



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学学报

2013
08
VOL.18

ISSN1006-8961
CN11-3758/TB



北魏



隋



初唐

数字文化遗产研究 P968

综述

尺度不变特征变换算子综述

刘立, 詹茵茵, 罗扬, 刘朝晖, 彭复员 885

图像处理和编码

改进的H.264帧间模式选择算法

任克强, 张旭光, 罗会兰 893

均质雾天下摄像机动态标定的车速估计

宋洪军, 陈阳舟, 郜园园 901

基于人脸五官特征的空域差错掩盖算法

张江鑫, 谢晋, 邝万坤 913

采用圆谱-傅里叶矩的图像区域复制粘贴篡改检测

秦娟, 李峰, 向凌云, 殷蓓蓓 919

图像分析和识别

基于视频感知哈希的视频篡改检测与多粒度定位

朱映映, 文振焜, 杜以华, 邓良太 924

视觉注意模型的道路监控视频关键帧提取

刘云鹏, 张三元, 王仁芳, 张引 933

格拉斯曼流形上的半监督判别分析

姜伟, 陆瑶, 杨炳儒 944

精神疲劳实时监测中多面部特征时序分类模型

陈云华, 张灵, 丁伍洋, 严明玉 953

图像分类中的概率乘积核函数

杨赛, 赵春霞 961

轮廓整体结构约束的壁画图像相似性度量

唐大伟, 鲁东明, 杨冰, 许端清 968

应用色彩空间聚类方法实现道路建模

向宸薇, 王拓, 于舰 976

图像理解和计算机视觉

杂乱背景和摄像机移动下的时空兴趣点检测

刘长红, 陈勇, 王明文 982

形状和颜色混合不变量在图像检索中的应用

公明, 曹伟国, 李华 990

计算机图形学

结合统计分类和边缘特征的最优视图提取

潘翔, 陈敖, 章国栋 1004

医学图像处理

局部熵驱动下的脑MR图像分割与偏移场恢复耦合模型

张建伟, 杨红, 陈允杰, 方林, 詹天明 1011

改进的模糊聚类实质肺结节3维分割

李翠芳, 聂生东, 王远军, 孙希文, 郑斌 1019

遥感图像处理

高光谱海岸带区域分割的活动轮廓模型

王相海, 金弋博 1031

改进的基于非局部均值的极化SAR相干斑抑制

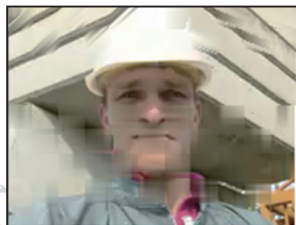
赵忠民, 赵拥军, 牛朝阳 1038

小波变换和稀疏表示相结合的遥感图像融合

刘婷, 程建 1045

《中国图象图形学报》论文摘要写作要求

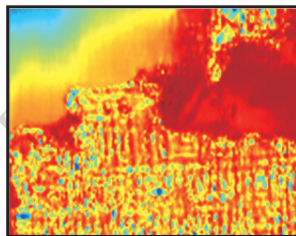
1054



基于人脸五官特征的空域差错掩盖算法 (第913页)



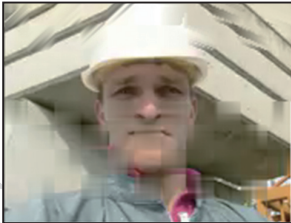
应用色彩空间聚类方法实现道路建模 (第976页)



改进的基于非局部均值的极化SAR相干斑抑制 (第1038页)

CONTENTS

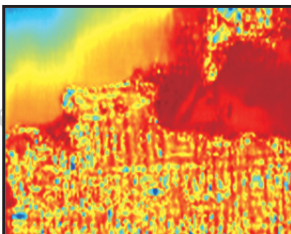
JOURNAL OF IMAGE AND GRAPHICS



Spatial error concealment algorithm based on the facial features (P913)



Road modeling using color spatial clustering (P976)



Improved polarimetric SAR speckle filter based on non-local means (P1038)

Review

- Summarization of the scale invariant feature transform
Liu Li , Zhan Yinyin, Luo Yang, Liu Chaohui, Peng Fuyuan 885

Image Processing and Coding

- Improved inter frame mode selection algorithm for H.264
Ren Keqiang, Zhang Xuguang, Luo Huilan 893
- Vehicle velocity estimation based on camera calibration under homogenous fog
Song Hongjun, Chen Yangzhou, Gao Yuanyuan 901
- Spatial error concealment algorithm based on the facial features
Zhang Jiangxin, Xie Jin, Kuang Wankun 913
- Detection of image region copy-move forgery using radial harmonic Fourier moments
Qin Juan, Li Feng, Xiang Lingyun, Yin Changming 919

Image Analysis and Recognition

- Video forgery detection and multi-granularity location based on video perceptual hashing
Zhu Yingying, Wen Zhenkun, Du Yihua, Deng Liangtai 924
- Key frame extraction based on the visual attention model for lane surveillance video
Liu Yunpeng, Zhang Sanyuan, Wang Renfang, Zhang Yin 933
- Semi-supervised discriminant analysis on Grassmannian manifold
Jiang Wei, Lu Yao, Yang Bingru 944
- Time-series classification model based on multiple facial feature for real-time mental fatigue monitoring
Chen Yunhua, Zhang Ling, Ding Wuyang, Yan Mingyu 953
- Probability product kernels in image classification
Yang Sai, Zhao Chunxia 961
- Similarity metrics between mural images with constraints of the overall structure of contours
Tang Dawei, Lu Dongming, Yang Bing, Xu Duanqing 968
- Road modeling using color spatial clustering
Xiang Chenwei, Wang Tuo, Yu Jian 976

