

文章编号: 2095-2163(2021)12-0047-07

中图分类号: G255.71

文献标志码: A

数字人文视域下云锦图像本体识别模型构建及应用研究

卢璐¹, 袁高腾², 梅嘉¹, 肖文美¹, 李木子³

(1 南京林业大学 图书馆, 南京 210037; 2 河海大学 计算机与信息学院, 南京 211100;

3 南京林业大学 继续教育学院, 南京 210037)

摘要: 云锦是中国丝织工艺的最高成就,是中华民族和全世界珍贵的历史文化遗产,蕴含着丰富的历史人文和社会经济价值,享有“锦中之冠”的美誉。为了分析不同材料云锦图像上的差异,并利用计算机技术分辨不同材料的云锦,完成云锦图像的自动识别,本研究项目收集了多个不同样本的云锦,获取云锦图像,并对云锦图像进行去噪,依据云锦类别设置标签;使用不同云锦图像的兴趣区域,并采用5个尺度、8个方向的Gabor滤波器组对兴趣区域进行滤波,提取Gabor纹理特征;对特征进行采样,并对采样后的特征进行分析;最后,使用贝叶斯、KNN等分类器分类,学习分类器对采样特征进行分类、分析,完成了12种云锦图像的分类。实验结果表明,使用KNN分类器对云锦特征进行分类,12个类别的云锦图像分类精度为56.31%。表明不同类别的云锦图像能够在纹理上进行区分,计算机自动识别工艺繁复非遗图像是可行的。

关键词: 图像识别; 纹理分析; 数字人文; 南京云锦

Research on construction and application of Yunjin image ontology recognition model in digital human visual field

LU Lu¹, YUAN Gaoteng², MEI Jia¹, XIAO Wenmei¹, LI Muzi³

(1 Library of Nanjing Forestry University, Nanjing 210037, China; 2 College of Computer and Information, Hehai University, Nanjing 211100, China; 3 College of continuing education, Nanjing Forestry University, Nanjing 210037, China)

[Abstract] Yunjin is the highest achievement of my country's silk weaving craftsmanship. It is a precious historical and cultural heritage of the China and the world. It contains rich historical, cultural and socio-economic values, and enjoys the reputation of the crown of the silk. In order to analyze the differences in the Yunjin images of different materials, use computer technology to distinguish the Yunjin of different materials, and complete the automatic recognition of the Yunjin images, multiple different samples of Yunjin and its images are collected. Later denoise on the Yunjin image is conducted, and the labels is set according to the category of the Yunjin. The different interest area of Yunjin images is obtained, and the Gabor filter group with 5 scales and 8 directions is used to filter the interest area to extract the Gabor texture feature. Then, the features are sampled and the sampled features are analyzed. Finally, Bayesian, KNN and other machine learning classifiers is used to classify and analyze the sampling features to complete the classification of 12 kinds of Yunjin images. The results show that the KNN classifier can be used to classify the Yunjin features, and the classification accuracy of the 12 categories of Yunjin images is 56.31%. It shows that different types of Yunjin images can be distinguished by texture, and it is feasible for the computer to automatically recognize intangible cultural heritage images with complicated processes.

[Key words] image recognition; texture analysis; digital humanities; Nanjing Yunjin

0 引言

近年来,得益于“人文记忆”研究的深入推进,面向图像的人文计算领域得到了越发广泛的关注。许慎在《说文解字》中提到“画成其物,随体诘屈”,意为文字和图像共同构成了文化表达、交流和传承的载体^[1]。南宋郑樵认为,图文互文极具史料价值,曾在《通志》中写到“置图于右,索象于图。”^[2]。

司马相如《子虚赋》:“众物居之,不可胜图”^[3]中的“图”即为事物的形象,文化内蕴的表征,而“象”为“圣人之意”,需要深度分析。欧文·潘诺夫斯基对图像和文化关系作了相当完备的阐述^[4]。贡布里希认为,文化思想传播离不开图像。文字描述可以做思之抵达与言之切近,极具现实主义和浪漫主义色彩,而图像突破了不同文化语言的藩篱,避免文字表述间离性,拓宽了思维层面对文化分析的基本

基金项目: 2021年度江苏省社会科学基金项目(21TQC006); 2021年度江苏省高校哲学社会科学基金项目(2021SJA0135)。

作者简介: 卢璐(1993-),女,硕士,馆员,主要研究方向:数字人文、信息情报;袁高腾(1993-),男,博士研究生,主要研究方向:图像处理与模式识别、计算机视觉;梅嘉(1983-),女,硕士,馆员,主要研究方向:数据挖掘;肖文美(1982-),女,硕士,馆员,主要研究方向:文化推广;李木子(1988-),男,硕士,馆员,主要研究方向:视觉交互。

