

# 基于 SVM 的图像分类算法与实现

张淑雅, 赵一鸣, 李均利

ZHANG Shu-ya, ZHAO Yi-ming, LI Jun-li

宁波大学 数字技术与应用软件研究所, 浙江 宁波 315211

Institute of DSP and Software Techniques, Ningbo University, Ningbo, Zhejiang 315211, China

E-mail: zhang\_shu\_ya@yahoo.com.cn

ZHANG Shu-ya, ZHAO Yi-ming, LI Jun-li. Algorithm and implementation of image classification based on SVM. Computer Engineering and Applications, 2007, 43(25): 40-42.

**Abstract:** This paper presents an algorithm to classify old black-and-white photographs and color photographs. The basic idea of the algorithm is that, first defines some indexes according to the different characteristics between them, then uses these indexes to find out images with obvious characteristics by advance classification, at last, classifies other images by SVM to finish the work of image classification, and the indexes above are used as learning parameters. The experimental results show that the definition of indexes is reasonable and the effect is satisfying.

**Key words:** old black-and-white photographs; color photographs; SVM; image classification

**摘 要:** 提出了一种对人类视觉中的黑白老照片和彩色照片进行分类的方法, 其基本思想是, 首先在分析两者各自特点的基础上, 定义了一些指标特征; 其次利用这些指标对图像进行预分类, 找出其中特征明显的图像; 然后使用支持向量机(SVM)的算法对这些指标进行学习来区分剩余图像, 最终达到对图像分类的效果。实验结果表明指标定义是合理的, 效果也较满意。

**关键词:** 黑白老照片; 彩色照片; SVM; 图像分类

**文章编号:** 1002-8331(2007)25-0040-03 **文献标识码:** A **中图分类号:** TP391.41

## 1 引言

现今图像分类是很多领域都需要研究的问题, 应用广泛。特别是图像的合理分类对提高基于内容的图像检索结果的准确性具有十分重要的作用。20 世纪 70 年代, 图像分类主要依靠人工标注的方法。随着计算机技术和数字化图像技术的发展, 图像数据库的规模越来越大, 人们逐渐将研究的重点转移到基于图像内容分析的自动分类研究上。近年来, 许多研究人员对图像的分类进行了深入研究, 并提出了各种解决方案。如 Huang 等人<sup>[1]</sup>通过获得图像颜色的空间相关性, 使用分类树对图像进行分类, 文献[2]中, 使用了颜色直方图和  $k$  近邻分类器来区分室内和室外图像, Vailaya 等人<sup>[3]</sup>使用 Bayesian 分类器, 颜色直方图和边缘方向直方图特征来区分日落/森林/山脉和城市/风景图像。

从不同的角度出发, 图像可以分为不同的类别。本文研究的重点是如何对视觉上的黑白老照片和彩色照片进行区分。这里首先要明确人们视觉上的黑白彩色与计算机存储中的黑白彩色之间的区别, 人们视觉上的黑白照片在计算机中往往是以彩色图片模式保存的。黑白老照片主要是黑白摄影时期的摄影作品, 由于年代久远, 日晒光照, 受潮等原因, 呈现不同程度的

老化, 表面出现泛黄, 泛绿, 泛褐等现象, 而彩色摄影照片中也有大量的黑白背景, 以及色彩不鲜明的情况。这些情况的存在, 是分类算法的难度所在。

本文首先在比较分析黑白老照片和彩色照片各自特点的基础上, 定义了一些指标特征。其次利用这些指标对两者进行逐级预分类, 找出其中特征明显的图像, 然后再将上述指标作为 SVM 学习参数来区分剩余图像, 这里分别考虑了图像的全局特征和局部特征, 实验结果保持了较高的准确性。

## 2 SVM 简介

SVM(Support Vector machine, 支持向量机)方法<sup>[4]</sup>是建立在统计学习理论的 VC 维理论和结构风险最小原理基础上的, 根据有限的样本信息在模型的复杂性和学习能力之间寻求最佳的折衷, 以期获得最好的推广能力。

