

## 视觉词汇的主成分线性编码方法\*

艾浩军<sup>1,2</sup>, 张敏<sup>1,2</sup>, 方禹<sup>1,2</sup>, 赵梦蕾<sup>1,2</sup>, 李泰舟<sup>1,2</sup>, 王红霞<sup>3</sup>

<sup>1</sup>(武汉大学 计算机学院, 湖北 武汉 430072)

<sup>2</sup>(武汉大学 国家多媒体软件工程技术研究中心, 湖北 武汉 430072)

<sup>3</sup>(武汉理工大学 计算机科学与技术学院, 湖北 武汉 430063)

通讯作者: 艾浩军, E-mail: aihj@whu.edu.cn, http://www.whu.edu.cn

**摘要:** 针对视觉物体分类中视觉词汇局部线性编码缺少显著性检验和共线性分析的问题, 提出了主成分线性编码方法, 选择与特征点具有最强线性相关性的  $K$  近邻视觉单词, 采用主成分多元线性回归方法以解决视觉单词的共线性问题, 从而减小编码系数的偏差和不稳定, 提高视觉物体分类的精度. 依据图像量化结果的稀疏性是影响分类精度的重要因素, 进一步对主成分线性编码得到的量化结果做稀疏性分析并进行能量正则化处理, 提高分类效率. 实验结果表明, 与已有方法相比, 平均分类正确率提高了 1% 以上.

**关键词:** 视觉词袋; 共线性; 主成分回归; 特征点归并; 能量正则化

中文引用格式: 艾浩军, 张敏, 方禹, 赵梦蕾, 李泰舟, 王红霞. 视觉词汇的主成分线性编码方法. 软件学报, 2013, 24(Suppl. (2)): 42-49. <http://www.jos.org.cn/1000-9825/13022.htm>

英文引用格式: Ai HJ, Zhang M, Fang Y, Zhao ML, Li TZ, Wang HX. Principal component linear coding for visual words. Ruan Jian Xue Bao/Journal of Software, 2013, 24(Suppl. (2)): 42-49 (in Chinese). <http://www.jos.org.cn/1000-9825/13022.htm>

### Principal Component Linear Coding for Visual Words

AI Hao-Jun<sup>1,2</sup>, ZHANG Min<sup>1,2</sup>, FANG Yu<sup>1,2</sup>, ZHAO Meng-Lei<sup>1,2</sup>, LI Tai-Zhou<sup>1,2</sup>, WANG Hong-Xia<sup>3</sup>

<sup>1</sup>(School of Computer Science, Wuhan University, Wuhan 430072, China)

<sup>2</sup>(National Engineering Research Center for Multimedia Software, Wuhan University, Wuhan 430072, China)

<sup>3</sup>(School of Computer Science and Technology, Wuhan University of Technology, Wuhan 430063, China)

Corresponding author: AI Hao-Jun, E-mail: aihj@whu.edu.cn, <http://www.whu.edu.cn>

**Abstract:** By means of significant test and co-linearity analysis, this paper proposes principal component linear encoding which selects the  $K$ -nearest neighbor visual word with the strongest linear correlation. The multiple linear regression method based on principal component is used to solve weak and instable coding caused by the visual words' co-linearity problem, improving the accuracy of the visual object classification effectively. Recognizing that the scarcity of the image quantify plays an important roles in the classification accuracy, the study analyzes the scarcity of the quantitative results obtained by the principal component linear encoding and then processes it with energy regularization to improve the classification efficiency further. The experimental results demonstrate that this method increases the recognition rate average over 1% than existing algorithms.

**Key words:** bag of visual words; co-linearity; principal component regression; feature points merging; energy regularization

Leung 在 2001 年将文本信息检索中词袋 (bag-of-words, 简称 BoW) 模型<sup>[1,2]</sup>应用于基于纹理特征的图片识别中. 由此而产生的视觉词袋 (bag-of-visual-words, 简称 BoV) 模型作为基于码本的图像分类方法, 是目前应用最广泛的视觉物体分类框架之一. 码本又称为视觉词汇表, 是包含所有图像类别局部特征的紧致集合. 视觉词汇编码研究以视觉词汇的组合表达物体局部特征的最佳方法, 是影响视觉分类的基本要素. 研究表明, 稀疏编码空间金

\* 基金项目: 国家科技支撑计划 (2012BAH35B03)

收稿时间: 2012-06-15; 定稿时间: 2013-07-22

字塔模型 ScSPM<sup>[3]</sup>中稀疏性权重分配方法在线性 SVM 分类器中表现很好,甚至优于使用传统权重分配的非线性核 SVM 分类器.局部限制性编码方法 LLC<sup>[4]</sup>更进一步地表明局部性在保留稀疏性的同时更接近编码的本质要求.

