

# 基于 CCA 的图像语义特征提取的分析与研究\*

韩昌刚<sup>1</sup>, 郭玉堂<sup>2</sup>

(1. 安徽大学 计算机科学与技术学院, 合肥 230601; 2. 合肥师范学院 计算机科学与技术系, 合肥 230061)

**摘要:** 为了提高图像语义特征提取的精确度, 克服目前大部分图像语义特征提取算法中, 因图像特征提取不当, 导致特征参数不能全面反映图像语义的问题, 提出了一种基于典型相关分析(CCA)的特征融合的图像语义特征提取方法。该方法首先采用圆形对称邻域取代传统的矩形邻域的方法, 对局部二值模式(LBP)纹理特征进行了改进, 然后采用高维小样本下典型相关分析对可伸缩颜色描述算子的颜色特征和改进的 LBP 纹理特征进行特征融合。实验结果表明, 所提出的方法明显提高了图像语义特征提取的精确度, 能有效地建立图像的低层特征与语义特征间的一致性。

**关键词:** 图像语义; 典型相关分析; 局部二值模式; 特征参数; 特征融合

**中图分类号:** TP391.41      **文献标志码:** A      **文章编号:** 1001-3695(2012)05-1938-05

**doi:**10.3969/j.issn.1001-3695.2012.05.090

## CCA-based analysis and research of image semantic feature extraction

HAN Chang-gang<sup>1</sup>, GUO Yu-tang<sup>2</sup>

(1. School of Computer Science & Technology, Anhui University, Hefei 230601, China; 2. Dept. of Computer Science & Technology, Hefei Normal College, Hefei 230061, China)

**Abstract:** For the purpose of have a better accuracy of image semantic feature extraction, overcome the problem that in most image semantic feature extraction algorithms, due to the improper extraction of image semantic feature, lead to the problem of feature parameters can not fully reflect the image semantic, this paper proposed an image semantic extraction algorithm based on canonical correlation analysis and feature fusion. In the proposed method, using the circular symmetric neighborhood, instead of the traditional method of rectangular neighborhood firstly, improved the local binary patterns (LBP) texture feature descriptor. Then, the work was feature fusion between the scalable color descriptor color feature and improved LBP texture feature using canonical correlation analysis under high dimension small sample. Experimental results show that the proposed method significantly improves the accuracy of image semantic feature extraction, creates the consistency between low-level features and high-level semantic effectively.

**Key words:** image semantic; canonical correlation analysis; local binary patterns; feature parameters; feature fusion

## 0 引言

随着计算机科学技术的进步, 尤其是图像处理技术的发展, 人类获得的图像信息越来越多, 对图像处理的要求也越来越强烈。在众多的图像中, 为了获得所需要的图片, 就要求提供好的图像语义特征, 以便准确地检索出要找的图片信息。

图像检索技术发展至今, 形成了几种主要的技术, 包括基于文本、基于内容和基于语义的图像检索。

基于文本的图像检索需要手工对图片库进行语义标注, 这就需要花费大量的人力、物力, 同时标注还带有很强的主观性。基于内容的图像检索(CBIR)<sup>[1]</sup>就克服了基于文本检索的局限性。CBIR 主要利用图像的纹理、颜色、形状、轮廓和对象的空间关系来对需要的图像进行检索。后来人们发现, 在检索中, 需要的是图像的语义信息, 而不是图像的颜色、纹理等信息, 这样就形成了基于语义的图像检索, 使得检索尽可能地贴近大家对图像内容的理解, 从而可以按照用户自己对图像的理解来进

行图像检索。

在进行图像语义的检索中, 正确地提取图像的语义特征是图像检索的一个关键方面。图像语义的提取过程就是把获得的图像的低层特征转换为高层语义的过程。在这个过程中, 大家熟知的语义鸿沟问题就是由于转换的不准确造成的。比如用含有蓝天、白云的图片做样图可能搜索到大海、蓝色旗帜等颜色相似的图像, 就是由于低层特征转换为高层语义时, 没有考虑到具体情况, 转换不准确造成的。

