

基于曲率方向特征的图像情感语义分类^①

Image Emotional Semantic Classification Based on Curvature Direction Features

于 瀛 陈 勇 (重庆师范大学 数学与计算机科学学院 重庆 400047)

摘要: 利用图像的低阶特征对高阶情感语义“静感”和“动感”进行分类。提出了一种新的特征向量——边缘点曲率方向向量,作为图像分类特征,然后利用支持向量机完成分类的工作。对油画、中国画等艺术图像进行了分类实验,实验结果表明该方法能够取得良好的效果。

关键词: 图像分类 情感语义 特征提取 曲率方向向量 支持向量机

从上世纪 90 年代初开始,基于内容的图像检索(Content-based image retrieval, CBIR)技术在图像检索和信息检索领域受到广泛关注,并逐渐成为图像检索的一种主流技术。依据语义的复杂性,一般将语义分为三个层次:从原始图像出发,第一层为低层特征层,包含了颜色、形状、纹理等图像的视觉特征及其组合,这个层次还没有利用图像的语义信息,基于内容的图像检索目前就处于该层次;第二层为对象语义层或语义模板层,需要根据一定的逻辑推理和识别以获得图像中包含的对象类别以及对象的空间位置关系,或者通过语义模板匹配判断图像属于某些语义类;第三层为抽象的语义层,包含了场景语义、行为语义、情感语义等,需要对所描述的对象和场景的含义进行高层推理,将图像内容和抽象概念联系起来。

情感语义作为图像语义中的最高层的语义,通常用形容词来描述,涉及到人的认知模型、文化背景和美学标准等多方面因素。基于情感语义的图像检索主要是通过图像的低层特征来反映高层情感语义,由于图像中信息的丰富性和人的主观因素(经历、心态、目标等)的多样性,因而情感图像检索的研究具有相当难度。同时,情感图像检索的研究与人类的视觉情感信息处理密切相关,是实现真正的“以人为主导”的信息技术的一个核心环节,具有重要的理论意义^[1-4]。

虽然有研究者根据一些已有的人类认知规则,将

低阶图像信息结合起来,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^[5],但已反映的情感在人类的情感范畴中占的比例还很低。因此,基于情感语义的图像分类和检索还有很大的改进空间。

在图像的特征描述中,轮廓能够具体的反映出图像的边界和区域的特征。轮廓有其自身的美学表现价值,图像中某些轮廓特征可以刺激人产生感性认识,轮廓也因此包含有情感语义。已有研究者根据图像轮廓提取的特征来对图像进行语义分类,例如,边缘直方图^[6]和改良后的边缘直方图^[7]以及线条加权方向长度向量^[8]等。本文主要研究艺术图像所表达的“静感”与“动感”这一组截然不同的情感的分类问题。提出了一种新的特征向量——边缘点曲率方向向量,使用支持向量机完成这组情感语义的分类。

1 特征提取

1.1 特征提取的依据

边缘点曲率的不同表现了不同的情感。曲率小的点对应比较直而平滑的曲线,曲率大的点对应频繁摆动的曲线。而直而平滑的线条往往表示静感、安定、永恒,与之相反,频繁摆动的线条多表示有动感,运动,混杂。

图 1 中,(a)和(b)表示“静感”的情感,(c)和(d)表示“动感”的情感。(a)和(b)中的边缘点多数曲率

^① 基金项目:国家自然科学基金项目(60703035);重庆市教育委员会科学技术研究项目(KJ070801)

收稿时间:2008-08-18

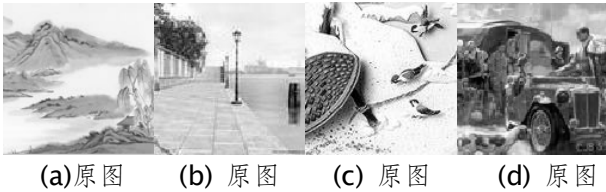


图 1 “静感”和“动感”的图例

较小,方向多为水平和垂直,总体呈现出缓和的状态,而(c)和(d)中边缘点曲率较大,方向多为倾斜的,所呈现出的状态不稳定。据此,可以根据边缘点的曲率和方向对图像中表达的“静感”和“动感”情感进行分类。