收稿日期: 2021-09-13

视域。

海德格尔阐述的“世界图像”^[5]、米歇尔提出的“图像表征”^[6]、“图像转向”^[7]都在文化思考的逻辑关系上,代入对图像本质的阐释。詹姆斯·吉布森构建了现代社会对视觉文化交互方式^[8]。西方学者认为,图像是特定文化存在和发扬的根底,在流传中可以直观呈现出本体的表象,蕴含的文化记忆从隐匿到澄明,揭示图像传递的情感和背后的隐喻,^[9]通过图像表征的敞开域可以进行分析、分类和存储。

虽然图像信息丰富,但是大部分用户仍处于“图像孤岛”的困境,难以在第一时间进行图像资源的最优选择。多米尼克·洛佩斯论述的“图像再现”^[10]的理论中提到,如果进行图像的优化需要着重关注图像“面”的差异导致的辨识模式上的差异,遂提出“识别的动态性”。

本文所涉“云锦图像”主要突出的是其作为“视觉史料”的身份,包括纹样、花色、材质、纹理、组织结构等类别特征^[11]。目前,国内外很多作为史料展现的图像数字人文项目初见成效。如《蒙娜丽莎》数字化存储^[12]、欧盟博物馆藏品分类^[13]、敦煌艺术图像规范标准^[14]、故宫博物院文物服务库、中世纪艺术索引等。但大部分图像项目局限于“冰冻式”的静态采集、存储与展现,只做了主题、时间和地域等大体分类,只有少数针对图像和语义结合进行了实体识别,但没有实现大规模应用,鲜有对图像本体纹样和组织结构深度挖掘和内源性关联研究。

云锦是中国丝织工艺的最高成就,迄今已有1600多年的手工织造历史,享有“锦中之冠”的美誉。云锦的纹样题材具有托物言志深厚情感,作品与材质上的美术元素表现了创作者的气韵风骨^[15],使其纹样图必有意,意必吉祥,需要借助图像识别分类技术来探究源文化形态的本真状态,展现历史绵延性和强韧性,为云锦图像后续应用打下基础。

1 相关工作

云锦图像数据识别与存储涉及艺术、人文、情报、计算机等多个学科领域。笔者主要探究图像内在纹理特征提取和动态性演变,因此主要涉及两方面的知识背景:图像识别分类研究现状与云锦图像纹样特征提取研究现状。

1.1 图像识别分类研究现状

1.1.1 图像识别方法

目前,常见图像识别方法有:

(1)设置模糊集识别方法。常用来识别医学图像,分类匹配率精准度较高,但过于依赖图像的特征;

(2)人工神经网络方法。自动拟合非线性模型,避开人为操作主观性影响,客观且简化,但对图像样本、硬件性能和相关参数解释都要求很高,很难广泛应用;

(3)统计识别方法和识别结构方法。通过特征向量的数学提取或者子模型逐层识别还原,都可以对图像进行简单分类,但无法识别图像整体结构。

1.1.2 织物图像识别研究进展

织物的图像识别和分类是目前图像学领域的难关。由于织物具有纹理繁复、织线材质多样和组织结构复杂等特性,现面临数据预处理时颜色失真、模糊、噪声滤除困难、局部特征不易提取等困境。相关学者通过不同技术构建实例模型,李敏等^[16]利用快速傅里叶变换和形态滤波,处理小提花织物纹样的周期特点,淡化一般性纹路,着重加工纹样特殊点,提高疵点检测精确度。张开生等^[17]通过机器视觉算法和卷积循环神经网络(CRNN)技术,对磁性纤维加密图像进行特征提取。由于女红作品背景复杂繁多,存在图像特征不易提取且难以分割的问题,包萍萍等^[18]采用阈值和字符图像分割的技术,将女书图像的背景进行去噪处理,取得了不错的滤除效果。

1.2 云锦图像纹样特征提取研究现状

图像纹理是人们视觉系统对自然界物体表面现象的一种感知,可描述和区分不同物体的重要特征。针对不同类别的云锦图像,可以采用频谱法提取图像的纹理特征,其本质是将基本纹理基元以及在图像区域中的不同形式出现的“副本”,用在不同尺度和方向上的子波能量分布表示出来。本文主要采用Gabor纹理特征分析图像,选用了8个方向、5个尺度的Gabor函数,对图像进行滤波提取图像特征,并对特征进行分析、分类。