### 2.1 线性 SVM

SVM 的主要思想是建立一个超平面作为决策曲面, 使得正例和反例之间的隔离边缘被最大化。对于二维线性可分情况, 令  $H$  为把两类训练样本没有错误地分开的分类线,  $H_1, H_2$  分别为过各类中离分类线最近的样本且平行于分类线的直线,

**基金项目:** 国家自然科学基金(the National Natural Science Foundation of China under Grant No.60672072); 浙江省自然科学基金(the Natural Science Foundation of Zhejiang Province of China under Grant No.Y106505); 宁波市科技攻关项目(the Special Science and Technology Foundation of Ningbo of China under Grant No.2005B100016)。

**作者简介:** 张淑雅(1982-), 女, 硕士生, 研究方向: 图形图像与虚拟现实技术; 赵一鸣(1958-), 男, 教授, 主要研究领域: 计算智能, 虚拟现实, 图像处理等; 李均利(1972-), 男, 博士后, 教授, 主要研究领域: 图像度量, 图像质量评价, 图像编码, 图像匹配与检索, 医学图像处理等。

它们之间的距离叫做分类间隔。所谓最优分类线就是要求分类线不但能将两类正确分开,而且使分类间隔最大。在高维空间,最优分类线就成为最优分类面。

设线性可分样本集为 $(x_i, y_i), i=1, 2, \dots, n, x \in R^d$ , 即 $x$ 是 $d$ 维特征向量,  $y \in \{+1, -1\}$ 是类别标号,  $d$ 维空间线性判断函数的一般形式为 $g(x)=w \cdot x+b$ , 分类面方程为:

$$w \cdot x + b = 0 \quad (1)$$

式中 $w$ 为权向量,  $b$ 为分类阈值。要求分类面对所有样本正确分类, 就是要求它满足:

$$y_i [(w \cdot x_i) + b] - 1 \geq 0, i=1, 2, \dots, n \quad (2)$$

满足上述条件且使 $\|w\|^2$ 最小的分类面就叫做最优分类面,  $H1, H2$ 上的训练样本点, 也就是使式(2)中等号成立的样本点称作支持向量。解这个最优化问题后得到的最优分类函数是:

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^n a_i y_i (x_i \cdot x) + b \right\} \quad (3)$$

## 2.2 非线性 SVM 与核函数

在学习样本是线性不可分, 但却是非线性可分的情况下, 可以通过非线性变换把学习样本变换到高维空间, 使其在高维空间里是线性可分的。用核函数 $K(x, y)$ 代替原来的点积 $(x \cdot y)$ , Mercer定理指出, 核函数 $K(x, y)$ 通过与其相联系的非线性变换 $\phi$ 隐含地把特征向量映射到高维特征空间, 使得学习样本成为线性可分的。常用的核函数有:

- (1) 多项式核函数 $K(x, y) = [(x \cdot y) + 1]^d$ ;
- (2) 径向基函数 $K(x, y) = \exp(-\sigma \|x - y\|^2)$ ;
- (3) Sigmoid 函数 $K(x, y) = \tanh(k(x \cdot y) - \mu)$ 等。

选定核函数后, 最优分类函数成为:

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^n a_i y_i K(x_i, x) + b \right\}$$

## 3 算法实现

### 3.1 指标特征定义

从web处获得了RGB颜色空间的黑白老照片和彩色照片各1000幅, 通过比较分析, 定义了一些指标特征, 通过提取这些特征进行分类。

定义1(灰度差异度) 将 $d$ 的平均值定义为灰度差异度 $D$ , 其中 $d$ 为图像中各像素RGB三值与其平均值之差平方和的平方根, 用公式表示为:

$$D = \frac{\sum_{i=1}^N d_i}{N}, N \text{ 为图像中像素个数} \quad (4)$$

$$d = \sqrt{(R - \text{average})^2 + (G - \text{average})^2 + (B - \text{average})^2}$$

$$\text{average} = \frac{R + G + B}{3}$$

RGB三值相等时, 图像表现为纯灰色, 纯色的黑白照片即为此类, 三值越接近, 图像越接近无色彩, 三值差异较大时, 图像呈现出一定的色彩。

图1给出了实验中部分图像的灰度差异度, 为显示清晰, 这里只随机选择了黑白老照片和彩色照片各50幅。保存较好的老照片, RGB三值彼此接近, 因而 $D$ 值较小, 而色彩鲜明的彩色照片, RGB三值各异,  $D$ 值较大, 当然中间存在一个 $D$ 值交叉区域。

