

利用小波和矩进行基于形状的形象检索

姚玉荣 章毓晋

(清华大学电子工程系, 北京 100084)

摘要 形状是图象中目标的重要特征, 基于形状的形象检索近来在基于内容的图象库系统管理和应用中得到越来越多的重视。现已研制的系统存在两个问题, 一是性能的不稳定性; 二是相对于平移、旋转和尺度变换的变化性。针对以上问题, 该文提出了一种新的基于形状的形象检索算法。此算法先对亮度图象进行小波模极大值变换以得到多尺度的边界图象, 再利用7个不变矩提取每一尺度边界图象的特征, 所有尺度上的矩共同组成图象的特征向量。图象的相似度用图象特征向量的归一化加权欧氏距离表示。用服装图象数据库进行试验得到的结果表明该算法能较好地描述图象的形状及空间分布信息, 并具有较好的平移、尺度、旋转不变性。用更为复杂的花卉图象数据库进行试验得到的结果也证明了该算法的优越性。

关键词 形象检索 形状特征 小波 不变矩

中图分类号: TP391.3 **文献标识码:** A **文章编号:** 1006-8961(2000)03-0206-05

Shape-Based Image Retrieval Using Wavelet and Moment

YAO Yu-rong, ZHANG Yu-jin

(Department of Electronic Engineering, Tsinghua University, Beijing 100084)

Abstract Shape is an important feature of objects in image and shape-based image retrieval has obtained more and more attentions in recent research on content-based management and utilization of image database system. Although several systems have been developed, two main shortcomings are still existed. The first is that the performance is not stable. The second is that the variance with respect to translation, scaling, and rotation. To cure the above problems, this paper presents a novel shape-based image retrieval algorithm. The algorithm first transforms the luminance image with wavelet modulus maximum to get multi-scale edge images, then employs a set of seven invariant moments to extract the features of image. Consequently, each image is characterized by a multi-scale moment vector in feature space. Similarity is given by the Euclidean distance between two images' normalized moment vectors. Experimental results on clothes image database show that this algorithm can well capture the shape and spatial information of image and it is invariant with respect to translation, scaling and rotation of objects. In addition, the algorithm is also tested with more complicated flower images in a database; the experimental results further verify the effectiveness of the algorithm.

Keywords Image retrieval, Shape feature, Wavelet, Invariant moments

0 引言

近年来, 由于多媒体数据库的大量应用, 使得基于内容的形象检索成为一个活跃的研究领域。对基

于特征的形象检索系统来说, 图象的特征提取是关键模块。虽然很多形象检索系统已经建立起来了, 很多特征提取方法(基于颜色、形状、纹理等等)也已经被采用, 但以下两个问题仍旧存在: 基于图象的空间分布或形状特征的检索技术还不成熟;

平移、尺度、旋转变换不变性的问题仍没有得到很好的解决。针对以上两个问题, 本文采用小波变换模极大值(以下简称小波模极大值, wavelet modulus maxima)的方法来提取图象的多尺度目标边缘信息, 然后采用多尺度的不变矩作为特征来度量图象中目标的形状相似性。

1 算法介绍

1.1 小波模极大值

由于小波是处理多尺度可视化信息的强有力的数学工具, 近来有些图象检索系统利用小波变换后的系数为特征来解决利用图象的空间分布和形状特征进行相似性检索的问题^[1]。但是通常的通过规则抽样进行离散小波变换获得的小波系数缺少平移不变性^[2]。这个问题可以用自适应(不规则)的抽样方案来解决。Mallat^[3]建议用小波模极大值来描述信号的奇异性, 而信号的奇异点通常是信号中具有重要物理意义的点。对图象来说, 小波模极大值描述的是图象中目标的多尺度边界。小波模极大值是在对多尺度小波变换进行不规则抽样的基础上得到的, 用它可以得到平移不变性, 这样就克服了上面所说的小波变换的缺陷^[4]。

