文献[3]在形状检索过程中使用这些傅里叶系数之间的欧基里德距离来判断形状相似性, 具有较好的检索效果。文献[4]使用一种归一化的傅里叶描述子进行分层次的形状检索, 在实验中也取得了较为满意的效果。

目前, 基于形状特征的图像检索大致分为两类: 基于图像对象区域的局部(Local)形状特征检索和基于全局(Global)形状特征的检索。前者是指在特征提取阶段, 提取图像中相关对象的形状区域或边界等完整描述, 然后利用该区域及其特征进行图像检索, 该方法要求对图像进行图像分割操作, 其检索规模和应用范围受到一定限制^[5,6]; 而后者则无须进行图像分割, 直接利用能反映图像形状的全局特征量来进行图像检索^[7], 比较适合于大型的通用图像数据库。

在本文中, 我们针对国家高技术“863”项目“智能化电脑刺绣编程系统”中所需图标、饰物和商标等图案的检索需求, 研究基于 Hu 形状不变矩的图像全局形状特征提取, 实现了

IMS 形状检索算法。对该算法, 我们分别进行了相应的图像检索实验, 并提供了其图像检索实例, 效果很好。

1 基于形状不变矩的 IMS 检索算法

1.1 形状不变矩特征提取方法

作为一组关于形状统计值, 矩不变量的表示形式有多种, 如 Legendre 矩、Zernike 矩等。我们对图像的形状特征提取采用 Hu 矩表示, 其在图像像素数目一定的情况下, 具有对图像的旋转(Rotate)、平移(Shift)和尺度变化(Scale)的不变性^[9]。

定义 1 图像的 $(p+q)$ 阶矩。给定二维连续函数 $f(x, y)$, 式(1)定义了其 $(p+q)$ 阶矩:

$$M_{pq} = \int \int x^p y^q f(x, y) dx dy, \quad p, q = 0, 1, 2, \dots \quad (1)$$

按照式(1), 一幅大小为 $M \times N$ 的二维图像 $\{f(i, j) \mid i=0, \dots, M; j=0, \dots, N\}$, 其 $(p+q)$ 阶矩定义如式(2)所示:

$$M_{pq} = \sum_{i=1}^M \sum_{j=1}^N f(i, j) \cdot i^p j^q, \quad p, q = 0, 1, 2, \dots \quad (2)$$

p 和 q 可取所有的非负整数值, 它产生一个矩的无限集。根据帕普利斯(Papoulis)唯一性定理, 这个无限集完全可以确定二维图像函数 $f(i, j)$ 本身。其零阶矩只有一个, 即 $M_{00} = \sum_{i=1}^M \sum_{j=1}^N f(i, j)$ 。对于二值图像, 令其背景值为 0, 形状区域内值为 1, 则零阶矩表示该形状区域的面积。每一个图像矩除以零阶矩得到的商具有形状的尺度变化无关性。

定义 2 图像的 $(p+q)$ 阶中心矩。为了保证形状特征的位置不变性, 还必须计算中心矩, 即以物体的质心为原点计算图像的不不变矩值。由零阶矩和一阶矩可以算出图像的质心 (i, j) :

$$(i, j) = (M_{10}/M_{00}, M_{01}/M_{00})$$

则该图像矩阵的 $(p+q)$ 阶中心矩计算公式如式(3)所示:

$$m_{pq} = \sum_{i=1}^M \sum_{j=1}^N f(i,j) \cdot (i-i)^p (j-j)^q, \quad p, q = 0, 1, 2, \dots \quad (3)$$

如果对中心矩根据形状区域面积进行归一化, 即将 m_{pq} 用 $m_{pq}/M_0^{(p+q)/2}$ 代替, 则得到的不变矩可以满足尺度无关性。这里的尺度为形状区域的尺度, 而不是整个图像的尺度。

定义 3 图像的 Hu 不变矩。计算图像的七个面积归一化中心矩 $\{m_{11}, m_{02}, m_{20}, m_{21}, m_{12}, m_{03}, m_{30}\}$ 。Hu 不变矩即是关于这七个矩的函数, 满足形状的平移、旋转和尺度不变性。其计算如式 (4) 所示:

$$\begin{aligned} u_1 &= m_{20} + m_{02} \\ u_2 &= (m_{20} - m_{02})^2 + 4m_{11}^2 \\ u_3 &= (m_{30} - 3m_{12})^2 + (3m_{21} + m_{03})^2 \\ u_4 &= (m_{30} + m_{12})^2 + (m_{21} + m_{03})^2 \\ u_5 &= (m_{30} - 3m_{12})(m_{30} + m_{12}) [(m_{30} + m_{12})^2 - 3(m_{21} + m_{03})^2] + \\ &\quad (3m_{21} - m_{03})(m_{21} + m_{03}) [3(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2] \\ u_6 &= (m_{20} - m_{02}) [(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2] + 4m_{11}(m_{30} + m_{12}) \\ &\quad (m_{21} + m_{03}) \\ u_7 &= (3m_{12} - m_{30})(m_{30} + m_{12}) [(m_{30} + m_{12})^2 - 3(m_{21} + m_{03})^2] + \\ &\quad (3m_{21} - m_{03})(m_{21} + m_{03}) [3(m_{30} + m_{12})^2 - (m_{21} + m_{03})^2] \end{aligned} \quad (4)$$