基于视觉词汇的线性编码方法得到的图像量化结果具备线性可分性,也就是说,该方法与线性核 SVM 分类器结合也可以获得优异的分类正确率.本文提出一种主成分线性编码方法,解决了视觉单词的共线性问题,减小编码系数的偏差和不稳定性,从而提高了视觉物体分类的精度.

## 1 局部特征点的线性编码

李氏导数平滑性是分类函数的一个重要特性,可以描述局部特征点编码误差对视觉物体分类的影响.若函数  $f(x)$  满足李氏导数平滑,则对于  $\forall x, x' \in R^d$ , 满足:

$$|f(x) - f(x') - \nabla f(x')(x - x')| \leq \frac{\beta}{2} \|x - x'\|^2 \quad (1)$$

其中,  $\beta$  为常数.

线性核的 SVM 分类器可看作多个线性分类器的求和叠加,因而满足李氏倒数平滑的条件,那么  $f(x)$  可使用  $f(x') + \nabla f(x')(x - x')$  逼近,其中  $x'$  为局部特征点  $x$  的量化表示.这表明图像分类精度受局部特征点编码误差控制,因而需要解决在图像分类中局部特征点编码误差的最小化这一关键问题.

局部特征点的线性编码采用以特征点为因变量,以  $K$  近邻视觉单词为自变量的多元线性回归模型.基于视觉词汇表  $CB = \{C_i, i = 1, \dots, N\}$ , 图像局部特征点  $X$  线性编码问题的数学模型如式(2)所示,其中,  $\beta_i$  为量化的权重系数,  $\varepsilon$  是不可观测的随机变量.

$$\begin{cases} X = \sum_{i=1}^N \beta_i C_i + \varepsilon \\ \varepsilon \sim N(0, \sigma^2) \\ \sum_{i=1}^N |\beta_i| = 1 \end{cases} \quad (2)$$

记  $C = [C_i] (i = 1, \dots, N)$ ,  $\beta = [\beta_i] (i = 1, \dots, N)$ , 则模型的矩阵形式描述为

$$\begin{cases} X = \beta^T C + \varepsilon \\ \varepsilon \sim N(0, \sigma^2) \end{cases} \quad (3)$$

训练图片的局部特征点使用  $K$ -均值方法聚类,聚类中心即作为视觉单词词汇表.各划分中向量的均值是局部特征的线性组合,因而  $K$ -均值方法求得的视觉词汇表是线性的.也就是说,在线性编码模型中,视觉单词词汇表  $CB$  为图像局部特征点空间  $\{X\}$  的线性基底.因此,称这类视觉词汇表是局部特征点的线性视觉词汇表.

## 2 线性编码的显著性检验

局部特征点与视觉单词的线性相关系数可以在全局尺度上作为  $K$  近邻线性编码的回归自变量,在实际应用中,仍然需要进一步检验局部特征点  $X$  与视觉单词  $C_i$  是否确有线性关系.在多元线性回归模型中,采用假设检验的方法对线性编码中自变量的选择和回归做显著性检验.

### 2.1 线性编码中视觉单词的选择

在视觉词袋模型的局部特征点的权重分配过程中,传统的硬直方图<sup>[5]</sup>分配方法选择距离局部特征点最近的视觉单词并分配权重 1,其余分配权重 0;核函数的权重分配方法<sup>[6]</sup>分别选择  $K$  近邻和全部视觉单词并使用核函数分配权重;局部线性编码<sup>[7]</sup>中选择  $K$  近邻为自变量做多元线性回归,得到的线性系数作为分配权重.依据流形学习理论,特征点局部流形同胚于欧式空间, $K$  近邻点对局部特征点具有更强的线性相关性.在基于特征点的线性编码的视觉词袋模型中,需为局部特征选择与其具有高度线性相关性的视觉单词进行编码.通过以上的定性分析,可初步判断  $K$  近邻视觉单词对特征点具有显著的线性影响.

