

# 球面调和描述子在图像形状匹配中的应用

章志勇<sup>1</sup> 杨柏林<sup>1</sup>

**摘要** 球面调和分析是球面信号分析的重要工具, 本文采用球面调和描述子对图像进行形状匹配. 首先把平面图像映射到球面上, 然后计算球面图像的球面调和描述子, 最后通过球面调和描述子得到平面图像的相似性. 实验结果表明, 球面调和描述子可以较好地匹配图像的形状相似性, 并且具有旋转和比例不变性.

**关键词** 球面调和分析, 形状匹配, 形状描述子

**中图分类号** TP391.41

## Spherical Harmonic Descriptor for Image Shape Matching

ZHANG Zhi-Yong<sup>1</sup> YANG Bai-Lin<sup>1</sup>

**Abstract** Spherical harmonics is of importance for spherical signal analysis. This paper matches shape similarity among images by a spherical harmonic descriptor. First, images are mapped into spherical images, then the similarities are matched among spherical images by a spherical harmonic descriptor. The experimental results show that the spherical harmonic descriptor is not only precise in image shape matching but also invariant to rotation and scale transforms.

**Key words** Spherical harmonics, shape matching, shape descriptor

### 1 引言

图像形状相似性比较是模式识别与图像检索的重要研究内容, 目前有多种图像形状相似性比较算法, 其中较常用的方法是利用函数描述子计算图像的形状特征. 函数描述子根据其计算的图像特征可以分为轮廓描述子和图像描述子. 轮廓描述子通过计算图像的轮廓特征来比较图像的形状相似性, 这类描述子主要用于二值图像的形状比较. 目前, 常用的轮廓描述子是傅立叶描述子 (Fourier descriptor, FD)<sup>[1,2]</sup>, 傅立叶描述子具有旋转和比例不变性的特点. 图像描述子是利用整个图像的形状特征来比较图像的形状相似性, 图像的形状特征包括图像的轮廓特征和轮廓内部区域的特征. 几何矩不变量描述子<sup>[3]</sup>在图像的形状比较中得到了应用, 但是几何矩不是正交矩, 并且计算量很大. Legendre 描述子<sup>[4]</sup>是正交描述子, 具有比例不变性的优点, 但是它不具有图像旋转不变性. Zernike 描述子<sup>[5,6]</sup>是正交描述子, 具有比例和旋转不变性, 它被 MPEG7 标准选为图像形状匹配描述子. 但是, Zhang 认为 Zernike 描述子不能有效地比较图像的径向特征, 并且提出了通用傅立叶描述子 (Generic Fourier descriptor, GFD)<sup>[7]</sup>. 文

献 [8] 提出了 Fourier-Mellin 描述子, Fourier-Mellin 描述子具有图像旋转不变性, 但是 Fourier-Mellin 描述子计算复杂, 并且累积误差很大. 球面调和分析 (Spherical harmonics) 是球面信号分析的重要工具, 它在三维网格的计算中得到了广泛的应用, 但是球面调和分析与球面信号的坐标系方向相关. Michael<sup>[9]</sup>基于球面调和分析, 提出了一种与坐标系方向无关的球面调和描述子, 并且把它用于三维模型的特征计算. Marcin<sup>[10]</sup>基于同样的具有旋转不变性的球面调和描述子, 提出了一种三维模型相似性比较的方法. 球面调和分析在平面图像的处理中应用较少, 特别是在图像的相似性比较领域中应用很少, 我们可以把平面图像映射为球面图像, 然后采用球面调和描述子来比较平面图像的相似性.

### 2 球面调和描述子

如图 1 所示, 球面调和分析是建立在球面坐标系上, 其计算如下

$$a_{n,m} = \int_0^{2\pi} \int_0^\pi \sin(\theta) f(\theta, \varphi) \bar{Y}_n^m d\theta d\varphi \quad (1)$$

$$f(\theta, \varphi) = \sum_{n=0}^{\infty} \sum_{m=-n}^n a_{n,m} Y_n^m(\theta, \varphi) \quad (2)$$

$$Y_n^m(\theta, \varphi) = \sqrt{\frac{(2n+1)(n-m)!}{4\pi(n+m)!}} P_n^m(\cos(\theta)) e^{im\varphi} \quad (3)$$

