

一种基于 Zernike 矩形状检索的新算法

郭丹 闫德勤 吴晓婷 刘胜蓝

(辽宁师范大学计算机与信息技术学院 大连 116081)

摘要 高维 Zernike 矩作为图像检索的形状特征描述子,具有描述图像区域细节信息的能力,能够全面有效地表征图像的内容。但是高维的矩存在着“维数灾难”的问题,不仅使算法的复杂度增大,而且会增加不必要的信息,造成主要信息混淆,影响对图像的描述。提出了流行学习的方法来处理冗余的数据信息。在通过拉普拉斯图保持局部样本数据不变的情况下,引入了全局算法来保证样本的整体性。考虑到信息之间的相关性而影响投影的准确率,对其进行 Schur 特征值分解,得到正交基向量,从而使数据重构相对容易,并且 Zernike 矩的旋转不变性仍能保持下来,使检索得到的图像更加符合人眼视觉效果。该方法在检索性能上优于传统的局部保持投影方法,检索效果有明显的提高。

关键词 Zernike 矩,图像检索,主成分分析,局部保持投影,Schur 分解

中图分类号 TP301.6 **文献标识码** A

New Algorithm of Zernike Moments Features for Shape-base Image Retrieval

GUO Dan YAN De-qin WU Xiao-ting LIU Sheng-lan

(Department of Computer and Information Technology, Liaoning Normal University, Dalian 116081, China)

Abstract As shape feature descriptors, high dimension zernike moments have the function of describing the detail information of image region, which exist “dimension disaster”. This will result to increase the complexity of the algorithm and unnecessary information which make major information confused, and will affect describing the content of the image. A new algorithm based on Manifold method was proposed to realize dimension deduction in image data. Under the condition of Laplace figure keeping local sample data, overall algorithm was introduced to ensure the integrity of the sample. Considering the influence of the correlation between information on projection accuracy, schur eigenvalue decomposition was made to obtain the orthogonal vectors. This can make the data reconstruction relatively easier, and the rotation invariant of Zernike moment can still keep down, then making the image retrieval accords with the human visual effect. This method is superior than LPP in the retrieval performance, and retrieval results are significantly improved.

Keywords Zernike moment, Image retrieval, PCA, LPP, Schur decomposition

1 引言

随着 Internet 日益普及,特别是随着多媒体技术、存储技术、压缩技术和网络带宽等技术的不断发展,图像的使用日益广泛。一方面,网上信息(包括许多数字图片)每年在以指数形式增长;另一方面,很多图像数据库,包括遥感图像、医学图像、多媒体、数字图书馆等应用数据库也连接到互联网供用户检索。随着应用的不断增加,大量图像数据库检索问题得到人们的广泛关注。要想有效管理和利用这些庞大的图像资源,就必需有高效率的图像检索技术。

提取图像特征是进行图像检索的首要前提。图像基本特征包括颜色、纹理、形状和物体间方位关系等。自然界的物体主要靠形状来区别,因为对具有相同颜色特征的物体而言,其形状可能不同,而且形状表示了图像中有意义的区域或相关

对象。于是,在图像检索领域中,对图像的形状特征进行恰当的描述就显得尤其重要。形状通常被认为是一条封闭的轮廓曲线所包围的区域,所以形状的描述涉及到一条封闭边界以及这个封闭边界所包围的区域。通常形状的描述可以分为基于轮廓和基于区域两大类。其中,基于轮廓的形状描述技术已相当成熟,不足之处是没有考虑形状区域内的像素。基于区域的形状描述方法主要有区域的面积、形状的纵横比、区域拓扑、区域频谱等。

但基于矩的描述方法更适合图像的形状特征,因为它们不仅仅计算轮廓上的像素点,还计算构成形状区域的所有像素点^[1]。目前常用的矩主要有几何矩、中心矩、正交不变矩等,而 Zernike 矩就是正交不变矩中比较成熟并被广泛使用的一种复数矩^[2]。Zernike 矩作为一种基于区域的形状描述子,具有良好的旋转不变性和易构造高阶矩等特点,使其能够

到稿日期:2009-12-04 返修日期:2010-02-11 本文受国家自然科学基金(60372071),中国科学院自动化研究所复杂系统与智能科学重点实验室开放课题基金(20070101),辽宁省教育厅高等学校科学研究基金(2008344)资助。

郭丹(1985-),女,硕士生,主要研究方向为图像检索和降维等,E-mail:guodan0526@163.com;闫德勤(1962-),男,博士,教授,主要研究方向为模式识别、图像检索、数据挖掘和信息安全等;吴晓婷(1985-),女,硕士生,主要研究方向为模式识别、数据降维;刘胜蓝(1984-),男,主要研究方向为数据降维和模式识别等。

很好地提取图像内部的形状信息。但是也存在一些不足,由于高阶 Zernike 矩能够描述图像的细节信息,因此为了提高检索精度,我们尽量去构造高阶矩作为形状描述子,从而导致特征数据维度过高。对于高维数据,要理解隐藏其中感兴趣的模式或知识,并不是所有的维度都是必要的或重要的。因此所提取的大量图像特征中有可能存在冗余信息,从而导致图像处理结果的精确度下降,还会导致图像处理算法的复杂度过高。