云锦历史悠久、品种繁多,目前大体上分为“库缎、库锦、库金、妆花”4类。同一件云锦作品上,纹样题材的选取、色彩的搭配、构图板式的风格都有定位。根据缠枝、寓意和均衡进行不同材质丝线的色相纯净度的调和,糅合变化更细腻绚丽,最终达到逐花异色,轩昂华丽的效果。云锦整个制作过程,贯穿了文化内涵的开发^[19],丰富了云锦图像美感。云锦图像和其它织物一样,都是由不同组织点构造,根据浮沉交织规律的不同,形成各异的组织结构。如,平纹是由经线和纬线相互一浮一沉的交织而成;斜纹是无数纬组织点围绕一个经组织点,反之亦然;缎纹

各点是有规律散布且经纬线仅一个组织点。随着经纬纱重叠变化,不同织线沉浮规律交错,织成的图像色影丰富。织线材质不同,有的是孔雀羽毛和鸵鸟羽毛加捻而成,与金银线交织而成的织物层次多变。

云锦图像遵循“立象以尽意”的艺术特征^[20],从中提取的不同纹样处于相互缠绕、勾连和均衡之中,特征提取起来十分繁难。目前,四大品种结构分类见表 1。

表 1 云锦品种、纹样和组织结构分类表

Tab. 1 Classification of Yunjin based on varieties, patterns and organizational structure

品种维度	库缎	库金	库锦	妆花
用途	成衣缎料	镶滚衣边、帽边、裙边和垫边	金彩线装饰、大花纺织锦面料	佛经经面、帷幔、服装、帐子的装饰,桌围、椅披、龙袍、蟒袍和扇盖的织成
技艺	按位换花本、单面单织、小梭局部挖花盘织	通梭织造、地部组织勾勒轮廓	金线通梭、丢阳缝	挖花盘织、彩绒纬管、通经断纬、彩绒纬纤管
分类	妆金库缎、妆彩库缎、本色花库缎、金银点库缎和地花两色库缎	曲水纹锦、冰梅锦、串菊锦	二色金库锦、彩色库锦、天华锦、芙蓉妆	妆花绸、妆花绒、妆花罗、妆花布、妆花纱、妆花绢和金宝地
组织结构	经纬同色、经纬异色、金银线装饰局部	金线或银线交织、花满地少	金线和彩纬交织、满地规矩、交错重叠	重纬组织、地纬加织彩纬
代表作品	云龙纹、缠枝牡丹金八宝、梅兰竹菊金银点花库缎、彩妆蝴蝶	方永泰本机真库金、织金陀罗经	红地莲二色金库锦、天华锦、芙蓉妆	四则八吉牡丹莲妆花缎、云锦大云龙妆花缎、寸蟒缎、鲤鱼戏水金宝地

2 数据来源与研究方法

2.1 数据来源

本文从南京市云锦博物馆收集了 12 组尺寸为 8 000 * 6 000 的云锦图像,根据博物馆官方花本集,对云锦类别进行标记,包括手工段和电机段两大部分,共计 12 个类别。本文主要对 12 个不同类别的云锦图案进行分析研究。其中包括:八宝底纹料、四合如意填花料、新百花料、普天同庆料、喜字并蒂莲料牡丹锦料、凝香富贵料、气象万千料和电机的龙纹料等。其中,八宝底纹料又包含多个颜色。如粉橘、深土和紫等。

本文对上述不同的云锦料进行收集,并完成云锦图案的分类识别。收集的云锦图像数据如图 1 所示。



(a) 牡丹锦料

(b) 喜字并蒂莲料

图 1 不同类别云锦图像

Fig. 1 Different types of Yunjin images

本文主要对云锦图像进行纹理分析,提取不同类别云锦图像纹理特征,通过对不同图像特征的分析,构建云锦类别识别模型。构建步骤如下:

Step 1 收集不同图像和不同品种的云锦(来自云锦研究所官方花本集),利用扫描器收集扫描云锦图像,并对云锦图像划分子区域,构建云锦图像数据集。

Step 2 依据云锦类别,对图像添加标签构建数据样本;使用小波分析对云锦图像进行去噪,完成图像的预处理;使用 Gabor 滤波器对云锦图像的子区域进行滤波,得到云锦图像的纹理特征。

Step 3 使用机器学习中不同的分类器分析图像特征,对纹理特征进行分类,构建分类模型。

2.2 研究方法

本文提出的模型由数据预处理、纹理分析、特征选择、机器学习分类等部分组成。首先,对图像进行去噪,获取兴趣区域,并设置兴趣区域标签;其次对兴趣区域进行纹理分析,提取不同的纹理特征,依次比较不同的图像纹理,选择合适的纹理特征,并进行特征选择以降低特征的维度,压缩特征大小;然后对比不同的分类算法,包括 SVM、KNN 等分类器对云锦图像特征进行自动分类。模型结构如图 2 所示。