收稿日期 2005-11-21 收修改稿日期 2006-6-28  
Received November 21, 2005; in revised form June 28, 2006  
浙江省科技计划项目 (2005C13021) 资助  
Supported by Technological Project of Zhejiang Province (2005C13021)  
1. 浙江工商大学计算机与信息工程学院 杭州 310035  
1. College of Computer Science and Information Engineering, Zhejiang Gongshang University, Hangzhou 310035  
DOI: 10.1360/aas-007-0683

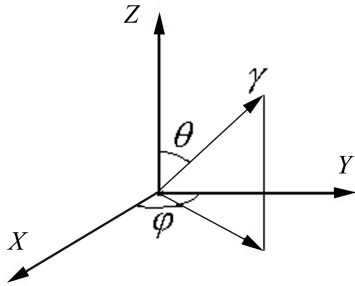


图 1 球面坐标系

Fig. 1 Spherical coordinate

式中,  $P_n^m(x)$  是 Legendre 多项式,  $n$  是球面调和函数系数的频带宽度,  $m$  是频带宽度下的频率分量. 球面调和函数具有一些重要的数学性质 [9,10], 设子空间  $V_l$  是如下一些分量所张成的空间

$$V_l = \text{span}(Y_l^{-l}, Y_l^{-l+1}, \dots, Y_l^l) \quad (4)$$

那么有:

- 1) 球面调和函数是一组基函数, 它们满足正交归一性质.
- 2) 固定频率  $l$  的球面调和函数构成了一个旋转变换下的不变子空间.
- 3) 每个固定频率  $l$  的系数的模的平方和是旋转不变量, 也就是说, 具有旋转不变性.

球面调和描述子  $SHD$  如式 (5) 所示, 其中,  $\|SH_n(\theta, \varphi)\|$  是  $SH_n(\theta, \varphi)$  的二阶范数.

$$SHD = \left\{ \frac{\|SH_1(\theta, \varphi)\|}{\|SH_0(\theta, \varphi)\|}, \frac{\|SH_2(\theta, \varphi)\|}{\|SH_0(\theta, \varphi)\|}, \dots \right\} \quad (5)$$

$$SH_l(\theta, \varphi) = \sum_{t=-l}^l a_{l,t} Y_l^t(\theta, \varphi) \quad (6)$$

从球面调和函数的数学性质可以看到, 球面调和描述子对球面信号具有旋转和比例不变性. 从球面调和描述子的物理意义上看, 球面调和描述子的每个分量是每个频带的能量总和与基带的能量总和的比例, 因此球面调和描述子对球面信号具有旋转和比例不变性.

### 3 图像检索实验

球面调和描述子可以用于比较球面图像的形状相似性, 并且具有坐标系旋转和比例不变性. 对于平面图像的形状相似性比较, 我们首先可以把平面图像映射为球面图像, 然后采用球面调和描述子来比较平面图像的相似性.

### 3.1 平面图像的球面映射

平面图像到球面图像可以采用不同的映射方法, 我们比较分析了两种映射方法, 一种是直接把平面图像映射到球面, 得到球面图像; 另外一种是先对平面图像进行极坐标采样, 然后把采样映射到球面上得到球面图像.

#### 3.1.1 平面图像到球面的直接映射方法

把平面图像直接映射到球面上, 平面图像的高宽两个方向分别作为球面的纬度线和经度线. 如图 2 所示, 平面图像的每一行对应球面的纬度线, 每一列对应球面的经度线. 对于直接映射方法, 球面图像绕  $Z$  轴旋转相当于平面图像的平移反绕变化. 如图 3 所示, 采用直接映射方法把 Lena 图像映射到球面上, 然后对球面图像绕  $Z$  轴旋转得到 Lena 图像的平移反绕图像, 这种图像的平移反绕现象在周期性的图像中比较常见. 对于平面图像到球面图像的直接映射, 平面图像不需要进行像素点采样就可以直接映射到球面上, 平面图像的像素信息可以完整地保存在相应的球面图像中.