针对这一问题,需要采用一种有效的方法来降低数据维数。最近提出许多用于数据降维的流行学习算法并且成功地应用于人脸识别中,取得了很好的实验效果。在此类方法中,主成分分析(PCA)是使用较广的方法^[4]。但是,最近许多研究表明,图像数据中可能存在非线性结构,而 PCA 仅考虑了数据的全局结构,因此它们不能有效处理数据的流形结构。近来,已经提出了许多流形学习方法来发现数据中的非线性流形结构,如 ISOMAP^[5]、局部线性嵌入(LLE)^[6]和拉普拉斯映射图(LE)^[7]。虽然在大多情况下这些算法更能反映数据的本质特性,但是也存在一些缺点,就是对于新来的样本点不能进行有效的处理,而且没有直接的转换矩阵,不如线性降维算法简单、直观。为了克服上述缺点,He 等人提出了局部保持投影(LPP)方法并成功应用到人脸识别中^[8],该方法能够保持图像数据的局部结构信息。对于图像检索这个很耗时间的应用来说,数据极为庞大,用户都希望尽快得到结果。在保证准确度的前提下,选取相应计算量小的,以便在有效地缩短检索时间的情况下获得相对精确的检索效果。LPP 方法更适合在图像检索中使用。Lu Ke, He Xiaofei 将 LPP 直接应用到图像检索中^[9],取得了一定的效果,但是它只考虑到图像的局部结构信息,忽视了数据的整体性,具有一定的局部限制性。为了更好地刻画图像所表达的信息,我们采用局部和全局相结合的算法正交地引入 PCA 的拉普拉斯图映射(OLP),不仅保留了图像数据的局部非线性,而且保留了数据集绝大部分结构特征的紧凑子空间。为使每个分量能够充分发挥其作用,并且不受其他信息的干扰,我们又对其进行 Schur 分解,以这种简单易行的方式获得正交基向量,使图像检索的效果更加精确有效。在 1008 幅 14 类图像数据库上进行了实验,实验结果验证了算法的有效性,并且其明显优于传统的 LPP 算法。

2 Zernike 矩及其形状特征库

2.1 Zernike 矩描述子

Zernike 提出了一组定义在极坐标下($x^2 + y^2 \leq 1$)上的复数多项式 $V_m(x, y)$, 表达如下:

$$V_m(x, y) = V_m(\rho, \theta) = \exp(jm\theta)$$

式中, n 为正整数或零, m 为满足条件, $n - |m|$ 是偶数, 且 $n \geq |m|$, (ρ, θ) 是点 (x, y) 的极坐标表示, ρ 表示从圆心到点 (x, y) 的向量长度, θ 为向量 ρ 沿逆时针方向与 x 轴的夹角, $f(x, y)$ 为向量长度, $f(x, y)$ 等于 0 或 1。 R_m 为正交的径向多项式, 且有 $R_{n, -m}(\rho) = R_m(\rho)$, 定义如下:

$$R_m(\rho) = \sum_{k=0}^{(n-|m|)/2} (-1)^k \frac{(n-k)!}{k! \left(\frac{n+|m|}{2} - k\right)! \left(\frac{n-|m|}{2} - k\right)!} \rho^{n-2k}$$

Teague 以 Zernike 正交多项式为基础, 给出了二维函数

$f(x, y)$ 的 Zernike 矩的定义:

$$Z_m = \frac{n+1}{n} \iint_{x^2+y^2 \leq 1} V_m^*(x, y) f(x, y) dx dy$$

式中, $V_m^*(x, y)$ 表示 $V_m(x, y)$ 的共轭。对于一幅离散数字图像, 它的 m 重 n 阶 Zernike 矩表示为:

$$Z_m = \frac{n+1}{n} \sum_x \sum_y f(x, y) V_m^*(\rho, \theta), x^2 + y^2 \leq 1 \quad (1)$$

2.2 Zernike 矩的提取与特征库的形成

2.2.1 Zernike 矩的提取

因为 Zernike 矩被定义在极坐标系 (ρ, θ) 内, 并且 $|r| \leq 1$, 计算 Zernike 多项式需要对图像坐标 (i, j) , $i, j = 0, 1, 2, \dots, N-1$ 进行线性变换, 并映射到合适的圆形区域内, $(x, y) \in R^2$ 。现有两种常用的转换方法^[3], 如图 1 所示。在此, 我们选择第二种 Zernike 矩的提取方法, 将图像转换到极坐标下的单位圆中, 即将目标的重心(形心)作为极坐标的圆心, 以圆心到目标区域内最外像素点的距离为半径(取最远距离作为半径, 使得区域中所有像素都落在单位圆内, 避免像素信息丢失), 对目标区域内的像素重新采样到单位圆内。相比之下, 第二种方法较第一种有一定的优势, 因为第一种忽略了目标的边缘像素, 这样会引起一定的误差。直角坐标到极坐标的转换为 $x = \rho \cos \theta, y = \rho \sin \theta$, 其中 $\rho = \sqrt{x^2 + y^2}, \theta = \arctan \frac{y}{x}$ 。计算中, 要注意直角坐标系中像素点所在的象限, 因为 $\theta = \arctan \frac{y}{x}$ 定义在区间 $[-\frac{\pi}{2}, \frac{\pi}{2}]$ 内。最后计算出各阶 Zernike 矩, 取其幅值作为图像的描述子, $|A_m|$ 作为其目标描述向量。本文主要取前 36 阶矩。