Image Understanding and Computer Vision

- Spatio-temporal interest point detection in cluttered backgrounds with camera movements
Liu Changhong, Chen Yong, Wang Mingwen 982
- Application of combined shape and color invariants to image retrieval
Gong Ming, Cao Weiguo, Li Hua 990

Computer Graphics

- Best view selection based on statistical classification and edge features
Pan Xiang, Chen Ao, Zhang Guodong 1004

Medical Image Processing

- Brain MR image segmentation and bias correction model based on local entropy
Zhang Jianwei, Yang Hong, Chen Yunjie, Fang Ling, Zhan Tianming 1011
- Segmentation of sub-solid pulmonary nodules based on improved fuzzy C-means clustering
Li Cuifang, Nie Shengdong, Wang Yuanjun, Sun Xiwen, Zheng Bin 1019

Remote Sensing Image Processing

- The active contour model for segmentation of coastal hyperspectral remote sensing image
Wang Xianghai, Yin Yibo 1031
- Improved polarimetric SAR speckle filter based on non-local means
Zhao Zhongmin, Zhao Yongjun, Niu Chaoyang 1038
- Remote sensing image fusion with wavelet transform and sparse representation
Liu Ting, Cheng Jian 1045

中图法分类号: TP391.4 文献标识码: A 文章编号: 1006-8961(2013)08-0968-08

论文引用格式: 唐大伟, 鲁东明, 杨冰, 许端清. 轮廓整体结构约束的壁画图像相似性度量[J]. 中国图象图形学报, 2013, 18(8): 968-975.
[DOI:10.11843/jig.20130811]

轮廓整体结构约束的壁画图像相似性度量

唐大伟, 鲁东明, 杨冰, 许端清

浙江大学计算机科学与技术学院, 杭州 310027

摘要: 图像分类是计算机视觉的重要研究领域, 选择一种特征建立图像间的相似性度量是图像分类的关键问题。鉴于壁画图像自身的特点, 轮廓特征是能够表达壁画图像语义的重要特征。研究表明轮廓可以作为图像的重要特征进行图像的识别和分类, 但以往研究往往是通过两两最相似轮廓间的 chamfer 距离来计算图像间的相似性, 或者对轮廓建立局部描述符, 聚类生成词典, 用统计直方图的方式描述图像特征, 然后用支持向量机(SVM)进行图像分类。这些方法都忽略了轮廓间的整体结构关系, 缺乏对所有轮廓的整体性描述, 而现实中一幅图像的语义更多的是一种整体上的语义。基于轮廓整体结构的图像间相似性度量方法, 图像间轮廓的相似度计算要受到与其他轮廓空间结构关系的约束, 由此生成的相似度更能够表达两幅图像的整体相似程度。实验结果表明本文方法在壁画图像的分类上相对于没有整体结构约束的方法精度有所提高。

关键词: 轮廓特征; 整体结构; 相似度; 图像分类

Similarity metrics between mural images with constraints of the overall structure of contours

Tang Dawei, Lu Dongming, Yang Bing, Xu Duanqing

College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China

Abstract: Image classification is an important research field of computer vision, and the key problem of which is to select a type of feature and establish the similarity metrics between images. In view of the mural image characteristics, the contour feature plays an important role in expressing the mural image semantics. Many studies have shown that the contours can be used as an important feature in image recognition and classification. However, previous studies tend to use the chamfer distance between each pair of the most similar contours to compute the similarity between images, or build local descriptors for each contour, clustering into codebook, and describe the image features as histograms, then do the classification using SVM. However, these methods ignore the overall structure between the contours, lack of the overall view of the all contours, while in reality the semantics of an image tend to be more of a holistic semantic. In this paper, we study the similarity metrics between images based on the overall structure of contours, the calculation of contour similarity is subject to the constraint of the space structure relations with other contours, the generated similarity are more able to express the overall similarity between two images. The experimental results show that our method improved accuracy compared to others in mural image classification.

Key words: contour features; overall structure; similarity; image classification

收稿日期: 2012-10-25; 修回日期: 2013-01-14

基金项目: 国家重点基础研究发展计划(973) 基金项目(2012CB725305)

第一作者简介: 唐大伟(1984—), 男, 浙江大学计算机科学与技术专业在读博士研究生, 主要研究方向为模式识别、图像处理。E-mail:

tangdawei@zju.edu.cn

0 引言

图像分类是计算机视觉重要的研究领域,其核心问题主要在两个方面:1)图像特征的描述;2)分类器的设计。图像特征是对图像自身属性的客观描述,直接关系到图像分类的准确度。

一个好的图像特征应该具备以下属性:通过该特征的描述,能够准确地表达同类图像之间的相似性,以及不同类图像之间的相异性。本文将通过描述敦煌壁画中不同类别艺术形象图像特征,从而建立壁画图像之间的相似性度量机制,最终实现壁画图像的自动分类。壁画图像是人工作品,与自然图像相比有着自身的特点:1)基本组成是线描和色块;2)随意性较高,相同类别的图像可能具有较大的视觉差异性。由特点1),轮廓和区域更加能够表达壁画图像的语义特征,然而区域内部纹理和色彩变化多端,对于壁画图像来说,区域特征反而不如轮廓特征的判别性强。从人类认知角度看,鉴赏一幅壁画,获得的语义信息是整幅壁画整体信息的总和,是多个局部特征按照特定的结构组合产生了图像的语义。从壁画的整体性角度出发,基于轮廓特征的整体结构建立两幅壁画图像间的相似性度量,进而实现壁画图像分类。

