

文章编号:1001-9081(2007)02-0453-03

## 基于颜色特征的图像情感分类

李海芳, 贺 静, 焦丽鹏

(太原理工大学 计算机与软件学院, 山西 太原 030024)

(sxlhf123@163.com)

**摘要:**分析了服装图像的颜色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。利用服装图像中的低阶特征实现高阶情感语义(穿着“优雅”与“俗丽”)的分类。

**关键词:**服装图像;情感;概率神经网络;颜色直方图

**中图分类号:** TP391.41 **文献标识码:** A

### Image emotional semantic classification based on color features

LI Hai-fang, HE Jing, JIAO Li-peng

(College of Computer and Software, Taiyuan University of Technology, Taiyuan Shanxi 030024, China)

**Abstract:** This paper analyzed the strong relationship between color features and human sensations and showed how high-level emotional representation of painting can be inferred from perceptual level features suited for the particular classes (elegance vs. flaring). Classification was performed by Probabilistic Neural Network (PNN).

**Key words:** image of clothing; affective; Probabilistic Neural Network (PNN); color histogram

## 0 引言

图像中蕴涵着情感信息,如何有效模拟人观察图像后引起的情感感觉,实现基于语义的图像检索,是一个崭新且有挑战性的前沿课题<sup>[1]</sup>。图像情感语义分类可以提高语义图像检索的精确度和效率,是语义图像检索的关键技术之一。

在目前的图像语义分析中,研究人员根据图像低层特征进行简单的语义分类,但很少涉及情感语义范畴。虽然文献[2]根据一些已有的人类认知规则,将低层图像信息结合起来,转化为情感层次的词组,对图像和视频进行了高层语义描述。但是总的来说,由于情感研究涉及到心理学、生理学等多个方面,基于情感的图像分类和检索还处于探索和起步阶段。

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“优雅”与“俗丽”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选择适当的分类算法。色彩在服饰中是最响亮的视觉语言<sup>[3]</sup>,正常情况下当人们注视一件服装时,首先给人们的视觉以强烈刺激的也恰恰是服装的色彩,本文将采用颜色作为分类的特征;概率神经网络(Probabilistic Neural Networks, 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<sup>[4]</sup>。本文采用概率神经网络作为分类算法。

完成图像情感语义分类的步骤如图 1 所示。

图中的数据源为服装图像;特征提取首先分析低阶特征与高阶语义的对应关系,选择颜色直方图作为图像特征;选择适当的分类器首先对其进行训练,然后对服装图像进行分类,

最后对实验结果进行验证。

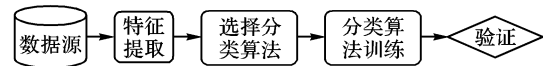


图 1 图像情感分类

## 1 图像特征提取

### 1.1 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<sup>[5]</sup>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,而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,每个色彩都有不同的特性,都有独特的色彩感情与个性表现。

色彩是有感情的:类似色的组合,由于色相差异小,明暗对比弱,给人稳重、优雅的感觉。它适合于成熟、稳重、内向的人。又如红对紫,米黄对紫红等的组合,其之间较鲜明的对比,它适合于年轻、活跃,外向的人<sup>[6]</sup>。颜色特征与情感的关系中,色彩出现的频率与主色调的色差与服装风格有着明显的联系,颜色质朴单一且色差小的服装让人感觉优雅,特别以类似的颜色组合,即色环上 45° 以内的邻近色组合。这种配合能获得稳重、优雅的效果。而颜色鲜明多变的服装给人感觉俗丽(如图 2, 图 3)。图 2(a)的服装图像只有黑色、灰色几种临近色的组合,给人一种稳重优雅的感觉,对应的颜色直方图如图 2(b),横轴表示色彩等级,纵轴表示在某一个颜色等级上具有该颜色的像素在整幅图像中所占的比例,由直方图可以看出图像只有几种相邻色在整幅图中占有比例;对比图 2(a),图 3(a)的服装图像的颜色是多种艳丽颜色的组合,对应的颜色直方图如图 3(b),由图可看出图中包含了多种颜色

收稿日期:2006-08-30;修订日期:2006-11-03 **基金项目:**山西省自然科学基金资助项目(2006011030)

**作者简介:**李海芳(1964-),女,山西太原人,副教授,博士,主要研究方向:信号与信息处理、数据挖掘、图形图像处理;贺静(1983-),女,山西太原人,硕士,主要研究方向:图形图像处理;焦丽鹏(1981-),男,山西太原人,硕士,主要研究方向:图形图像处理。