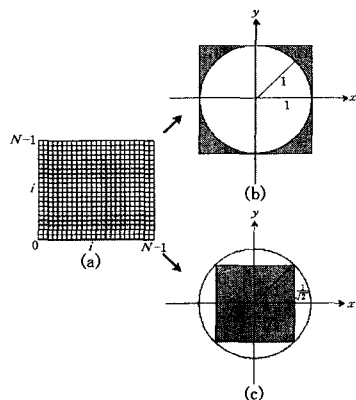


图 1 Zernike 矩提取方法

在计算 Zernike 矩之前要对图像进行归一化操作^[4], 因为 Zernike 矩本身具有旋转不变性。因此, 为了获得鲁棒性的图像特征, 图像 $f(x, y)$ 应该在计算 Zernike 矩之前进行相应的变换。为了获得平移不变量, 要将图像 $f(x, y)$ 的重心移到单位圆的中心:

$$\{g(x_T, y_T); x_T = x - \bar{x}, y_T = y - \bar{y}\}$$

式中, $g(x_T, y_T)$ 是新的平移函数; (\bar{x}, \bar{y}) 是图像函数 $f(x, y)$ 的重心, 计算如下:

$$\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}$$

式中, m_{01}, m_{10} 和 m_{00} 分别是图像函数 $f(x, y)$ 的 $(0+1), (1+0)$ 和 $(0+0)$ 阶的几何矩, 它们分别定义为:

$$m_{10} = \iint x f(x, y) dx dy$$

$$m_{01} = \iint yf(x, y) dx dy$$

$$m_{00} = \iint f(x, y) dx dy$$

尺度不变量是在上述图像平移之后的基础上对其进行归一化的,新的目标区域的 $m_{00} = \beta$, β 是预先定义的值。所以新的尺度函数 $h(x_s, y_s)$ 定义为:

$$\left\{ h(x_s, y_s) : x_s = x \sqrt{\frac{\beta}{m_{00}}}, y_s = y \sqrt{\frac{\beta}{m_{00}}} \right\}$$

平移和尺度不变性可通过下式获得:

$$o(x, y) = f\left(\frac{x}{a} + \bar{x}, \frac{y}{a} + \bar{y}\right)$$

式中, $a = \sqrt{\beta/m_{00}}$, $o(x, y)$ 是新的经过平移和尺度缩放的函数。

2.2.2 矩特征库的形成

Zernike 矩的基是正交径向多项式,因此可以保证所提取的特征相关性小,冗余性小,抗噪声能力强;Zernike 矩具有旋转不变性,经过相关改进后可以具备平移不变性以及更好的比例不变性。一幅图像的形状特征可以用一组 Zernike 矩集合很好地表示,低阶矩描述的是一个图像目标的整体形状,高阶矩描述的是图像目标的细节。在式(1)中,可以根据 z_m 中 m 的变化计算出图像不同阶的 Zernike 矩;相应地,在 m 值确定的情况下,又可以根据 n 值的变化计算 m 阶的各个 Zernike 矩值,而这些矩值就可以被看作具体图像的 Zernike 矩特征。把计算出的 Zernike 矩的各阶矩值,按照 m 递增次序并且在 m 确定的情况下,再按照 n 递增次序排列,则可以组成相关图像形状特征的 Zernike 矩特征集 z_m 。如表 1 所列的 Zernike 矩的前 36 阶矩,是在考虑 $m > 0$ 时的情况。若再计算 $m < 0$ 时的情况,那么 Zernike 矩的个数就会翻倍。因此,要对这些高维的 Zernike 矩的特征向量作降维处理,以便后续分析和处理。

表 1 Zernike 矩及其个数

| n | Zernike 矩 | | | | 矩个数 | | |
|----|------------|------------|------------|------------|------------|-------------|---|
| 0 | $z_{0,0}$ | | | | 1 | | |
| 1 | $z_{1,1}$ | | | | 1 | | |
| 2 | $z_{2,0}$ | $z_{2,2}$ | | | 2 | | |
| 3 | $z_{3,1}$ | $z_{3,3}$ | | | 2 | | |
| 4 | $z_{4,0}$ | $z_{4,2}$ | $z_{4,4}$ | | 3 | | |
| 5 | $z_{5,1}$ | $z_{5,3}$ | $z_{5,5}$ | | 3 | | |
| 6 | $z_{6,0}$ | $z_{6,2}$ | $z_{6,4}$ | $z_{6,6}$ | 4 | | |
| 7 | $z_{7,1}$ | $z_{7,3}$ | $z_{7,5}$ | $z_{7,7}$ | 4 | | |
| 8 | $z_{8,0}$ | $z_{8,2}$ | $z_{8,4}$ | $z_{8,6}$ | $z_{8,8}$ | 5 | |
| 9 | $z_{9,1}$ | $z_{9,3}$ | $z_{9,5}$ | $z_{9,7}$ | $z_{9,9}$ | 5 | |
| 10 | $z_{10,0}$ | $z_{10,2}$ | $z_{10,4}$ | $z_{10,6}$ | $z_{10,8}$ | $z_{10,10}$ | 6 |

3 降维及其引进 PCA 的局部投影的改进

3.1 降维的定义

降维技术是希望通过一个映射找到数据的低维表示。假设我们得到 D 空间上的一个实向量样本集 $\{p_i\}_{i=1}^N$, 降维问题基于如下的基本假设:高维数据实际位于一个维数比数据空间的维数小得多的流形上。降维的目的就是获得这一流形的低维的坐标表示。

我们可以定义降维问题的模型为 (P, M) , 其中,

(1) $P = \{t_i\}_{i=1}^N$ 是 D 维空间中的数据集合(一般是 R^D 的一个子集);

$$(2) \text{ 映射 } F: \begin{matrix} P \rightarrow X \subset R^E \\ P_i \mapsto x_i = F(p_i) \end{matrix} \quad E \ll D$$

称 D 维数据集 P 到 E 维数据集 E 嵌入映射。之所以能对高维数据进行降维,是因为数据的原始表示常常包含大量冗余:

(1) 有些变量的变化比测量引入的噪声还要小,因此可以看作是无关的;

