

基于特征融合的图像情感语义分类

呼克佑, 贺静, 焦丽鹏

HU Ke-you, HE Jing, JIAO Li-peng

太原理工大学 计算机与软件学院, 太原 030024

College of Computer and Software, Taiyuan University of Technology, Taiyuan 030024, China

E-mail: huyan_ky@163.com

HU Ke-you, HE Jing, JIAO Li-peng. Image semantic classification based on multi-features. Computer Engineering and Applications, 2008, 44(20): 181-184.

Abstract: The methods that based on color or color-space features for image semantic classification are simple and efficiently, but no consider the shape feature of the object in the image, so its classification result always not ideally. The paper uses cloth image as date, proposes a new image classification method which base on multi-features including the color feature and the shape feature. The relationship between high-level emotional and level features is strongly. The paper brings forward the relationship between high-level emotional representation and perceptual level features based on the clothing image. Classification is performed by PNN (Probabilistic Neural Networks). The experiment indicates that this method is great effective in image semantic classification and improves the image semantic classification precision

Key words: image of clothing; multi-features; Probabilistic Neural Networks(PNN); semantic classification

摘要: 基于颜色或颜色-空间信息的图像分类方法, 由于没有考虑图像中所含目标对象的形状特征, 分类效果不够理想, 以服装图像作为数据源, 提出并设计了颜色-边缘方向角二维直方图, 将图像的颜色特征与形状特征融合起来进行图像分类。图像中的低阶可视化特征与高阶情感概念之间有着密切的关联, 分析了服装图像的颜色和形状的融合特征与情感之间的相关性, 采用概率神经网络作为分类算法来完成情感语义分类, 实验结果表明, 该方法的分类精度有了明显的提高

关键词: 服装图像; 特征融合; 概率神经网络; 语义分类

DOI: 10.3778/j.issn.1002-8331.2008.20.055 **文章编号:** 1002-8331(2008)20-0181-04 **文献标识码:** A **中图分类号:** TP391

1 引言

随着多媒体和网络技术的迅速发展, 图像正在成为多媒体的主要表现形式, 图像中蕴涵着丰富的情感信息, 如何有效地模拟人观察图像后所引起的情感感觉, 实现基于语义的图像检索, 是一个崭新且有很大挑战性的前沿课题^[1]。基于图像的低阶特征将图像库按照情感语义进行合理的分类, 会大大提高基于语义的图像检索的性能。

服装是人体的外包装, 是生活的必须。随着生活水平的提高, 服装设计的目的不仅仅为防寒遮丑, 而在于满足和适应人们不断增长着的物质和精神生活对服装的需求。人们在挑选服装时, 主要是根据服装所能表达的情感语义对其进行选择, 所以根据服装情感语义对服装图像进行分类有很大的实用价值。本文是根据服装图像所表达的“成熟”与“清纯”这一组截然不同的情感而对其进行分类。

图像情感语义分类的关键任务是分析低阶特征与高阶语义的对应关系, 选择恰当的特征以及适当的分类算法。本文将

采用颜色与形状的融合作为分类的特征; 采用概率神经网络作为分类算法。

2 图像特征提取

2.1 低阶特征与高阶语义的对应关系

图像特征的选择是任何一个分类系统均需要解决的一个重要的环节, 其主要目标是获得最优、最显著有用的特征。对于服装, 服装设计师们^[2]总结出其情感表达最重要的三大要素是色彩、造型和面料, 它们共同确立了整件服装的风格。在形成服装状态的过程中, 最能够创造艺术氛围、感受人们心灵的因素是服装的色彩, 人们经常根据配色的优劣来决定对服装的取舍, 来评价穿着者的文化艺术修养。因此, 色彩是构成服装的重要因素之一。除色彩之外, 服装的外形轮廓能给人们深刻的印象, 在服装整体设计中造型设计属于首要的地位。当面对任何一种服装形式时, 视觉感知了它的“外形”——形状和轮廓。这些形状和轮廓可使服装的复杂形体变得单纯简洁, 它用精练的

基金项目: 山西省自然科学基金 2006011030 (the Shanxi Natural Science Foundation of China under Grant No.2006011030)。

