

# 融入模糊理论的 SVM 在图像情感识别中的应用研究

陈俊杰 张大炜 李海芳

(太原理工大学计算机与软件学院 太原 030024)

**摘要** 引入了将模糊理论融入 SVM 的改进理念,即 FSVM,并实现了一种利用 FSVM 作为分类器且对图像逐层进行分类直至情感语义层面的分类系统,其难点在于建立从图像的低阶图像特征到高阶语义特征之间的映射关系,以及如何选取适合的隶属度函数来确立测试图片的具体语义类别。实验结果表明,本系统在图像情感识别中确实具有简单、快速、高效等特点,从而证明本系统将图像的语义分类提升到情感层面是成功的。

**关键词** 模糊理论,支持向量机,隶属度,图像情感识别,情感语义

中图法分类号 TP391 文献标识码 A

## Application Research of SVM Introjecting Fuzzy Theory in Image Affective Recognition

CHEN Jun-jie ZHANG Da-wei LI Hai-fang

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

**Abstract** This paper introduced FSVM, which introjects fuzzy theory to SVM, achieves a classification system which classifies image layer by layer to affective semantic level by FSVM, and proposed one kind of image affective semantics classification method. The difficulty is to establish a mapping from image features to image affective semantics and how to select fitting membership function to test image semantic class. The experimental result shows that the system is simple, fast, effective, and so on, therefore our system is proved to be successful in promoting the image semantic classification to affective semantic level.

**Keywords** Fuzzy theory, SVM, Membership, Image affective recognition, Affective semantic

## 1 引言

随着计算机技术的突飞猛进和互联网的广泛应用,大量图片、图像等半结构化和非结构化数据涌到了我们的面前,如何有效地组织、管理和检索这些大规模的图像数据库已经成为当前的一个热门话题<sup>[1,2]</sup>。众所周知,绝大多数的图像中蕴含着丰富的情感语义,如何处理这些图片数据,从这些更为繁杂的非传统的数据中挖掘出能为人们所用的知识,从其中识别出相应的情感呢?

带着以上这些问题,先来审视一下当看到一幅图像时是如何感知该图像所蕴含着的情感的。当双眼注视到一幅图像时,首先会根据图片的轮廓、色彩等视觉特征判断出该图像是一种什么样的事物,与此同时也会根据图像的整体布局、颜色、明暗等特征感知到其间所蕴含着的情感信息。

能否让计算机模仿人的感知过程,也能准确地识别出蕴含于图像之中的情感呢?首先,人类的眼睛和大脑可以识别出图像的轮廓、色彩等视觉特征,而当今的计算机技术也能支持计算机识别出图像的纹理、形状、颜色等低阶特征。其次,人类能根据图像的轮廓、色彩等视觉特征判断出图像是怎样的事物,我们也可以令计算机根据图像的低阶特征并通过分类或聚类等相关算法把图像分门别类——最终确定图像是怎

样的事物。再者,人类在判明事物类型之后,会根据图像的整体布局、颜色、明暗等特征感知到图像所表达的情感,同样可以令计算机做到与人相近的地步,这也是工作所在:建立从图像的低阶特征——形状、纹理、颜色、轮廓等到图像的高阶特征——情感语义之间的映射关系。

人类的情感是极其复杂且界定极不精确的,在处理复杂度极高、精确度极低的情感语义层面问题时,复杂度和精确度的权衡问题不可避免地摆在了我们面前。正如系统理论的先驱 Zadeh 在不相容原理(Law of Incompatibility)中所指出的:“随着复杂性的增加,精确的陈述失去意义而有意义的陈述失去精度”。如何处理这些不精确、不确定的信息呢?融入模糊理论的 SVM——FSVM(Fuzzy Support Vector Machine)是解决这一问题的一柄利器。本文将使用 FSVM 作为图像情感语义分类器来处理复杂度极高、精确度极低的情感语义分类问题。

## 2 模糊理论

模糊理论作为一种处理不精确性和不确定性信息的理论工具,其应用已经深入到自然科学和社会科学的各个领域,并被认为是 21 世纪的全球十大热门技术之一。现实世界是复杂多变的,世界的复杂性通常是因含糊不定而引起的。对事