线性相关系数描述了向量的线性相关程度,在计算局部特征点  $X$  与视觉单词  $C_i$  的线性关系中,以  $X$  为应变量,  $C_i$  为自变量建立一元线性回归模型,如式(4)所示.

$$X = \beta_0 + \beta_1 C_i \quad (4)$$

在一元线性回归模型中,线性相关系数  $r$  定义为

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \quad (5)$$

当  $r=0$  时,  $\beta_1=0$ , 这说明  $X$  和  $C_i$  之间不存在线性关系;绝大多数情况下  $0 < |r| < 1$ , 此时,  $X$  和  $C_i$  之间存在一定的线性关系,当  $r < 0$  时称  $X$  与  $C_i$  负相关,当  $r > 0$  时称  $X$  与  $C_i$  正相关,  $r$  的绝对值越大表明线性相关度越强.

为了探寻局部特征点与视觉单词的距离和线性相关系数之间的关系,对所有视觉单词按其局部特征点的欧式距离升序排序,并计算两者的线性相关系数.按照距离的大小排序后,视觉单词与局部特征点的线性相关系数如图 1 所示.

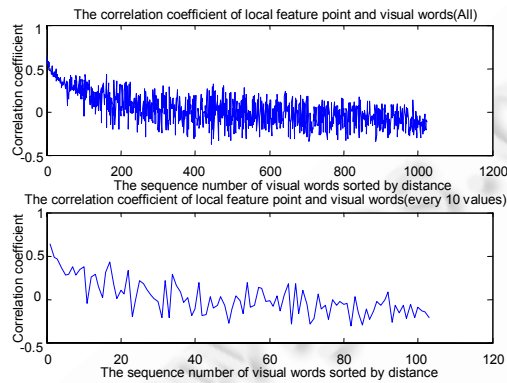


Fig.1 Scattergram of local feature point against linear correlation coefficient of visual words

图 1 局部特征点与视觉单词线性相关系数分布图

系数如图 1 所示.

从图 1 中可得到以下结论:(1) 相关系数值随距离的增加呈下降趋势,在距离较近时相关系数值下降较快,在 200 个视觉单词以后趋于平缓,距离近的视觉单词与局部特征点有较大的正相关系数;(2) 距离较近的视觉单词与局部特征点的相关系数大多为正数,而距离较远的相关系数大多为负数,这说明近邻视觉单词与局部特征点正相关,而远邻视觉单词与局部特征点负相关;(3) 距离远的视觉单词与局部特征点的线性相关系数绝对值较小,说明其线性相关程度比较低.由此可以认为选择  $K$  近邻视觉单词作为局部特征点线性回归模型的自变量是合理且必须的.

## 2.2 回归模型的显著性检验

### 2.2.1 回归模型的显著性检验

在以  $K$  近邻视觉单词  $C_1, \dots, C_i, \dots, C_K$  为自变量,局部特征点  $X$  为应变量的多元线性回归模型中,若  $X$  与  $C_1, \dots, C_i, \dots, C_K$  并不存在线性关系,那么模型中所有的回归系数  $\beta_i$  均为 0.因而为检验整个回归模型是否显著,做出如下假设  $H_0$ :

$$H_0: \beta_1 = \dots = \beta_K = 0 \quad (6)$$

假设  $H_0$  成立,为寻求合适的统计量以检验假设,这里先介绍多元线性回归模型中的平方和分解公式,如式(7)所示.其中  $S_T = \sum (x_i - \bar{x})^2$  为总偏差平方和,  $S_e = \sum (x_i - \hat{x}_i)^2$  为残差平方和,  $S_R = \sum (x_i - \bar{x})^2$  为回归平方和.

$$S_T = S_e + S_R \quad (7)$$

当假设  $H_0$  成立时,可证明  $\frac{S_R}{\sigma^2} \sim \chi^2(m)$ ,  $\frac{S_e}{\sigma^2} \sim \chi^2(n-m-1)$ , 从而统计量  $F = \frac{S_R/m}{S_e/(n-m-1)} \sim F(m, n-m-1)$ , 其中,  $n$  为自变量维数,  $m$  为自变量个数.  $F$  为选择的用于检验假设  $H_0$  的统计量,在给定的显著水平  $\alpha$  下,若  $F \geq F_{1-\alpha}(m, n-m-1)$ , 则拒绝  $H_0$ , 即认为回归效果显著,否则接受  $H_0$ , 即认为回归效果不显著.