作者简介: 呼克佑 (1963-), 男, 讲师, 主要研究领域为: 信号与信息处理, 数据挖掘; 贺静 (1983-), 女, 在读硕士, 主要研究领域为: 智能信息处理, 数据挖掘; 焦丽鹏 (1984-), 男, 在读硕士, 主要研究领域为: 智能信息处理, 数据挖掘。

收稿日期: 2007-10-10 **修回日期:** 2008-01-21

语言说出了形状的某些基本性质。综上所述,在对服装图像进行特征提取时,颜色特征与形状特征的融合能很好的表达服装风格的特征。本章主要介绍色彩与形状因素与服装风格的关系,服装风格以“成熟”与“清纯”为例。

“成熟”与“清纯”这两种服装风格截然不同,本章通过实例分析颜色与形状的综合因素对服装风格的影响。

成熟风格的衣服比较保守,不太受流行左右,追求严谨而高雅,文静而含蓄,是以高度和谐为主要特征的一种服饰风格,正统的西式套装是成熟的典型代表。从造型元素角度讲,成熟风格多用线造型,线造型多表现为分割线和少量装饰线,衣领处多为V字形折线且与W型的尖角底襟自然接合。服装轮廓多H形,而O形则相对较少。色彩较统一,多以灰色、深蓝、酒红、墨绿、宝石蓝、紫红色等沉静高雅的古典色为主,如图1。图1(a)为一件较正统的西式套装,给人一种稳重的感觉,它对应的颜色直方图为图1(b);由图可知图像中只包含了灰色及与灰色调相近的颜色;对应的边缘图为图1(c)所示,服装上部衣领处的线条于其他风格的服装对比比较特别。

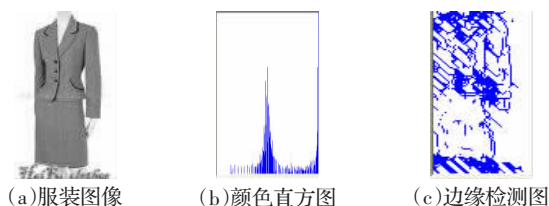


图1 成熟风格的图像特征

清纯风格的衣服主要表现女孩的天真、可爱、纯洁及文静之美,强调简约、自然、青春的风格。从造型元素的角度讲,造型简洁,线条流畅。清纯风格衣服的外形线较少,以上紧下松的曲线为主,轮廓多为A形,且领口处多表现为折叠的曲线如U字形曲线。色彩柔和,一般采用亮丽素雅的色彩,以白色为主,或浅蓝色与白、浅黄、浅绿、浅紫、浅红等色彩搭配,黄与蓝色搭配,绿与浅蓝色、紫色等搭配,均可表现女孩的天真、可爱、纯洁及文静之美,如图2。图2(a)为一套白色为主色调的裙子,给人一种清纯、恬静的感觉。对应的颜色直方图如图2(b)所示,由图可知图像以白色为主;对应的边缘图如图2(c)所示,由图可知造型简洁,线条流畅,轮廓为A形,且领口处表现为U字形曲线。

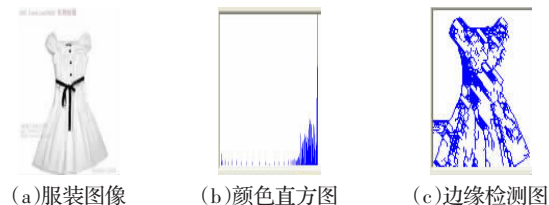


图2 清纯风格的图像特征

以上通过图例的方式,分析了颜色和轮廓对这两种风格的影响,当然服装风格的划分还有很多种,鉴于各种因素,本文暂且以此两种风格为标准对服装图像进行分类。

2.2 特征的融合表示

由于颜色是图像的重要底层特征之一,所以,目前的图像分类技术主要基于颜色特征的。而颜色直方图描述的是图像的颜色统计信息,没有考虑图像中所含对象的形状特征,所以单纯基于颜色的图像分类方法存在许多不足。实际图像中都可能