## 1 相关研究工作

作为当今的研究热点, 图像语义特征提取吸引了众多计算机学者的关注。为了快速精准地寻找出目标图像, 就需要提供出色的图像语义特征提取算法, 因此各种特征提取算法应运而生。

文献[2]提出可伸缩颜色描述子, 利用图像的颜色特征来提取图像的语义特征, 实现了对 256 级颜色特征的可变长描

**收稿日期:** 2011-09-17; **修回日期:** 2011-10-24      **基金项目:** 安徽省自然科学基金资助项目(11040606M134); 安徽省高校自然科学基金重点资助项目(KJ2009A150)

**作者简介:** 韩昌刚(1986-), 硕士研究生, 主要研究方向为图形处理(hanchanggang@foxmail.com); 郭玉堂(1962-), 教授, 硕士, 主要研究方向为模式识别与图像处理。

述,降低了计算复杂度。

文献[3]提出边缘直方图,给定一幅图像,把它分成  $4 \times 4$  的子图,把边缘类型分为五类:垂直的、水平的、 $45^\circ$ 斜角、 $135^\circ$ 斜角、无方向的。对于每个边缘类型,分配一个二阶矩阵来描述它。

文献[4]提出边缘方向直方图,使用边缘检测来检验图像边缘特征。常用的边缘检测方法主要有 Robert、Sobel、Prewitt 和 LOG 算子。使用梯度幅值和梯度方向来反映图像的边缘纹理特征。

文献[5]提出用距描述来表征图像的形状特征。其定义了两个二阶不变矩和五个三阶不变矩共七个不变矩,加上图像的均值、方差、平滑度、三阶矩、一致性、熵六个纹理统计量,组成了距描述和统计描述的 13 维全局特征向量。

文献[6]提出局部二值模式(LBP)。LBP 是一种灰度范围内的纹理描述方式,使用环形分割邻域。其利用符号函数,实现了灰度级的不变性,并对每一个 LBP 值赋予一个二进制权值。

本文提出使用局部二值模式(LBP)和可伸缩颜色描述子来表征图像的低层语义特征,再使用高维小样本下的典型相关分析对选取的特征进行融合,使用融合后的语义特征来表示图像的语义,建立起图像低层特征与高层语义之间的关联。实验结果表明,所提出的方法明显提高了图像语义特征提取的精确度。

## 2 混合特征提取

### 2.1 基于可伸缩颜色描述子的颜色特征提取

图像的颜色空间有多种表示方式,本文选用 HSV 颜色空间,即色调  $H$ 、饱和度  $S$ 、亮度  $V$ 。由于获得的图像是在 RGB 空间中的,所以首先要将图像颜色空间转换为 HSV 空间。

如果使用 HSV 颜色直方图来描述图像的整体颜色特征,所用到的特征向量就会非常多,因此在计算之前,要对 HSV 空间进行量化,使用 HSV 空间上的 256 级(bin)直方图。由于  $H$  分量的描述能力与人的视觉最接近,把  $H$  均匀划分为 16 级,每个等级 22.5 度, $S$  均匀划分为 4 级, $V$  均匀量化为 4 级,每个等级 0.25。这样使用的颜色直方图级数为  $16 \times 4 \times 4 = 256$  级。

本文使用的是可伸缩颜色描述子<sup>[2]</sup>。一般的 256 bins(直方条)中,每一条都要使用 11 bit 来表示,这样存储空间就非常大。显然,如果用尽可能少的比特来表示一幅图像,检索速度就会大大提高。这就是可伸缩颜色描述子。

为了实现可伸缩颜色描述子,需要使用 HAAR 变换,如图 1 所示。在 HAAR 变换前,把直方图每个 bin 的表示由 11 bit 非均匀量化为 4 bit。HAAR 变换的主要作用就是减少描述颜色直方图的比特数。HAAR 变换的基本操作是和操作与差分操作。经过一次和操作,直方图的 256(bins)数就会减半,迭代一次减半。和操作得到低通系数,差分操作得到高通系数。HAAR 变换的高通(差分)系数包含了大部分的图像细节信息。