图 2 模型结构整体框架

Fig. 2 The overall framework of the model structure

如图 2 所示,数据预处理包括图像数据的读取、图像去噪等;纹理分析包括 Gabor 滤波、特征采样组成;分类模型包括参数优化、特征选择、特征筛选等组成,构建好分类特征后使用多种分类器对特征进行比较分类,对比不同方法的分类精度和效率。

2.3 数据预处理

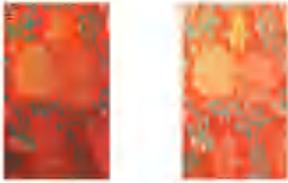
数据预处理的主要目的是完成图像的去噪,以及获取兴趣区域,方便构建分类模型。由于云锦图

像是扫描仪扫描后的图像,受相机镜头和云锦样本的影响,图像会存在噪音。因此需要对原始图像进行降噪处理^[21]。本文的主要目标是判断云锦图像类别,考虑到不同云锦图像花纹有不同的边缘特征,去噪的目的是为了消除图像上离散点,因此采用中值滤波的方式对图像进行滤波去噪^[22]。

中值滤波抑制噪声,其基本原理是将云锦图像某点相邻域中各点值的中值代替该点,让周围的像素值接近真实值,从而消除孤立的噪声点的去噪方法,是非线性信号处理技术^[23]。该方法是用一个二维滑动模板 W ,将图像在 W 内的像素值按大小进行排序,生成单调上升(或下降)的二维数据序列。相关公式如下:

$$g(x,y) = \text{med}\{f(x-k, y-j), (k,j \in W)\} \quad (1)$$

式中, W 为二维模板; f 为原始图像; g 为处理后的图像。中值滤波去噪后的图像如图 3 所示。



(a) 原始图像 (b) 去噪后图像

图 3 中值滤波去噪

Fig. 3 Median filter denoising

如图 3 所示,图像(b)中细微的点较图(a)有所减少。完成图像的去噪任务之后,根据专业人员的标记,选择合适的兴趣区域。兴趣区域的大小为 $100 * 100$,并依据标签构建数据集。由于一张云锦图像上有多个兴趣区域,需要构建合适的云锦图像数据。经过筛选,本文共采集到 12 个类别 576 幅云锦图像。图 4 为筛选的部分云锦图像数据。

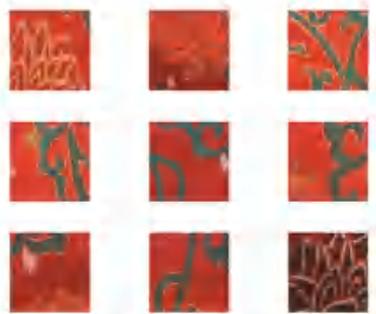


图 4 云锦图像兴趣区域

Fig. 4 Interest area of Yunjin image

3 云锦图像纹理识别

3.1 Gabor 纹理分析

特征分析由图像纹理分析、特征处理、特征选择

等内容组成。主要包括对 12 个类型的云锦图像进行特征比较,根据不同样本类别之间的可分性,对特征进行排序、降维,得到关键的几个特征,并使用常用的分类器对特征进行分类,调节分类器参数,得到分类结果,以验证不同特征的分类效果。

以 Gabor 小波为基础的多分辨率滤波器纹理分析称为 Gabor 纹理分析。二维 Gabor 滤波器能很好地描述哺乳动物大脑视觉皮层中可接收的信息域的分布,两者在空间、频域上均具有相似的局部特点。因此在对不同类别的云锦图像使用 Gabor 滤波器有一定的优势。

在空间域中使用 Gabor 滤波器组,对云锦图像的兴趣区域进行滤波分解。一个二维 Gabor 滤波器是一个正弦平面波调制的高斯核函数。Gabor 滤波器的脉冲响应,可以定义为一个正弦平面波乘以高斯函数。由于乘法卷积性质,Gabor 滤波器脉冲响应的傅里叶变换是其调和函数的傅里叶变换和高斯函数傅里叶变换的卷积。Gabor 滤波器由实部和虚部组成,二者相互正交。二维 Gabor 函数公式为:

$$g(x,y) = \frac{1}{\sqrt{2\pi\sigma_x\sigma_y}} \exp\left\{-\frac{1}{2}\left(\frac{x-\omega}{\sigma_x}\right)^2 - \frac{y^2}{\sigma_y^2} + 2\pi\omega xj\right\} \quad (2)$$