含有一定的目标对象,其形状特征构成了图像的另一种基本底层特征,它也是人类视觉系统进行图像识别的关键信息之一。本文设计了一种颜色-边缘方向角二维直方图,将图像的颜色与形状特征完美地融合起来。

2.2.1 颜色特征的量化

颜色空间是颜色的表示方法,RGB颜色空间是现在数字图像广泛采用的颜色空间,但它与人眼的感知差别很大^[3]。这里采用了HSV(Hue, Saturation, Value)色彩空间,它不仅能贴近人对图像的色彩理解,而且在匹配颜色和判断一种颜色是否相似于另一种颜色时,HSV颜色空间被视为明智的选择。它由3个分量组成,分别代表色调H(hue)、饱和度S(saturation)和亮度V(value)。

首先将图像 r, g, b 值转换为 h, s, v 值($h \in [0, 360], v \in [0, 1], s \in [0, 1]$),公式如下:

$$\text{Max}=\max(r, g, b)$$

$$\text{Min}=\min(r, g, b)$$

$$v=0.299r+0.587g+0.114b$$

$$s=\begin{cases} 0, & \text{if Max}=0 \\ (\text{Max}-\text{Min})/\text{Max} & \text{else} \end{cases}$$

$$h=\begin{cases} 0, & \text{if Max}=\text{Min} \\ 60(g-b)/(\text{Max}-\text{Min}) & \text{if Max}=r \text{ and } g>b \\ 360+60(g-b)/(\text{Max}-\text{Min}) & \text{if Max}=r \text{ and } g<b \\ 60[2+(b-r)/(\text{Max}-\text{Min})] & \text{if Max}=g \\ 60[4+(r-g)/(\text{Max}-\text{Min})] & \text{else} \end{cases} \quad (1)$$

进而,根据人眼对颜色的感知特性,对HSV空间进行32级非均匀量化^[4]。量化方法如下:

(1)对于亮度 $v<0.2$ 的颜色认为是黑色, $Cod=0$ 。

(2)对于饱和度 $s<0.1$ 且亮度 $v>0.2$ 的颜色按亮度划分为3种灰度。分别为:深灰(0.2,0.5)、浅灰(0.5,0.8)和白色(0.8,1), $Cod=1,2,3$ 。

(3)其它颜色认为是彩色,将其划分为成28种彩色, $Cod=4+4 \times h+2 \times s+v$ 。

(4)对色度 $h[0, 360]$ 划分为赤、橙、黄、绿、青、兰、紫7种颜色。门限分别为:[20,45,75,165,200,270,330], h 分别取值为0,1,2,3,4,5,6。

(5)对于亮度 v 划分为:暗色(0.2,0.5)和明色(0.5,1), v 取值0,1。

(6)对于饱和度 s 划分为:淡色(0.1,0.45)和浓色(0.45,0.1), s 取值0,1。

按照上述方法将HVS颜色空间量化成 $4+7h \times 2s \times 2v=32$ 种代表色,有效地压缩了颜色特征,且能更好地符合人眼对颜色的感知特性^[5]。每种颜色用一个代号表示,得到一个新的颜色空间,记为:

$$HC=\{C_0, C_1, C_2, \dots, C_{31}\} \quad (2)$$

2.2.2 形状特征的计算

边缘作为图像最基本的特征之一,包含着有价值的目标边界信息,反应了图像中物体的形态,是视觉感知的重要线索。由于本文的服装图像是比较规则的目标,边缘方向角直方图会有明显的周期性;所以本文利用边缘方向角直方图表示图像的边缘。Jain^[6]提出的边缘方向角直方图方法,具有尺度不变性,能够较好地描述图像边缘方向角的分布,并且在图像中有局部噪声的情况下鲁棒性较好。具体实现方法如下:使用canny算子

($\sigma=1$, 高斯滤波器宽度=4)对图像进行边缘检测, 得到的边缘图像中每个边缘像素都具有一个梯度模值 $M(x, y)$ 和一个方向梯度相角值 $\theta(x, y)$:

$$M(x, y) = \left[\left(\frac{\partial f(x, y)}{\partial x} \right)^2 + \left(\frac{\partial f(x, y)}{\partial y} \right)^2 \right]^{\frac{1}{2}} \quad (3)$$

$$\theta(x, y) = \arctg \left(\frac{\partial f(x, y)}{\partial y} / \frac{\partial f(x, y)}{\partial x} \right) \quad (4)$$