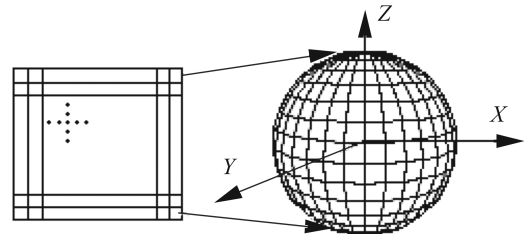


图 2 球面图像直接球面映射

Fig. 2 Direct mapping of image

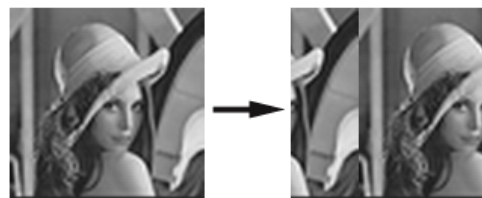


图 3 平移反绕图像

Fig. 3 Image roll back

#### 3.1.2 平面图像到球面的极坐标映射方法

首先对平面图像进行极坐标采样, 然后把采样映射到球面上. 如图 4 所示, 首先对平面图像进行极坐标采样, 然后把采样同心圆由里到外映射为球面的纬度线. 对于极坐标映射方法, 球面绕  $Z$  轴旋转相当于平面图像的旋转变化, 如图 5 所示. 对于平面图像的极坐标采样, 离极坐标中心越远, 采样密度越低, 为了尽可能充分地保留图像的信息, 极坐标的采

样密度可以根据图像的大小进行调整, 使得每个采样点的大小小于图像像素的大小.

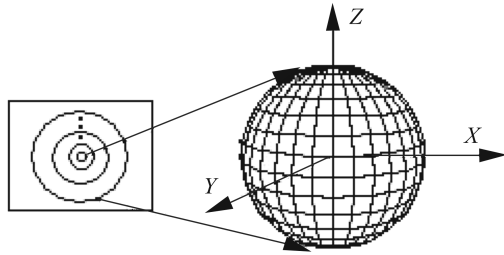


图4 平面图像极坐标映射  
Fig. 4 Polar mapping of image

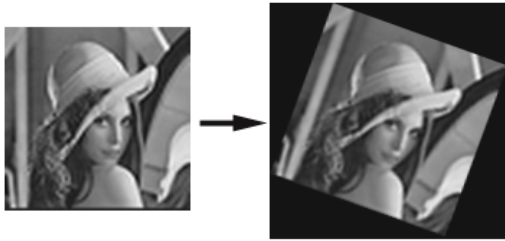


图5 图像旋转变化  
Fig. 5 Image rotation

### 3.2 球面调和描述子频带宽度实验

从球面调和描述子的计算公式 (5) 可以看到, 我们需要对球面调和描述子选择一个合适的频带宽度, 使得球面调和描述子能较准确地衡量信号的特征. 我们对图 6(a)(见下页上方) 所示的球面信号进行球面调和, 然后使用不同频带宽度的球面调和系数进行信号重构. 如图 6 所示, 当频带宽度  $l = 32$  时, 重构的信号与原球面信号差别很小.

我们对 52 种不同的球面信号进行了频带宽度  $l = 32$  的重构, 然后计算重构的球面信号与原球面信号的误差, 误差计算如式 (7) 所示

$$diff = \frac{1}{N} \sum_{i=1}^N \frac{|OriSig - Sig_{32}|}{|OriSig|} \quad (7)$$

式中,  $OriSig$  是原球面信号,  $Sig_{32}$  是采用频带宽度为 32 的球面调和系数重构的球面信号,  $N$  是球面信号采样个数. 我们对 52 种不同的球面信号进行重构实验统计, 当频带宽度为 32 时, 平均误差  $diff$  不超过 8%. 由分析结果可以看出, 由于球面调和系数  $l = 32$  的频带系数可以较为精确地描述球面信号, 因此本算法中的频带宽度  $l$  选择为 32.

### 3.3 球面调和描述子检索实验

我们对球面调和描述子, GFD, Zernike 描述子进行了对比检索实验, 实验采用的机器配置为 P3

600M, 256M 内存. 对于球面调和描述子, 它的计算步骤如下:

步骤 1. 把平面图像映射为球面图像.