(2) 有些变量与其他变量有很强的相关性(例如是其他变量的线性组合或是其他函数依赖关系),可以找到一组新的不相关的变量。

这些冗余信息在处理时常常消耗过多的系统开销,在许多情形下可以从一定程度上剔除这些冗余信息,从而获得更加经济的表示方式。

3.2 特征向量的维数约减新算法

在判定一幅图像的类别时,要考虑许多特征,如颜色直方图、纹理、边缘、明暗、区域分布等等。若要全部列出,也许可以有几十上百个特征。本文中以提高维 Zernike 矩作为形状特征描述子的特征就有 36 个,为了减少算法的复杂度,用少量的几个综合特征代替原来的许多特征是有实际意义的。作为降维的经典算法,主成分分析的基本思想就是把数据投影到方差最大的方向,使得样本数据能够在此方向保持数据整体的稳定性,保留了重要的、相对多的鉴别信息,由此获得图像数据的主要特征,从而实现图像数据的降维。但是它仅仅考虑了数据的全局结构,对于图像数据这种非线性结构,作为线性算法的主成分分析来说,也很难保持图像的局部流形结构。针对此问题,以及对各种算法的研究,选取局部保持投影算法 LPP 和主成分分析算法 PCA 相结合。虽然同样是线性算法,但是它是继承了拉普拉斯特征映射(LE)算法的思想,即基于最近邻图来建立映射,因而具有一般线性降维算法所不具备的流形学习能力^[6]。

3.2.1 对主成分分析(PCA)的研究

为了方便计算,下面对 PCA 进行重写:

对于 n 个样本点 $X_{D \times n} = \{x_1, x_2, \dots, x_n\} \in R^D$, 其中第 i 个有:

$$\begin{cases} y_{11} = u_{11}x_{11} + \dots + u_{1D}x_{1D} = u_1^T x_1 \\ \vdots \\ y_{D1} = u_{D1}x_{11} + \dots + u_{DD}x_{1D} = u_D^T x_1 \end{cases} \Leftrightarrow y_i = U^T x_i$$

式中, $U_{D \times n} = [u_1, u_2, \dots, u_n]$, 进而可以写成矩阵形式 $Y = U^T X$, u 是样本协方差阵的特征向量矩阵。进而找到 $\check{U}_{d \times n}$, 使 $\check{Y} = \check{U}^T X$, 其中 $\check{Y} = \{y_1, y_2, \dots, y_n\} \in R^d$ 。设样本的协方差阵 $C_{D \times D} = \frac{1}{n-1} \sum_{j=1}^n (x_j - \bar{x})^T$, 其中 $\bar{x} = \frac{1}{n} \sum_{j=1}^n x_j$ 。为了简写协方差阵,我们把已经中心化的样本记为 $\hat{X} = [\hat{x}_1, \dots, \hat{x}_n]$, 其中 $\hat{x}_j = x_j - \bar{x}$, 由此得样本的协方差矩阵为: $C_{D \times D} = \frac{1}{n-1} \hat{X} \hat{X}^T$, 而 $\bar{x} =$

$\frac{1}{n} \sum_{j=1}^n x_j = \frac{1}{n} X e$, 因此有:

$$\begin{aligned} \hat{X} &= [\hat{x}_1, \dots, \hat{x}_n] = [x_1 - \bar{x}, \dots, x_n - \bar{x}] = X - \bar{x} e^T = X - \\ &\frac{1}{n} X e e^T = X \left(I - \frac{1}{n} e e^T \right) = X J, J = I - \frac{1}{n} e e^T \end{aligned}$$

因此,协方差矩阵可以写为 $C = X J J^T X^T$ 。又 $J J^T = J$, 所以 $C = X J X^T$ 。PCA 的优化函数为:

$$\begin{cases} \max & u^T C u \\ s. t & u^T u = 1 \end{cases} \Rightarrow \begin{cases} \max & u^T X J X^T u \\ s. t & u^T u = 1 \end{cases} \quad (2)$$

3.2.2 新算法 OLP 的描述

我们建立一个具有 m 个顶点的权图 G 。如果节点 x_i 是 x_j 的 K -近邻或者 x_j 是 x_i 的 K -近邻,则在节点 x_i 和 x_j 之间用一条边连接。按上述规则构成的权图可以很好地模拟内嵌在高维空间中的低维流形。大量研究表明,人类的感觉(听觉、视觉)往往表现为非线性的流形结构,这就使得图像数据这样的视觉产物含有大量的非线性特性。因此,在图像检索中基于上面的权图来建立映射可以获得良好的效果。

令 x_i 和 x_j 为一个局部 K -邻域结构中里的两个点集,并且满足 $\|x_i - x_j\| < K$ 。定义一个集合 U^k ,使 $J^k = \{\|x_i - x_j\| < K\}$ 。当 x_i 和 x_j 投影到 u 方向时,得到新的点集 y_i 和 y_j 。因此数据样本的局部分布可以定义为:

$$M_L(u) \triangleq \frac{1}{2} \frac{1}{N_{L, i, j \in J^k}} \sum (y_i - y_j)^2 \propto \frac{1}{2} \sum_{i, j \in J^k} (y_i - y_j)^2 \quad (3)$$

式中, N_L 是满足 $\|x_i - x_j\| < K$ 的样本对的个数。

定义邻接矩阵 W , 满足下面的条件:

$$W_j = \begin{cases} 1, & \text{如果 } x_j \text{ 是 } x_i \text{ 的近邻或者 } x_i \text{ 是 } x_j \text{ 的近邻} \\ 0, & \text{其它情况} \end{cases}$$

很显然邻接矩阵 W 是对称的,因此式(3)可以为:

$$\begin{aligned} M_L(u) &= \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N W_j (y_i - y_j)^2 \\ &= \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N W_j (u^T x_i - u^T x_j)^2 \\ &= u^T \left[\frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N W_j (x_i - x_j)(x_i - x_j)^T \right] u \\ &= u^T S_L u \end{aligned}$$