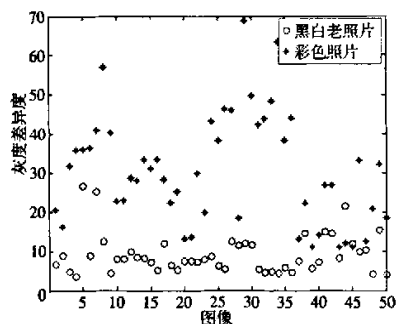


图1 部分图像的 $D$ 值比较

分别考虑了4种情况:

(1) 计算完整图像的全局 $D$ 值, 记为 $D_e$ 。这一步可分离出整体上老化程度非常不明显的黑白老照片和色彩十分鲜明的彩色照片。

(2) 比较图像全局 $D$ 值和除去黑白纯色后 $D$ 值的差值, 记为 $D_s$ 。部分经过处理的彩色照片背景变成纯黑白色, 这时计算得到的差值较大。

(3) 计算将图像分块后各分图像的 $D$ 值, 将分图像的 $D$ 最大值记为 $D_b$ 。对于有一定老化程度的老照片和整体上颜色分布不均匀即某局部色彩鲜艳的彩色照片, 计算得到的全局 $D$ 值可能相差不大, 影响判断, 这时就很有必要考虑局部特征。

(4) 比较各分图像之间 $D$ 值的差值, 将其中差值最大值记为 $D_d$ 。老照片整体上颜色分布较均匀, 而彩色照片在局部特征不够明显的情况下, 局部 $D$ 值间却可能存在较大差异。

定义2(颜色丰富度) 将图像中包含的颜色数定义为颜色丰富度 $E$ , 用公式表示为:

$$E = \sum_{i=0}^{255} \begin{cases} 1, & \text{如果 } c_i \text{ 不等于 } 0 \\ 0, & \text{否则} \end{cases} \quad (5)$$

这里, 首先将颜色空间转换成了HSV, 它与人的感觉相符合, 由3个分量组成, 分别代表色调 $H$ 、饱和度 $S$ 和亮度 $V$ 。将这个变换表示为:

$$T: RGB \rightarrow HSV \quad (6)$$

其次, 把HSV空间均匀地量化成256个颜色, 其中 $H$ 被量化成16个等级,  $S$ 和 $V$ 分别被量化成4个等级。为了计算方便, 把HSV的三维空间转换成一维的特征空间。将这个过程表示为:

$$Q: HSV \rightarrow C \quad (7)$$

其中 $C = \{c_i | i=0, 1, 2, \dots, 255\}$ ,  $c_i$ 表示第 $i$ 种量化颜色。

与图1类似, 图2给出了部分图像的颜色丰富度。彩色照

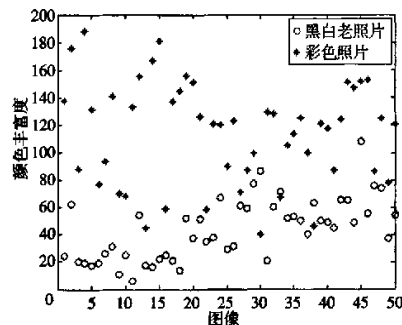


图2 部分图像的 $E$ 值比较

片包含丰富的颜色特征,而黑白老照片色彩较单调,因此部分彩色照片虽然  $D$  值不明显,但  $E$  值可能很大,考虑全局  $E$  值,记为  $E_w$ 。

**定义3 (灰度退化度)** 若图像中像素的  $RGB$  三值彼此间两两差值始终在一定范围内波动,则定义符合该条件的像素个数  $n$  在像素总个数  $N$  中所占的比例为灰度退化度  $R$ ,用公式表示为:

$$R = \frac{n}{N} \times 100 \quad (8)$$

图3给出了部分图像的灰度退化度。黑白老照片年代久远,因不同的老化原因, $RGB$  三值中某值发生退化,使得计算得到的  $R$  值相对较高。这里考虑图像的全局  $R$  值和分块后各分图像的局部  $R$  值最小值,这是因为部分彩色照片中可能存在一些颜色类似于老照片,导致全局灰度退化度较大,影响判断,将两者分别记为  $R_w$  和  $R_s$ 。