下面给出 HAAR 函数的公式:

$$\text{har}(2^p + n, t) =$$

$$\begin{cases} \sqrt{2^p} & \text{if } \frac{n}{2^p} \leq t < \frac{(n+1)}{2^p} \\ -\sqrt{2^p} & \text{if } \frac{(n+1)}{2^p} \leq t < \frac{(n+2)}{2^p} \\ 0 & \text{else if} \end{cases} \quad (1)$$

HAAR 函数是定义在  $[0, 1]$  区间的函数。其中:  $p = 1, 2, \dots; n = 0, 1, 2, \dots, 2^p - 1$ 。



图1 HAAR变换

HAAR 变换公式为

$$\begin{bmatrix} Ha(0) \\ Ha(1) \\ \vdots \\ Ha(N-1) \end{bmatrix} = \frac{1}{N} [\text{har}_{2^p}] \begin{bmatrix} f(0) \\ f(1) \\ \vdots \\ f(N-1) \end{bmatrix} \quad (2)$$

其中:等式左边是变换矩阵序列;  $\text{har}_{2^p}$  是  $2^p$  阶哈尔矩阵;等式右边矩阵是时间序列。

总的来说,由直方图得到的过程步骤为:首先 256 级颜色直方图非线性量化,然后经过 HAAR 变换,变换后再进行线性量化,最后得到哈尔变换系数。通过这些系数来进行图像匹配,就可以实现图像检索,检索速度会大大提高。

### 2.2 基于改进的 LBP 算子的纹理特征提取

局部二值模式(LBP)<sup>[6]</sup>首先是用来作图像局部特征比较的,传统的 LBP 方法首先通过阈值来标记中心点像素与其邻域像素之间的差别。LBP 是人脸识别中经常使用的一种方法。

最早的 LBP 算子是定义在  $3 \times 3$  邻域上的,它使用中心像素值来计算八邻域像素值,然后再乘以邻域的权值得到结果。下面给出一个例子,如图 2 所示。

example		
6	4	3
7	5	2
8	9	5
threshold(10010111) weights		
1	0	0
1	*	0
1	1	1

LBP=1+16+32+64+128=241  
C=(6+7+8+9+5)/5-(4+3+2)/3=4

1	2	4
128	*	8
64	32	16

图2 LBP算子的一个实例

使用二维分布 LBP 和局部对比测量值  $C$  作为该邻域的特征。但是为了使得到的特征更有代表性,矩形邻域的大小是可以改变的,可以使用  $4 \times 4$  或  $5 \times 5$  的邻域。相比矩形邻域,圆形对称邻域的效果更好。下面的假设都是基于圆形对称邻域的<sup>[6]</sup>。

假设  $P$  为圆形对称邻域除了中心点的像素数目, $R$  为圆形对称邻域的半径, $g_c$  表示圆形对称邻域的中心点像素值, $g_p = (p=0, \dots, p-1)$  表示圆形对称邻域的其他点像素值。使用统计方法来描述图像特征。 $T$  表示  $P+1$  级图像像素的联合分布:

$$T = t(g_c, g_0, \dots, g_{p-1}) \quad (3)$$

由于  $g_c$  的值是  $g_p = (p=0, \dots, p-1)$  的平均值,可以从  $g_p$  得来,再假设  $g_c$  与图像特征无关。所以  $T$  可以改写为

$$T = t(g_c, g_0 - g_c, \dots, g_{p-1} - g_c) = t(g_c) t(g_0 - g_c, \dots, g_{p-1} - g_c) \quad (4)$$

再则  $g_c$  描述的是整体特征,图像邻域整体与  $g_c$  无关,仅

与  $g_p = (p = 0, \dots, p - 1)$  有关。同时为了实现灰度级的不变性,考虑在  $T$  中使用符号函数  $s$ 。化简上式得到:

$$T = t(g_0 - g_c, \dots, g_{p-1} - g_c) = t(s(g_0 - g_c), \dots, s(g_{p-1} - g_c)) \quad (5)$$