步骤 2. 计算球面信号频带宽度  $l \leq 32$  的球面调和系数, 然后计算球面调和描述子.

步骤 3. 计算图像之间的相似度, 相似度计算如式 (8) 所示, 相似度越大, 图像的相似性越好.

$$Sim = 1 - K * \sum_{l=0}^{32} \sqrt{64-l} |SHD_1(l) - SHD_2(l)| \quad (8)$$

图像的信息主要集中在球面调和系数的低频和中低频率处, 系数  $\sqrt{64-l}$  用于体现不同频带下的加权,  $K$  为常数.

在实验中, 我们采用了 MPEG7 CE-Shape 1 和 MPEG7 CE-Shape 2 图像数据库, MPEG7 CE-Shape1 图像数据库中包含了 52 类, 共 856 幅图像. MPEG7 CE-Shape 2 图像数据库包含了 36 类, 共 758 幅图片. 我们在这个图像数据库的基础上进行了图像的检索实验, 并且根据实验结果计算查全率 (Recall) 与查准率 (Precision) 的对比图, 查全率和查准率的定义如下

$$Recall = \frac{\text{relevant correctly retrievad}}{\text{all relevant}} \quad (9)$$

$$Precision = \frac{\text{relevant correctly retrieved}}{\text{all retrieved}} \quad (10)$$

其中,  $\text{relevant correctly retrievad}$  是正确检索的相关图像,  $\text{all relevant}$  是所有的相关图像,  $\text{all retrieved}$  是所有的检索结果. 查全率表示返回的正确检索图像占整个相关图像的比例, 这个指标衡量检索系统返回正确结果的能力. 查准率表示在所有检索返回结果中, 正确检索的图像的比例, 这个指标衡量检索返回结果的精确性.

#### 3.3.1 球面调和描述子检索实验

我们利用图像数据库对球面调和描述子, GFD, Zernike 描述子进行了检索实验. 在球面调和描述子的实验中, 分别采用直接映射方法和极坐标映射方法把平面图像映射到球面上. 实验结果如图 7 所示, 从图 7 可以看到, 当采用极坐标映射方法时, 球面调和描述子可以取得较好的检索效果. 当采用直接映射方法时, 检索的准确性比较低, 这是由于直接映射方式不具有旋转不变性. 图 8 是球面调和描述子的一组检索结果.

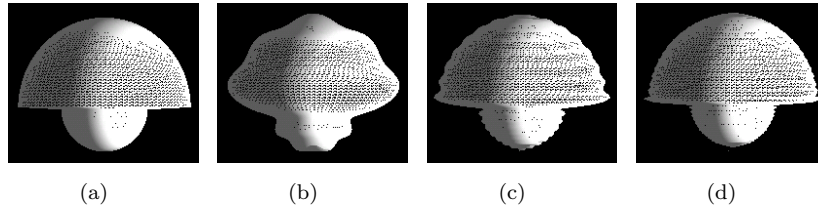


图 6 球面信号重构 ((a) 原球面信号; (b)  $l = 16$  的重构信号; (c)  $l = 32$  的重构信号; (d)  $l = 64$  的重构信号)  
 Fig. 6 Spherical signal reconstruction ((a) The original Spherical Signal; (b) The reconstruction of  $l = 16$ ; (c) The reconstruction of  $l = 32$ ; (d) The reconstruction of  $l = 64$ )

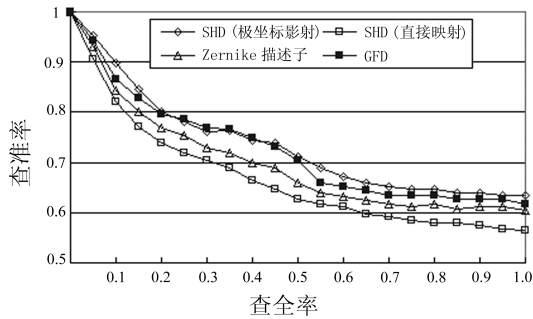


图 7 不同描述子的图像检索结果  
 Fig. 7 Performance evaluation of our approach, Zernike descriptor, and GFD

