

基于 MDS 的统计形状聚类

张春燕, 汤进, 赵海峰, 罗斌

(安徽大学 计算智能与信号处理教育部重点实验室, 安徽 合肥 230039)

摘要:针对物体轮廓曲线,采用新的模与相位保持的傅里叶描述方法,将连续的轮廓曲线降维归一化为 R^M 空间中的点,引入 DTW(Dynamic Time Warping)相似度,使用多维尺度分析(MDS),对多类物体基于形状进行聚类。实验结果表明,在合适的相似度下,该方法能够对不同的形状序列进行聚类。

关键词:DTW 相似度;多维尺度分析;统计形状聚类

中图分类号:TP391.4

文献标识码:A

文章编号:1673-629X(2007)03-0058-04

Statistical Shape Classification Based on MDS

ZHANG Chun-yan, TANG Jin, ZHAO Hai-feng, LUO Bin

(Ministry of Education Key Lab. of Intelligent Computing & Signal Processing, Anhui University, Hefei 230039, China)

Abstract: A new normalized Fourier descriptor based on objects contour is defined, which keeps the amplitude and phase information of Fourier coefficients. Finding meaningful low - dimensional embedding in a high - dimensional shape space. Dynamic time warping (DTW) distance is used to measure the degree of similarity between two shapes. Shape classification is based on multidimensional scaling. Experimental results indicate that the approach provides satisfied classification.

Key words: DTW distance; MDS; statistical shape classification

0 引言

自动识别任意物体的形状是项复杂的工作,这是因为当物体发生平移、旋转、尺度变换甚至扭曲、遮挡后,使计算机快速、准确地识别任意物体的形状是一件相当困难的事情。基于物体形状的识别,通常是给定任意物体的图像,通过图像分割得到物体所在的区域,然后对该区域的边界进行轮廓跟踪,提取出该物体的形状轮廓,进行识别。文中只针对物体轮廓是平面、闭合、简单曲线标准形状库,进行聚类。

1 形状的预处理

在不同位置、不同角度、不同姿态或不同伸缩比例下,同一个对象能够形成多种不同的形状。需要将其降维归一化到 R^M 空间统一处理,找到高维空间点的规律,定义合适的相似度,就可以准确地描述形状之间的距离,再将图像通过多维尺度分析(MDS),降维

为低维空间中的点,达到聚类的目的。

1.1 形状的表达

在计算机中物体的形状有多种表示方式:边界曲线上各点的坐标序列,边界上的点到物体重心的距离以及边界上的点随弧长变化的切线方向和曲率等。基于不同的形状表示方式,许多种形状识别方法已经被提出,如:基于边界特征点^[1],不变矩^[2,3],傅里叶描述子^[4]和自回归模型^[5]等方法。在一维曲线的轮廓形状识别中,需要从这些轮廓线的坐标出发来讨论问题:即跟踪轮廓线并提取轮廓坐标,然后再利用这样得到的信息来识别二维曲线形状。计算机进行边缘搜索得到的轮廓是直角坐标系中的一些离散的点,在形状分析时需要用到一维处理方法,不便使用曲线形状分析。需要将坐标从二维降至一维但要保留二维轮廓形状信息。基本思想是:假定物体的形状是一条封闭的曲线,沿边界曲线上所有点的序列为: $\{x(k), y(k); k = 0, 1, \dots, n-1\}$, 用复数的形式表示为: $z(k) = x(k) + jy(k)$, ($j = \sqrt{-1}$), 这样,形状 z 可由点序列 $z = (z(1), z(2), \dots, z(N-1))$ 表示。则将一维序列的离散傅里叶系数定义为:

$$Z(m) = \sum_{k=0}^{n-1} z(k) \exp[-j2\pi\omega k/n] = R_m e^{j\theta_m} \quad (1)$$

收稿日期:2006-05-09

基金项目:国家自然科学基金(60375010);安徽省人才开发基金(2001Z021);安徽省教育厅项目(2006KJ053B)