到稿日期:2008-09-23 返修日期:2008-12-04 本文受国家自然科学基金(60773004),山西省自然科学基金(2006011030,2007011050)资助。  
陈俊杰(1956—),男,博士,教授,博士生导师,主要研究方向为数据库与智能信息处理,E-mail:chenjj@tyut.edu.cn;张大炜(1980—),男,硕士研究生,主要研究方向为智能信息处理、情感计算;李海芳(1964—),女,博士研究生,教授,主要研究方向为智能信息处理。

物了解得越多,对其理解的程度就越高,其复杂性也就越低。随着复杂性的下降,复杂系统在建模时根据某种计算方法所得到的精度就显得更有用(或更有意义)<sup>[3]</sup>。系统复杂性与解决方法的对应关系曲线如图1所示。

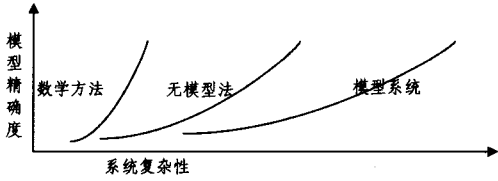


图1 系统复杂性与解决方法的对应关系图

在采用模糊技术处理个体元素对标准模糊集的识别时,即待识别的对象是明确的元素,而标准类型是模糊的时候,需要采取隶属度函数原则对其进行分类。隶属度函数的设计是整个模糊算法的关键,不同的隶属度函数会对算法的处理结果以及算法实现的难易程度产生不同的影响,这就要求隶属度函数必须能客观、准确地反映系统中样本存在的不确定性。在对复杂度不断攀升的问题进行分类时,传统的“非此即彼”的二元隶属关系已无法满足,这时隶属度的取值范围已不是 $\{0,1\}$ ,而是 $[0,1]$ 。在实际应用中经常用到3类隶属函数,如图2所示。

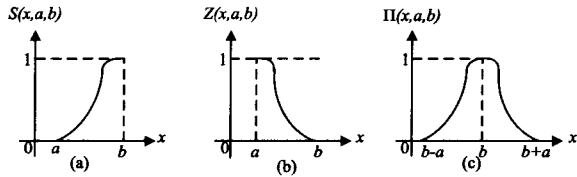


图2 3种常用隶属函数

(1) S函数(偏大型隶属函数)

$$S(x;a,b) = \begin{cases} 0, & x \leq a \\ 2\left(\frac{x-a}{b-a}\right)^2, & a < x \leq \frac{a+b}{2} \\ 1-2\left(\frac{x-b}{b-a}\right)^2, & \frac{a+b}{2} < x \leq b \\ 1, & b < x \end{cases} \quad (1)$$

$a, b$ 为待定集合的参数。这种隶属函数用于表示如年老、热、高等表示偏向大的一方的模糊现象,如图2(a)所示。

(2) Z函数(偏小型隶属函数)

这种隶属函数可用于表示像年轻、冷、矮等偏小的一方的模糊现象,如图2(b)所示。

$$Z(x;a,b) = 1 - S(x;a,b) \quad (2)$$

(3) Pi函数(中间型隶属函数)

$$\Pi(x;a,b) = \begin{cases} S(x;b-a,b), & x \leq b \\ Z(x;b,b+a), & x > b \end{cases} \quad (3)$$

这种隶属函数可以用于表示像中年、适中、平均等趋于中间的模糊现象,如图2(c)所示。

### 3 支持向量机

支持向量机 SVM(Support Vector Machine)是 Vapnik 等人在 20 世纪六、七十年代开始致力于统计学习理论方面研究的基础上,90 年代中期提出,进而不断发展和成熟起来的。以图 3 为例简单介绍一下 SVM 的工作原理:在图 3(a)中假设有实心点和空心点两类样本且这两类样本(训练集)是线性可分的,其中  $L$  是分类线, $L_1, L_2$  分别为经过各类中离分类线