若 $\theta(x, y)$ 在整个平面上的积分为 1, 并且它在 x 或 y 为无限远处收敛到 0, 则定义 $\theta(x, y)$ 为二维平滑函数。现定义两个小波函数 $\psi(x, y)$ 和 $\psi^p(x, y)$ 为

$$\begin{aligned} \psi(x, y) &= \frac{\partial \theta(x, y)}{\partial x} \\ \psi^p(x, y) &= \frac{\partial \theta(x, y)}{\partial y} \end{aligned} \quad (1)$$

这样, 图象 $f(x, y)$ 的小波变换的两个分量在尺度为 s 时的定义为

$$\begin{aligned} W_s^1 f(x, y) &= f * \psi_s(x, y) \\ W_s^2 f(x, y) &= f * \psi_s^p(x, y) \end{aligned} \quad (2)$$

对于二进制小波变换, 就有

$$\begin{aligned} \begin{pmatrix} W^{2^j} f(x, y) \\ W^{2^j} f(x, y) \end{pmatrix} &= 2^j \begin{pmatrix} \frac{\partial}{\partial x} (f * \theta^j)(x, y) \\ \frac{\partial}{\partial y} (f * \theta^j)(x, y) \end{pmatrix} \\ &= 2^j (\nabla * \theta^j)(x, y) \end{aligned} \quad (3)$$

可以看出, 式(3)中小波变换的两个分量正比于梯度矢量 $(\nabla * \theta^j)(x, y)$ 的两个分量。在任一尺度 2^j , 梯度矢量的模等于

$$M^{2^j} f(x, y) = \sqrt{|W^{2^j} f(x, y)|^2 + |W^{2^j} f(x, y)|^2} \quad (4)$$

梯度矢量与水平轴的夹角为

$$A^{2^j} f(x, y) = \arg \left(W^{2^j} f(x, y) + W^{2^j} f(x, y) \right) \quad (5)$$

$f * \theta^j(x, y)$ 上变化剧烈的点是沿着梯度方向 $A^{2^j} f(x, y)$ 上模 $M^{2^j} f(x, y)$ 为局部极大值的那些点。我们只需记录下这些模极大值点的位置以及相应的模 $M^{2^j} f(x, y)$ 和角度 $A^{2^j} f(x, y)$ 的大小即可。

Mallat 已经证明^[3]:

- (1) 小波变换模极大值方法具有平移不变性;
- (2) 如果信号本身的傅立叶变换是带限的且小波函数 ψ 是紧支撑的, 则小波变换模极大值的表示是完备的。

受文献[5]的启发, 我们选择一维小波函数 $\psi(x)$ 的傅立叶变换满足

$$\Psi(\omega) = i\omega \left[\frac{\sin(\omega/4)}{\omega/4} \right]^4 \quad (6)$$

在二维情况下, 我们选择的是可分离的滤波器, 使得滤波器在一维时与式(6)中给出的小波相对应。对二维图象的小波分解就变成了分别按行和按列对图象做一维小波分解。

从视觉上看, 图象的小波变换的模极大值点位于图象的边界上(如图 1 所示), 也就是说, 变换的结果可以检测出图象的边界。容易看出检测出的边界图象是不同尺度上的多分辨率图象。图 1 中第一排的单幅图象是原始图象, 大小是 160×120 。第二排是 7 个尺度的模图象 $(M^{2^j} f)_{1 \leq j \leq 7}$, 尺度从左向右增加, 黑色对应的象素值为 0, 白色对应高的象素值。第三排显示的是模 $(M^{2^j} f)_{1 \leq j \leq 7}$ 的局部极大值点的位置。第四排给出的是模值大于给定阈值的那些局部模极大值点的位置。

1.2 不变矩

有了图象的小波变换表示之后, 下一步就是定义相似性度量。直接在小波变换域中度量两个图象的相似性既困难又不方便。区域的 7 个不变矩^[6]具有我们所期望的平移、尺度、旋转不变性, 所以这里我们采用 7 个不变矩来表示小波变换后的多尺度边界图象的特征。

值得注意的是, 通常所说的矩(如文献[6]中的定义)指的是区域的矩, 而这里图象小波变换的模极大值点所对应的是边界(曲线), 即是边界的矩。区域矩的公式不能直接用来计算曲线矩, 所以此处必须对不变矩的计算公式进行修正。