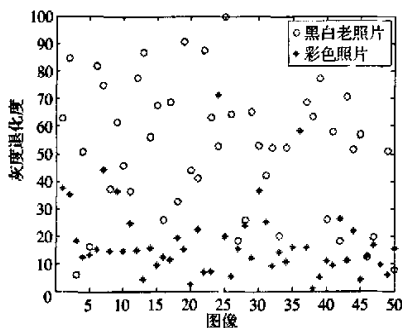


图3 部分图像的  $R$  值比较

通过对上述1000幅黑白老照片和1000幅彩色照片的统计分析,为每一个指标设置了上限阈值或下限阈值,并对其进行归一化,用1代表彩色照片,-1代表黑白老照片。当指标  $D_w$ 、 $D_s$ 、 $D_b$ 、 $D_g$  和  $E_w$  不小于上限阈值时,记为1,这5个指标中除  $D_s$  外不大于下限阈值时,记为-1,因为  $D_s$  小,并不表示图像是老照片的概率大,故不对其设置下限阈值。指标  $R_w$  和  $R_s$  则相反,不小于上限阈值记为-1,不大于下限阈值记为1。各指标位于阈值中间的其他值相应折算成-1与1之间的值,  $D_s$  取值范围是[0,1]。实验中所用上限和下限阈值如表1所示。

表1 各指标特征阈值对应表

	$D_w$	$D_s$	$D_b$	$D_g$	$E_w$	$R_w$	$R_s$
上限	49	5	60	44	170	70	50
下限	1	-	6	4	8	10	10

### 3.2 预分类

对于图像库中大量的黑白老照片和彩色照片,首先,对图像提取前面描述的7个指标进行逐级预分类,通过判断某个指标一层一层刷选出其中特征明显的图像,考虑到效率问题,先使用全局指标,再使用局部指标,即依次为  $D_w$ 、 $E_w$ 、 $R_w$ 、 $D_s$ 、 $D_b$ 、 $D_g$  和  $R_s$ 。若计算得到某一指标值为1或-1,则自动将图像归入相应类别。要注意的是,某些指标不能用于单独判断,应与其他指标结合使用。

### 3.3 SVM 分类

对于剩余图像,利用SVM来区分。用作学习参数的图像特征的选择很重要,颜色特征是图像分类中应用最为广泛的视觉

特征,其中颜色直方图<sup>[5]</sup>方法简单方便,具有平移、旋转、尺度不变性,但它只对颜色出现的频率进行统计,因而区分能力不高。这里我们采用上述提出的指标组成7维向量作为SVM学习参数。

实验总体流程图如图4所示。

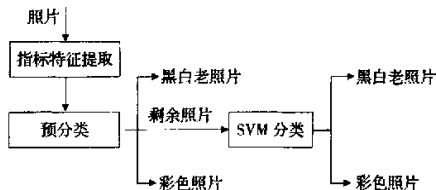


图4 实验总体流程图

## 4 实验结果及分析

从web处另外获取了黑白老照片和彩色照片各1000幅,用上述提出的方法进行了测试。采用分类正确率来评估分类效果,定义如下:

$$\text{分类正确率} = \frac{\text{正确区分的图像数}}{\text{区分出来的图像数}} \times 100\%$$

预分类实验结果如表2所示。从表2可见,初期分类将大量特征较鲜明的图像区分出来,其中黑白老照片和彩色照片各889和783幅(其中包括错误区分的各1和3幅),减轻了后期的分类工作,这里各指标阈值的设置对结果有重要影响。

将剩余黑白老照片和彩色照片各取出50幅作为SVM训练样本,对其余图像再次分类,使用的核函数、参数取值及结果如表3、4、5所示。要注意的是,SVM是针对预分类中无法处理的剩余图像,所以其识别正确率比预分类要略低。