式中: ω 为高斯函数的复调制频率。

$g(x,y)$ 的傅里叶变:

$$G(u,v) = \exp\left\{-\frac{1}{2}\left(\frac{u-\omega}{\sigma_u}\right)^2 - \frac{v^2}{\sigma_v^2}\right\} \quad (3)$$

式中: $\sigma_u = \frac{1}{2\pi\sigma_x}$, $\sigma_v = \frac{1}{2\pi\sigma_y}$ 。

以 $g(x,y)$ 为母小波,通过对 $g(x,y)$ 进行适当尺度变换和旋转变换,可以得到自相似的一组滤波器,称为 Gabor 滤波器,即:

$$g(x,y) = a^{-m}g(x',y'), a > 1, m, n \in Z \quad (4)$$

式中: $x' = a^{-m}(x\cos\theta + y\sin\theta)$; $y' = a^{-m}(-x\cos\theta + y\sin\theta)$; $\theta = \frac{n\pi}{k}$; a^{-m} 为尺度因子; k 为方向的数目。

通过改变 m 和 n ,便可以得到一组方向和尺度都不同的滤波器组。

Gabor 滤波器组的非正交性,意味着滤波后的图像存在冗余信息。为了消除这些冗余信息,在设计滤波器时要确保 Gabor 滤波器组的响应在频谱上半峰幅值能相互接触,并且互不重叠。可采用以下公式减少冗余信息:

$$\sigma_u = \frac{(a-1)U_h}{(a+1)\sqrt{a\ln 2}} \quad (5)$$

$$\sigma_u = \tan^{-1} \frac{\omega}{2k} - \frac{2 \ln \frac{\omega}{U_h}}{2 \ln 2} - \frac{(2 \ln 2)^2 \sigma_u^2}{U_h^2} \frac{1}{2} \quad (6)$$

式中: $a = \frac{\omega}{U_l} \frac{1}{S-1}$; U_h 和 U_l 分别为高频和低频的中心频率; S 为尺度; k 为方向个数。公式(3)中 $\omega = U_h$ 。

图 5 为一组 $a = \sqrt{2}$ 、 $n = 0$ 、 $k = 8$ 、 $m = 0, 1, 2, 3 \dots 8$ 的 Gabor 小波函数。由图可以得出, 二位的 Gabor 滤波器能够在多个方向上获取特征。



图 5 Gabor 小波函数提取特征 1

Fig. 5 Feature extraction 1 from Gabor wavelet function

图 6 为一组, $a = \sqrt{2}$ 、 $k = 8$ 、 $m = 0$ 、 $n = 0, 1, 2, 3, 4$ 的 Gabor 小波函数。由图可见, Gabor 滤波器能够在不同尺度下获取图像特征。

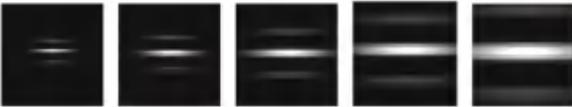


图 6 Gabor 小波函数提取特征 2

Fig. 6 Feature extraction 2 from Gabor wavelet function

依次选择由 5 个尺度、8 个方向组成的 40 组 Gabor 滤波器组, 依次对云锦图像兴趣区域进行滤波。滤波过程如图 7 所示。

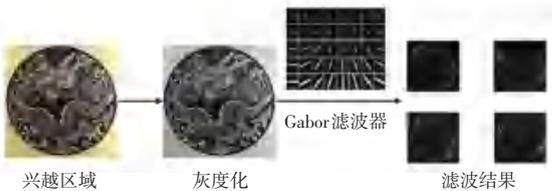


图 7 滤波过程图

Fig. 7 Filtering process diagram

3.2 云锦特征分析

在对云锦图像兴趣区域的滤波后, 得到相应的特征图像, 需要对其进行分析。由于每个云锦图像都有 40 组特征图像, 每张图像的大小为 $100 * 100$, 依次将图像展开, 得到大小为 $1 * 400\ 000$ 的特征序列。为了减少特征维度, 以 1 000 为单位依次对特征序列进行采样, 所得到的采样效果如图 8 所示。