基于轮廓形状的图像特征表达已有广泛的研究。形状特征具有不受光照、色彩和纹理影响等优点。近些年比较流行的形状特征描述有基于采样点的描述和基于轮廓线段的描述两类。Belongie 等人^[1]提出的形状上下文方法通过统计形状上的采样点与形状上的其他点两两之间的角度和距离,在 log-polar 空间建立 2 维直方图表示当前采样点的信息,形状上下文特征融合了采样点的周围环境信息,具有尺度和形变不变性。Donoser 等人^[2]提出基于弦角的描述符,其思想是该采样点与其他采样点组成的弦,与这两点之间的采样点(与后采样点距离为 Δ)与后采样点组成的弦所称的角度来描述采样点特征,具有平移、旋转和尺度不变性。文献^[3]提出了形状树,表达了不同层次精度的形状信息,将局部和全局的形状特征很好地融合在了一起。基于轮廓线段的描述方法也有很多,Shotton 等人^[4]提出将轮廓线段作为基本特征,通过 Chamfer 匹配^[5]寻找形状最相似轮廓线段达到目标形状的检测。同样,Riemenschneider^[6]通过先匹配相似轮廓线段,然后

通过已经匹配的轮廓组合来形成目标位置的预测,从而实现目标物体检测。Ferrari 等人^[7]使用相邻的相对较直的轮廓线段组合作为局部特征,叫做 kAS,然后将该局部特征聚类生成词典作为特定类的轮廓描述符。Ravishankar 等人^[8]使用短的线段来近似目标物体的外部轮廓,与 kAS^[7]不同,他们使用弯曲的曲线线段代替相对较直的线段,这样能够更好地表达线段间的相异性。然而,目前基于轮廓特征的图像分类与目标检测都是在相似轮廓的匹配基础上,通过两两最相似轮廓的相似性度量来描述图像间的相似性,而并没有考虑同一幅图像内部轮廓自身的显著性以及轮廓间的结构关系属性。

从人类认知角度出发,结合壁画图像自身的特点,提取图像的轮廓特征作为基本特征,然后建立图像间的相似性度量机制,进而实现壁画图像分类。在相似性度量过程中,在考虑轮廓线段自身形状属性的基础上,融合轮廓的显著性信息以及轮廓间的整体结构信息,从而形成全局性的轮廓特征描述。在轮廓间整体结构关系的约束下,不断地在测试图像中查找与样本图像轮廓的最相似轮廓,每一轮的查找结果将作为下一轮查找的先验条件,有别于 Chamfer 匹配^[5]方法只寻找形状相似性最大的轮廓而不考虑整体结构关系。在轮廓整体结构的约束下,能够同时融入整体的轮廓线段,具有更高的相似性,最终计算得到的相似度也更加能够反应图像整体的相似度,更加符合人类的视觉感受。而且本文相似度的计算方法与图像尺寸、旋转均无关,具有较好的鲁棒性。

1 图像间相似度计算

1.1 轮廓检测与轮廓间相似度计算

首先采用 Canny^[9]或者 Pb^[10]检测图像的轮廓,然后对所有检测到的轮廓进行预处理。在预处理阶段主要做两方面的工作:1)截断,将有较大折点出现的轮廓线段,从折点处将该轮廓线段一分为二;2)连接,对于相邻的首尾相接的两条轮廓线段,如果首尾连接可以得到更长的平滑的轮廓线段,则将这两条轮廓线段首尾连接组成一条轮廓线段。图1展示了对轮廓线段的两种预处理方式。

根据以上规则,多次执行预处理后,最终得到该图像的较平缓的轮廓线段集合。

综合文献^[2]和文献^[3]的方法来描述轮廓线段,并以此计算两条轮廓线段的相似度。

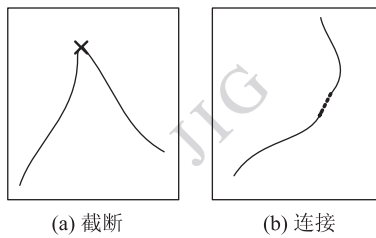


图1 轮廓线段的预处理

Fig. 1 Preprocessing of contour segments

假设有轮廓线段 l_i , $\{p_1, p_2, \dots, p_m\}$ 为 l_i 上弧长等距离的采样点, l_i 可表示为 $\widehat{p_1 p_m}$ 。采用 $\angle p_1 p_{m/2} p_m$ 来描述 l_i 的形状特征, 其中 $p_{m/2}$ 表示 l_i 的中点, 这仅仅是对 l_i 形状非常粗糙的描述, 为了更加准确地对 l_i 形状进行描述, 将 l_i 从中点 $p_{m/2}$ 处分为 $\widehat{p_1 p_{m/2}}$ 和 $\widehat{p_{m/2} p_m}$, 然后用同样的方式描述 $\widehat{p_1 p_{m/2}}$ 和 $\widehat{p_{m/2} p_m}$, 同样的, 可以继续将 $\widehat{p_1 p_{m/2}}$ 和 $\widehat{p_{m/2} p_m}$ 向下细分, 最终形成一个树形的轮廓形状描述结构^[1], 如图2所示。