表2 预分类结果

图像类别	区分出图像比例	正确率
黑白老照片	88.90%	99.89%
彩色照片	78.30%	99.62%
平均	83.60%	99.76%

表3 线性核函数分类结果

图像类别	正确率
黑白老照片	86.89%
彩色照片	77.84%
平均	80.26%

表4 多项式核函数分类结果

图像类别	正确率			
	$d=1$	$d=2$	$d=3$	$d=4$
黑白老照片	86.89%	83.61%	81.97%	68.85%
彩色照片	77.84%	82.04%	86.83%	91.02%
平均	80.26%	82.46%	85.53%	85.09%

表5 径向基核函数分类结果

图像类别	正确率			
	$\sigma=2$	$\sigma=3$	$\sigma=4$	$\sigma=5$
黑白老照片	81.97%	80.33%	77.05%	72.13%
彩色照片	88.02%	89.22%	91.02%	93.41%
平均	86.40%	86.84%	87.28%	87.72%

通过表3、4、5可见,后期的SVM分类综合利用各指标有效地区分出剩余图像,选用不同的核函数对识别率有不同的影响,而核函数的参数变化时,结果也会略有不同,需要根据实验来选择合适的参数。实验结果保持了较高的准确性,证明了指标定义是合理的。作为比较,将图像特征改为前面提及的HSV颜色空间中的256维颜色直方图,采用相同的训练集和测试集,仍用上述的SVM进行测试,发现结果略差,得到最高的平均正确率为75.44%。

(下转86页)

$\theta_i$  的单次调整步长不能超过导致以上 4 种情况发生的最小值, 具体可采用试探法根据式(8)(11)(12)进行计算确定。

## 5 实例分析

以一维空间回归问题为例, 对受噪声干扰的 sinc 函数进行建模, 即  $y_i = \sin(x_i)/x_i + \xi_i$ , 其中单变量  $x_i$  为  $[-8.94, 9.00]$  区间上均匀分布的 300 个点, 噪声分为三段, 均服从均值为 0 的正态分布, 方差分别为 0.04, 0.25, 0.09。

实验首先通过标准支持向量回归对其进行训练, 再根据训练结果中的支持向量分布情况对  $\varepsilon$  进行调整, 调整时采用第 4 节所描述在线训练算法, 观察能否提高精度并改善支持向量分布。标准支持向量回归参数可通过交叉校验方法得到, 核函数选择高斯核, 参数分别为  $\sigma=1$ ,  $C=0.1$ ,  $\varepsilon=0.4$ , 本文时间窗调整算法中各参数选择为  $\rho=0.5$ ,  $\gamma=0.2$ ,  $v=0.3$ 。结果如表 1 所示。

表 1 各数据段  $\varepsilon$  调整前后对比

		[-8.94, -3.00]	(-3, 3]	(3, 9)
噪声方差		0.04	0.25	0.09
调整前 $\varepsilon$		0.40		
$\varepsilon$ 调整后	支持向量比例	7%	51%	27%
	均方差(MSE)	0.005 5		
$\varepsilon$ 调整后		0.278	0.480	0.400
$\varepsilon$ 调整后	支持向量比例	20%	38%	27%
	均方差(MSE)	0.004 8		

由表 1 可见, 采取标准支持向量回归时, 由于各数据段的噪声方差不同, 因此支持向量的分布也有很大差异, 大部分的支持向量集中在噪声较大的数据段, 通过支持向量分布的后验信息, 对各段  $\varepsilon$  进行自动调整, 在保持全部数据的整体稀疏性的同时, 也使支持向量分布较为均匀, 增强了回归结果的泛化能力, 拟合精度也得以提高。

(上接 42 页)

SVM 是新的且非常有效的分类方法, 本文利用 SVM 对剩余图像进行分类, 从实践上证明了该分类方法的可行性和有效性。这里我们将同样的训练集和测试集在 BP 网络上进行了测试, 图像特征采用上述 7 个指标, 网络共 3 层, 第 1 层 7 个神经元, 第 3 层 1 个, 第 2 层隐藏层分别设置不同的神经元个数, 得到平均正确率如表 6 所示。