对于区域 $f(x, y)$ 来说, 若给它一个尺度变换 $x = \alpha x, y = \alpha y$, 它的矩就要乘以 $\alpha^l \alpha^l \alpha^2$, 其中因子

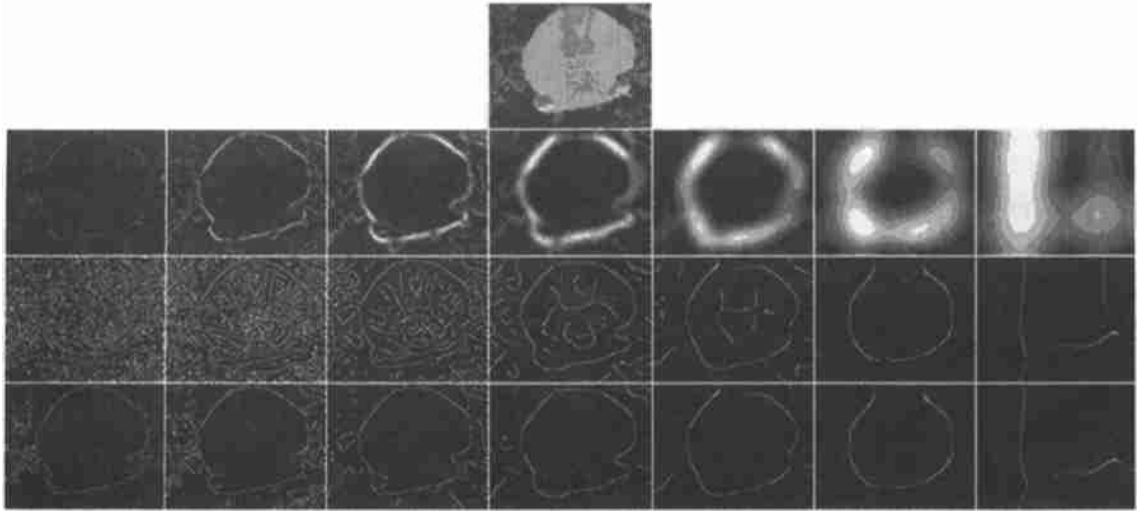


图1 图象小波变换的模极大值点

α^2 是由于尺度变化而带来的目标面积变化所引起的 $f(x, y)$ 的中心矩就变成了 $u_{pq} = u_{pq} * \alpha^{p+q+2}$. 归一化矩定义为

$$\eta_{pq} = \frac{u_{pq}}{(u_{00})^{\gamma}} \quad (7)$$

为了使得归一化矩对尺度变换是不变的, 必须有

$$\eta_{pq} = \eta_{pq}, \text{ 即 } \frac{u_{pq} * \alpha^{p+q+2}}{(u_{00} * \alpha^2)^{\gamma}} = \frac{u_{pq}}{(u_{00})^{\gamma}}, \text{ 所以可推出}$$

$$\gamma = (p + q + 2)/2 \quad (8)$$

对于曲线来说, 尺度的变化造成周长的变化, 变化因子相应是 α , 而不是 α^2 . 此时的中心矩就变成了 $u_{pq} = u_{pq} * \alpha^{p+q+1}$. 同样, 为了使归一化矩具有尺度

变换不变性, 必须有 $\eta_{pq} = \eta_{pq}$, 即 $\frac{u_{pq} * \alpha^{p+q+1}}{(u_{00} * \alpha)^{\gamma}} = \frac{u_{pq}}{(u_{00})^{\gamma}}$, 所以可推出

$$\gamma = p + q + 1 \quad (9)$$

请注意式(8)与式(9)之间的不同 这里我们使用了修正后的不变矩的计算公式(9).