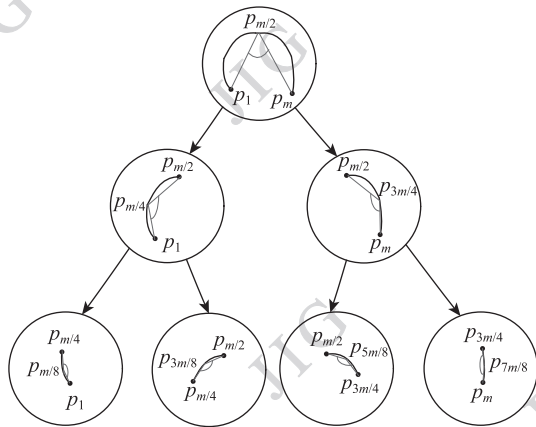


图2 树形轮廓描述结构

Fig. 2 Tree-shaped contour description

假设轮廓线段 l_i 的树形形状结构总共有 L 层, 分别为 $0, 1, 2, \dots, L-1$, 那么可以用一个特征向量 U 来表示该轮廓的形状, 即

$$U = [U^0, U^1, \dots, U^l, \dots, U^{L-1}]^T \quad (1)$$

$$U^l = \left(S \left(\widehat{p_1 p_{m/2^l}} \right), S \left(\widehat{p_{m/2^l} p_m} \right), \dots, S \left(\widehat{p_{m/2^{l-1}} p_{(l+1)m/2^l}} \right), \dots, S \left(\widehat{p_{(2^l-1)m/2^l} p_m} \right) \right) \quad (2)$$

$$S \left(\widehat{p_{m/2^l} p_{(l+1)m/2^l}} \right) = \angle p_{m/2^l} p_{(2^l+1)m/2^{l+1}} p_{(l+1)m/2^l} \quad (3)$$

同样, 对另一条轮廓线段 l_j 做同样的描述后得到特征向量 V , 则计算 l_i 和 l_j 之间的相似距离为

$$d_1(l_i, l_j) = \sum_{i=1}^N \lambda_i (U_i - V_i)^2 \quad (4)$$

式中, λ_i 为树形结构中每层形状特征的权重, N 为最终轮廓线段特征向量的元素个数, U_i, V_i 分别为两条最终轮廓线段特征向量的元素。

通过该方法得到的两条轮廓线段的相似性度量机制与轮廓的尺度、旋转、位置均无关, 能够很好地描述两条轮廓线段的形状相似程度。

1.2 融合显著度的轮廓相似度

浏览一幅画时, 往往画中某一部分能够更加引起人们的注意, 称这部分为显著特征, 显然, 显著特征更能够描述一幅画的语义。根据文献[11]中对显著区域的描述, 给轮廓线段的显著度这样描述: 相对越长的平滑的连续线段具有更高的显著度。将这个显著度作为条件参与到轮廓相似性的计算中。

站在图像的角度来比较两条轮廓线段的相似性, 一幅图像中的一条轮廓线段 l_i 与另一幅图像中的轮廓线段 l_j 相似的条件除了这两条轮廓形状相似之外, 还必须在各自所属图像中具有相似的重要性, 即具有相似的显著度。

假设 l_i 和 l_j 分别为图像 I_1 和 I_2 中两条轮廓线段, 在各自图像中的显著度分别为 p_1 和 p_2 , 则得到融合显著度的图像间轮廓线段相似距离

$$d(l_i, l_j) = \alpha d_1 + \beta d_2 \quad (5)$$

式中, α 和 β 分别为各部分的权重, 通过后面学习的过程得到, d_1 为形状相似距离, 由式(4)计算得到, d_2 为显著度的相似距离, 即

$$d_2 = (p_1 - p_2)^2 \quad (6)$$

显著度 p 为

$$p = \frac{|l|}{SIZE(I)} \quad (7)$$

式中, $|l|$ 为线段长度, $SIZE(I)$ 为一个描述图像大小的函数, 可以实现图像间轮廓相似性与图像的尺寸无关。

1.3 融合轮廓间整体结构的轮廓相似度

图像的语义更多是全局语义, 比较两幅图像时, 更倾向于从整体的表现去比较。鉴于此, 在比较两幅图像的轮廓特征时, 对轮廓间的整体空间结构关系进行融合, 两条轮廓线段的相似度不仅仅与自身的形状和显著度有关, 同时与当前轮廓线段和图像中其他轮廓线段的结构关系有关。如图3所示, l'_1, l'_2 是图像 I_2 中与 l_1, l_2 形状最相似的轮廓线段, 在此条件下, 即使在形状属性上 l'_2 与 l_3 具有最高的相似度, 但是在融合轮廓间的整体结构关系后, l'_3 会取代 l'_2 成为与 l_3 最相似的轮廓线段, 因为在已经存在的 l'_1, l'_2 是 l_1, l_2 的最相似轮廓条件下, l_1, l_2 与 l_3 的结构关系同 l'_1, l'_2 与 l'_3 的结

构关系也需要具有一定的相似性。

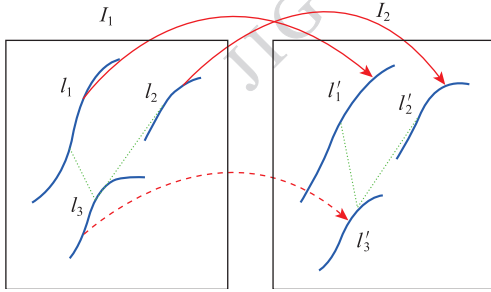


图 3 轮廓整体结构约束下的轮廓相似度

Fig. 3 Contour similarity with the overall structure constraint

对同一幅图像中的两条轮廓线段 l_i 和 l_j , 采用 3 个维度: l_i 和 l_j 的距离、长度比、方向角的差值来描述 l_i 和 l_j 的空间结构, 即

$$S_{ij} = \left(\frac{D_{ij}}{SIZE(I)}, \frac{|l_j|}{|l_i|}, |dir_j - dir_i| \right) \quad (8)$$