其中, $\theta \in (-90^\circ, 90^\circ)$, 而边缘像素点的方向梯度相角与该点的边缘线条方向垂直, 所以线条方向为: $\alpha(x, y) = \theta(x, y) + 90^\circ$, $\alpha \in (-0^\circ, 180^\circ)$ 。

根据图像的边缘方向角梯度相角图, 可以计算出线条方向 α 的分布直方图。由于人们对线条角度的感知具有一定的模糊性, 因此以 degree 度为单位对线条方向直方图进行量化, 并归一化。

2.2.3 颜色-边缘方向角二维直方图

由于颜色直方图和边缘方向角直方图代表的是图像的全局统计信息, 为了把不同颜色的同一种物体或相同颜色的不同物体区分开来, 本文融合图像的颜色特征与形状特征, 定义了颜色-边缘方向角二维直方图: 设 $HC = \{<C_0, C_1, C_2, \dots, C_{31}>\}$ 是量化的颜色取值空间, 将向量角 HA 的取值区间 $[0, 2\pi)$ 分成 L 等份, 每份宽度为 $\frac{2\pi}{L}$, 记为 $A_j, j=1, 2, \dots, L$ 。如果统计图像中属于第 C_i 种颜色, 且梯度方向角处于 A_j 的像素点数为 $C_i(j)$, 然后对其进行归一化, 得:

$$H_i(j) = \frac{C_i(j)}{\sum_{m=1}^L C_i(m)} \quad (5)$$

其中 $i=0, 1, \dots, 31, j=1, 2, \dots, L$, 即为第 C_i 种颜色的梯度方向角归一化直方图, 它具有光照不变性、位移不变性和尺度不变^[7]。为了使其具有旋转不变性, 对其再进行离散傅立叶变换, 然后计算每一个变换分量的模值, 将变换后的模值直方图作为图像最终的梯度方向角直方图, 记为 $FH_i(j)$ 。最后将这 32 个 L 维的向量 $FH_i(j)$ 排在一起, 就得到颜色-边缘方向角二维直方图, 记为:

$$H(i, j) = (FH_0(j), FH_1(j), \dots, FH_{31}(j))^T \quad (6)$$

经过实验发现当对 L 取 16 时, 检索的效率和精度综合达到最佳。所以在本文后面的对比实验中, L 取 16。

3 概率神经网络

概率神经网络(Probabilistic Neural Network, PNN)是 Specht 博士^[8]于 1989 年提出的一种神经网络模型。PNN 是一种基于统计原理的神经网络模型, 在分类功能上与最优 Bayes 分类器等价, 同时它是完全前向的计算过程, 因此, 具有训练时间短且不易收敛到局部最小的优点, 这种算法对于多幅图像的分类能得到较好的分类效果^[9]。所以本文采用概率神经网络作为分类算法, 实验达到了理想的效果。

概率神经网络的结构如图 3。

该网络结构可以分为 4 层, 第 1 层是输入层, 输入矢量的维数为 n , 其传递函数是线性的, 接受输入向量并进行规格化; 第 2 层是模式层, 它完成一个给定类的权矢量 W_j 与模式矢量 X 之间的点乘积, 即 $Z_j = X \cdot W_j (j=1, \dots, L)$, 同时在输出层到求

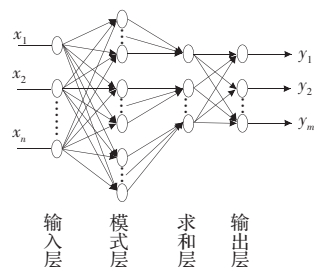


图3 概率神经网络

和层之前完成对 Z_j 的非线性映射, 传递函数采用 $g(Z_j) = \exp[(Z_j - 1)/\sigma^2]$, Z_j 为该层第 j 个神经元的输入, σ^2 为均方差; 第 3 层是求和层, 第三层称之为累加层, 它具有线性求和的功能输出式为: $\sum \exp[-(X - W_j)^T (X - W_j) / (2\sigma^2)]$, 这一层的神经元数目与欲分的模式数目相同。第 4 层输出层是竞争层, 每个神经元分别对应于一个数据类别, 它接收从求和层输出的各类概率密度函数, 最大的神经元输出为 1, 即对应的那一类为待识别的样本模式类别, 其他神经元的输出全为 0。