且相邻色在图中所占比例相差甚远。

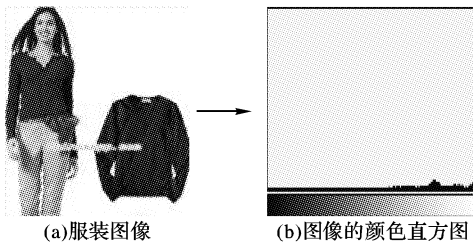


图2 风格优雅的图片特征

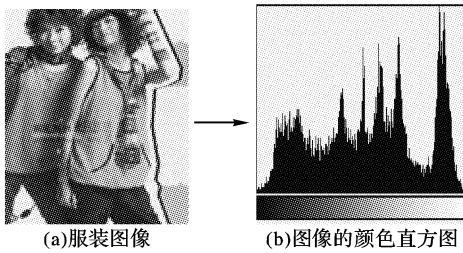


图3 风格俗丽的图片特征

### 1.2 颜色特征提取

颜色直方图是在许多图像检索系统中被广泛采用的颜色特征。它所描述的是不同色彩在整幅图像中所占的比例,而并不关心每种色彩所处的空间位置,即无法描述图像中的对象或物体。本文的目的是根据颜色在服装图像中所占的比例大小来判断服装的风格,鉴于此我们在进行颜色特征提取时采用颜色直方图。

#### 1.2.1 颜色空间

颜色空间是颜色的表示方法,RGB 是现在数字图像广泛采用的颜色空间,但它并不直接与色调、饱和度和亮度等人的主观感觉相对应。我们采用了 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与另一种颜色时,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)、亮度 V (value)。我们采用了如下的从 RGB 到 HSV 的变换方法:

$$\begin{aligned}
 \text{Max} &= \max(R, G, B) \\
 \text{Min} &= \min(R, G, B) \\
 V &= 0.299R + 0.587G + 0.114B \\
 S &= \begin{cases} 0, & \text{if Max} = 0 \\ (\text{Max} - \text{Min}) / \text{Max} & \text{else} \end{cases} \\
 H &= \begin{cases} 0, & \text{if Max} = \text{Min} \\ 60(G - B) / (\text{Max} - \text{Min}) & \text{if Max} = R \text{ and } G > B \\ 360 + 60(G - B) / (\text{Max} - \text{Min}) & \text{if Max} = R \text{ and } G < B \\ 60[2 + (B - R) / (\text{Max} - \text{Min})] & \text{if Max} = G \\ 60[4 + (R - G) / (\text{Max} - \text{Min})] & \text{else} \end{cases}
 \end{aligned} \tag{1}$$

其中,  $R, G, B, S, V \in [1, 0], H \in [0, 360]$

#### 1.2.2 颜色量化

在 HSV 空间中,  $H$  从  $0^\circ$  到  $360^\circ$  变化时,色调依次呈现为红、橙、黄、绿、青、蓝、紫、粉。

本文将  $H$  分量根据视觉对颜色的心理感觉,分为不相等

间隔的 8 份,饱和度  $S$  分成 2 份,亮度  $V$  分成 1 份,并根据色彩的不同范围进行量化,当  $V$  足够小 ( $V < 0.15$  时),视觉感知的颜色基本上接近黑色,可以忽略  $H$  的影响,仅需一个量化值就可以表示。当  $S$  足够小 ( $S < 0.1$  时),视觉感知的颜色基本上接近灰度图像,可以忽略  $H$  的影响,仅需四个量化值就可以表示。量化后的色调、饱和度和亮度值分别为  $H, S, V$ :

$$\begin{aligned}
 H &= \begin{cases} 0 & \text{if } h \in (330, 360] \cup [0, 25] \\ 1 & \text{if } h \in (25, 41] \\ 2 & \text{if } h \in (41, 75] \\ 3 & \text{if } h \in (75, 156] \\ 4 & \text{if } h \in (156, 201] \\ 5 & \text{if } h \in (201, 272] \\ 6 & \text{if } h \in (272, 285] \\ 7 & \text{if } h \in (285, 330] \end{cases} \\
 S &= \begin{cases} 0 & \text{if } s \in (0.1, 0.65) \\ 1 & \text{if } s \in [0.65, 1] \end{cases} \\
 V &= 0 \quad \text{if } v \in (0.15, 1]
 \end{aligned} \tag{2}$$

按照以上的量化级,把 3 个颜色分量合成为一维特征向量:

$$L = HQ_s Q_v + SQ_v + V \tag{3}$$

其中:  $Q_s$  和  $Q_v$  分别是分量  $S$  和  $V$  的量化级数,取  $Q_s = 2, Q_v = 1$ 。因此,式(3)可以表示为:

$$L = 2H + S + V \tag{4}$$

这样,量化后的 3 个分量  $H, S, V$  依式(4)合为一体,根据式(4), $L$  的取值范围为  $[0, 1, \dots, 15]$ , 计算  $L$  获得 16 柄的一维直方图。

#### 1.2.3 颜色直方图

颜色直方图<sup>[7]</sup>最早用于基于颜色的图像检索。色彩直方图空间  $H$  的定义:

$$\begin{aligned}
 H &= | (h[0], h[1], \dots, h[L_k], \dots, h[15]) | \\
 \sum_{k=1}^n h[L_k] &= 1, 0 \leq h[L_k] \leq 1
 \end{aligned} \tag{5}$$