### 2.2.2 单个回归系数的显著性检验

若整个线性回归模型的检验显著,仅仅说明回归系数  $\beta_1, \dots, \beta_i, \dots, \beta_K$  不全为 0.在这种情况下,可能存在其中某一个或者几个回归系数  $\beta_i = 0$ , 这表明局部特征点  $X$  与视觉单词  $C_i$  无显著线性关系,或者与其他视觉单词

$C_j (j \neq i)$  存在共线性关系,导致该变量可从回归方程中删除.因而在整个模型回归效果显著时,仍需逐个检查每个视觉单词  $C_i$  对局部特征点  $X$  是否具有显著的线性影响.这相当于做如下假设检验:

$$H_{0i}: \beta_i = 0, i = 1, 2, \dots, m \quad (8)$$

由  $\hat{\beta}_i \sim N(\beta_i, c_{ii}\sigma^2)$ ,  $\frac{S_e}{\sigma^2} \sim \chi^2(n-m-1)$ , 选择统计量  $t_i = \frac{\hat{\beta}_i}{S\sqrt{c_{ii}}} \sim t(n-m-1)$  以检验假设  $H_{0i}$ , 其中  $n$  为自变量维数,  $m$  为自变量个数.值得注意的是,统计量  $t_i$  的值已经消除了单位的影响,是一个无量纲的量.在给定的显著水平  $\alpha$  下,若由样本值算的  $|t_i| \geq t_{1-\alpha/2}(n-m-1)$ , 则拒绝  $H_{0i}$ , 即认为视觉单词  $C_i$  对局部特征点  $X$  具有显著线性影响;否则接受  $H_{0i}$ , 认为视觉单词  $C_i$  对局部特征点  $X$  无显著线性影响.

在实验中,对若干局部特征点与其  $K$  近邻视觉单词的多元线性回归模型做显著性检验,置信度  $\alpha$  设置为 95%, 经过分析得出基本一致的结论:整个多元线性回归模型的回归效果显著,但一些单个回归系数的回归效果不显著.不显著的回归系数对应的视觉单词被分配较小的权重系数.少数情况下,权重系数甚至为负数.表 1 为一组回归模型的显著性检验结果.

Table 1 Significant test of the linear regression model

表 1 线性回归模型的显著性检验

$T$ 值	回归系数	是否显著	$F$ 值	是否显著
2.99	0.392	是	21.14	是
1.02	0.146	否		
3.17	0.338	是		
1.16	0.162	否		
0.59	0.089	否		

视觉单词  $C_i$  对特征点  $X$  无显著线性影响产生的可能原因有两个:(1)  $C_i$  与  $X$  无线性关系;(2)  $C_i$  对  $X$  的线性影响被其他视觉单词替代了.基于线性相关系数的分析可知,  $K$  近邻视觉单词对局部特征点有较强的线性影响,所以可以排除第 1 个原因.基于以上分析,判断视觉单词之间由于存在共线性问题导致部分视觉单词对特征点的线性影响被其他视觉单词替代,以至于单个回归系数假设检验不显著,从而影响分类结果.

### 3 主成分线性编码

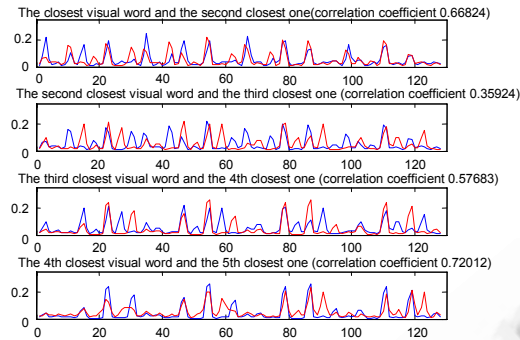
在多元线性回归模型的参数估计方法中,通常使用最小二乘法求解,最小二乘法具有无偏性、均方差最小等优良特性,但当自变量存在共线性时,最小二乘估计的性质会变坏.当增加或删除一个自变量时,回归系数变化很大,甚至改变符号,某些回归系数的正、负号与实际情况相反;某些回归系数的绝对值特别大.在局部特征点的局部线性编码模型中,视觉单词的共线性也会导致同样的问题.