作者简介:张春燕(1976-),女,安徽铜陵人,博士研究生,讲师,主要研究方向为统计模式识别;罗斌,教授,主要研究方向为模式识别。

$$m = -N/2, \dots, -1, 0, 1, \dots, N/2 - 1$$

其中 R_m 和 θ_m 分别是经过离散傅里叶变换后得到的第 m 个系数的幅度和相位。目前,许多基于轮廓的傅里叶描述子不使用傅里叶变换后的相位系数,因为相位不具有旋转不变性,会随着图像旋转或轮廓起始点不同而产生位移。但是如果简单地丢弃相位信息,也会丢失形状的某些特征,因此如果不考虑相位信息,就会带来一些描述错误。因此,应保留傅里叶变换后的相位信息,并做归一化处理。

1.2 边缘的降维

考虑形状的能量大多集中在低频部分,傅里叶变换的高频分量很小且容易受到高频噪声的干扰,由于傅里叶变换的高频分量对应一些细节,而低频分量对应总体形状,所以在应用中只取 M 维低频分量的傅里叶系数来近似描述边界的形状,这样有 2 个优点:1) 可以对形状进行平滑,滤掉部分噪声;2) 大大减少运算量。但是 M 选取多少合适呢?这个需要在表示的准确和有效性之间做出选择, M 选择大一些,获得的系数多,傅里叶逆变换后得到的物体轮廓点也多。可以通过这些系数所代表的能量大小来考察,然后选取一个经验数值,定义能量函数:

$$E(X) = \sum_{m=-\frac{M}{2}, m \neq 0}^{\frac{M}{2}-1} |Z_m|^2$$

以 MPEG7-SETB 为例(见图 1),计算 $E(M)/E(N)$,结果如图 2 所示。

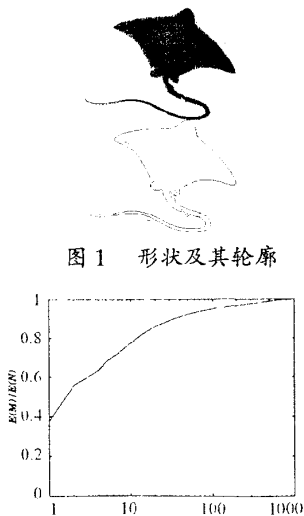


图 1 形状及其轮廓

从图中可以看到当选择 $m \in [16; 64]$ 时,能量比 $E(M)/E(N)$ 的变化范围在 84% ~ 93%。可以根据形状的复杂程度,选择一个较小的 M ,比如 $M = 64$,表示该形状。以图 1 的形状为例,通过 canny 算子提

取边缘后得出形状边界像素点集序列为 1164 个采样点,降维后仅剩 64 点能代表形状的信息图(见图 3(c)),因此在保留形状特征的同时大大降低了傅里叶变换的运算量,并消除了曲线离散化采样引起的误差。

1.3 归一化

在频域中对 Z 进行归一化处理,解决平移、缩放、旋转的问题。傅里叶变换是一一对应的映射,一组描述符对应唯一的一个边界轮廓图像与形状的尺度、方向和曲线的起始点位置有关。为了识别具有旋转、平移和尺度不变性的形状,需要在频域对傅里叶变换的系数进行归一化^[6,7],见表 1。

表 1 归一化傅里叶系数

不变性	修正系数
平移	$Z(0) = 0$
缩放	$R(m) = R(m)/R(1)$
旋转	$\theta(m) = \theta(m) - \frac{\theta_{-1} + \theta_1}{2}$
	$\theta(m) = \theta(m) + m \frac{\theta_{-1} + \theta_1}{2}$

在物体形状的频域做降维和归一化处理后,再使用傅里叶逆变换:

$$\hat{z}(i) = \frac{1}{M} \sum_{m=-\frac{M}{2}}^{\frac{M}{2}-1} \hat{Z}(m) \exp[-j2\pi im/M] \quad (2)$$

重新恢复物体形状,这样就得到抽样的空域归一化的物体形状轮廓,该轮廓有 M 个点,取这 M 个轮廓点作描述形状的特征向量。从 MPEG-7 形状库 SETB 中分别选择了复杂程度不一样的形状:狗(dog-1.jpg)(见图 3(a)),五瓣花(device0-10.jpg)(见图 3(b)),蝙蝠鱼(kk-471.jpg)(见图 3(c)),取 $M = 64$ 进行降维归一化。