式中, $S_L = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N H_j (x_i - x_j)(x_i - x_j)^T$, 由于矩阵 W 是对称的,因此有如下推导:

$$\begin{aligned} S_L &= \frac{1}{2} \left(\sum_{i=1}^N \sum_{j=1}^N W_j x_i x_i^T + \sum_{i=1}^N \sum_{j=1}^N W_j x_j x_j^T - 2 \sum_{i=1}^N \sum_{j=1}^N W_j x_i x_j^T \right) \\ &= \left(\sum_{i=1}^N D_i x_i x_i^T - \sum_{i=1}^N \sum_{j=1}^N H_j x_i x_j^T \right) \\ &= (XDX^T - XHX^T) = XLX^T \end{aligned} \quad (4)$$

式中, $X = (x_1, x_2, \dots, x_N)$; D 是对角矩阵,其对角线上的元素是矩阵 H 的列向量的和, $D_i = \sum_{j=1}^N H_j$ 。 $L = D - H$ 是拉普拉斯矩阵。这样引入全局算法 PCA, 结合式(2)和式(4), 得到如下目标函数:

$$M(U) = \begin{cases} \min & u^T XLX^T u / u^T XJX^T u \\ s. t & u^T XDX^T u = 1 \end{cases}$$

求 $\min(u^T XLX^T u / u^T XJX^T u)$, 即求 $\min(u^T X(L - J)X^T u)$ 。再应用 Lagrange 乘子法, 首先构造拉格朗日方程如下:

$$L_L(u)_{\text{OLP}} = \min\{u^T X(L - J)X^T u - \lambda(u^T XDX^T - 1)\}$$

对拉格朗日方程求导:

$$\frac{\partial L_L(u)_{\text{OLP}}}{\partial u} = u^T X(L - J)X^T u - \lambda(u^T XDX^T - 1) = 0$$

约束优化问题(3)的求解变为求以下广义特征值问题:

$$X(L - J)X^T u = \lambda XDX^T u$$

那么基向量 $\{u_1, u_2, \dots, u_d\}$ 是矩阵, 因为矩阵 $(XDX^T)^{-1} X(L - J)X^T$ 是前 d 个特征值所对应的特征向量。因为 $(XDX^T)^{-1} X(L - J)X^T$ 不是对称的, 所以基向量 $\{u_1, u_2, \dots, u_d\}$ 不一定是正交的, 这使得数据重构很困难。为了解决这一问题, 本文应用 Schur 分解定理^[6], 以一种简单的方式来获得

正交基向量。当被分析的矩阵是非对称矩阵时, Schur 分解定理可以很好地代替特征分析法。

定理 1(Schur) 设 A 为 n 阶方阵, $\lambda_1, \lambda_2, \dots, \lambda_n$ 是 A 的特征值, 不论它们是实数还是复数, 总存在相似矩阵 U , 使得 $A = UTU^H$, 其中 T 为三角矩阵, 对角线上的元素是 $\lambda_1, \lambda_2, \dots, \lambda_n$ 。

依据定理, 当 A 为实 n 阶方阵时, 它的 Schur 分解是 $A = UTU^H$, 这里 U 就是正交矩阵。为了获得正交基向量, 我们并没有对 $D^{-1}X(L - J)X^T$ 进行特征分析, 而是对 $D^{-1}X(L - J)X^T$ 进行 Schur 分解。假设 $D^{-1}X(L - J)X^T$ 的分解是 UTU^T , u_1, u_2, \dots, u_n 是矩阵 U 的所有列向量。

假设 u_1, u_2, \dots, u_d 是 $D^{-1}X(L - J)X^T$ 前 d 个最小特征值所对应的 Schur 分解向量, 那么 u_1, u_2, \dots, u_n 可以最小化目标函数 $M(U)$ 。

证明: 因为 u_1, u_2, \dots, u_d 是 $D^{-1}X(L - J)X^T$ 前 d 个最小特征值所对应的 Schur 分解向量, 所以有 $D^{-1}X(L - J)X^T u_i = \lambda_i u_i, i = 1, 2, \dots, d, \lambda_i$ 是 $D^{-1}X(L - J)X^T$ 的第 i 个最小特征值。从上述可以得出:

$$\begin{aligned} [u_1, u_2, \dots, u_d]^T X(L - J)X^T [u_1, u_2, \dots, u_d] \\ = [u_1, u_2, \dots, u_d]^T X(L - J)X^T [u_1, u_2, \dots, u_d] \\ \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_d) \end{aligned}$$

因此可以有:

$$\begin{aligned} M(u_1, u_2, \dots, u_d) &= \\ \frac{[u_1, u_2, \dots, u_d]^T X(L - J)X^T [u_1, u_2, \dots, u_d]}{[u_1, u_2, \dots, u_d]^T D [u_1, u_2, \dots, u_d]} \\ &= \prod_{i=1}^d \lambda_i = \min M(U) \end{aligned}$$

于是, 选取 u_1, u_2, \dots, u_n 这一组新的 Zernike 矩作为特征向量, 进行图像检索。

4 实验结果

为了检验本文方法的性能, 我们选取 1008 幅、14 个种类、每类含有 72 幅带有不同旋转角度的图像库作为实验对象。提取图像的 Zernike 矩作为图像的形状特征, 运用欧式距离度量法进行图像间的相似性度量。应用本文提到的算法 OPL 与传统的 LPP 作对比, 显示前 40 个图像的返回结果。我们从不同的角度来评判本文方法的有效性。