#### 3.1 视觉单词的共线性分析

如前文所述,在多元线性回归模型中,某些单个系数回归效果不显著,并且已经排除了线性无关的可能性,也就是说,  $K$  近邻视觉单词之间存在较强的共线性.共线性指的是,若存在常数  $a_0, a_1, \dots, a_m (m \geq 2)$  使得  $a_1x_1 + a_2x_2 + \dots + a_mx_m \approx a_0$  成立,则称自变量  $x_1, x_2, \dots, x_m$  是多重共线性的或共线性的.也可以表示为,至少存在 1 个  $x_i$  可

以由其他变量线性表示,即  $x_i \approx \left( a_0 - \sum_{j \neq i} a_j x_j \right) / a_i (a_i \neq 0)$ .

可以通过线性相关系数来分析视觉单词间的共线性问题.图 2 描述了第  $i$  近邻视觉单词与第  $i+1$  近邻值的分布,相邻视觉单词分布非常接近,而且具有最大 0.72 之高的线性相关系数.表 2 为一组  $K$  近邻视觉单词间相关系数的详细数据,数据进一步表明了视觉单词之间存在高度线性相关性,因而导致模型中某些单个回归系数不显著.

Fig.2 The correlation of  $K$  neighbours visual words图2  $K$  近邻视觉单词的相关性Table 2 The correlation coefficient of  $K$  neighbours表2  $K$  近邻的相关系数

	近邻 1	近邻 2	近邻 3	近邻 4	近邻 5
近邻 1	1.00	0.36	0.67	0.48	0.49
近邻 2	0.36	1.00	0.04	0.19	0.34
近邻 3	0.67	0.04	1.00	0.35	0.34
近邻 4	0.48	0.19	0.35	1.00	0.22
近邻 5	0.49	0.34	0.34	0.22	1.00

### 3.2 共线性问题解决

多元线性回归模型参数的最小二乘估计为  $\tilde{\beta} = (C'C)^{-1}C'X$ , 理论上可以证明  $\tilde{\beta}$  是  $\beta$  的最小方差无偏估计, 即在所有无偏估计中拥有最小方差的无偏参数估计. 在局部特征点的线性编码过程中, 由于视觉单词间的共线性导致模型参数估计结果非常不稳定, 相似的局部特征的编码结果经常出现较大的差别, 致使图像的量化过程不可控并导致分类精度降低.

在局部特征点线性编码的实验中, 采用岭回归方法(ridge regression)<sup>[8]</sup>解决视觉单词共线性问题的效果并不佳. Massy 在 1965 年提出了主成分回归(principal component regression)<sup>[9]</sup>, 在上述诸方面近似于岭估计, 但处理数据的方法不一样.

局部特征点主成分线性编码方法的计算步骤描述如下:

#### (1) 计算 $K$ 近邻视觉单词的主成分

采用主成分分析的方法, 计算得到  $K$  近邻视觉单词  $\{C_i\} (i=1, \dots, K)$  的主成分向量  $\{Z_j\} (j=1, \dots, K)$ ,  $Z_j = \sum P_{ij}C_i$ . 按照方差大小对主成分变量  $Z_j$  进行排序, 以一定比例或者既定个数去掉方差太小的  $Z_j$ , 留下方差较大的  $Z_j$  作为主成分向量. 若记  $C = [C_i]$ ,  $Z = [Z_j]$ ,  $P = \{P_{ij}\}$ , 则主成分分析的矩阵形式为

$$Z = PC \quad (9)$$

#### (2) 计算主成分回归系数

主成分分析方法求得  $K$  近邻视觉单词的主成分向量两两正交, 其相互之间没有任何共线性. 因而以主成分向量为自变量, 以局部特征点为应变量, 采用最小二乘方法求得的主成分回归系数是稳定的.

$$\tilde{\alpha} = (Z'Z)^{-1}Z'X = \lambda Z'X \quad (10)$$

#### (3) 计算 $K$ 近邻视觉单词的回归系数

在计算得到主成分回归系数后, 根据主成分转化矩阵  $P$  还原得到原自变量的回归系数, 即

$$\tilde{\beta} = P\tilde{\alpha} \quad (11)$$

称  $\tilde{\beta}$  为原模型的主成分参数估计.

#### 4 特征点归并与能量正则化

特征点归并方法和能量正则化处理是影响视觉物体分类精度的两个重要因素.本节对这两种方法进行阐述和分析,并基于本文提出的主成分编码方法和视觉单词形状描述模型使用这两种方法以提升分类效率.