表 6 BP 网络分类结果

隐藏层神经元个数	正确率
2	84.21%
4	85.09%

表 7 实验整体分类结果

图像类别	正确率
黑白老照片	98.11%
彩色照片	98.53%
平均	98.32%

SVM 算法最终转化为凸二次优化问题, 得到的解是全局最优解; BP 网络方法得到的解由于可能存在局部最优解, 因此 SVM 方法解决了 BP 网络方法中无法避免的局部极值问题, 在识别率上较优一些。

本实验整体分类结果如表 7 所示, 其中 SVM 分类选择上面较好的实验结果。除去训练样本, 实际测试黑白老照片和彩色照片各 950 幅。

本文的分类算法, 会将一部分彩色照片归为黑白老照片。观察这些图像, 发现它们在视觉效果上和黑白老照片很接近, 色彩十分单调。而黑白老照片错判为彩色照片的原因主要是老

## 6 结语

本文提出了一种时间窗内  $\varepsilon$  可调的在线支持向量回归方法, 根据各时间窗内支持向量比例的后验信息对  $\varepsilon$  进行调整, 适应于噪声水平随时间变化的回归分析问题, 改善了支持向量分布, 提高回归泛化能力及拟合精度。分析了  $\varepsilon$  调整时对解的影响, 在此基础上给出了  $\varepsilon$  调整的加权支持向量回归在线训练算法, 避免重复求解凸优化问题, 这一在线算法亦可应用于其它类型的  $\varepsilon$  可变的支持向量回归。通过实例分析, 这种方法能适应噪声时变的系统, 回归效果较好。(收稿日期: 2007 年 1 月)

## 参考文献:

- [1] Vapnik V N. The nature of statistical learning[M]. New York: Springer-Verlag, 1995.
- [2] Vapnik V N. Statistical learning theory[M]. New York: Wiley, 1998.
- [3] Smola A J, Schölkopf B A. A tutorial on support vector regression[J]. Statistics and Computing, 2004, 14(3): 199-222.
- [4] Schölkopf B, Smola A J, Williamson R C, et al. New support vector algorithms[J]. Neural Computation, 2000, 12(5): 1207-1245.
- [5] Van Gestel T, Suykens J, De Moor B, et al. Volatility tube support vector machines[J]. Neural Network World, 2000, 10(1): 287-297.
- [6] Tay F E H, Cao L J.  $\varepsilon$ -descending support vector machines for financial time series forecasting [J]. Neural Processing Letters, 2002, 15(2): 179-195.
- [7] 宋晓峰, 陈德钊, 胡上序. 结构可调的支持向量回归估计[J]. 控制与决策, 2003, 18(6): 698-702.
- [8] Smola A J, Murata N, Schölkopf B, et al. Asymptotically optimal choice of  $\varepsilon$ -loss for support vector machines[C]//Proceedings of the International Conference on Artificial Neural Networks, Perspectives in Neural Computing. Berlin: Springer, 1998: 105-110.
- [9] Ma J, Theiler J, Perkins S. Accurate on-line support vector regression[J]. Neural Computation, 2003, 15(11): 2683-2703.

化不均匀, 该问题有待进一步研究解决。

## 5 结论和进一步工作

本文提出的指标考虑到了图像的全局和局部特征, 有效地对黑白老照片和彩色照片进行了分类, 实验结果证明了指标的合理性, 也从实践上证明了 SVM 分类方法的可行性和有效性。核函数的选择很重要, 如何确定适当的核函数及其参数还有待进一步研究。(收稿日期: 2007 年 5 月)

## 参考文献:

- [1] Huang J, Kumar S R, Zabih R. An automatic hierarchical image classification scheme[C]//Proceedings of the Sixth ACM International Conference on Multimedia, Bristol, England, 1998: 219-228.
- [2] Szummer M, Picard R. Indoor-outdoor image classification[C]//Proceedings of the IEEE International Workshop on Content-based Access of Image and Video Database, 1998: 42-51.
- [3] Vailaya A, Figueiredo M A T, Jain A K, et al. Image classification for content-based indexing [J]. IEEE Trans. Image Process, 2001, 10(1): 117-130.
- [4] Theodoridis S, Koutroumbas K. Pattern Recognition[M]. 2nd. USA: Elsevier Science, 2003.
- [5] Swain M J, Ballard D H. Color indexing[J]. International Journal of Computer Vision, 1991, 7(1): 11-32.