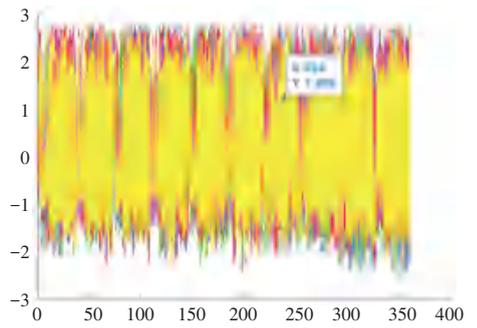
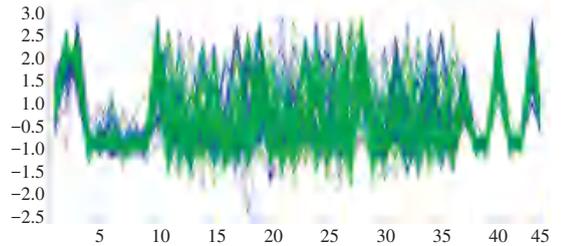


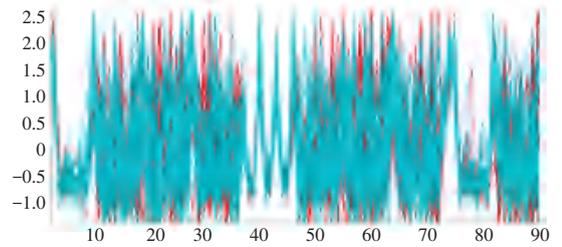
图 8 特征采样图

Fig. 8 Feature sampling diagram

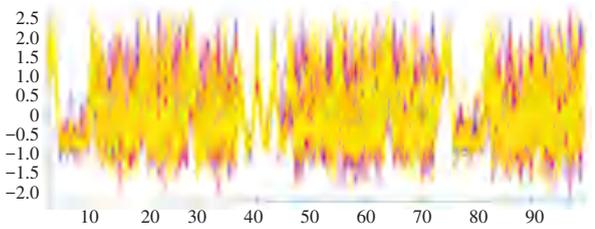
如图 8 所示, 由于有 12 个不同类别的云锦, 在图像上显示较为复杂。为了对其进行有效区分, 以两类为一组在图上进行比较, 如图 9 所示。初步分析可知, 不同类别的云锦图案是能够通过 Gabor 特征进行分辨的。



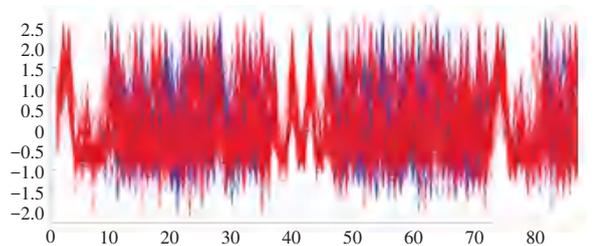
(a) 八宝底纹料和四合如意填花料



(b) 新百花料和普天同庆料



(c) 喜字并蒂莲料和牡丹锦料



(d) 凝香富贵料和气象万千料

图 9 特征对比

Fig. 9 Feature comparison

3.3 云锦特征选择

由于高维数据存在很多冗余或不相关的特征,因此需要对特征进行选择。通过分析特征的可分性,筛选出关键特征,完成特征的选择。

本文主要对特征的可分性进行判据排序,完成特征的筛选。对400个属性的重要性进行排序,采用基于统计的可分性判据,计算每个属性的相对重要性,并选择不同数量的关键特征。其方法如下:

3.3.1 对特征数据进行归一化处理

由于采用不同尺度的 Gabor 滤波器提取特征,则不同尺度下特征数值相差过大。因此,需要对特征数据进行归一化处理,减少实验误差。选用离差标准化方法将特征值映射到 $-1 \sim 1$ 之间,转换公式如下:

$$y = 2 * \frac{x - x_{\min}}{x_{\max} - x_{\min}} - 1 \quad (7)$$

式中: x 为原始的特征值; x_{\min} 为所有特征值的最小值; x_{\max} 为所有特征值最大值。

3.3.2 针对不同属性

由于同类物体各样本之间存在共性,因此类内样本间属性距离应该比不同类样本间属性距离小。根据类内间距离对特征进行排序,计算公式如下:

$$D(x, \{a^i\}) = \sqrt{\frac{\sum_{i=1}^K [\sum_{k=1}^n (x_k - a_k^i)^2]}{K}} \quad (8)$$

式中: K 为样本数量; $\{a^i\}$ 为点集; a_k^i 为点集中 i 个点的第 k 个分量。

为了筛选出关键特征,需依次计算每个属性点到点集的距离,并依据距离值对属性进行排序。排序公式如下:

$$idx = \sqrt{(1 - e^{-\frac{dist}{\beta}})^2} \quad (9)$$

式中: β 为设置的特征权重因子, $dist$ 为特征点到点集的距离。 idx 值越大,表明该属性分类效果越好。

3.4 云锦特征分类

通过对云锦图像的纹理分析发现,不同类别的云锦,纹理特征上存在差异。为了有效的区分不同类别的云锦图像,本文使用贝叶斯分类、k近邻分类两种分类器对特征进行分类,并通过对分类器的调参获取较为精准的分类精度。