基于矩的形状描述可以保存原有形状的信息, 矩 M_{pq} 也足以重建原始函数 $f(x, y)$ 。但是, Hu 不变矩的不变性是在连续情况下, 假设图像像素数目在线性变换过程中保持不变的条件下进行的。而在离散情况下, 图像变换是不连续的, 除平移变换外, 旋转和尺度变换均会导致图像像素数目的变化, 从而使计算结果产生误差。由于图像的大小会影响不变矩特征值, 因此, 在进行图像相似性匹配时, 应首先将图像库中的图像规范为相同尺寸大小的图像。对于一幅彩色 RGB 图像, 计算其 Hu 不变矩特征量的过程如图 1 所示。

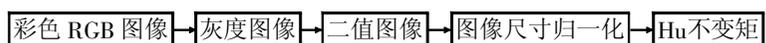


图 1 彩色图像 Hu 不变矩计算流程

得到二值图像后, 就可以实现提取其不变矩特征的算法。考虑到矩的计算代价较大, 因此, 可以预先使用算法提取出其矩不变量特征, 存放于图像的形状特征索引库中, 然后提供给图像检索算法执行图像查询。

定义 4 图像离心率 e 。为提高图像检索的准确度, 可以从 Hu 不变矩定义图像的离心率 (Eccentricity) 特征量 e , 按式 (5) 计算:

$$e = \frac{(m_{20} - m_{02})^2 + 4m_{11}^2}{(m_{20} + m_{02})^2} \quad (5)$$

离心率特征 e 表示了图像的最大轴向与图像的最小轴向的比率, 同样也满足物体平移、旋转和尺度不变性。将 Hu 不变矩特征量 $u_1 \sim u_7$ 和离心率 e 合并, 就形成了图像的全局形状特征向量 S 。该特征代表的是一种全局的形状特征。

1.2 IMS 形状检索算法

IMS 形状检索算法的基本思想就是用图像的 Hu 不变矩 $u_1 \sim u_7$ 和离心率 e 作为图像的形状特征索引, 使用适当的相似性距离定义, 计算出两幅图像的相似性距离。当距离足够小时, 就认为两幅图像是相似的。IMS 形状检索算法步骤如下:

(1) 用户选择待查询的示例图像, 提交检索请求。

(2) 计算示例图像的形状特征 S , 并从图像数据库中获取图像的形状特征索引。

(3) 计算图像数据库中各图像 (形状特征 $S = \{S_i | i = 1, 2, \dots, 8\}$) 与示例图像 (形状特征 $S = \{S_i | i = 1, 2, \dots, 8\}$) 之间的相似性距离 $dist$, 如式 (6) 所示:

$$dist = \sum_{i=1}^8 \frac{|S_i - S_i|}{|S_i|} \cdot w_i \quad (6)$$

其中, w_i 为各个特征分量所具有的权重。由于高阶矩跟形状的关联较小, 在进行基于内容的图像检索时, 重点使用低阶矩, 其权重较大, 如一阶矩可以用来计算区域的中心, 二阶矩也就是惯量矩, 可以用来计算形状的主轴。另外, 对于电脑刺绣中所需的图标、饰物和商标等图案, 离心率 e 的检索效果较好, 其权重也相对较大。

此外, 为了保证式 (6) 所计算出的相似性距离的有效性, 该距离使用示例图像的形状特征值进行了归一化。

(4) 若距离小于指定阈值, 即 $dist < \theta$, 则将对应图像加入图像检索结果集中。

(5) 将图像检索结果集图像排序输出。

(6) 算法结束。

2 实验结果与分析

我们首先在 260 幅图像的图像库中对本文中基于 Hu 不变矩的 IMS 形状检索算法进行了实验。图像库中包括电脑刺绣图标、饰物和商标等图案, 图像尺寸大小不等, 以 JPEG 文件格式存取, 均对其 128×128 像素大小的缩略图进行不变矩特征提取。IMS 形状检索算法的图像检索实例如图 2 所示。