最后为了得到 LBP 的值,把二进制权值  $2^p$  赋给每一个  $s(g_p - g_c)$  得到 LBP:

$$LBP_{p,R} = \sum_{p=0}^{p-1} s(g_p - g_c) 2^p \quad (6)$$

圆形(环形)分割可以保持空间信息和区域的旋转不变性。为了实现 LBP 算子的灰度级一致性和旋转不变性。新定义一个量  $U$ :

$$U(LBP_{p,R}) = |s(g_{p-1} - g_c) - s(g_1 - g_c)| + |s(g_{p-2} - g_c) - s(g_0 - g_c)| + \sum_{p=2}^{p-1} |s(g_p - g_c) - s(g_{p-2} - g_c)| \quad (7)$$

根据  $U$  值化简 LBP:

$$LBP_{p,R} = \begin{cases} \sum_{p=0}^{p-1} s(g_p - g_c) 2^p & \text{if } (U \leq 2) \\ p + 1 & \text{else if} \end{cases} \quad (8)$$

算法 1 LBP 特征提取算法。

Input: 输入图片  $I$ 。

Output:  $1 \times 18$  的 LBP 直方图数据。

a) 把  $I$  分割成  $M$  ( $M$  不宜太大) 个块。

b) 设定圆形对称邻域的大小为:  $P = 16, R = 2$ 。根据式(7),计算每个块的 LBP 直方图值。

c) 把  $M$  个块得到的 LBP 值组合起来得到描述图像特征的全局 LBP 值,输出。

这就是得到的最后的 LBP 结果,与此同时,改变  $P, R$  的值可以获得更好的 LBP 图像特征来用于图像比较。

### 2.3 各种特征分析与比较

对于本文提出的五种特征参数,它们在描述图像语义时有各自的优缺点,表 1 给出了各种特征参数的分析与比较。

表 1 图像特征参数的应用与优点比较

特征参数	适用于描述图像的范围和优点
颜色直方图	由于使用图像大多是彩色图像,因此颜色特征在描述图像特征时非常重要。也就是对于彩色图像效果很好
边缘方向直方图	有时目标和背景的色彩相近,或者图像的色彩信息较差,这时基于色彩的图像特征就不足以描述图片的特征。使用边缘方向直方图效果就很好
边缘直方图	边缘特征对于边界区分明显的图像效果更好。在进行图像与图像之间的匹配时具有很重要的作用。因为它描述了边缘的空间分布信息
局部二值模式	LBP 是一种灰度范围内的纹理描述方法,具有良好的推广性,可以描述亮点、暗点、纹理细节等一些局部微小的特征,具有单调变换不变性和旋转不变性。在人脸识别领域应用取得了成功
距描述	矩是形状的全局描述,具有自带的识别和过滤噪声的能力,可以实现尺度和旋转不变性

表 1 给出了各种特征参数的适用范围和优点,有时一种特征参数并不能很好地描述图像的语义特征,需要结合不同特征参数的优点来描述提取图像特征,才能取得好的效果。下面介绍一种基于高维小样本下的特征参数融合方法,为使用两种特征描述图像语义提供了一种策略。

### 3 基于高维小样本下的 CCA 特征融合

#### 3.1 典型相关分析(CCA)

典型相关分析又叫典型分析,是分析两组变量之间关系的

一种方法,对于有联合分布关系的两组向量,要分析其中一组变量和另外一组变量之间的关系。

研究者可能有两组很大的变量,但是它只想得到这两组变量中相关度最高的变量之间的关系,这个时候使用 CCA 就是一种很好的方法。举个例子,有关人体的两组变量,一组是形态指标(身高、坐高、体重、胸围、肩宽、盆骨宽),另一组是机能指标(脉搏、收缩压、舒张压、肺活量),就可以使用 CCA 来研究这两组变量之间的关系。下面给出具体分析。