贝叶斯分类的基本思想为:对于给出的待分类项,求解在此项出现的条件下,各个类别出现的概率,其最大者,就认为此待分类项属于该类别^[24]。

贝叶斯分类结果的混淆矩阵如图10所示。

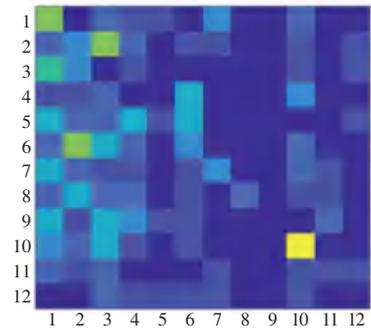


图10 贝叶斯分类结果

Fig. 10 Bayesian classification results

如图10所示,12个类别的云锦,总体分类结果最高在38%左右。表明 Gabor 纹理特征能够区分不同类别的云锦。

其次,选择K近邻分类(KNN),对12种类别的云锦纹理特征进行分类。KNN算法是有监督学习分类算法,通过计算测试样本和所有训练样本的距离。根据最近的 K 个训练样本的类别,通过多数投票的方式进行预测。该算法具有不需要参数,训练时间短等优点。

利用 Gabor 纹理分析获取图像特征,并通过特征筛选,获取不同数量的特征,并通过调节 KNN 算法的 K 值,验证最佳分类精度。KNN 分类结果的混淆矩阵如图11所示。

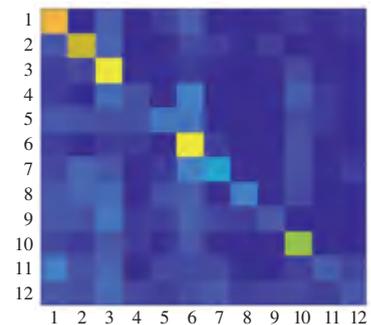


图11 KNN 分类结果

Fig. 11 KNN classification results

如图11所示, K 值为4时精度最高,分类精度为56.31%。由此可见,使用KNN算法对 Gabor 纹理特征进行分类是有效的。

本文对不同类别的云锦进行纹理分析,并完成了纹理特征的分类。由于不同类别的云锦料之间的差异较大,因此分类结果也存在较大误差。因此,对12类云锦料的分类精度只在57%左右,但也证实了不同的云锦料在图像上存在差异。

4 结束语

本文通过使用 Gabor 滤波器对不同材质肌理、色彩和纹样意象的云锦进行滤波,以及纹理特征增强,精细化纹理疏密、虚实、冷暖对比,抑制同质化纹理特征,提高了不同云锦料的分类精度。结果表明,不同材料的云锦图像,可以利用纹理间的差异进行区分。

本文算法实现了首次运用到云锦这类工艺繁复意蕴深厚的非物质文化遗产的图像资源上,取得不错的实验精度。但在准确度和扩展度方面仍有较大的提升空间。在后续研究中,考虑扩展更细粒度的肌理纹样及样本规模,进一步实现实体识别优化。同时,将结合云锦图像资源的语义信息抽取技术,以便于实现本体知识元的关联和推理,更好展现云锦魅力,讲好中国故事,为探索基于数字人文视域下纹理分类研究提供新的研究思路和方法。

参考文献

- [1] 俞悦. 画成其物哀其事——从汉文字“女”谈中国女性文化[J]. 名作欣赏,2019(20):163-164.
- [2] 郑樵撰. 通志二十略[M]. 王树民,点校.北京:中华书局,1995.
- [3] 夏征农. 辞海(第六版彩图本)[M]. 上海:上海辞书,2009:2287.
- [4] 郭小川. 西方美术史研究评述[M]. 黑龙江:黑龙江美术出版社,2003:288.
- [5] 海德格尔. 林中路,孙周兴译[M]. 上海:上海译文出版社,2008:86,89.
- [6] 米歇尔 W J T. 图像转向. 先锋学术论丛——文化研究第3辑[M]. 天津:天津社会科学院出版,2002:17.
- [7] 米歇尔 W J T. 图像理论[M]. 北京:北京大学出版社,2006:3.
- [8] See James J. Gibso. The Perception of the Visual World [M].

Cambridge: The Riverside Press,1950.