本文的情况是一个二分类问题, 当输出为 $y_1=1, y_2=0$ 时, 代表风格为“成熟”的服装图像; 当 $y_1=0, y_2=1$ 时, 代表风格为“清纯”的服装图像。

4 实验步骤及结果分析

首先收集 200 幅服装图像, 选择情感形容词对, 精心挑选一组形容词词对(“成熟”与“清纯”), 针对此词对对服装图像进行基于情感语义的分类。本文选择 4 位在校大学生参与, 每个人对所有图像中的服装风格做出评价, 将共同被评价为“成熟”的和“清纯”的图像挑出, 而去除风格模糊不清的图像, 最后精心挑选 120 幅图像, 其中“成熟”为 65 幅, “清纯”为 55 幅。这些将作为实验的数据源。实验以“分类准确率”作为评价准则。

图像情感语义分类包括如下过程:

(1) 图片格式转换及统一图片大小。由于收集图片的格式与大小不一, 这样势必影响实验的效果, 因此首先将图片统一调整为 bmp 格式, 并且将图片大小都调整为 384×256 或 256×384 大小。

(2) 图像特征提取。利用 2.2 小节中的方法计算每幅图像的颜色-边缘方向二维直方图矢量 $H(i, j) = (FH_0(j), FH_1(j), \dots, FH_{31}(j))^T, j=1, 2, \dots, L$ 作为下一步神经网络的输入矢量, 矢量维数为 $2, L$ 的取值为 16 时分类效果最好, 如表 1。

表 1 不同 L 下图像分类性能

L 的值	4	8	16	32	64
准确率/%	80	84	86	83	79

(3) 利用训练集(各类别抽出一定数量的图像作为训练集)对概率神经网络进行训练。将图像的颜色-边缘方向二维直方图作为神经网络的输入矢量, 用户的分类作为神经网络的目标矢量, 对神经网络进行训练。训练出的神经网络就能自动对输入的图像进行分类, 过程如图 4。

(4) 使用各类别余下的图像作为测试集对神经网络的评估结果进行验证。对于任何新的待分类图像可直接作为已训练原型系统的输入, 针对不同的图像语义分类问题根据输入的图像

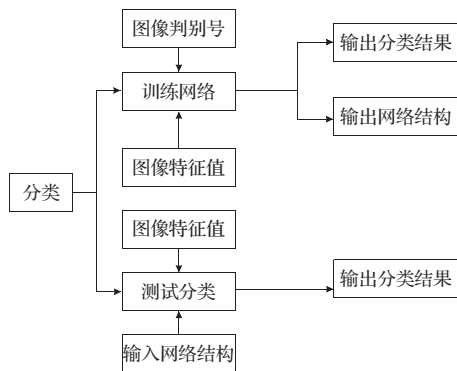


图4 神经网络分类器流程图

特征可以得到图像所属语义类别。将其输出与该图的用户结果进行比较,检验判断的正确率。实验部分结果如图5所示。



(a)“成熟的”服装图像

(b)“清纯的”服装图像

图5 服装图像分类结果

实验达到了85.75%的准确率,结果表明,不论是跟利用颜色特征的图像情感语义分类相比,还是跟利用形状特征的相比,都具有更好的分类效果。如图6。

5 结论

针对颜色直方图不考虑图像中所含对象的形状特征等缺点,本文设计了颜色-边缘方向角二维直方图,它把图像的颜色特征与形状特征很好的融合在一起,作为图像特征的衡量依据。实验结果表明,本文的方法明显优于单纯基于颜色直方图的方法或基于形状的方法,并且分类的结果非常符合人眼的感