对于两组变量的线性组合:

$$Xa = \sum_{i=1}^p a_i x_i \\ Yb = \sum_{i=1}^q b_i y_i \quad (9)$$

其相关系数为

$$\rho(a, b) = \frac{a'S_{12}b}{(a'S_{11}ab'S_{22}b)^{\frac{1}{2}}} \quad (10)$$

首先要考虑的问题就是什么时候  $\rho(a, b)$  取得最大值。假设  $X$  是一组  $m \times p$  的数据,  $Y$  是一组  $m \times q$  阶的数据,  $Z = (X, Y)$ 。中心化  $Z$ , 计算:

$$S = \frac{1}{n} Z'Z = \frac{1}{n} \begin{pmatrix} X'X & X'Y \\ Y'X & Y'Y \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \quad (11)$$

这是  $X$  与  $Y$  的协方差阵。下面给出几个基本定义。

定义 1 令  $K = S_{11}^{-1/2} S_{12} S_{22}^{-1/2}$ ,  $N_1 = KK'$ ,  $N_2 = K'K$ , 对  $K$  进行奇异值分解得

$$K = (\alpha_1, \alpha_2, \dots, \alpha_k) D (\beta_1, \beta_2, \dots, \beta_k)' \quad (12)$$

其中:  $\alpha_i$  是  $N_1$  对应于  $\lambda_i$  的特征向量,  $\beta_i$  是  $N_2$  对应于  $\lambda_i$  的特征向量。  $D$  是  $N_1$  或  $N_2$  的奇异值组成的对角矩阵, 即  $D = \text{diag} \{ \sqrt{\lambda_1}, \sqrt{\lambda_2}, \dots, \sqrt{\lambda_k} \}$ 。

如果令

$$a_i = S_{11}^{-1/2} \alpha_i, b_i = S_{22}^{-1/2} \beta_i, i = 1, \dots, k \quad (13)$$

称  $a_i$  与  $b_i$  是  $X, Y$  的典型向量,  $\varphi_i = Xa_i$  与  $\phi_i = Yb_i$  是  $X, Y$  的典型变量,也就是 CCA 要求解的结果,它是  $X$  或  $Y$  各列的线性组合,是最终结果。  $\rho_i = \sqrt{\lambda_i}$  就是  $X, Y$  的第  $i$  个典型相关系数。

$$a_i'S_{11}a_i = b_i'S_{22}b_i = 1, a_i'S_{11}a_j = b_i'S_{22}b_j = 0$$

$$i = 1, \dots, r - 1, 1 \leq r \leq k \quad (14)$$

定义 2 在式(14)的条件下  $f(a, b) = a'S_{12}b$ , 在  $a = a_r, b = b_r$  条件下取得最大值  $\rho_r$ 。

通过以上的分析,给出 CCA 算法的步骤:

算法 2 CCA 求解算法。

Input:  $X$  为一组  $m \times p$  的数据,  $Y$  为一组  $m \times q$  阶的数据。

Output:  $\varphi_i, \phi_i$  为所要求的典型变量。

a) 首先将  $X, Y$  标准中心化,根据式(11)计算  $Z$  的协方差矩阵  $S$ 。

b) 计算  $K, N_1, N_2$  的值,假设  $\text{rank}(N_1) = \text{rank}(N_2) = k$ , 求得  $N_1, N_2$  的非零特征向量  $\alpha_1, \alpha_2, \dots, \alpha_k, \beta_1, \beta_2, \dots, \beta_k$ , 以及非零特征值  $\lambda_1, \lambda_2, \dots, \lambda_k$ 。

c) 根据式(13)计算  $X, Y$  的第  $i$  对典型向量  $a_i, b_i$ 。

d) 根据  $\varphi_i = Xa_i, \phi_i = Yb_i$  计算  $X, Y$  的第  $i$  对典型变量和  $X, Y$  的典型相关系数,从大到小排列如下:  $\sqrt{\lambda_1} \geq \sqrt{\lambda_2} \geq \dots \geq \sqrt{\lambda_k}$ 。