##### 4.1 特征点归并方法

特征点归并方法是指在计算图像特征点的编码结果后,如何以特征点编码结果进行表示.常用的特征点归并方法包括求和归并、最大值归并、平方和平均归并、绝对值平均归并等,其中求和归并和最大值归并使用最为广泛.求和归并将图像区域内特征点的编码结果做向量加法,将计算得到的和向量作为图像的量化结果.最大值归并则是将图像区域内所有特征点的编码结果按维数取最大值得到的向量作为该区域的量化结果.在视觉词袋模型系列方法中,多数图像量化方法采用求和的归并方法.这种归并方法描述了图像上包含各视觉单词的个数,本身具备较强的物理意义.然而在 LLC 方法中,最大值归并方法量化结果的分正确率非常高,其中一个重要的原因是最大值归并方法的量化结果分布不稀疏,所以在 SVM 分类其中具有很好的效果.

##### 4.2 编码正则化

在图像的量化过程中,编码系数的分配往往集中于少数的视觉单词,从而导致图像的量化结果为稀疏向量,其示意图如图 3(c)所示.稀疏向量以欧式距离为相似性度量的分类效果很不理想,因而在图像量化结果的稀疏性是影响分类正确率的重要因素.能量正则化方法对图像量化结果做如式(12)的处理,处理后的结果如图 3(d)所示.

$$f(z) = \text{sign}(z) \times |z|^\alpha \quad (12)$$

实验结果表明,能量正则化方法降低了图像量化结果的稀疏性,从而提高了视觉物体分类的精度.在 LLC 方法中,求和归并方法主要是因为编码得到的结果比最大值归并方法更为稀疏,导致其分类精度较低.采用如上所述的能量正则化处理,求和归并方法的分类精度与最大值归并方法相当,在某些图像库上甚至好于最大值归并方法.

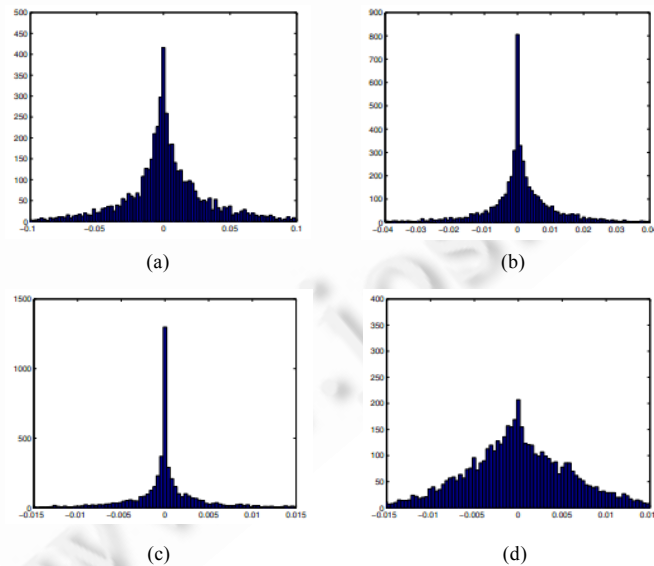


Fig.3 Energy regularization process

图3 能量正则化处理

## 5 实验结果与分析

Caltech-101 图像分类中常用的通用图像数据库如图 4 所示,包括飞机(airplane)、汽车(car)、人脸(face)、摩托车,每类图片 400~700 张图片不等,共 2 156 张图片.文献[10]实验中所使用的图像数据库,本文使用同样的图像数据库以便进行实验结果的比较.图 4 给出了部分实验样本.

局部特征点的量化结果使用 SPM 算法以包括局部特征点的空间位置信息.SPM 这 3 层的子区域个数分别为  $1 \times 1, 2 \times 2, 4 \times 4$ .实验中分别在 3 种归并方法下比较主成分线性编码(以下称 PLC)和局部线性编码方法(以下称 LLC)的分类识别效果,3 种归并方法包括和归并、最大值归并、正则和归并方法.分类结果见表 3.