1.2 边缘点曲率方向向量

1.2.1 图像边缘提取和边缘点的方向计算

本文采用了多尺度小波变换^[9]来提取图像边缘。用小波变换将图像按尺度分解成不同空间不同频率的子图像。图像经过小波变换后被分割成四个频带:水平、垂直、对角线和低频。图像的边缘信息主要集中在低频中。通过定义两个小波 $\psi^1(x,y)$ $\psi^2(x,y)$ 和实现多尺度图像边缘的提取:

$$\psi^1(x,y) = \frac{\partial \theta(x,y)}{\partial x} \quad \psi^2(x,y) = \frac{\partial \theta(x,y)}{\partial y} \quad (1)$$

其中函数 $\theta(x,y)$ 满足条件:

$$\iint \theta(x,y) = 1 \quad \lim_{x \rightarrow \infty} \theta(x,y) = 0 \quad (2)$$

当尺度为 2^j 时,图像经小波变换由两部分组成:

$$W_{2^j}^1(x,y) = f * \psi_{2^j}^1(x,y) \quad W_{2^j}^2(x,y) = f * \psi_{2^j}^2(x,y) \quad (3)$$

在每个尺度下,梯度向量的模为

$$M_{2^j} f(x,y) = \sqrt{|W_{2^j}^1 f(x,y)|^2 + |W_{2^j}^2 f(x,y)|^2} \quad (4)$$

梯度向量角可以表示为:

$$A_{2^j} f(x,y) = \arctan\left(\frac{W_{2^j}^2 f(x,y)}{W_{2^j}^1 f(x,y)}\right) \quad (5)$$

当尺度为 2^j 时,在 $A_{2^j} f(x,y)$ 的方向上,使 $M_{2^j} f(x,y)$ 具有局部最大值的点即为边缘点。

边缘点的方向和梯度向量正交,因此边缘点的方向可以定义为:

$$D_{2^j} f(x,y) = A_{2^j} f(x,y) + \frac{\pi}{2} \quad (6)$$

在边缘提取的同时,依据公式(6),我们即可得到边缘点的方向。这里将边缘点按方向分为 4 类: $[-75^\circ, -15^\circ)$, $[-15^\circ, 15^\circ)$, $[15^\circ, 75^\circ)$, $[75^\circ, 105^\circ)$, 这样一个图像的边缘点方向向量可表示为 $D=[D1, D2, D3, D4]$, 其中 $D1, D2, D3$ 和 $D4$ 分别代表图像中出现边缘点的方向为 $[-75^\circ, -15^\circ)$, $[-15^\circ, 15^\circ)$, $[15^\circ, 75^\circ)$ 和 $[75^\circ, 105^\circ)$ 的点的个数,这里 $D1$ 和 $D3$ 表示倾斜方向, $D2$ 和 $D4$ 分别表示水平和垂直的方向。

本文采用紧支撑的二次样条小波,连续微分进行边缘检测。选取 4 个不同尺度 ($j=0, j=1, j=2, j=3$) 进行比较。图 2 给出了图 1(a)在这四个尺度下的边缘。



图 2 不同尺度下图 1(a)的边缘

1.2.2 边缘点曲率向量的提取

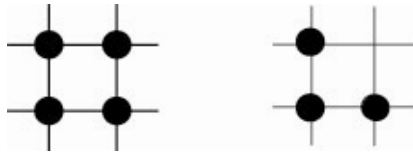
从边缘点矩阵 C 中提取曲率向量 S 分为以下二个步骤完成:

步骤一: 将得到的边缘进行细化。细化的定义依赖于击中击不中变换,采用结构元素对处理边缘^[10]。利用细化结构对 $B=(E,F)$ (见图 3) 对边缘 C 进行细化的定义为:

$$C \otimes B = C \setminus (C * B) \quad (7)$$

利用结构元素对 B^1, B^2, \dots, B^k , 迭代产生输出序列 $C^1 = C \otimes B^1, \dots, C^k = C^{k-1} \otimes B^k$, 随着迭代进行,得到的图像也不断细化,最终结果将稳定下来。可以选择相同的结构元素对,通常使用八个方向的结构对进行细化,使得细化后的图像更加对称。

步骤二: 利用一个 3×3 的滤波模板(见图 4)对细化后的边缘点矩阵进行滤波得到矩阵 C' 。 C' 中每个非零元素值可转换为一个二进制串,其中‘1’的个数为 8-邻接范围内点的个数(包括中心点),‘1’在串中