最近的样本且平行于分类线的直线, $L_1$  与  $L_2$  之间的距离( $2/\|W\|$ )为分类间隔。经过机器学习后会得到一个最优分类线,该线满足既可以将两类正确分开又可以使分类间隔最大,分类线方程为  $x \cdot y + b = 0$ ,对其归一之后使得样本集  $(x_i, y_i), i=1, 2, 3, \dots, n, x \in R^2, y \in \{+1, -1\}$ , 满足条件  $y_i[(w, x_i) + b] - 1 \geq 0, i=1, 2, 3, \dots, n$ 。满足上述条件且使得  $\|W\|$  最小的分类线(面)就叫做最优分类线(面),使得  $y_i[(w, x_i) + b] - 1 = 0$  的样本点就是支持向量,解之可得到最优分类函数:

$$f(x) = \text{sign}\left(\sum_{i=1}^n \alpha_i y_i (x_i, x) + b\right) \quad (4)$$

其中  $0 \leq \alpha_i \leq C, i=1, 2, 3, \dots, n$ 。以上是针对样本线性可分的情况,接下来根据图 3(b)介绍样本线性不可分但却非线性可分的情况,通过  $\Phi$  变换把样本变换到高维空间,使其在高维空间里成为线性可分的,但这将导致所谓的“维数危机”。为解决“维数危机”问题引入核函数的概念,利用核函数  $K(x, y)$  来代替之前提到过的点积  $(x \cdot y)$ 。引入核函数后的最优分类函数变成:

$$f(x) = \text{sign}\left(\sum_{i=1}^n \alpha_i y_i K(x_i, x) + b\right) \quad (5)$$

$$b = \frac{1}{N_{NSV}} \sum_{x_i \in JN} (y_i - \sum_{x_j \in J} \alpha_j y_j K(x_j, x_i)) \quad (6)$$

其中  $N_{NSV}$  为标准支持向量,  $JN$  为标准支持向量集,  $J$  为支持向量集。

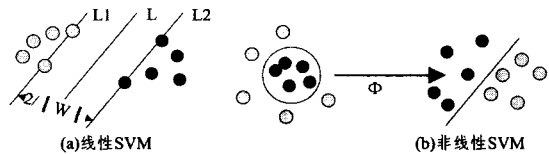


图3 SVM分类示意图

支持向量机固然有其出色的学习特性,一直以来都是机器学习的研究热点,在很多领域也得到了成功的应用。然而,随着该技术的不断发展,其诸多局限性也日益显露出来——若训练集的输出为模糊数,则传统向量机将无能为力<sup>[4]</sup>,并且传统的支持向量机在构造最优分类面时其样本是具有相同作用的。然而在解决实际问题时,样本中往往含有很多位于“临界面”附近的野值点,这些点将直接导致所得出的分类面并不是最优分类面。针对上面所述的问题,引入融合了模糊理论的支持向量机——FSVM。

### 4 模糊支持向量机

现实应用中处理分类问题对某些重要的要求会相对低一些,这些问题对于常规的支持向量机类需要非常高的分类精度,而对其他类分类精度则是很难解决的。针对这一问题,台湾学者 Lin 等人提出了模糊支持向量机(FSVM)<sup>[5]</sup>。模糊支持向量机相对于传统的支持向量机,除了样本特征以及类属性外还增加了隶属度元素。根据样本不同特征对分类的贡献多少赋以不同的隶属度,进而削减了野点对分类的负面影响,训练集由原来的  $(x_i, y_i)$  转换为  $(x_i, y_i, \mu(x_i)), i=1, 2, 3, \dots, n, x \in R^n, y \in \{+1, -1\}, \mu(x_i)$  为隶属度,相应的最优分类面判别函数和核函数分别为:

$$\tilde{f}(x) = \text{sign}\left(\sum_{x_i \in SV} \alpha_i y_i K(x_i, x) + b\right) \quad (7)$$

$$K(x, y) = \exp\left(-\frac{\|x - z\|^2}{2\sigma^2}\right) \quad (8)$$

其中  $0 \leq \alpha_i \leq \mu(x_i)C, i=1, 2, 3, \dots, n$ 。与传统的支持向量机相比 FSVM 最优分类面判别函数中  $\alpha_i$  的取值范围由原来的  $[0, C]$  变为  $[0, \mu(x_i)C]$ , 也就是说对  $\alpha_i$  的取值范围进行了模糊化处理, 这样处理的好处就在于最大程度地减弱野值对分类的负面影响, 从而令所得的分类面最优。