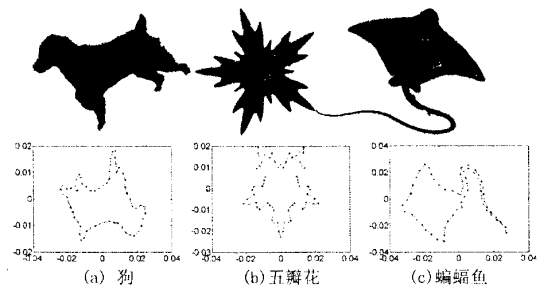


图 3 基于傅里叶的降维归一化图

2 形状空间的相似度

在识别中,距离及差异性(或相似性)度量方法是最基础的概念,度量方法的选择直接影响聚类结果的质量,可以说,许多聚类算法的差异性主要体现在差异性度量方法。目前,二维形状聚类 and 检索工作中已有很多的距离计算方法,来计算不同形状间的距离:任意

两个 M 维的形状边缘序 $z = \{z(0), z(1), \dots, z(M-1)\}$ 和 $z' = \{z'(0), z'(1), \dots, z'(M-1)\}$, 可以定义如下 3 个距离:

(1) Euclidean 距离度量, 是最常用的距离, 通过归一化傅里叶描述子可以计算任意两个形状 z 和 z' 间的相似程度, 用欧氏距离计算归一化傅里叶描述子间的形状差异为: $d_{\text{EU}}(z, z') = \sqrt{\sum_{i=0}^{M-1} (z(i) - z'(i))^2}$ (3) 但欧氏距离对噪声非常敏感, 且不能直接对傅里叶算子的模进行处理, 不能进一步对形状库聚类 and 匹配。

(2) 角分离度 (cos) 距离, 常用来度量两个样本向量各自方向上的单位向量之间的角度, 尤其是相对数值比较重要时, 可使用此度量, 以解决噪声和局部遮挡问题。

$$d_{\text{cos}}(z, z') = \frac{\sum_{i=1}^M z(i)z'(i)}{\sqrt{\sum_{i=1}^M z(i)^2 \sum_{i=1}^M z'(i)^2}} = \frac{\langle z(i), z'(i) \rangle}{\sqrt{\|z\|^2 \|z'\|^2}} \quad (4)$$

(3) DTW (Dynamic Time Warping) 距离度量是基于动态规划的思想, 解决了发音长短不一的匹配问题, 是语音识别中出现较早、较为经典的一种算法, 在语音识别中获得了很大的成功。Ross McConnell^[8]应用它来识别 SAR 图像, Hanqing Lu^[9]改进了动态规划算法, 提出动态匹配的概念, 将其用于图像相似度的度量。定义 $\text{tail}(z) = (z_1, \dots, z_{M-1})$, $d_{\text{rw}}(<, >) = 0$, $d_{\text{rw}}(z, < >) = d_{\text{rw}}(<, z') = \infty$, $d_{\text{rw}}(z, z')$ 定义为:

$$d_{\text{rw}}(z, z') = \sqrt{\delta(z(0), z'(0)) + \min \begin{cases} d_{\text{rw}}(z, \text{tail}(z')) \\ d_{\text{rw}}(\text{tail}(z), z') \\ d_{\text{rw}}(\text{tail}(z), \text{tail}(z')) \end{cases}} \quad (5)$$

式中, δ 是任意距离 (文中采用 EU 距离, $\delta(z(0), z'(0)) = |z(0) - z'(0)|^2$)。

Euclidean 可以看成是 DTW 距离的特殊情况, DTW 距离大大消除了形状对噪声的敏感性, 使基于形状的识别成为可能。

3 MDS 投影

多维尺度变换是用于分析距离矩阵或相异 (dissimilarity) 程度矩阵 (近邻矩阵) 以形成在降维空间中表示数据点的一类方法。通过低维空间, 通常是二维空间展示多个对象之间的联系, 利用平面距离来反映研究对象之间的相似程度, 产生一张能够看出这些对