式中, D_{ij} 表示两轮廓线段间的距离, 其值为 l_i 上的所有点到 l_j 的最短距离的均值, dir_i 和 dir_j 分别表示 l_i 和 l_j 的方向角。

假设有两幅图像 I_1 和 I_2 , I_1 包含 m 条轮廓线段, 记为 $\{l_1^1, l_2^1, \dots, l_m^1\}$, I_2 包含 n 条轮廓线段, 记为 $\{l_1^2, l_2^2, \dots, l_n^2\}$ 。根据式(8), 可得图像 I_1 和 I_2 的轮廓间结构关系 3 维矩阵 $S_1 \in \mathbf{R}^{m \times m \times 3}$, $S_2 \in \mathbf{R}^{n \times n \times 3}$, 然后将图像间两条轮廓线段的相似性计算转化为在融合轮廓间整体结构关系下的相似性计算。

采用一个映射函数 f 表示第 2 幅图像中与第 1 幅图像中的某条轮廓线段最相似的轮廓线段之间的映射关系, 如 $f(l_i^1) = l_j^2$ 表示图像 I_2 中与 I_1 中轮廓线段 l_i^1 最相似的轮廓段是 l_j^2 。

定义两幅不同图像 I_1 和 I_2 中两条轮廓线段 l_i^1 和 l_j^2 的相似性距离为

$$(l_i^1, l_j^2) = \alpha d_1 + \beta d_2 + \gamma d_3 \quad (9)$$

式中, d_1 是自身形状差异, 由式(4)计算得到, d_2 是显著度差异, 由式(6)计算得到, d_3 是轮廓整体性结构差异, 将其定义为与其他轮廓线段空间结构关系间的 L_2 距离, 即

$$d_3(l_i^1, l_j^2) = \sum_{t \neq i} \|S_1(t, i) - S_2(f(t), j)\|_2 \quad (10)$$

进一步可以得到两幅图像 I_1 和 I_2 之间的相似性距离

$$D(I_1, I_2) = \sum_i d(l_i^1, f(l_i^1)) \quad (11)$$

最关键的问题是如何确定映射 f , 设计了一个在自身形状属性和显著度为已知先验条件下的映射 f 的计算方法。首先初始化 f 为只考虑自身形状属性和显著度条件而不考虑整体结构性条件下的最相似轮廓线段映射, 然后采用迭代过程逐步融合整体结构关系, 具体的最相似轮廓线段的计算过程如下:

1) 输入两幅图像轮廓线段集合: $L^1 = \{l_1^1, l_2^1, \dots, l_m^1\}$, $L^2 = \{l_1^2, l_2^2, \dots, l_n^2\}$;

2) 初始化 f , 对 L^1 中任意的线段 l_i^1 , 令 $f_i = \arg \min_j (\lambda d_1(l_i^1, l_j^2) + (1 - \lambda) d_2(l_i^1, l_j^2))$;

3) 初始化相似性距离 $d = \sum_i (\lambda_1 d_1(l_i^1, l_j^2) + \lambda_2 d_2(l_i^1, l_j^2) + (1 - \lambda_1 - \lambda_2) d_3(l_i^1, l_j^2))$;

4) 根据前面确定的 f , 对所有的 i , 更新 f_i 为 $f_i = \arg \min_j (\lambda_1 d_1(l_i^1, l_j^2) + \lambda_2 d_2(l_i^1, l_j^2) + (1 - \lambda_1 - \lambda_2) \times d_3(l_i^1, l_j^2))$;

5) 重新计算相似距离 $d' = \sum_i (\lambda_1 d_1(l_i^1, l_j^2) + \lambda_2 d_2(l_i^1, l_j^2) + (1 - \lambda_1 - \lambda_2) d_3(l_i^1, l_j^2))$;

6) 如果 $|d - d'| > \delta$, 则 $d = d'$, 然后转到步骤 4), 否则转到步骤 7);

7) 输出 f 为两幅图像最相似轮廓线段的映射。

得到映射函数 f 之后, 根据式(11)可求得两幅图像的相似性距离, 注意到式(11)是不对称的, 即

$$D(I_1, I_2) \neq D(I_2, I_1)$$

因为算法是从 I_1 中的轮廓线段出发, 在 I_2 中寻找最相似轮廓线段, 但是这样可能会存在 I_2 中有某些轮廓线段在 I_1 中却无法找到相似的轮廓线段的情况。

根据文献[12]中对相似性的定义, 两幅图像 I_1 和 I_2 有较高的相似性的条件是: 图像 I_1 的特征都能在 I_2 中存在, 同时图像 I_2 的特征也都在 I_1 中存在。因此采取双向计算的策略, 在计算 $D(I_1, I_2)$ 的同时计算 $D(I_2, I_1)$, 然后用两者的均值作为图像 I_1 和 I_2 的相似性距离。

图像间相似度的计算方法融合了轮廓形状的相似性, 轮廓显著度相似性以及轮廓整体结构的相似性 3 部分。轮廓的形状属性和轮廓显著度均采用了统一标准化的计算方式, 对轮廓的大小、旋转、位置均具有不变性; 对于轮廓整体结构信息, 采用两两轮廓间的相对距离、长度比, 以及方向角的差异来描述, 而并没有考虑轮廓间的严格空间位置关系, 因此同样对于图像的尺寸、旋转具有不变性。