为了评价本文方法的效果, 采用“精确度 (precision)”和“检索效率 (recall)”作为图像相似性的评价准则。其中精确度 I 定义为检索结果队列中检索到的目标图像数与检索结果队列中所有图像数之比, 即 $I = Y/X$, 其中 Y 为目标图像数, X 为结果队列中所有的图像数。检索率 J 定义为检索结果队列中检索到的目标图像数与数据库中全部目标图像数之比, 即 $J = Y/F$, F 为数据库中的目标图像总数。由精确度和检索率的定义可知, $I \in [0, 1], J \in [0, 1]$ 。通过执行例子集中的各个查询来计算出查询的平均检索精度和查全率, 据此就能给出系统的检索性能评价。对本文提出的算法 OLP 和应用传统的 LPP 算法所得到的结果进行了对比。我们选取图示中的第一个图像作为样本图像, 从图像库中查询与样本图像相似的图像。通过表 2 可以看出, 经过本文改进的算法查准率和查全率都比较高, 并且效果明显好于传统的 LPP 算法。例如应用传统 LPP 算法的图示(见图 2)中, 与小盒差距较大的其他形状的图像都被检索出来。而经过改进的本文算法, 返回的相关图像全部是想要得到的目标图像。

表 2 LPP 和 OLP 检索结果的查准率和查全率

| LPP | | | | | | OLP | | | | |
|--------|-------|--------|--------|-------|-------|-------|--------|--------|-------|-------|
| 图像 | 图像数 F | 返回图像 X | 相关图像 Y | 精确度 P | 查全率 R | 图像数 F | 返回图像 X | 相关图像 Y | 精确度 P | 查全率 R |
| 招财猫 | 72 | 30 | 21 | 0.700 | 0.292 | 72 | 30 | 25 | 0.833 | 0.347 |
| 鸭子 | 72 | 30 | 20 | 0.667 | 0.278 | 72 | 30 | 22 | 0.733 | 0.306 |
| 三角圆 | 72 | 30 | 13 | 0.433 | 0.181 | 72 | 30 | 19 | 0.633 | 0.264 |
| 轿车 | 72 | 30 | 6 | 0.200 | 0.083 | 72 | 30 | 11 | 0.367 | 0.153 |
| ANCIN | 72 | 30 | 30 | 1.000 | 0.417 | 72 | 30 | 30 | 1.000 | 0.417 |
| 赛车 | 72 | 30 | 22 | 0.733 | 0.306 | 72 | 30 | 22 | 0.733 | 0.306 |
| 积木 | 72 | 30 | 13 | 0.433 | 0.181 | 72 | 30 | 25 | 0.833 | 0.347 |
| 药瓶 | 72 | 30 | 29 | 0.967 | 0.403 | 72 | 30 | 30 | 1.000 | 0.417 |
| TYLENO | 72 | 30 | 17 | 0.567 | 0.236 | 72 | 30 | 26 | 0.867 | 0.361 |
| 圆滑盒 | 72 | 30 | 6 | 0.200 | 0.833 | 72 | 30 | 30 | 1.000 | 0.417 |
| 小伞 | 72 | 30 | 20 | 0.667 | 0.278 | 72 | 30 | 30 | 1.000 | 0.417 |
| 小猪 | 72 | 30 | 13 | 0.433 | 0.181 | 72 | 30 | 14 | 0.467 | 0.194 |
| 杯子 | 72 | 30 | 29 | 0.967 | 0.403 | 72 | 30 | 29 | 0.967 | 0.403 |
| 敞篷车 | 72 | 30 | 10 | 0.333 | 0.139 | 72 | 30 | 13 | 0.433 | 0.181 |

再用排序方法作为评价指标。由图像库得到 14 类图像的每类图像总数 F , 由 14 类图像和检索结果得到分别用原算法和改进算法检索的 14 类图像检索结果中的相关图像数和相关图像的排序序号 P_i , 根据公式计算相关图像的平均排序 $AVRR = (1/R) \sum_{i=1}^R P_i$, 理想的相关图像平均排序 $IAVRR = 2/T$ 。因此得出传统 LPP 算法和本文改进算法的相关平均排序, 见表 3 和表 4。 $IAVRR$ 表示在理想情况下所有返回图像全部排列在最前面时的平均序号, $AVRR$ 表示返回图像排列在前面的程度, 其值越小越接近 $IAVRR$ 越好。因为图像总数是返回图像数的 2 倍, 所以 $AVRR/IAVRR$ 越小越接近 0.5 越好。从表 3 和表 4 可以看出, 改进算法的 $AVRR/IAVRR$ 更小, 更接近 0.5, 即相似图像排在前面。可见改进算法的检索效果优于原算法, 而且改进算法的检索速度比原算法快。

表 3 LPP 的检索结果的平均排序

| 图像 | F | Y | P_i | AVRR | IAVRR | AVRR/IAVRR |
|--------|----|----|--|--------|-------|------------|
| 招财猫 | 72 | 21 | 1-19,23,29 | 11.520 | 36 | 0.320 |
| 鸭子 | 72 | 20 | 1-3,4-6,8,11-18, 22-23,25-27 | 12.100 | 36 | 0.336 |
| 三角圆 | 72 | 13 | 1-3,6-14 | 7.385 | 36 | 0.205 |
| 轿车 | 72 | 6 | 1,2,11,13,20-22 | 15.000 | 36 | 0.417 |
| ANCIN | 72 | 30 | 1-30 | 15.500 | 36 | 0.431 |
| 赛车 | 72 | 13 | 1-5,8,10-12,16,17,25,29 | 11.000 | 36 | 0.306 |
| 积木 | 72 | 14 | 1,4,6,8,10,12,14-17,23,24 | 10.714 | 36 | 0.297 |
| 药瓶 | 72 | 29 | 1-13,15-30 | 15.552 | 36 | 0.432 |
| TYLENO | 72 | 17 | 1,4,5,7,8,11,13,14, 18,20,26,30 | 9.235 | 36 | 0.257 |
| 圆滑盒 | 72 | 6 | 1,2,4,5,13,27 | 8.667 | 36 | 0.241 |
| 小伞 | 72 | 20 | 1,3-5,7,9,11-15,17, 19,20,21,25-27,29,30 | 15.400 | 36 | 0.428 |
| 小猪 | 72 | 13 | 1-12,18 | 7.385 | 36 | 0.205 |
| 杯子 | 72 | 29 | 1-25,27-30 | 15.137 | 36 | 0.420 |
| 敞篷车 | 72 | 10 | 1,2,4,8,19,24,25,26,28,29 | 16.600 | 36 | 0.461 |