### 3.2 基于 CCA 的图像特征融合

设  $X = [x_1, x_2, \dots, x_n]^T$ ,  $x_i$  是 256 维的颜色特征向量;  $Y = [y_1, y_2, \dots, y_n]^T$ ,  $y_i$  是 18 维的 LBP 特征向量,  $n = 1000$  表示样本的大小。经过 CCA 求解算法得到典型变量  $U = \varphi, V = \phi$ , 典型相关系数  $\lambda$ 。其中:  $\lambda$  是  $1 \times 17$  维的向量,  $U = [u_1, u_2, \dots, u_n]^T, V = [v_1, v_2, \dots, v_n]^T, u_i, v_i$  都是 17 维的向量。

对于给定的图像  $I$ , 提取其颜色特征  $x_I$ , LBP 特征  $y_I$ 。经过 CCA 处理后的特征记为  $u_I, v_I$ 。任意两特征向量  $z_i, z_j$  之间的距离即相似度定义为  $f(z_i, z_j)$ , 使用直方图相交距离计算相似度, 此处  $z = u, v$ 。对于得到的  $U, V$  可以用来检索出相关的图像。

#### 算法 3 基于 CCA 图像语义特征的检索算法。

Input:  $U$  为  $1000 \times 17$  的数据,  $V$  为  $1000 \times 17$  的数据;  $X$  为  $1000 \times 256$  的可伸缩颜色描述子颜色特征,  $Y$  为  $1000 \times 18$  的 LBP 特征。

Output: 检索出的相关图片, 查准率  $P_1, P_2$ 。

a) 对于给定的测试图像  $I$ , 提取颜色特征  $x_I$ , LBP 特征  $y_I$ 。使用 CCA 算法处理后的特征为  $u_I, v_I$ 。先使用  $U$  计算给定图像与图像库中图像的相似度  $f_1 = f(u_I, u_i)$ 。

b) 使用  $V$  计算给定图像与图像库中图像的相似度  $f_2 = f(v_I, v_i)$ 。

c) 计算  $F = \lambda_1 f_1 + \lambda_2 f_2$  作为最后的相似度值, 其中  $\lambda_1 + \lambda_2 = 1$ 。

d) 根据  $F$  值给出相似度最小的 16 幅图片。查询五次, 计算给定图像的查准率  $P_1$ 。

e) 将  $U, V$  换成  $X, Y, u_I, v_I$  换成  $x_I, y_I$ 。重复 a) ~ d), 计算  $F$  值, 给出新检索出的 16 幅图片。计算给定图像的查准率  $P_2$ 。

本文中, 原先直接使用颜色直方图、LBP 值两种特征来度量图像之间的相似度。经过 CCA 处理后, 可以直接使用经过 CCA 处理后的特征来度量图像之间的相似度。实验表明, 使用 CCA 后的效果比原先不使用 CCA 的效果提高了很多。

## 4 实验结果与分析

从 coral 图像库中选取 25 类图像, 每类 40 张, 共 1 000 张图像来检索。每类随机抽取五幅作为实例图像。每次验证查准率共完成  $25 \times 5$  次查询。

### 4.1 图像相似测度

检索首先就要比较图像的相似度。提取图像的特征后选择合适的相似性度量方法进行计算。例如, 给出一个查询图像实例, 根据查询实例的特征计算图像库中任一图像与查询图像的相似度, 然后给出相似度最大的前  $R$  个图像。

相似性度量方法有欧氏距离、余弦距离、直方图相交距离。欧氏距离是使用最广泛的相似性度量方法; 余弦距离就是两个向量之间夹角的余弦, 它不受坐标轴旋转、放大、缩小的影响; 直方图相交距离是一种特殊的 L1 距离, 通常用来对图像的纹理特征进行度量, 其简单快捷, 可以有效抑制图像噪声对检索结果的影响。本文使用直方图相交距离进行相似测度计算。

### 4.2 实验结果与分析

首先比较这五种描述图像特征的平均查准率。这五种描述图像的特征分别是颜色直方图、边缘方向直方图、边缘直方图、局部二值模式(LBP)、纹理距描述。选用五类图片来作比较, 分别是 horse 类、car 类、tiger 类、sun 类、flower 类。得到的查准率如图 3 所示。