象相关性的感知图, 进行聚类。给定一个数据矩阵, 构造 $n \times n$ 阶相异程度矩阵 D 和距离测度, 在 R^e 空间中找到 n 个 M 维点 z_1, \dots, z_n 的结构, 其中 $z_i = (z_i(0), \dots, z_i(M-1))$, 为 M 维的形状空间中的点, 使点对之间的距离在一定意义上接近于相异程度。该方法能找到点的坐标和空间 e 的维数 $e \leq M$ 。给定 M 维空间中的 n 个点: z_1, \dots, z_n , 计算每个点对之间的距离 d_{ij} , 构造距离矩阵 $D = (d_{ij})_{n \times n}$, 在 e 维空间中确定一组点的坐标, 可以通过对 $n \times n$ 阶矩阵 T 进行分解得到。 $T = ZZ^T$, 由点 z_i 和 z_j 之间的距离为 d_{ij} , 则:

$$T_{ij} = -\frac{1}{2} [d_{ij}^2 - d_{i.}^2 - d_{.j}^2 + d_{..}^2] \quad (6)$$

其中, $d_{i.} = \frac{1}{n} \sum_{j=1}^n d_{ij}^2$, $d_{.j} = \frac{1}{n} \sum_{i=1}^n d_{ij}^2$,

$d_{..} = \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n d_{ij}^2$, $Z = [z_1, \dots, z_n]^T$ 是 $n \times p$ 阶坐标矩阵。因为 T 是实对称矩阵, 所以可以写成 $T = U\Lambda U^T$, 其中, U 的列是 T 的特征向量, Λ 是特征值为 $\lambda_1, \dots, \lambda_n$ 的对角矩阵。因此, 取 $Z = U\Lambda^{1/2}$ 为坐标矩阵。可以对特征值排序: $\lambda_1 \geq \dots \geq \lambda_n = 0$; $\lambda_{p+1} = 0$, 要在降维空间中表示数据, 则仅使用那些与最大特征值关联的特征向量。选择 r 个特征, 以使 $\sum_{i=1}^r \lambda_i < k$, $\sum_{i=1}^n \lambda_i < \sum_{i=1}^r \lambda_i$, 于是可取 $Z = [u_1, \dots, u_r] \text{diag}(\lambda_1^{1/2}, \dots, \lambda_r^{1/2}) = U_r \Lambda_r^{1/2}$ 作为 $n \times r$ 阶坐标矩阵, 其中 Λ_r 是 $r \times r$ 阶对角矩阵, 其对角线元素为 λ_i 。试验中取 $r = 2$, 将 M 维的形状空间中的结果投影到 2 维的空间中。

4 实验与结果分析

进行对象实验的对象数据集为 MPEG-7 中 SETB(1400) 数据库, 数据库中共有 700 个物体形状轮廓, 每个对象有 20 幅不同角度、不同大小的轮廓图像, 整个数据库共有 1400 幅图像, 向量化图像后, 取 M 的值为 64, 则观察数据的维数有 1400×64 。选定 3 类图像共 60 幅图像, 见图 4。首先使用傅里叶降维归一化后, 分别用欧式距离、角分离度距离、DTW 距离计算距离矩阵, 由多维尺度分析方法投影到 R^2 中, 图 5 为三种距离下的 MDS 聚类投影结果。

实验结果显示: Euclidean 距离能很好地对五瓣花聚类, 但是其他两类会产生混淆, 因为五瓣花只有局部信息不同, 在傅里叶降维归一化下, Euclidean 距离会很近, 但其他两类聚类效果不好, 是因为 Euclidean 距离不能消除边缘轮廓的变形; 角分离度距离则不能描述形状间的差异, 没有把三类分开; DTW 距离下的 MDS 将这三类形状聚合的最好, 既对局部不敏感, 又能消除轮廓的变形, 是形状分析中合适的相似度。