在采用模糊技术处理时, 隶属度函数的设计是整个模糊算法的关键所在, 其必须能客观、准确地反映系统中样本存在的不确定性<sup>[6]</sup>。本文选用的是基于 S 函数的隶属度函数。

## 5 实验结果及分析

### 5.1 创建图片库

为了验证算法的有效性, 本文选取中国科学院心理研究所心理健康重点实验室编制的中国情感图片系统 (Chinese

Affective Picture System, CAPS) 中的 852 幅图片作为原始图片。首先对每幅原始图像进行平移、旋转等生成统一尺寸为  $433 \times 316$  的标准图并入库; 接下来取出图库中 470 幅图片请大一学生对图片库中图片进行人为分类作为数据源, 最终确定室外图片共 241 幅。室外图片分为 3 类, 分别是人造景观类 95 幅、自然风光类 134 幅、其他类别 12 幅。人造景观又分为 3 类: 巧夺天工 67 幅、残垣断壁 18 幅、其他类别 10 幅; 自然风光类分为海洋风光 36 幅、山脉风光 19 幅、森林风光 11 幅、花草树木 50 幅、其他 18 幅。室内图片共 229 幅, 分为事物 158 幅、人物 65 幅、其他 6 幅。其中事物又下分为家装 110 幅、食物 26 幅、其他 22 幅; 人物分支下又分为面部表情 55 幅、其他 10 幅。具体分类如图 4 所示。实验将以“分类准确率”作为评价准则。

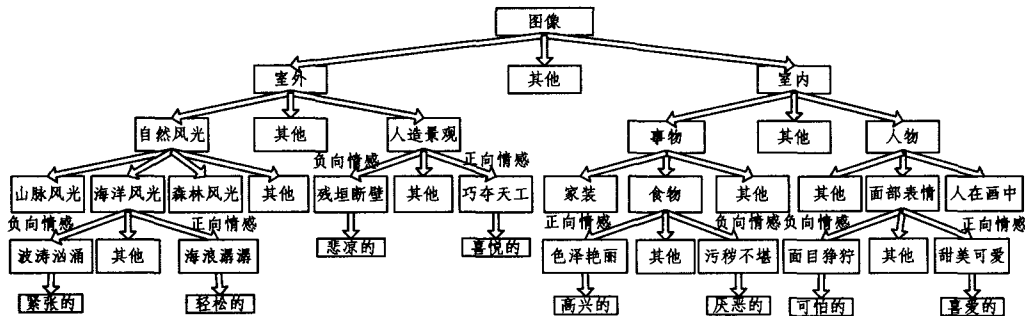


图 4 预期图像分类图

### 5.2 分级图像的分类结果

#### 5.2.1 室内与室外

我们认为在一幅图片中天空不可见 (图片上半部分无成片蓝色)、无绿色、有人物时则可判断该图像为室内图像; 在一幅图片中天空可见、有绿色、没有人物时, 则可判断该图像为室外图像。本文据此依据蓝绿相关向量 (BGCV) 对图片进行室内、室外的分类, 如表 1 所列。

表 1 室内与室外分类结果

类别	特征	训练集		测试集	
		数量	准确率	数量	准确率
室外	BGCV	241	88%	300	86%
室内	BGCV	229	83%	300	84%

#### 5.2.2 自然景观与人造景观

经过上一级分类的训练集和测试集共计 541 幅室外图像作为本层分类的图像数据库, 其中包括自然风光 234 幅、人造景观 258 幅。采用边缘方向直方图 (EDH) 和边缘平行线 (EP) 相结合作为特征输入。表 2 所列为 FSVM 与分类器 LVQ (Learning Vector Quantization) 分类效果的比较, 不难看出 FSVM 具有较高的准确率, 从而说明 FSVM 具有较好的学习和分类能力。