1.3 特征向量归一化

由图象得到的所有矩(每层7个)共同组成了一个特征向量 特征向量中的每个特征元素都具有不同的物理意义, 而且它们的幅度也可能大不相同, 因此在以欧氏距离为度量时, 就会产生很大的偏差 必须通过特征归一化来消除这种偏差

我们将特征向量记为 $F = [f_1, f_2, \dots, f_j, \dots, f_N]$, N 是特征元素的个数, 用 I_1, I_2, \dots, I_M 代表图象库中的 M 幅图象 对图象 I_i , 其相应的特征向量记为 $F_i = [f_{i,1}, f_{i,2}, \dots, f_{i,j}, \dots, f_{i,N}]$ 这样由图象库中的 M 幅图象, 就可以得到一个 $M \times N$ 的特征矩阵

$F = f_{i,j}$, 其中 $f_{i,j}$ 是 F_i 的第 j 个特征元素 F 的每一列是长度为 M 的特征序列, 表示为 F_j . 设 F_j 是高斯序列, 计算其均值 m_j 和标准差 σ_j , 然后利用式(10)就可以将原序列归一化为 $N(0, 1)$ 分布的序列

$$f_{i,j} = \frac{f_{i,j} - m_j}{\sigma_j} \quad (10)$$

1.4 算法总结

现在将算法总结如下:

- (1) 对图象做小波分解, 得到多尺度的模图象;
- (2) 记录下小波变换域中模是局部极大值并且模大于事先设定的阈值的那些点, 这样就得到了多尺度的边界图象;
- (3) 对每一尺度的边界图象计算出它的7个不变矩, 所有尺度上的不变矩共同组成这个图象的特征向量;
- (4) 对特征向量进行归一化;
- (5) 查询时计算两幅图象的归一化特征向量的欧氏距离作为两幅图象的相似度

需要指出的是, 此处我们只考虑了图象的空间分布及形状信息, 如果是彩色图象, 我们只对图象的亮度通道进行处理

2 实验及结果

为测试新算法的性能并与其它算法比较, 我们进行了一系列实验 首先, 我们生成了10个简单的几何图形来测试我们的算法 这些几何图形有三角形、正方形、长方形及圆形, 它们分别有不同的平移、尺度或旋转变化 图2(a)是对一个三角形的检索结