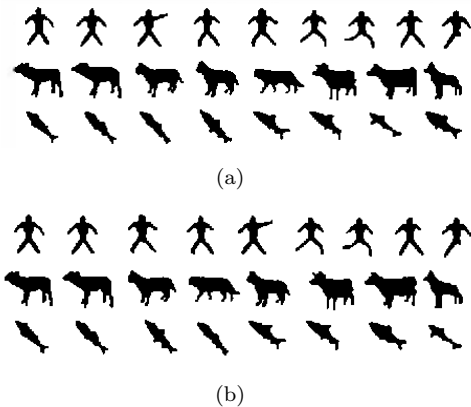


图 8 球面调和描述子图像结果 ((a) 极坐标采样映射方法的检索结果; (b) 直接映射方法的检索结果)  
 Fig. 8 Image retrieval result of spherical harmonic descriptor ((a) Retrieval results with polar mapping; (b) Retrieval results with direct mapping)

在实验中, 四种方法的时间花费统计结果如表 1 所示. 从实验结果可以看到, 球面调和描述子的计算时间要慢于 GFD 方法, 这是由于快速傅立叶算法的计算复杂度要低于球面调和描述子的计算复杂度. 但是由于我们只需要计算球面调和和分析系数的前 32 个频带系数, 因此计算的速度依然是可以接受的. 球面调和描述子的检索时间要好于 Zernike 描述子和 GFD, 这是因为检索的时间主要花费在相

似度计算上, 而相似度计算的时间取决于描述子的特征向量维数. 在实验中, 球面调和描述子的维数是 32, Zernike 描述子的维数是 64, GFD 的维数是 128.

表 1 特征计算和检索时间  
 Table 1 The time of character calculation and retrieval

时间	SHD (极坐标映射方法)	SHD(直接映射方法)	Zernike 矩	GFD
特征计算	1.28	1.22	1.37	0.89
检索	1.43	1.43	2.98	4.82

从整个检索实验分析可以看到, 当采用极坐标映射方法时, 球面调和描述子的检索准确性要好于 GFD 的检索准确性, 并且在检索试验中所用的总时间要少于 GFD.

### 3.3.2 平移反绕变换和旋转变换检索实验

对图像进行平移反绕变换, 然后采用直接映射方法把平移反绕的图像映射为球面图像, 最后采用球面调和描述子比较它们的相似性, 实验结果如图 9 所示. 从图 9 可以看到, 虽然球面调和描述子只采用了前 32 个频带的球面调和和分析系数, 但是 32 个频带宽度带来的特征损失是可以接受的. 实验结果表明, 采用直接映射方法, 球面调和描述子具有较好的平移反绕鲁棒性, 其检索偏差在合理的范围内.

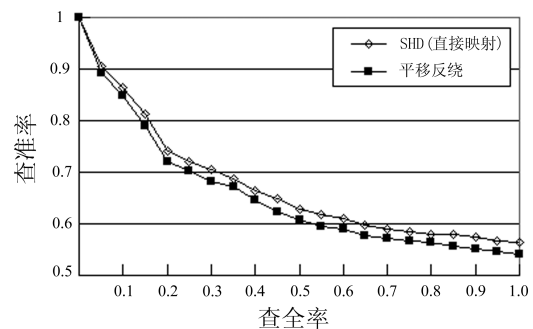


图 9 平移反绕变换检索实验  
 Fig. 9 Retrieval result of image roll back

对图像顺时针旋转一个随机角度, 然后采用极坐标映射方法把旋转图像映射为球面图像, 最后采

用球面调和描述子比较它们的相似性, 实验结果如图 10 所示. 从图 10 可以看到, 前 32 个频带宽度带来的信号特征损失是可以接受的, 球面调和描述子具有较好的旋转变换鲁棒性, 其检索偏差在合理的范围内.