表 2 自然风光与人造景观分类结果

类别	特征	数量	准确率 (LVQ)	准确率 (FSVM)
自然风光	EDH/EP	234	90.3%	92.2%
人造景观	EDH/EP	258	91.4%	93%

#### 5.2.3 海洋风光

承接上一级分类的结果, 在自然景观类的基础上继续细分, 分为山脉风光、海洋风光等, 表 3 所列以海洋风光类为例说明。

表 3 海洋风光情感分类结果

类别	特征	数量	情感类别	情感语义	准确率
海浪潺潺	EDH/EP	24	正向	轻松的	92.3%
波涛汹涌	EDH/EP	26	负向	紧张的	91.4%

运用 EDH 与 EP 作为特征输入, 用 FSVM 为分类器, 将海洋风光继续分为正向情感类以及负向情感类。如“海浪潺潺”类, 其情感语义为“轻松的”; “波涛汹涌”类, 其情感语义为“紧张的”。分类结果如图 5 所示。

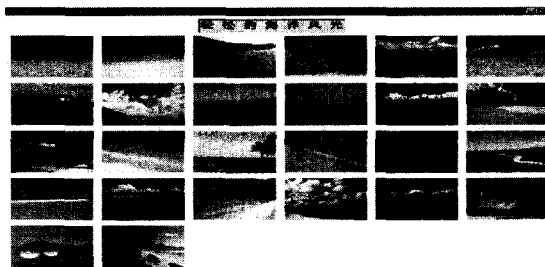


图 5 海洋风光的图像情感语义分类结果

**结束语** 本文对前人的工作进行了大胆创新。实验结果表明这些创新是较为成功的, 但可以预见的是仍然存在着相当大的革新和拓展空间。据此, 我们将在认真总结所取得研究成果的同时, 在下一阶段的研究工作中将决策机制构建入本系统。我们乐观地相信, 构建了决策机制的 FSVM 会在图像情感识别中有更为优异的表现。

## 参考文献

[1] Semeulders A W M, et al. Content-Based image retrieval at the end of early years[J]. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2000, 22(12): 1349-1380

习的结果总是使聚类区内各神经元的权重向量保持向输入向量逼近的趋势。从而使具有相近特性的输入向量聚集在一起。这种结构的网络能从输入信息中找出规律,并根据这些规律相应地调整网络,使得以后的输出与之相适应。

SOM网络由输入层和竞争层组成。输入层对应一个高维的输入向量,竞争层由一系列二维网格节点构成,输入点与输出点为全连接,并通过连接权量化连接强度。设在第 $t$ 次学习后,输入模式 $X$ 激活第 $j$ 个神经元,则模式 $X$ 与第 $j$ 个神经元间的距离最小。设它们之间的距离 $d_j(t+1)=x_i-u_{ij}(t+1)$ ,由于 $x_i-u_{ij}(t+1)=(1-\eta(t))[x_i-u_{ij}(t)]$ ,而 $0<\eta(t)<1$ ,因此有 $d_j(t+1)<d_j(t)$ 。这样当再次输入模式 $X$ 时,它与神经元 $j$ 之间的距离就会逐渐减小,神经元 $j$ 获胜的机会也逐渐增大。通过这种迭代,更新获胜单元及其邻近区域的权值,使得输出节点保持输入向量的拓扑特征。因此,最终的获胜神经元可以表示对该输入模式的分类。

车牌区域的精确定位实际上是一个二分类问题,SOM网络能够很好地解决这类问题。网络的训练样本可以分为两种:标准车牌区域和伪车牌区域。为了区别二者的不同,在上节中给出了4种车牌区域的特征向量,令SOM网络的输入向量 $Input=\{\delta_i, D_i, H_i, P_i\}$ ,输入到神经网络中进行训练、学习;对于竞争层的结构,通过大量的试验选择 $3\times 2$ 网格结构。这样,神经网络就会自动地将 $N$ 个样本矢量按照其本质的属性分别归属到与之对应的类别中,判断其是否为车牌区域,这样就可以精确地定位出车牌区域。

## 5 仿真结果与分析

为验证算法的可行性,本文选取了200幅 $640\times 480$ 大小、在不同光照和时间下的图像对算法进行测试。其中120幅用来训练SOM网络,另外80幅图片用来测试。算法在Matlab7.04环境下编程实现。