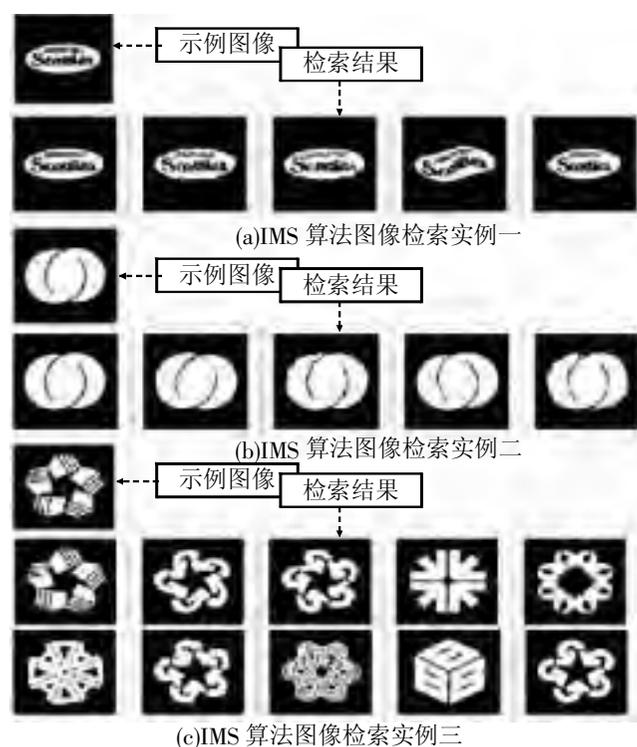


图 2 IMS 算法的图像检索实例

从图 2(a) 可以看出, IMS 形状检索算法对于图像的扭曲形变具有不变性, 并对图像的基本形状特性具有鲁棒性, 在具有一定形变干扰的情况下, 仍能得出较好的图像检索结果。图 2(b) 中的图像检索结果则表明, IMS 算法对于图像的旋转具有不变性。图 2(c) 中的检索实验则表明, 虽然检索结果排列的顺序有时与人的主观视觉判断不尽一致, 但与输入相似的图像基本都检索出来了, 检索精度和效果都较好。

3 小结

本文针对国家高技术 863 项目“智能化”(下转第 260 页)

的不利影响,而且基于水平和垂直边缘检测结合旋转投影设计了一种新的车牌倾斜校正算法。此算法克服了目前已有算法在车牌污迹严重时检测精度下降的缺点,不仅完成了车牌的倾斜校正,而且能够进行车牌上下位置精定位。实验结果表明,所提出的算法简单易行,不仅能够抑制光照不均,提高图像的对比度,而且在车牌图像有污迹、字符粘连和断裂以及车牌边框不明显时都能有效地进行倾斜校正,为后续的字符分割奠定了良好的基础。



图 9 水平倾斜校正后车牌的垂直边缘

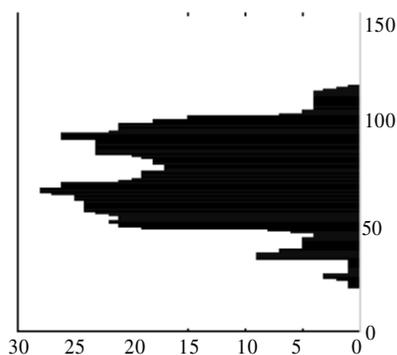


图 10 水平倾斜校正后垂直边缘的水平投影

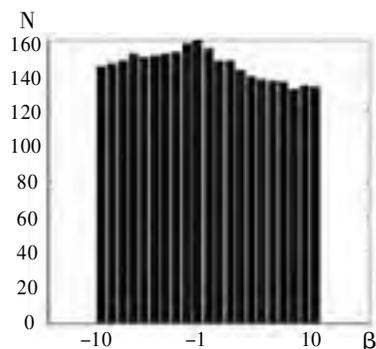


图 12 B-N 直方图



图 11 上下位置精分割



图 13 校正后车牌

参考文献:

[1] Demetrios Michalopoulos, et al. An Error Back-propagation Artificial Neural Networks Application in Automatic car License Plate Recognition[C]. *Developments in Applied Artificial Intelligence*, 2002. 1-8.

[2] Paolo Comelli, Paolo Ferragina, Mario Notturmo Granieri, et al. Opti-

cal Recognition of Motor Vehicle License Plates[J]. *IEEE Transactions on Vehicular Technology*, 1995, 44(4): 790-799.

[3] Jenn-Kwei Tyan, et al. A Character Segmentation Algorithm for Recognition of Vehicle License Plate[C]. *Part of the SPIE Conference on Mobile Robots XIV*, 1999. 12-21.