1.4 噪声轮廓线段的处理

实际计算中, 图像不可避免地会存在噪声影响,

壁画图像更是如此。因此在实际寻找最相似轮廓线段过程中,会有大量的噪声轮廓参与计算从而造成算法鲁棒性下降。

为了降低噪声的影响,允许部分轮廓线段不参与计算。在计算映射函数 f 的过程中,假设某条轮廓线段 l_i^1 的最相似轮廓线段为 l_j^2 ,即 $l_j^2 = f(l_i^1)$ 。假如 $d(l_i^1, l_j^2) > \tau$ (τ 为阈值),则规定 $f(l_i^1) = \emptyset$ 。这种情况下, l_i^1 将被看做噪声线段而不会参与到相似性的计算中。但是为了避免只有很少量轮廓线段参与计算而使得两幅图像有较高相似度从而产生较大误差,对噪声的出现采取惩罚机制,使用非噪声轮廓线段像素所占比例作为惩罚因子,得到两幅图像的相似距离

$$D(I_1, I_2) = P^{-1} \left(\sum_{i, f(l_i^1) \neq \emptyset} d(l_i^1, f(l_i^1)) \right) \quad (12)$$

式中, p 为惩罚因子,即

$$p = \frac{\sum_{i, f(l_i^1) \neq \emptyset} |l_i^1|}{\sum_i |l_i^1|} \quad (13)$$

这样,噪声轮廓相对越多,则两图像间具有相对更大的相似距离。

1.5 度量归一化与权重学习

轮廓线段长度各不相同,每幅图像中轮廓线段的数量也不同,因而轮廓线段间形状相似性的比较以及图像间相似性比较的度量标准不同。在轮廓形状属性的比较中,由于长度因素导致形状层次的不同,采用加权平均的方式达到度量的统一;同样,对两幅图像间相似性比较,采用平均每对轮廓线段的相似度来表示两幅图像间的相似度。这样,最终图像间相似距离为

$$D(I_1, I_2) = \rho^{-1} \left(\frac{1}{N} \sum_{i, f(l_i^1) \neq \emptyset} d(l_i^1, f(l_i^1)) \right) \quad (14)$$

式中, N 为第1幅图像中所有参与计算的非噪声轮廓线段的数量。

对于两幅图像间相似度计算,通过最优化公式(9)中的权值 α, β, λ 进行学习,使得对于 D 的计算能够达到最小化同类别图像间距离的效果,同时最大化不同类别图像间的距离。

将式(14)中 $d(l_i^1, f(l_i^1))$ 重新写为 $d(l_i^1, f(l_i^1)) = \sum_{i=1}^3 \omega_i d_i = \omega^T D$,其中 $\omega = [\alpha, \beta, \lambda]^T$, $D = \{d_i | i=1, 2, 3\}$ 。

假设有一幅图像 I , T 是与图像 I 同类别图像, F 是与图像 I 不同类别图像,则有

$$D(I, F) > D(I, T) \quad (15)$$

$$\frac{1}{N_F} \sum_i d(l_i, f_{I \rightarrow F}(l_i)) > \frac{1}{N_T} \sum_i d(l_i, f_{I \rightarrow T}(l_i)) \quad (16)$$

$$\omega^T \left(\frac{1}{N_F} \sum_i D_{I \rightarrow F}(l_i) \right) > \omega^T \left(\frac{1}{N_T} \sum_i D_{I \rightarrow T}(l_i) \right) \quad (17)$$

$$\omega^T \left(\frac{1}{N_F} \sum_i D_{I \rightarrow F}(l_i) - \frac{1}{N_T} \sum_i D_{I \rightarrow T}(l_i) \right) > 0 \quad (18)$$

在训练阶段,构造 T 组数据 $x_1, x_2, \dots, x_T$, $x = \frac{1}{N_F} \sum_i D_{I \rightarrow F}(l_i) - \frac{1}{N_T} \sum_i D_{I \rightarrow T}(l_i)$,可以采用支持向量机(SVM)^[13]的最大化边界准则进行优化并获得最优的 ω 值,即

$$\min_{\omega, \xi} \left(\frac{1}{2} \omega^T \omega + C \sum_{i=1}^T \xi_i \right) \quad (19)$$

$$\text{s. t. } \omega^T x_i \geq 1 - \xi_i, \xi_i \geq 0, \forall i = 0, 1, \dots, T$$

由此得到能够最优化结果的权值。

2 实验及分析

2.1 实验图像库

所选的实验素材出自《敦煌石窟艺术》画册经过扫描得到的艺术图像,共涵盖了29个洞窟的壁画图像,覆盖北周、隋、初唐、盛唐、中唐、晚唐、五代和元8个朝代,艺术形象丰富,风格各异。

选取典型的飞天、佛像、供养人3种艺术形象进行分类。由于壁画历史年代久远,不同时期的同种艺术形象具有截然不同的构造,因此飞天形象中又分为北周、隋、初唐、西魏4个时期,实验详细的壁画图像资料情况见表1。图4展示了不同时期飞天的代表性图像。其他艺术形象由于数量较少,暂不进行分析。