(a) E (b)F
图 3 细化结构元素对 B=(E,F)

的位置反映了相邻点间的位置关系。若滤波后 C' 中某点的值为 322，转换为二进制串后是 '101000010'，其表示相邻点间的位置关系如图 5 所示。

2 ⁰	2 ¹	2 ²
2 ⁷	2 ⁸	2 ³
2 ⁶	2 ⁵	2 ⁴

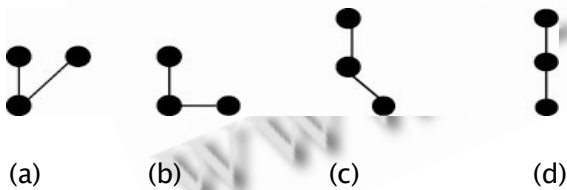
图 4 滤波模板

0	1	0
0	1	0
1	0	0

图 5 二进制串 '101000010' 对应的位置

为了简化问题的分析和处理，在这里只考虑 3 个点的情况。3 个点不同的位置关系反映了由这三点组成的曲线的曲率大小。考虑到对称性，可将 3 点间的位置关系归结为 4 类，如图 6。图 6 中，(a) 和 (b) 表示曲率偏大的情况，(c) 和 (d) 反映曲率偏小的情况。

一幅图像的曲率 S 将是一个四维向量，S=[S1, S2, S3, S4]，其中 S1, S2, S3 和 S4 分别表示图 6 中的 (a), (b), (c) 和 (d) 所示的 4 类曲率情形在图像的所有边缘点中出现的次数。



(a) (b) (c) (d)
图 6 3 点间不同的位置关系

1.2.3 边缘点曲率方向向量

我们提出的曲率方向向量由曲率向量和方向向量共同构成，表示为 SD=[S, D]，其中 S=[S1, S2, S3, S4], D=[D1, D2, D3, D4]。下面我们将曲率方向向量作为图像的高阶情感语义“静感”和“动感”进行分类的特征向量。分类器将采用支持向量机(SVM)来实现。

2 支持向量机简介

本文选择支持向量机作为分类器。支持向量机不仅有统计学习理论[11]作为其推广性的保证，而且在很多实际问题中获得了优于其他方法的结果。下面对支持向量机的基本原理做简要介绍。

在线性可分的情况下，就会存在一个超平面使得训练样本完全分开，该超平面可描述为：

$$w \cdot x + b = 0 \tag{8}$$

其中，“·”是点积，w 是 n 维向量，b 为偏移量。

最优超平面是使得每一类数据与超平面距离最近的向量与超平面之间的距离最大的这样的平面。最优超平面可以通过解下面的二次优化问题来获得：

$$\min \Phi(w) = \frac{1}{2} \|w\|^2 \tag{9}$$

满足约束条件：

$$y_i(w \cdot x_i + b) \geq 1, i = 1, 2, 3, \dots, n. \tag{10}$$

在特征数目特别大的情况，可以将此二次规划问题转化为其对偶问题：

$$\max w(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j (x_i \cdot x_j) \tag{11}$$

满足约束条件：

$$\sum_{i=1}^n \alpha_i y_i = 0, \alpha_i \geq 0, i = 1, 2, \dots, n \tag{12}$$

对于线性不可分的情况，可以把样本 X 映射到一个高维特征空间 H，并在此空间中运用原空间的函数来实现内积运算，这样将非线性问题转换成另一空间的线性问题来获得一个样本的归属。根据泛函的有关理论，只要一种核函数满足 Mercer 条件，它就对应某一空间中内积，因此只要在最优化分类面上采用适当的内积函数就可以实现这种线性不可分的分类问题。

本文选择的是高斯径向基核函数 (RBF)，其具体形式如下：

$$K(x, x_i) = \exp\left\{-\frac{\|x - x_i\|^2}{2\sigma^2}\right\} \tag{13}$$

RBF 核函数中仅有一个参数 σ^2 ，它隐含地改变着映射函数 $\Phi(x)$ ，从而控制着特征空间的性能。

3 仿真实验与分析

本文收集了包括油画、中国画等在内的艺术图像 150 幅，请 8 位年龄不同的人进行“静感”和“动感”的分类，分辨不清的图片从图像库中去除。最后得到

由 100 幅图像构成的图片库。

为了验证本文提出的分类方法的有效性，将图片库进行 3 组划分，分别是：(1)训练集 70 张，测试集 30 张；(2)训练集 60 张，测试集 40 张；(3)训练集 50 张，测试集 50 张。每组划分随机从图片库中选择 10 次，取均值表示(见表 1)。