图4 三类图像序列

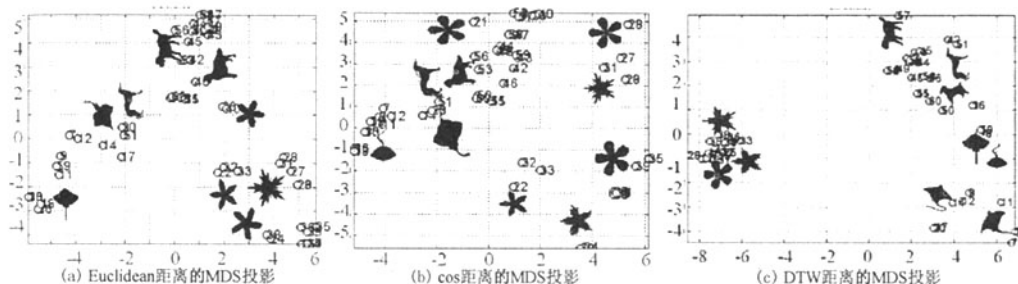


图5 三种距离的MDS投影

参考文献:

- [1] Chang C C, Hwang S M, Buehrer D J. A shape recognition scheme based on relative distances of feature points from the centroid[J]. Pattern Recognition, 1991, 24 (11): 1053 - 1063.
- [2] Teague M R. Image analysis via the general theory of moments [J]. Optical Society of America, Journal, 1980, 70(8): 920 - 930.
- [3] Dudani S A, Breeing K J, McGhee R B. Aircraft identification by moment invariants[J]. IEEE Trans Computers, 1977, C-26(1): 39 - 46.
- [4] Zahn C T, Roskies R Z. Fourier Descriptors for Plane Closed Curves[J]. IEEE Trans Computers, 1972, C-21: 269 - 281.
- [5] Dubois S R, Glanz F H. An auto regressive model approach to two - dimensional shape classification[J]. IEEE Trans on PA-MI, 1986, 8: 55 - 56.
- [6] Wallace T P, Wintz P A. An Efficient Three - Dimensional Aircraft Recognition Algorithm Using Normalized Fourier Descriptors[J]. Computer Graphics and Image Processing, 1980, 13: 99 - 126.
- [7] Bartolini I, Ciaccia P, Patella M. WARP: Accurate Retrieval of Shapes Using Phase of Fourier Descriptors and Time Warping Distance[J]. IEEE Trans Pattern Anal Mach Intell, 2005, 27 (1): 142 - 147.
- [8] McConnell R, Kwok R, Curlander J C, et al. Ψ -s correlation and dynamic time warping: Two methods for tracking ice floes in SAR image[J]. IEEE Trans Geoscience and Remote Sensing, 1991, 29(6): 1004 - 1012.
- [9] Lu Hanqing, Songde M A. Representation of feature point and error analysis[R]. Beijing: Institute of Automation, Chinese Academy of Sciences, 1993.

(上接第 57 页)

4 结束语

文中讨论的行列并行计算的二维离散小波变换优化算法,在具有并行结构的处理器上实现时具有较大应用价值,不仅提高了运算速度,同时减少了运算时的内存需求。对于包含图像处理和多媒体功能的嵌入式产品开发具有很强的实用价值。

参考文献:

- [1] Sweldens W. The lifting scheme: a construction of second generation wavelets[J]. SIAM J Math Anal, 1997, 29(2): 511 - 546.
- [2] Sweldens D W. Factoring wavelet transforms into lifting steps [J]. Fourier Anal Appl, 1998, 4(3): 247 - 269.
- [3] Chrysafis C, Ortega A. Line based reduced memory, wavelet image compression[J]. IEEE Transaction on Image Processing, 2000, 9(3): 378 - 389.
- [4] 孟 军. 基于提升算法离散小波变换的静止图像编码系统实现[D]. 南京: 东南大学, 2003.
- [5] 李忠锋. 基于 C6000 DSP 的 JPEG2000 图像压缩技术研究 [D]. 武汉: 华中科技大学, 2004.
- [6] ISO/IEC15444 - 1 (ITU Recommendation T. 800) JPEG 2000 Image Coding System[S]. 2003.
- [7] 许宪东, 安 波, 王亚东. 基于 JPEG/JPEG2000 的静态图像编码技术[J]. 哈尔滨师范大学学报, 2004(3): 69 - 71.
- [8] 夏 怡. 第二代小波在 JPEG2000 中的应用[J]. 计算机与现代化, 2005(9): 21 - 24.