表 4 OLP 检索结果的平均排序

| 图像 | F | Y | P_i | AVRR | IAVRR | AVRR/IAVRR |
|-------|----|----|-------------------------------|--------|-------|------------|
| 招财猫 | 72 | 25 | 1-18,23,24-29 | 13.720 | 36 | 0.381 |
| 鸭子 | 72 | 21 | 1-21 | 11.520 | 36 | 0.320 |
| 三角圆 | 72 | 19 | 1-18 | 9.000 | 36 | 0.250 |
| 轿车 | 72 | 11 | 1-5,7,10,14,19,20-22,24 | 13.818 | 36 | 0.384 |
| ANCIN | 72 | 30 | 1-30 | 15.500 | 36 | 0.431 |
| 赛车 | 72 | 22 | 1,4-6,9-14,16,17, 19-23,25-30 | 16.682 | 36 | 0.463 |
| 积木 | 72 | 25 | 1-19,22-25,29,30 | 13.720 | 36 | 0.381 |

| | | | | | | |
|--------|----|----|---------------------------------------|--------|----|-------|
| 药瓶 | 72 | 30 | 1-30 | 15.500 | 36 | 0.431 |
| TYLENO | 72 | 26 | 1-21,23,25-28 | 13.846 | 36 | 0.385 |
| 圆滑盒 | 72 | 30 | 1,2,4,5,13,27 | 15.500 | 36 | 0.431 |
| 小伞 | 72 | 30 | 1-30 | 15.500 | 36 | 0.431 |
| 小猪 | 72 | 14 | 1,3,5,7,8,10,12,15,16, 21,22,24,26,28 | 14.143 | 36 | 0.393 |
| 杯子 | 72 | 29 | 1-23,25-30 | 15.207 | 36 | 0.422 |
| 敞篷车 | 72 | 13 | 1-8,10,15,17,19,26 | 9.461 | 36 | 0.263 |

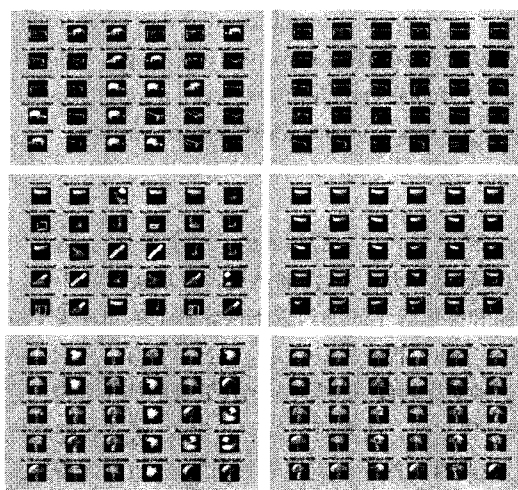


图 2 LPP 和 OLP 检索结果对比

而且,若再做仔细精确判断,如图 3 所示,虽然检索出来的图像的总数大概相近,但是因为图像显示的顺序是按照图像间的相似距离由小到大排序的,所以,越往前排的图像,越与目标图像相似。从直观上看,经过本文算法所得到的图像相比传统 LPP 检索得到的图像更加符合人眼视觉效果。从 Zernike 矩的角度来观察,通过上述的效果对比,我们也可以看到此算法更有效地保持了 Zernike 矩的旋转不变性,对于不同方向的、大小不一的同类图像,检索得更加精确些。从检索时间上来看,OLP 用时更短,提高了检索的效率。

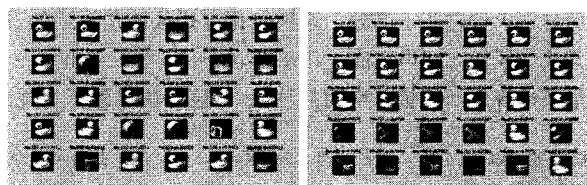


图 3 LPP 和 OLP 的检索结果精确对比

(下转第 286 页)

| | | | | | |
|----|-------|---------|-------|-------|-------|
| 2 | 大岩桐 | 56/56 | 80.27 | 70.73 | 73.24 |
| 3 | 杜鹃 | 105/105 | 85.29 | 74.06 | 75.86 |
| 4 | 凤信子 | 65/65 | 88 | 80.97 | 82.62 |
| 5 | 荷花 | 60/60 | 71.25 | 63.42 | 65.38 |
| 6 | 鸡冠花 | 59/59 | 71.19 | 65.67 | 67.07 |
| 7 | 喇叭花 | 87/87 | 80.92 | 70.39 | 72.36 |
| 8 | 玫瑰 | 68/68 | 74.26 | 60.34 | 63.50 |
| 9 | 蒲公英 | 30/30 | 98.67 | 97.33 | 97.87 |
| 10 | 三色堇 | 49/49 | 88.98 | 83.30 | 84.95 |
| 11 | 石蒜 | 144/144 | 93.82 | 90.52 | 90.92 |
| 12 | 水仙花 | 85/85 | 98.29 | 97.70 | 97.82 |
| 13 | 天堂鸟 | 95/95 | 94.05 | 90.70 | 91.29 |
| 14 | 向日葵 | 124/124 | 89.07 | 85.0 | 85.57 |
| 15 | 郁金香 | 80/80 | 83 | 75.22 | 76.78 |
| 16 | Total | — | 86.34 | 81.45 | 83.82 |