表 1 比较 3 组划分分类的正确率(j=1)

划分类别	正确率
1 (%)	81.25
2 (%)	89
3 (%)	71.67

按第 2 组划分比较不同尺度下图像分类的结果(表 2)。实验结果表明，当 j=1 时分类效果最好，因此推荐采用尺度 j=1 进行轮廓提取。

表 2 比较不同尺度下分类的正确率

小波尺度	j=0	j=1	j=2	j=3
正确率(%)	78	89	81	84
“静感”类正确率(%)	75.86	88.55	80.74	80.77
“动感”类正确率(%)	80.95	89.62	81.13	87.50

本文比较了新方法与线条方向直方图在尺度 j=1 下的正确率(表 3)。从表 3 中不难看出，新方法在“静感”和“动感”类中都取得比较好的结果。

表 3 j=1 时两种方法正确率比较

	线条方向直方图	曲率方向向量
正确率 (%)	78	89
“静感”类正确率(%)	79.31	88.55
“动感”类正确率(%)	80.95	89.62

实验结果表明，本文提出的方法是有效的，但仍存在一定的误差。造成误差的主要原因有 2 点：(1)为保证算法的效率，在提取曲率向量时，只考虑了存在 3 个点的情况，忽略附近其它的点；(2)本文对“静感”和“动感”的分类是基于图像的轮廓特征，而轮廓特征并不能完全表现出这组高阶情感语义的特征，例如，有些“静感”类的曲率变化大，而“动感”类的曲率变化却相对平缓，致使分类结果有误差。

4 结语

情感研究涉及到心理学、生理学等多门学科，而且人类感知和思维也极为复杂，所以情感语义分类有很多深层次的工作需要研究。如何在图像的低阶特征与图像的高阶情感语义之间建立联系是情感语义的分类具有挑战性的课题。本文根据图像中边缘点曲率和方向与高阶情感语义组“静感”和“动感”之间的联系，采用了边缘点曲率方向向量作为图像特征，实现了这组情感语义的分类，达到了 89% 的分类正确率，取得了较好的效果。

但是，考虑到情感本身所具有的复杂性，本文的分类方法也存在着一定的局限性。计算曲率时考虑更多种可能的情况，或者加入颜色、纹理等图像特征，都有望进一步提高评估正确率。

参考文献

- Asano T. Progress on 'Kansei' measurement technology and its application to visual inspection. INT J JPN SPREC ENG, 1998, 32(2): 85 - 89.
- Nagamachi M. Kansei engineering. Tokyo, Japan: Kaibundo Publishing, 1997.
- Picard RW. Synthetic emotion. IEEE Computer Graphics and Applications, 2000, 20(1):52 - 53.
- 王上飞. 感性信息处理在图像检索中的应用研究[博士学位论文]. 中国科学技术大学, 2002.5.
- Colombo C, Del Bimbo A, Pala A. Semantics in Visual Information Retrieval. IEEE Multimedia, 1999,6(3): 38 - 53.
- Vallaya A, Jaln AK, Zhang HJ. On image classification: City vs. landscape, in IEEE Workshop on Content-Based Access of Image and Video Libraries, Santa Barbara, CA, 1998:3 - 8.
- 王伟凝,余英林,张剑超.基于线条方向直方图的图像情感语义分类.计算机工程,2005(11):7 - 9.
- Wang WN, Yu YG, Zhang JC. A New Svm Based Emotional Classification of Image. Journal of Electronics(China), 2005(1):98 - 104.
- Mallat S, Zhong S. Characterization of Signals from Multiscale Edges. IEEE Trans. on Pattern Analysis Machine Intelligence, 1992:710 - 732.
- 侯学智,杨平,赵云松.CCD 图像的轮廓特征点提取算法.电子科技大学学报, 2004,(4):446 - 448.
- Nello Cristianini, John Shawe-Taylor.支持向量机导论.北京:电子工业出版社,2005.