表1 本文实验所选壁画图像

Table 1 Selected mural images in this experiment

类别	飞天				佛	供养人
	北周	初唐	隋	西魏		
数量	156	50	74	34	81	105

2.2 图像间相似度计算

根据1.5节的描述,对每一类图像,选取20幅

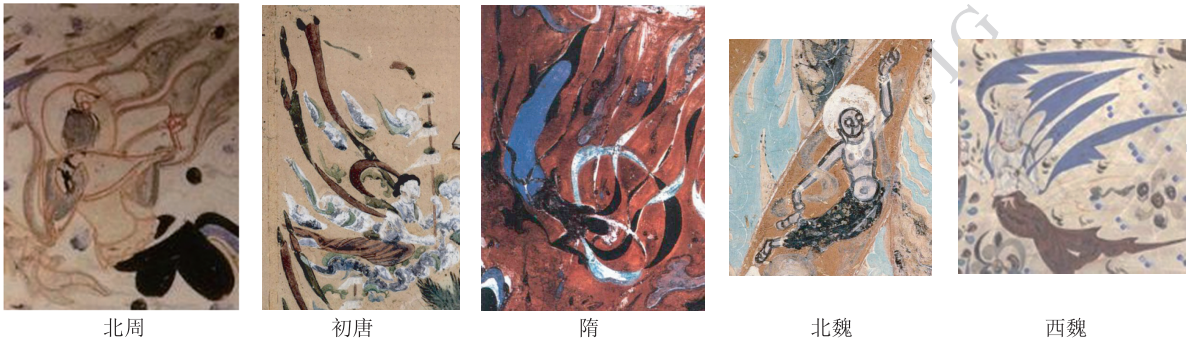


图 4 不同时期的飞天图像

Fig. 4 Flying Apsaras images of different periods

图像作为训练样本,对训练样本中的每一张图像,随机选取 10 幅同类别图像和 10 幅其他类别图像构造 10 组 x 值(详见 1.5 节),这样对所有的训练样本总共构造 200 组数据用于权重 ω 的训练,最终为每一类图像生成一个 ω 值。

然后用图像库中的剩余图像作为测试样本,在两两图像库之间求两两图像间的相似距离。采用双向相似距离计算然后取均值作为两幅图像的相似距离,最后将所有两两图像对的相似距离取平均值作为两个类别图像库的相似距离,计算结果如图 5 所示。

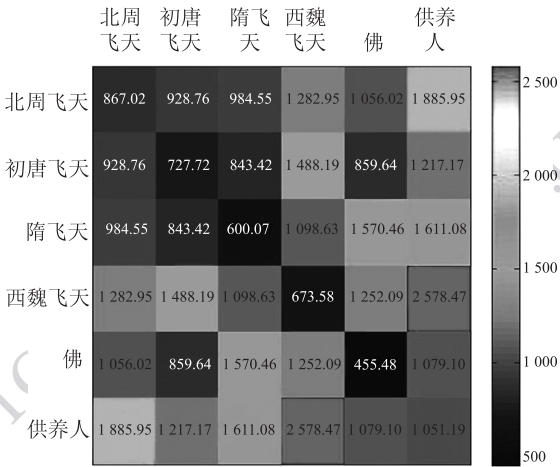


图 5 不同类别壁画图像库的相似距离

Fig. 5 Similarity distance of different mural image library

从图 5 中可以看出同类别的图像间的相似距离要小于不同类别图像间的相似距离,同时可以看出北周飞天、初唐飞天和隋代飞天 3 类之间有相对较小的相似距离,即具有较高的相似度。

2.3 壁画图像分类

通过定义的图像间相似性度量方法,可以实现

对壁画图像的分类,图像分类除了图像特征之外,分类器也是很重要的一个方面。近些年常用的分类方法有 SVM^[13]、K 最近邻(KNN^[14])、Boosting^[15]等。本文主要研究基于轮廓特征的图像间相似性度量方法,属于图像特征的范畴,因此不过多地对分类器进行讨论,只是选择目前比较流行的分类器实现图像分类来对特征有效性进行验证。分别采用 KNN 方法和 SVM 方法进行壁画图像分类,最后本文方法的实验结果分别与用 SHOTTON^[4]的不考虑整体结构关系的 chamfer 匹配以及 FERRARI^[7]的 PAS($k=2$ 时的 kAS)方法的实验结果做比较。

用 KNN 方法做分类时,分别选用 $K=10$ 和 $K=20$,然后取 K 最近邻中最多的图像类别作为测试图像的类别。实验中对每类图像随机选择 20 张作为样本图像,然后对剩余图像在样本库中求最近邻实现壁画图像的分类。在计算两两图像间的相似距离时,由于测试图像是未知的,因此选用统一的 ω 值,在验证每类壁画图像的正确分类率时,反复进行 50 次然后取均值得到平均正确分类率。表 2 和表 3 分别展示了 $K=10$ 和 $K=20$ 时用 KNN 方法在每类壁画图像库上的正确分类率。

表 2 KNN($K=10$)方法的分类结果
Table 2 Classification results of KNN ($K=10$)

图像库	本文方法	SHOTTON ^[4]	FERRARI ^[7]
北周飞天	62.20	46.86	55.75
初唐飞天	49.46	46.20	41.30
隋飞天	59.20	39.73	49.82
西魏飞天	85.28	75.50	78.00
佛	81.16	76.90	79.29
供养人	62.35	76.35	71.65

表 3 KNN (K = 20) 方法的分类结果

Table 3 Classification results of KNN (K = 20)

图像库	本文方法	SHOTTON ^[4]	FERRARI ^[7]	/%
北周飞天	65.00	58.76	59.52	
初唐飞天	59.29	51.30	49.40	
隋飞天	61.70	46.27	53.77	
西魏飞天	89.15	83.17	84.00	
佛	85.26	80.94	83.88	
供养人	69.74	83.89	73.55	