- [9] 刘云. 新媒体视域下生态文化图像艺术传播创新之维[J]. 汉字文化,2020(19):152-154.
- [10] 殷曼婷. 视觉“第二自然”的视知觉考察——多米尼克·洛佩斯的视觉识别论[J]. 河南社会科学,2018,26(8):25-31.
- [11] 冯尔康. 史学著作的图文配合与建构视觉史科学[J]. 学术月刊,2006,4.
- [12] ASMUS J F. Digital image processing in art conservation: new technology reveals secrets in classic works of art[J]. Byte, 1987, 12(3): 151-165.
- [13] MOURKOUSSIS N, WHITE M, PATEL M, et al. AMS -- Metadata for Cultural Exhibitions using Virtual Reality [C]// International Conference on Dublin Core and Metadata Applications. 2003: 193-202.
- [14] 吴健,俞天秀,张若识. 敦煌艺术图像数据库的建设[J]. 敦煌研究,2008(6):68-71.
- [15] 王弘泽. 女书艺术的文化与社会意义探微[J]. 美术教育研究, 2020(17):46-47.
- [16] 李敏,崔树芹,陈佳. 应用视觉显著性的小提花织物疵点检测[J]. 纺织学报,2016,37(12):38-42,48.
- [17] 张开生,王泽. 基于磁性纤维的加密纸张图像采集与识别研究[J/OL]. 中国造纸;1-7 [2021-06-09]. <http://kns.cnki.net/kcms/detail/11.1967.TS.20210524.1629.002.html>.
- [18] 包莘莘,李玲香,段华斌,等. 基于阈值与形态学的女红女书图像分割方法[J]. 湖南科技学院学报,2020,41(5):61-63.
- [19] 谢梦婷. 传统文化传播创新的路径探析——以综艺节目《遇见天坛》为例[J]. 汉字文化,2020(8):31-32.
- [20] 王全权,周碧琬. 论国产动画电影中传统文化的美学价值及其影响——以动画电影《大鱼海棠》为例[J]. 江苏大学学报(社会科学版),2018,20(4):17-21.
- [21] 陈宁,洪丰,白冰. 基于边缘特征和霍夫变换的斑马线识别方法[J]. 浙江科技学院学报,2019,31(6):476-483.
- [22] 张静,周佐. 脑部 MRI 图像自动分割算法的研究[J]. 信息技术,2018(6):76-80.
- [23] 陈海秀,成顶,胡祯林. 图像分析在液滴检测技术中的应用[J]. 传感器与微系统,2016,35(4):157-160.
- [24] 陈红松,王钢,宋建林. 基于云计算入侵检测数据集的内网用户异常行为分类算法研究[J]. 信息安全,2018(3):1-7.

(上接第46页)

算法通过对少数类、多数类以及合成样本进行加权处理,提高了 SVM 分类器的稳定性和准确率。实验结果表明论文提出的改进算法,相对于传统 SVM 和 SMOTE-SVM 能达到更高的分类预测精度,能对员工离职倾向进行有效地预测,从而能提前采取应对策略来降低员工离职意向和提升企业竞争力。

参考文献

- [1] 吴丹. 基于数据库知识发现的员工流失预测[J]. 科技与创新, 2019(14):16-19.
- [2] MARQUÉS A I, GARCÍAV, SÁNCHEZ J S. A literature review on the application of evolutionary computing to credit scoring[J]. Journal of the Operational Research Society, 2013, 64(9):1384-1399.
- [3] SHI Donghui, Joze fZurada, Waldemar Karwowski, et al. Batch and data streaming classification models for detecting adverse events and understanding the influencing factors[J]. Engineering

Applications of Artificial Intelligence, 2019, 85: 72-84.

- [4] Sotiris Kotsiantis. Combining bagging, boosting, rotation forest and random subspace methods[J]. Artificial Intelligence Review, 2011, 35(3): 223-240.
- [5] 董燕杰. 不平衡数据集分类的 Random-SMOTE 方法研究[D]. 大连:大连理工大学,2009.
- [6] 覃朗. 基于改进 SMOTE 非均衡支持向量机的建模与应用[D]. 南京:南京航空航天大学,2017.
- [7] KNAR-TEZEL S, LATECKI L J. Improving SVM Classification on Imbalanced Data Sets in Distance Spaces [C]//IEEE International Conference on Data Mining.IEEE, 2009:259-267.
- [8] ZIBA M, TOMCZAK J M, LUBICZ M, et al. Boosted SVM for extracting rules from imbalanced data in application to prediction of the post-operative life expectancy in the lung cancer patients[J]. Applied Soft Computing, 2014, 14(1):99-108.
- [9] HE H, GARCIA E A. Learning from imbalanced data[J]. IEEE Transactions on knowledge and data engineering, 2009, 21(9): 1263-1284.