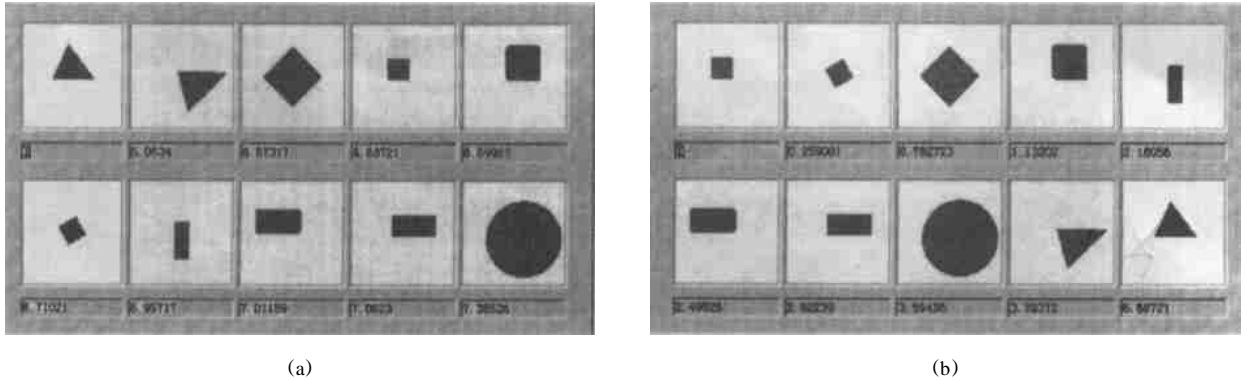


图 2 本文算法对几何图形图象的检索结果

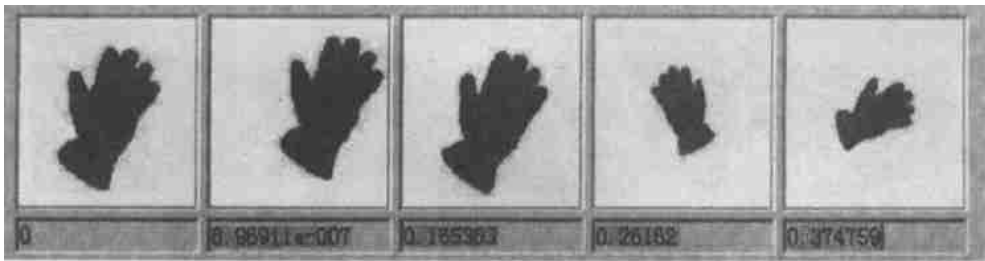


图 3 本文算法对服装图象库的检索结果

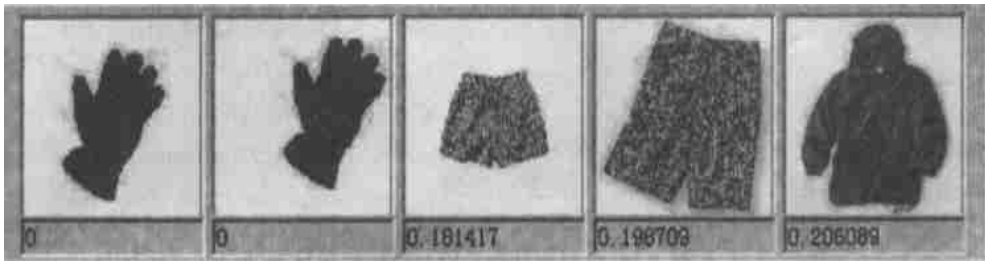


图 4 边界方向直方图算法对服装图象库的检索结果

果, 图 2(b) 是对一个正方形的检索结果 以下在显示图象检索结果时, 左上角是查询图, 检索结果按照图象的相似度从左到右, 从上到下进行排列, 结果图象下方对应的数字是相应图象与查询图之间的相似度 从图 2 可以明显看出, 无论相似的形状是否有平移、尺度或旋转的变化, 都能够得到较好的匹配 这说明我们的算法确实具有平移、尺度及旋转不变性

接下来我们用有 1 400 幅服装图象图片的图象库进行测试 图 3 是本文算法对一只手套的检索结果, 其中第一幅图象为查询图, 图象库中共有 5 只手套, 全部检出, 且排在前 5 位 图 4 是用常见边界方向直方图方法^[7]对同一只手套进行检索的结果 由于边界方向直方图方法不具有旋转变换不变性, 检索结果不令人满意, 另外 4 只手套中只有经平移变换的一只排在第 2 位, 其它 3 只未出现在前 5 位中

最后, 我们还对共有 500 幅图象的真实花草图象库进行了实验 图 5 是本文算法对 3 幅典型图象

的检索结果(最左边为查询图), 从检索结果可以看出对具有不同形态的单朵花、一大片花和一丛花可以区分得很好, 这说明本文算法能较好地描述图象的形状及空间分布信息

对于小波分解层数的选择, Mallat 证明对于一幅 $N \times N$ 的图象, 最大分解层数 $J = \log_2(N) + 1$. 实验结果显示出小波模极大值在分解层数大于 6 时就几乎没有什么区分能力了, 因此小波分解的层数以小于 6 为宜 对于不同的应用, 可以选择不同的小波分解层数 在我们的实验中, 由于服装图象库中的图象基本上是背景简单、目标明确的图象, 所以在对图象进行小波分解时可以采用较少的分解层数(1~ 2 层即可), 这样可以大大节省计算量; 而花草图象是较为复杂的自然图象, 对图象的小波变换的层数相应选择得较大(5~ 6 层), 这样可以较为充分地提取图象的形状信息