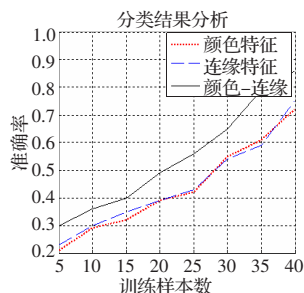


图6 分类结果对比图

知特性—即情感。由于情感研究涉及到心理学、生理学等学科,所以情感语义分类还处于探索和起步阶段。本文利用特征融合技术对服装图像中的“成熟”与“清纯”进行了简单的情感分类,并达到了较理想的效果,是对图像情感语义研究的有益探索。

参考文献:

- [1] 王伟凝,余英林.图像的情感语义研究进展[J].电路与系统学报,2003,8(5):101-109.
- [2] 张繁荣.论色感在服装设计中的应用[J].科技情报开发与经济,2005,15(16):129-130.
- [3] 王涛.基于颜色-空间特征的图像检索[J].软件学报,2002,13(10):2031-2036.
- [4] 孙君顶.基于颜色和形状特征的彩色图像检索方法[J].中国图象图形学报,2004,7(9):820-827.
- [5] 卢成武,种兰祥.融合颜色与形状特征的图像检索方法[J].微电子学与计算机,2006,23(2).
- [6] Vailaya A, Jain A K, Zhang H J. On image classification: city VS. landscape[C]//IEEE Workshop on Content-based Access of Image and Video Libraries, Santa Barbara, CA, 1998: 3-8.
- [7] 魏海.小波变换域内基于方向梯度相角直方图的图像检索算法[J].电路与系统学报,2001,6(2):20-25.
- [8] Specht D F. Probabilistic neural networks[J]. Neural Networks, 1990(3): 109-118.
- [9] 李朝锋,杨茂龙,许磊,等.概率神经网络与BP网络模型在遥感图像分类中的对比研究[J].国土资源遥感,2004,4: 11-13.

(上接29页)

电弧放电。实际解体分析:B相线圈钢压板接地铜线断、器身上杂物较多、垫块个别有位移。中压B相从下往上数第43、44饼间有明显放电且烧损严重,从下往上数第8段第2、3、4柄线圈有变形。确为电弧放电。

6 结语

DGA技术是电力变压器绝缘诊断的重要方法。变压器绝缘故障与油中特征气体组分含量以及特征气体组分比值密切相关。本文提出一种基于变长编码遗传算法的动态聚类算法,该算法可实现故障类别数的自动计算,求出故障聚类原型,完成故障聚类和诊断。综合应用特征气体组分含量和组分比值两种主要故障信息,故障划分正确率较高。今后的工作包括进一步研究油中特征气体组分含量以及特征气体组分比值两类数据的最佳匹配问题以及对算法参数的优选方法。

参考文献:

- [1] 何清.模糊聚类分析理论与应用研究进展[J].模糊系统与数学,1998,12(2):89-94.
- [2] 李俭,孙才新,陈伟根,等.灰色聚类与模糊聚类集成诊断变压器内部故障的方法研究[J].中国电机工程学报,2003,23(2):112-115.
- [3] 孙才新,郭俊峰,廖瑞金,等.变压器油中溶解气体分析中的模糊模式多层聚类故障诊断方法的研究[J].中国电机工程学报,2001,21(2):37-41.
- [4] 熊浩,张晓星,廖瑞金,等.基于动态聚类的电力变压器故障诊断[J].仪器仪表学报,2007,28(3):456-459.
- [5] Pal N R, Bezdek J C. On cluster validity for the fuzzy C-means model[J]. IEEE Trans Fuzzy Syst, 1995, 3(3): 370-379.
- [6] 董云影,张运杰,畅春玲.改进的遗传模糊聚类算法[J].模糊系统与数学,2005,19(2):128-133.
- [7] 陈舵,李雪,崔杜武,等.基于模糊度的聚类有效性函数[J].模式识别与人工智能,2008,21(1):34-41.
- [8] 陈舵,崔杜武,李雪,等.加权模糊聚类及其在电力变压器故障诊断中的应用[J].西安理工大学学报,2008(2).