从表 2 可以看到,我们提出的基于 SVM 多特征融合的花卉图像检索效果很理想。图像库中共有 15 类花卉图像,在返回 20 张图像的前提下,查准率达到 80% 以上的有 12 类,只有荷花、鸡冠花、玫瑰的查准率仅达到 70% 多;而查全率高达 80% 以上的共有 9 类;综合图像库中 15 类花卉,综合查准率高达 86.34%,综合查全率也达到 81.45%,而综合率达到了 83.82%。分析查准率没有达到 80% 的花卉种类,可能的原因如下:一是这些花卉的分割效果还不理想;二是这些花卉与其它类别的花卉比较相似;三是目前的特征不能完全描述这些花卉。

结束语 本文以植物花卉图像为研究对象,对植物花卉图像在区域分割、特征提取、基于内容的重复图像过滤以及基于 SVM 的植物花卉图像检索等问题进行了系统、深入和较为全面的研究。通过大量实验,验证了我们提出的重复图片过滤算法、基于 2RGB 混合颜色模型的自适应阈值分割算法的有效性。接着针对分割后的花卉图像提出了分块 Gabor、基于 HSV 颜色模型的加权不变矩、边缘 LBP 算子等特征提取算法。通过在花卉图像库中进行实验,验证了我们提出的若干算法的有效性。下一步的研究工作包括如何提取更适合花卉的特征以及如何结合相关反馈技术对分类结果进行改善等。

参 考 文 献

[1] 李向阳,庄越挺,潘云鹏. 内容的图像检索技术与系统[J]. 计算

机研究与发展,2001,38(3):345-354

- [2] 斯白露. 基于感兴趣区域的图像检索方法[D]. 北京:中国科学院,2002
- [3] 吴清锋. 基于内容的中草药植物故乡检索关键技术研究[D]. 厦门:厦门大学,2007
- [4] Canny J. A Computational Approach to Edge Detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8(11):679-698
- [5] 林开颜,吴军辉,徐立鸿. 彩色图像分割方法综述[J]. 中国图象图形学报,2005,10(1):1-10
- [6] 杜娟,李文锋. 基于金字塔连接算法的彩色图像分割[J]. 武汉理工大学学报,2006,28(1):112-115
- [7] Laurent I, Christof K, Ernst N. A Model of Saliency-based Visual Attention for Rapid Scene Analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20(11):1254-1259
- [8] Chen Guo. The Fisher criterion function method of image thresholding [J]. Chinese Journal of Scientific Instrument(仪器仪表学报),2003,24(6):564-567
- [9] 王国胜. 支持向量机的理论与算法研究[D]. 北京:北京邮电大学,2007
- [10] Stricker M A, Markus O. Similarity of Color Images[C]//SPIE Storage and Retrieval for Image and Video Databases. 1995, 2185:337-349
- [11] Chang E, Goh K, Wu Gang. CBSA: Content-based soft annotation for multimodal image retrieval using bayes point machines [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(1):26-38
- [12] Grigorescu S E, Petkov N, Kruizinga P. Comparison of Texture Features Based on Gabor Filters[J]. IEEE Transactions on Image Processing, 2002, 11(10):1160-1167
- [13] Ojala T, Pietikainen M, Maenpaa T. Multiresolution Gray-scale and Rotation Invariant Texture Classification with Local Binary Patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7):971-987
- [14] Maenpaa T, Ojala T, Pietikainen M, et al. Robust texture classification by subsets of local binary patterns [C]//Proc. 15th International Conference on Pattern Recognition. 2000:947-950

(上接第 251 页)

结束语 本文在 Zernike 矩的形状特征检索的基础上,提出了一种新的降维方法,将 LPP 中保持局部的拉普拉斯图中引入全局算法 PCA,使得保留样本局部结构的同时保证样本的相对完整性,而且有效地保持了 Zernike 矩本身具有的旋转不变性,提高了图像检索的精度。同时大大降低了时间和计算量,在实际应用中会有效提高工作效率。实验结果表明,本文方法不但有效,而且具有很好的检索性能,其检索结果能较好地接近人的视觉感知效果。

参 考 文 献

- [1] Ye Bin, Peng Jiaxiang. Improvement and Invariance Analysis of Zernike Moments Using as a Region-based Shape Descriptor [J]. Journal of Pattern Recognition and Image Analysis, 2002, 12(4):419-428
- [2] Belkasim S, Shridhar M, Ahmadi M. Pattern Recognition with Moment Invariants: A Comparative Study and New Results[J]. Pattern Recognition, 1991, 24(12):1117-1138
- [3] Chong Chee-way, Raveendran P, Mukundan R. A comparative a-

nalys of algorithms for fast computation of Zernike moments [J]. Pattern Recognition, 2003, 36:731-742

- [4] Chrjs H, Ding Q, He X F. Principal component analysis and effective K-means clustering[C]//Proceedings of the Fourth SI-AM International Conference on Data Mining. 2004
- [5] Tenenbaum J, de Silva V, Langford J. A global geometric framework for nonlinear dimensionality reduction[J]. Science, 2000, 290(12):2319-2323
- [6] 陈高曙,曾庆宁. 基于 LLE 算法的人脸识别方法[J]. 计算机应用研究,2007
- [7] Belkin M, Niyogi P. Laplacian Eigenmaps for Dimensionality Reduction and Data Representation[J]. Neural Computation, 2003, 15(6):1373-1396
- [8] He X, Niyogi P. Locality Preserving Projections [C]// Proc. Conf. Advances in Neural Information Processing Systems. 2003
- [9] 刘慧,袁文燕. 矩阵论及其应用 [M]. 北京:化学工业出版社,2003
- [10] Lu Ke, He Xiaofei. Image retrieval based on incremental sparse learning[J]. Pattern Recognition, 2005, 38(11):2047-2054