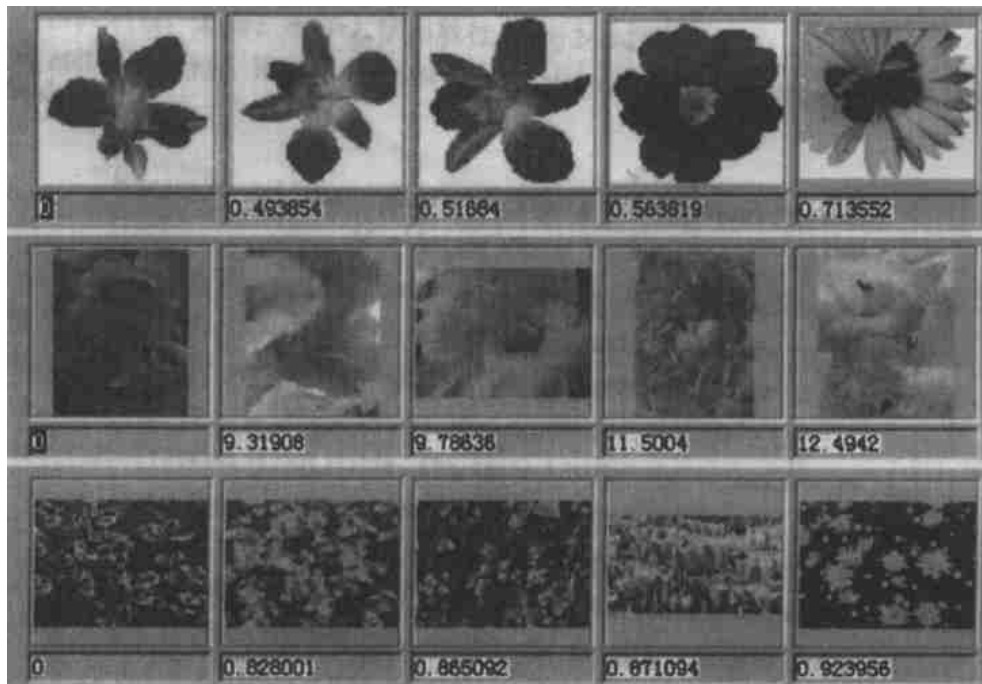


图 5 本文算法对花草图象库中 3 幅图象的检索结果

3 结论及讨论

本文针对目前基于形状的形象检索领域中存在的问题,提出了一种新颖的形状特征提取算法。此算法先对亮度图象进行小波模极大值变换,得到多尺度的边界图象,再利用 7 个不变矩提取每一尺度边界图象的特征。实验结果表明本文算法能较好地描述图象的形状及空间分布信息,并具有较好的平移、尺度、旋转不变性。

一个值得讨论的问题是如何使用小波分解后不同层上的信息。虽然经小波模极大值变换后图象的信息已经得到了压缩,但是仍然存在冗余。理想的情况应该是低分辨率上已经被使用过的信息不再被高分辨率所使用。我们在此采用 7 个不变矩来提取小波变换后每一层上的信息。也许可以采用其它特征函数或特征向量代替不变矩来描述图象的形状信息,使得检索的结果更能符合人类的视觉感受,当然要尽可能地保持平移、尺度、旋转变换的不变性。

参 考 文 献

- 1 Jacobs C E *et al*. Fast multi-resolution image querying. In: Proceeding of SIGGRAPH, 1995, 278~ 280
- 2 Mallat S. Wavelet for a vision. In: Proceeding of the IEEE, 1996, 33(4): 604~ 614

- 3 Mallat S, Hwang W L. Singularity detection and processing with wavelets. IEEE Trans Info Theory, 1992, 38(2): 617~ 643
- 4 Cvetkovic Z, Vetterli M. Discrete time wavelet extreme representation: Design and consistent reconstruction. IEEE Trans Signal Processing, 1995, 43(3): 681~ 693
- 5 Mallat S, Zhong S. Characterization of signals from multi-scale edges. IEEE Trans PAM I, 1992, 14(7): 710~ 732
- 6 边肇祺. 模式识别. 北京: 清华大学出版社, 1998, 248~ 274
- 7 Jain A K, Vailaga A. Image retrieval using color and shape. Pattern Recognition, 1996, 29(8): 1233~ 1244



姚玉荣 1972 年生, 1996 年于清华大学电子工程系获得工学学士和经济学学士双学位, 1999 年于清华大学电子工程系获得工学硕士学位。主要从事计算机图象处理和分析方面的研究。现在巨龙通信设备有限责任公司研究院工作, 从事光纤通信方面的研究与开发。



章毓晋 教授, 博士生导师。研究领域主要是由其倡导的图象工程(图象处理、图象分析、图象理解及其应用)学科, 已发表 100 余篇研究论文, 编著有《图象工程》(上、下册)和《图象分割》。