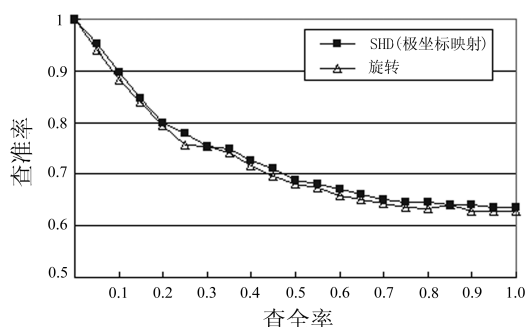


图 10 旋转变换检索实验

Fig. 10 Retrieval result of image rotation

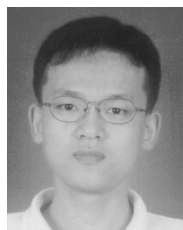
## 4 结论

球面调和描述子是一个三维的描述子, 它具有旋转和比例不变性. 本文把平面图像映射到球面上得到球面图像, 然后使用球面调和描述子比较球面图像的形状相似性, 最后得到平面图像的形状相似性. 本文分析比较了两种不同的平面图像到球面图像的映射方法: 一种是直接映射方法, 另一种是极坐标映射方法. 当采用直接映射方法时, 球面调和描述子对平面图像具有平移反绕不变性. 当采用极坐标映射方法时, 球面调和描述子对平面图像具有旋转不变性. 实验结果表明, 球面调和描述子可以较好地比较图像的形状相似性. 此外, 球面调和描述子是一个三维的描述子, 可以根据不同的使用场合, 选择不同的平面图像到球面图像的映射方法, 得到不同的平面图像相似性比较效果, 这使得球面调和描述子的使用可以非常灵活.

## References

- 1 Maciej M, Ahmad M O. Shape equalization of binary images using Fourier descriptor interpolation in low-level processing for image understand. In: Proceedings of IEEE Electrical and Computer Engineering. IEEE, 1995. 2: 1192~1195
- 2 Adpra M, Lew M S. IRUS: image retrieval using shape. In: Proceedings of IEEE International Conference on Multimedia Computing and System. IEEE, 1999. 2: 597~602

- 3 Hu M. Visual pattern recognition by moment invariants. *IRE Transactions on Information Theory*, 1962, 8(2): 179~187
- 4 Yun G E, Shu H Z, Luo L M. A new way of linear registration using the legendre orthogonal moment and application in 2-value image data. *Acta Electronica Sinica*, 2001, 29(1): 54~56
- 5 Zhang D S, Lu G J. Improving retrieval performance of Zernike moment descriptor on affined shapes. In: Proceedings of IEEE Multimedia and Expo. IEEE, 2002. 1: 205~208
- 6 Bin Y, Peng J X. Improvement and invariance analysis of Zernike moments. In: Proceedings of IEEE 2002 International Conference on Communications, Circuits and Systems and West Sino Expositions. IEEE, 2002. 2: 963~967
- 7 Zhang D S, Lu G J. Generic Fourier descriptor for shape-based image retrieval. In: Proceedings of IEEE International Conference on Multimedia and Expo. IEEE, 2002. 425~428
- 8 Stephane D. Robust and efficient Fourier-Mellin transform approximations for gray level image reconstruction. *Computer Vision and Image Understanding*, 2001, 83(1): 57~78
- 9 Michael K, Thomas F, Szymon R. Rotation invariant spherical harmonic representation of 3D shape descriptors. In: Proceedings of the 2003 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing. Aachen, Germany: Eurographics Association Press, 2003. 156~164
- 10 Marcin N, Klein R. 3D Zernike descriptors for content based shape retrieval. In: Proceedings of the Eighth ACM Symposium on Solid Modeling and Applications. Washington, USA: ACM Press, 2003. 216~225



章志勇 浙江工商大学讲师. 2005 年获得浙江大学计算机学院博士学位, 主要研究方向为数字图像处理, 虚拟现实技术. 本文通信作者.

E-mail: zzy@mail.zjgsu.edu.cn

(ZHANG Zhi-Yong Lecturer at Zhejiang Gongshang University. He received his Ph.D. degree from Zhejiang

University in 2005. His research interest covers digital image processing and virtual reality. Corresponding author of this paper.)



杨柏林 浙江大学 CAD & CG 国家重点实验室博士研究生. 主要研究方向为数字图像处理, 虚拟显示技术.

E-mail: ybl@cad.zju.edu.cn

(YANG Bai-Lin Ph.D. candidate at State Key Laboratory of CAD & CG, Zhejiang University. His research interest covers digital image processing and

virtual reality.)