从图 3 可以看出, 由于图像的彩色信息十分明显, 因此颜色直方图得到的效果最好。在每类中, 准确性都高于其他四种特征描述。同时可以看出, 最后一种情况使用纹理距描述来进行查询的效果非常差, 远远低于其他方法。分析该图还可以看出, 除了 car 类这一特殊情况外, 使用局部二值模式的效果比颜色直方图稍差一点, 但比其他特征的效果要好一些。可以看出, 五种情况的平均准确率并不高, 效果最好的颜色直方图的准确率才为 66.0000%, 而最差平均查准率仅为 11.5000%。为了解决这种不利的情况, 下面就提出解决这一问题的方法。

基于前面的分析, 直方图相交距离效果比较好, 基于此距离计算公式进行分析。在颜色直方图、边缘方向直方图、边缘直方图、局部二值模式(LBP)和纹理距离描述五种方法中, 由实验 2 可知, 颜色直方图与 LBP 的语义效果最好, 首先把这两种方法结合起来分析图像的语义特征。假设  $F = \lambda_1 f_1 + \lambda_2 f_2$ , 其中:  $f_i$  表示第  $i$  种语义特征描述方法,  $\lambda_i$  表示第  $i$  种语义特征描述方法的加权系数,  $F$  是混合语义特征。经过实验验证, 得到  $\lambda_1 = 0.7, \lambda_2 = 0.3$  取得的效果最好。

图 4 给出了在 horse、car、sun、plane、bear 五类图片上颜色直方图(COLORHIST)、LBP、直接使用这两种特征来混合提取特征(NCCA)、使用高维小样本下典型相关分析对提取的颜色直方图和 LBP 两种特征进行特征融合(CCA)四种情况下检索的查准率。

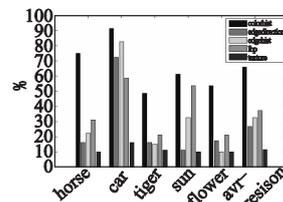


图3 五种特征参数查准率的比较

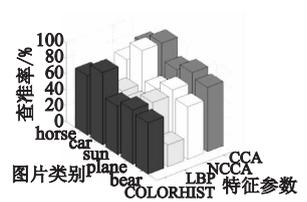


图4 COLORHIST、LBP、NCCA、CCA的比较

从图 4 中可以看出, 单独使用颜色直方图、LBP 的效果并不是太好; NCCA 的效果比单独使用一种特征参数的效果要好一点; 但是使用 CCA 对颜色直方图、LBP 两种特征参数进行特征融合后的效果比直接使用这两种特征混合提取语义特征的效果好了很多。

图 5 给出了使用 CCA 和不使用 CCA 两种情况的查准率。分析可知 CCA 比 NCCA 的准确性提高了很多, 具有较大的改善。图 5 在 horse、car、sun、plane、bear 五类图像中验证比较了使用 NCCA、CCA 方法的效果。分析可知, 对于 car 类的查询效果都比较好, 达到了 100%, 其他类使用 CCA 的查准率比不使用 CCA 的效果要好得多。对 plane、horse、bear、sun 四类图像的查准率都提高了 10% 以上, 使用 CCA 后的 horse、plane、bear 的查准率都超过了 80%。其中, Horse 类的准确性由 76.25% 提

高到了 85.0% ;plane 类的准确性由 75.0% 提高到了 87.5% ; bear 类的准确性由 66.25% 提高到了 81.25% ;sun 类的准确性由 63.75% 提高到了 78.75% 。

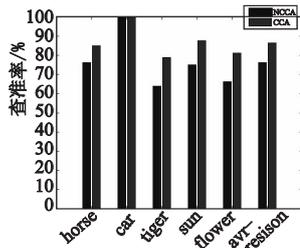


图5 NCCA与CCA的比较

图 6、7 给出了一次查询的效果。从图 6 中可以看出,在显示的 16 张图片中只查询出了 13 张相关图片,但是使用 CCA 后(如图 7 所示)查询出了 15 张相关的图片,准确率有了很大的提高。