[4] M Shridhar, et al. License Plate Recognition Using SKIPSM[C]. *Machine Vision and Three-dimensional Imaging Systems for Inspection and Metrology*, Proceedings of SPIE, 2001. 72-79.

[5] Rafael C Gonzalez, Richard E Woods. *Digital Image Processing, Second Edition*[M]. 2002. 85-86.

[6] Chen C-Y, Yu C-C, Hun Z-D. A 3-D Transformation to Improve the Legibility of License Plate Numbers[J]. *Journal of Forensic Sciences*, 2002, 47(3): 578-585.

[7] 富煜清, 沈巍, 黄心晔. 汽车牌照的提取方法研究[J]. *模式识别与人工智能*, 2000, 43(3): 345-348.

[8] 郝永杰, 刘文耀, 路烁. 畸变汽车牌照图像的空间校正[J]. *西南交通大学学报*, 2002, 37(4): 417-420.

[9] Hans A Hegt, et al. A High Performance License Plate Recognition System[C]. *IEEE International Conference on Systems, Man, and Cybernetics*, 1998. 4357-4362.

[10] 权伟, 郑南宁, 贾新春. 复杂背景下的车辆牌照字符提取方法研究[J]. *信息与控制*, 2002, 31(1): 25-29.

[11] 乐宁, 翁世修. 基于单个字符区域的倾斜字符校正技术的研究[J]. *图像识别与自动化*, 2002, (2): 1-7.

作者简介:

李文举(1964-),男,副教授,博士研究生,研究领域为图像处理、模式识别和智能控制;梁德群(1940-),男,教授,博士生导师,研究领域为低级视觉、并行结构和图像检测;崔连延(1957-),男,教授,博士研究生,研究领域为智能控制;毕胜(1964-),男,副教授,博士研究生,研究领域为图像处理、模式识别和工业检测。

(上接第 257 页) 电脑刺绣编程系统”中所需电脑刺绣图标、饰物和商标等图案的检索需求,研究基于 Hu 形状不变矩的图像全局形状特征提取,实现了 IMS 形状检索算法;并进行了相应的图像检索实验。实验结果表明,使用上述算法提取出的形状特征向量具有对平移、旋转和尺度变化的不变性,适合于进行图像形状的检索。该算法已应用于智能化电脑刺绣编程系统,其效果很好。

参考文献:

[1] Jain A K, Vailaya A. Image Retrieval Using Color and Shape[J]. *Pattern Recognition*, 1996, 29(8): 1233-1244.

[2] Fuhn C T, Roskies R Z. Fourier Descriptors for Plane Closed Curves[J]. *IEEE Transactions on Computers*, 1972, (21): 269-285.

[3] Liu Y, et al. Modified Fourier Descriptor for Shape Representation: A Practical Approach[C]. *Proc. 1st International Workshop on Image Database and Multimedia Search*, Amsterdam, the Netherlands, 1996.

[4] Liu W Y, Wang T, Zhang H J. A Hierarchical Characterization Scheme for Image Retrieval[C]. *Proc. 7th IEEE International Conference on Image Processing (ICIP 2000)*, Vancouver, Canada, 2000.

[5] Chen Y X, Wang J Z. A Region-based Fuzzy Feature Matching Approach to Content-based Image Retrieval[J]. *IEEE Transactions on*

Pattern Analysis and Machine Intelligence, 2002, 24(9): 1252-1267.

[6] Jiang J Z. Pathfinder: Multiresolution Region-based Searching of Pathology Images Using IRM[J]. *Journal of American Medical Informatics Association, Symposium Supplement*, 2000, 2000suppl: 883-887 (LCDS).

[7] Jain A K, Vailaya A. Shape-based Retrieval: A Case Study with Trademark Image Databases[J]. *Pattern Recognition*, 1998, 31(9): 1369-1390.

[8] Fu M K. Visual Pattern Recognition by Moment Invariants[J]. *IRE Transactions on Information Theory*, 1962, IT-8(2): 179-187.

[9] Elkasim S O, Shridhar M, Ahmadi M. Pattern Recognition with Moment Invariant: A Comparative Study and New Results[J]. *Pattern Recognition*, 1991, 24: 1117-1138.

作者简介:

黄勇(1979-),男,硕士研究生,主要研究方向为图像处理、机器学习;王崇骏(1975-),男,博士研究生,主要研究方向为图像处理、机器学习;王亮(1979-),男,硕士研究生,主要研究方向为数据库、电子商务;杭燕(1979-),女,硕士研究生,主要研究方向为智能图像处理、机器学习;陈兆乾(1940-),女,教授,博士生导师,主要研究方向为机器学习、知识工程、神经网络。