实验结果显示,SOM网络的训练时间为357.9220s,定位时间在0.7541~0.8360s之间,车牌的平均定位时间约为0.7865s。实验表明,不同背景环境的图片,定位时间不同。背景越复杂,定位时间越长。对于80幅测试图片,所提算法成功地检测到77幅,准确率为96.25%。识别失败的原因有两种,其中,1幅由于车牌遮挡造成定位失败;而另外两幅定位失败主要是因为车牌倾斜严重。此外,算法检测率与训练样本的选取也有很大的关系,若训练集合涵盖的范围足够广,且数量足够多,则检测效果会有很大程度提高。

实验结果表明,所提算法的定位准确率高,定位速度快,对于不同背景图片具有较强的检测稳定性。图6给出了一幅受污损车牌,结果表明算法可以对其实现精确定位;图7给出了一幅夜间车牌定位结果及其定位结果的反色图。从结果可以看出,由于算法采用了多特征的综合定位,算法受光照的影响很小。但是若车牌出现反光或退色严重等特例情况,算法

定位会失败。



图6 轻度污损车牌定位结果



图7 夜晚车牌的定位结果

**结束语** 车牌自动定位是汽车牌照自动识别系统中的关键,也是较难解决的问题。本文提出了一种新的车牌定位方法,该方法首先利用图像的垂直边缘信息和数学形态学滤波获得候选车牌区域,然后充分利用HSI空间的不同车牌特征信息,应用SOM网络准确地分割出车牌图像。4种不同特征互为补充,避免了传统方法检测标准单一造成误检的缺点。通过对80幅车牌图像进行仿真实验,其有效率可达96.25%。该方法从车牌图像的特征信息出发,充分利用车牌的先验知识,不但能突出车牌区域,而且还能滤除非车牌区域等特点,具有较强的适应性。

由于形态学结构元素的选取、车牌倾斜等原因,最后的定位图像中可能会包含少量的背景图片,这也是本文在今后的研究过程中要不断完善的部分。

## 参考文献

- [1] 赵启升,李存华.基于VC~(++)的车牌识别系统关键技术研究[J]. 计算机科学,2006,33(12):207-213
- [2] 苑玮琦,张亮.基于分块思想的汽车牌照定位算法研究[J]. 自动化学报,2007,33(7):768-770
- [3] Zheng Da-nian, Zhao Yan-nan, Wang Jia-xin. An efficient method of license plate location [J]. Pattern Recognition Letters, 2005,26(15):2431-2438
- [4] 李文举,梁德群,张旗,等.基于边缘颜色对的车牌定位新方法[J]. 计算机学报,2004,27(2):204-208
- [5] Park S H, Kim K I, Jung K, et al. Locating car license plates using neural networks [J]. Electronics Letters, 1999, 35 (17): 1475-1477
- [6] 刘晓芳,程丹松,刘家锋,等.采用改进HSI模型的车牌区域检测和定位方法[J]. 哈尔滨工业大学学报,2008,40(11):86-89
- [7] 曾丽华,李超,熊璋.基于边缘与颜色信息的车牌精确定位算法[J]. 北京航空航天大学学报,2007,33(19):1112-1116
- [8] 中华人民共和国公安部. GA36-1992 中华人民共和国机动车号牌

(上接第290页)

- [2] Rui Yong, Huang T S, et al. Image retrieval: past, present, and future [J]. Journal of Visual Communication and Image Representation, 1999, 10, 1-23
- [3] 高隽. 智能信息处理方法导论[M]. 北京:机械工业出版社, 2004:20-23

- [4] 杨志民,田英杰,邓乃扬.模糊支持向量分类机[J]. 计算机工程, 2005,31(20):25-32
- [5] Lin C F, Wang S D. Fuzzy Support Vector Machines [J]. IEEE Transaction on Neural Networks, 2002, 13(2):464-471
- [6] 张翔,肖小岭,徐光祐.模糊支持向量机中隶属度的确定与分析[J]. 中国图象图形学报,2006,11(8):1188-1192