用 SVM 方法进行分类时,使用的扩展高斯核函数^[13,16]为

$$K(I_1, I_2) = \exp\left(-\frac{1}{A}D(I_1, I_2)\right) \quad (20)$$

同样的,在分类时每类图像随机选择 20 幅图像作为正样本,然后随机选择 60 幅其余类别图像作为负样本,采用一对多方式训练多类 SVM,然后用库中剩余图像作为测试图像进行分类,反复进行 50 次取得平均正确分类率。在使用文献[4]和文献[7]中描述的特征进行分类时,采用其文献中定义的图像间相似性计算方法作为 SVM 核函数的主体。最终使用 SVM 的分类结果见表 4。

表 4 SVM 方法的分类结果

Table 4 Classification results of SVM

图像库	本文方法	SHOTTON ^[4]	FERRARI ^[7]	/%
北周飞天	69.26	59.57	67.35	
初唐飞天	62.46	56.35	57.20	
隋飞天	68.20	48.14	56.95	
西魏飞天	95.00	85.10	95.00	
佛	89.16	85.88	92.24	
供养人	78.35	87.84	81.20	

从实验结果可以看出,对于 4 个朝代的飞天和佛像 5 类数据,本文方法相比于没有考虑整体结构性的文献[4]和文献[7]方法都有了精度上的提高,对于供养人图像,由于供养人轮廓线条较少,本文方法并不具备优势。因此可以得出本文的方法在具有较多轮廓线条的壁画图像分类上更有优势。

3 结 论

提出了一种新的基于轮廓整体结构的图像间相似性度量方法,在计算两幅图像间轮廓相似性时,不仅考虑轮廓线段自身的形状属性,还考虑轮廓的显著度以及轮廓间的整体结构关系。此方法更能够体现图像轮廓的整体特征,也更加符合人类的视觉认知。本文方法最关键之处在于找到全局性的最相似轮廓线段的映射关系,采用一种迭代的方法,首先初始化为不考虑结构关系的最相似轮廓,然后慢慢融入轮廓间的结构关系。分别采用 KNN 和 SVM 分类方法进行了验证,实验结果表明本文方法对于壁画图像的分类相比于没有融合整体结构的方法有了改进。

本文方法以高层次的轮廓线段作为基本特征,从轮廓整体性方面建立图像间相似性度量,更加符合图像语义,但是另一方面,该方法比较依赖于底层特征检测器的输出,即图像轮廓的检测结果会直接影响图像间相似度的计算结果,因此需要对底层轮廓检测器的输出做预处理,使得轮廓更加符合需要。此外,该方法对噪声比较敏感,壁画图像中噪声很多,会对计算结果产生很大的影响,因此去噪工作也是非常重要的。文中给出了预处理和处理噪声的方法,但仍然有很大的改进余地,下一步的研究将根据目前方法的局限性,设计不依赖于底层检测器并且抗噪的更加鲁棒的特征表达,同时会建立更加高层的图像语义描述方法,实现对不同年代的同类别壁画形象建立统一的描述方式。

参考文献 (References)

[1] Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(4): 509-522.

[2] Donoser M, Riemenschneider H, Bischof H. Efficient partial shape matching of outer contours [C]//Proceedings of the 9th Asian Conference on Computer Vision. Xi'an, China: Springer, 2009, 5994:281-292.

[3] Felzenszwalb PF, Schwartz JD. Hierarchical matching of deformable shapes [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis, MN, USA: IEEE Computer Society, 2007, 1-8.

[4] Shotton J, Blake A, Cipolla R. Contour-based learning for object detection [C]// Proceedings of 10th IEEE International Confer-

- ence on Computer Vision. Beijing, China: IEEE Computer Society, 2005, 1:503-510.
- [5] Borgefors G. Hierarchical chamfer matching: a parametric edge matching algorithm [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1988, 10(6): 849-865.
- [6] Riemenschneider H, Donoser M, Bischof H. Using partial edge contour matches for efficient object category localization [C]// Proceedings of the 11th European Conference on Computer Vision. Crete, Greece: Springer, 2010, 6315: 29-42.
- [7] Ferrari V, Fevrier L, Jurie F, et al. Groups of adjacent contour segments for object detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(1): 36-51.
- [8] Ravishanker S, Jain A, Mittal A. Multi-stage contour based detection of deformable objects [C]// Proceedings of the 10th European Conference on Computer Vision. Marseille, France: Springer, 2008, 5302:483-496.
- [9] Canny J. A computational approach to edge detection [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1986, 8:679-714.
- [10] Martin D, Fowlkes C, Malik J. Learning to detect natural image boundaries using local brightness, color and texture cues [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(5): 530-549.
- [11] Song W, Toshiro K, Jeffrey M S, et al. Salient closed boundary extraction with ratio contour [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(4):546-561.
- [12] Simakov D, Caspi Y, Shechtman E, et al. Summarizing visual data using bidirectional similarity [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Anchorage, Alaska, USA: IEEE Computer Society, 2008, 1-8.
- [13] Christopher M B. Pattern Recognition and Machine Learning [M]. Berlin, Germany: Springer, 2006:326-345.
- [14] Zhang Hao, Berg A C, Maire M, et al. SVM-KNN: discriminative nearest neighbor classification for visual category recognition [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. New York, USA: IEEE Computer Society, 2006, 2:2126-2136.
- [15] Schapire R. A brief introduction to boosting [C]// Proceedings of the 16th International Joint Conference on Artificial Intelligence. Stockholm, Sweden: Morgan Kaufmann, 1999, 1401-1406.
- [16] Zhang J, Marszalek M, Lazebnik S. Local features and kernels for classification of texture and object categories: a comprehensive study [J]. International Journal of Computer Vision, 2007, 73(2):213-238.