图6 NCCA的检索结果



图7 CCA的检索结果

实验表明,使用高维小样本下典型相关分析对提取的颜色直方图、LBP 两种特征进行特征融合即 CCA 处理后的图像语义特征来查询,比不使用 CCA 的情况,查准率具有明显的改善和提高。

### 5 结束语

颜色和纹理是描述图像语义特征的重要方法。本文通过比较得出颜色直方图、LBP 特征参数在描述图像语义特征方面效果较好,提出了使用高维小样本下典型相关分析对这两种特征进行融合来描述图像的语义特征的算法,使查准率和查全率都有了较大提高。基于语义的图像检索是目前图像

检索发展的一个新的方向。这就需要对于给定的图像能够有效地提取图像的低层语义特征,建立低层特征与高层语义之间的关系,自动进行图像标注。进一步的工作是提出新的能够有效描述图像语义的算法,这是今后自动图像标注研究发展的趋势。

### 参考文献:

- [1] LIU Ying, ZHANG Deng-sheng, LU Guo-jun, *et al.* A survey of content-based image retrieval with high-level semantics [ J ]. *Pattern Recognition*,2007,40(1) : 262-282.
- [2] CHANG Jing-ying, LIAN C J, CHEN L G. Architecture and analysis of color structure and scalable color descriptor for real-time video indexing and retrieval [ C ]//Proc of the 5th Pacific Rim Conference on Advances in Multimedia Information Processing-Volume Part II . Berlin: Springer-Verlag. 2004:130-137.
- [3] RAHMAN S, NAIM S M, ALFAROOQ A, *et al.* Performance of MPEG-7 edge histogram descriptor in face recognition using principal component analysis[ C ]//Proc of the 13th International Conference on Computer and Information Technology. 2010:476-481.
- [4] LEVI K, WEISS Y. Learning object detection from a small number of examples the importance of good features[ C ]//Proc of IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society,2004:53-60.
- [5] ABDULLAH T N, OMAR K, NASRUDIN M F, *et al.* Enhancement of moment invariants calculation for Arabic handwriting recognition [ C ]//Proc of International Conference on Pattern Analysis and Intelligent Robotics. 2011:83-86.
- [6] XU Xian-chuan, ZHANG Qi. Medical image retrieval using local binary patterns with image Euclidean distance[ C ]//Proc of International Conference on Information Engineering and Computer Science. 2009:1-4.
- [7] LIAO Shu, CHUNG A C S. Texture classification by using advanced local binary patterns and spatial distribution of dominant patterns[ J ]. *IEEE Trans on Acoustics Speech and Signal Processing*,2007, 10(1) :1221-1224.
- [8] PEREIRA E T, GOMES H M, DeCARVALHO J M. Integral local binary patterns: a novel approach suitable for texture-based object detection tasks[ C ]//Proc of the 23rd SIBGRAPI Conference on Graphics, Patterns and Images. Washington DC: IEEE Computer Society, 2010: 201-208.
- [9] HAI Xian-wang. Local two-dimensional canonical correlation analysis [ J ]. *Signal Processing Letters*,2010,17(11) : 921-924.
- [10] AEAR E, ARSLAN S, YAZICI A, *et al.* Slim-tree and BitMatrix index structures in image retrieval system using MPEG-7 descriptors [ C ]//Proc of International Workshop on Content-Based Multimedia Indexing. 2008:402-409.
- [11] JEONG S, WONG C S, GRAY R M. Image retrieval using color histograms generated by Gauss mixture vector quantization[ J ]. *Computer Vision and Image Understanding*,2004,94 (1-3) :44-66.
- [12] ANDERSON T W. 多元统计分析导论[ M ]. 张润楚,译. 3 版. 北京:人民邮电出版社,2010:367-396.
- [13] CHEN Xiang-tao, CHEN Yu-juan. An improved edge detection in noisy image using fuzzy enhancement[ C ]//Proc of International Conference on Biomedical Engineering and Computer Science. 2010:1-4.