$h[L_k]$  表示第  $k$  种色彩的像素的频数。

度量两幅图像色彩直方图相似性的方法采用欧氏距离,欧氏距离的定义:

$$D(H_s, H_d) = [(H_s - H_d)^T (H_s - H_d)]^{1/2} \tag{6}$$

## 2 概率神经网络

PNN<sup>[8]</sup>采用 Parzen 窗的概率密度函数估计方法和贝叶斯决策规则,并将它们放入神经网络框架中,判断未知数据最大可能属于哪个已知数据集。当这种网络用于模式分类时,可以得到贝叶斯最优结果。PNN 的结构如图 4 所示。

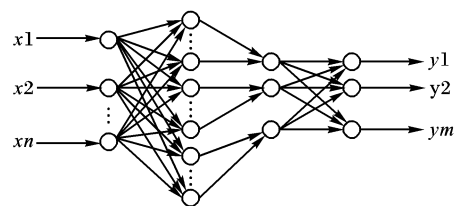


图4 概率神经网络

该网络结构可以分为四层,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  $Z_j = X \cdot W_j$ ,同时在输出层到求和层之前完成对  $Z_j$  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,  $\sigma^2$  为均方差;第三层是求和层,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。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“优雅”的服装图像;当  $y_1 = 0, y_2 = 1$  时,代表风格为“俗丽”的服装图像。

### 3 实验步骤及结果分析

首先收集 100 幅服装图像,选择情感形容词对,精心挑选一组形容词词对(“优雅”对“俗丽”),针对此词对对服装图像进行基于情感语义的分类,本文选择四位在校大学生参与,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“优雅”的和“俗丽”的图像挑出,而去除风格模糊不清的图像,最后精心挑选 70 幅图像,其中“优雅”为 40 幅,“俗丽”为 30 幅。这些将作为实验的数据源。利用图像的多特征融合来表示高阶的情感语义特征,其实验结果达到了较高的满意度。

图像情感语义分类包括如下过程:

1) 统一图像大小。为了提高计算速度,将图像调整为统一合适的大小。最终获得 70 幅图像。

2) 对图像可视化特征研究,提取特征向量。使用前文中的方法计算每幅图的颜色直方图,将此作为下步概率神经网络的输入矢量。

3) 利用训练集(35 幅)对概率神经网络进行训练。将图像的融合特征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,如图 5 所示。

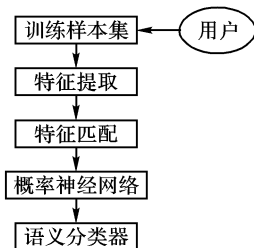


图 5 语义分类器

4) 使用余下的 55 幅图像作为测试集对神经网络的评估结果进行验证。根据输入的图像特征,神经网络自动对图像“优雅”和“俗丽”语义进行分类,将其输出与该图的用户结果进行比较,检验判断吻合率,它表明图像分类后的结果与原始评估之间的吻合程度。

5) 我们选用欧式距离的距离度量方法,实验达到了 75.23% 的准确率。实验部分结果如图 6 所示。

由于该算法只考虑了颜色对服装图像风格的影响,而影响服装风格的因素还包括服装的款式、长短、类型、门襟、面料、纹理、装饰、领、袖等,且人类情感具有较强的主观性和复杂性,因此该算法并不能完全模拟人对图像的感觉,理解图像

的高级语义。在试验结果中不可避免地与人类感知分类有一定差别。



(a) “优雅的”服装图像



(b) “俗丽的”服装图像

图 6 服装图像分类结果

### 4 结语

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。本文利用颜色特征与服装图像中优雅和俗丽风格的联系,对服装图像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。

本系统只是对服装图片中的优雅和俗丽这两种风格进行了分类,服装还有很多种其他的风格,我们可以在今后的学习中不断扩充情感语义类别,所以系统有很好的扩充性,由于情感本身具有主观性和复杂性,本文的方法还存在着很大的不足,有望进一步提高系统的复杂性和评估的正确率。

#### 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] COLOMBO C, DEL BIMBO A, PALA P. Semantics in Visual Information Retrieval [J]. IEEE Multimedia, 1999, 6 (3): 38 - 53.
- [3] 色彩——服装的特殊语言[EB/OL]. <http://hznz.zjol.com.cn/gb/node2/node68266/node222133/node222358/userobject15ai4427849.html>, 2006.
- [4] 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.
- [5] 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.
- [6] 服饰新视野百科知识[EB/OL]. <http://book.jsjy.net/book/C608B6FC0D381F803333/index.pdf>, 2006.
- [7] SWAIN MJ, BALLARD DH. Color indexing[J]. International Journal of Computer Vision, 1991, 7(1): 11 - 32.
- [8] SPECHT DF. Probabilistic Neural Networks[J]. Neural Networks, 1990, 3: 109 - 118.