Fig.4 Part of images in image library

图 4 图像库部分图片

Table 3 Experimental results of image library

表 3 图像库的实验结果

方法	和归并		最大值归并		正则和归并		Wang <sup>[10]</sup>
	LLC	PLC	LLC	PLC	LLC	PLC	
飞机	97.9	98.1	99.1	99.5	99.1	99.4	98.0
汽车	88.6	89.5	99.5	99.5	99.8	98.9	99.5
人脸	87.3	88.3	96.6	96.8	96.5	96.6	94.7
摩托车	98.0	97.7	99.4	99.6	99.6	99.3	96.0
平均	92.2	<b>93.7</b>	98.7	<b>98.8</b>	98.7	<b>98.8</b>	97.7

可以看出,和归并方式下,PLC 方法比 LLC 方法高出 1.5%,除摩托车外,其他 3 类物体的识别率都有提高;在最大值和正则和归并的方式下,在识别率已高达 98%的前提下,PLC 识别率比 LLC 仍有 0.1%的提高.

我们将本文算法与近年发表的相关结果进行了比较,见表 4.本文算法在使用基本局部特征点-网格划分和 SIFT 描述算子时,识别结果优于 Csurka<sup>[5]</sup>等人的算法.

Table 4 Comparison among the experimental results of different algorithms

表 4 本文算法与其他方法的实验结果比较

方法	本文算法	Wang <sup>[10]</sup>	Csurka <sup>[5]</sup>	Oplet	Fergus	Sivic <sup>[7]</sup>
飞机	99.5	98.0	96.3	88.9	90.2	95.3
汽车	99.5	99.5	97.7	91.1	90.0	98.1
人脸	96.8	94.7	92.7	93.5	96.4	94.0
摩托车	99.6	96.0	94	92.2	92.5	83.6
平均	<b>98.8</b>	97.7	96.6	89.7	-	-

总结上述实验结果,本文的算法具有很好的识别率,平均能够提高 1%以上.视觉词袋基本模型使用求和的归并方法,其物理意义非常显著,即这意味着该图像区域中包含各类特征点的数量.在求和的归并方法中,本文的算法明显好于一般的局部线性编码方法.这表明在局部特征点量化过程中,本文提出的主成分线性编码方法通过去除共线性,求解得到的系数更加有效和稳定.



## 6 结 论

本文针对视觉物体分类的词袋模型中的词汇量化问题,以线性回归模型的显著性检验和共线性分析为基础,提出了主成分线性编码方法,减小了词汇编码误差,然后比较编码输出的多种归并方法,通过分析编码结果的稀疏性对分类结果的影响,使用能量正则化方法提高了视觉物体分类的正确率。

### References:

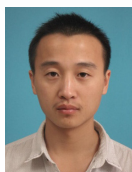
- [1] Joachims T. Text categorization with support vector machines: Learning with many relevant features. In: Nédellec C, Rouveirol C, eds. Proc. of the Machine Learning: ECML'98. Heidelberg: Springer-Verlag, 1998. 137–142.
- [2] Tong S, Koller D. Support vector machine active learning with applications to text classification. The Journal of Machine Learning Research, 2002,2:45–66.
- [3] Yang JC, Yu K, Gong YH, Huang T. Linear spatial pyramid matching using sparse coding for image classification. In: Proc. of the Computer Vision and Pattern Recognition. Miami, 2009. 1794–1801.
- [4] Wang JJ, Yang JC, Yu K, Lü FJ, Huang T, Gong YH. Locality-Constrained linear coding for image classification. In: Proc. of the Computer Vision and Pattern Recognition. San Francisco, 2010. 3360–3367.
- [5] Csurka G, Bray C, Dance C, Fan L. Visual categorization with bags of keypoints. In: Proc. of the Workshop on statistical learning in computer vision, ECCV. Prague, 2004. 1–22.
- [6] Van Gemert JC, Geusebroek JM, Veenman C, Smeulders AWM. Kernel codebooks for scene categorization. In: Forsyth D, Torr P, Zisserman A, eds. Proc. of the Computer Vision—ECCV 2008. Berlin, Heidelberg: Springer-Verlag, 2008. 696–709.
- [7] J. Sivic, B. C. Russell, A. A. Efros, A. Zisserman, and W. T. Freeman. Discovering objects and their location in images. In: Proc. of the Computer Vision and Pattern Recognition. 2005. 370–377.
- [8] Hoerl AE, Kennard RW. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 1970,12(1):55–67.
- [9] Massy WF. Principal components regression in exploratory statistical research. Journal of the American Statistical Association, 1965,60(309):234–256.
- [10] Wang YJ. Expressive methods of visual object based on remarkable local feature [Ph.D. Thesis]. Beijing: Beijing University of Technology, 2010 (in Chinese with English abstract).

### 附中文参考文献:

- [10] 王彦杰. 基于显著局部特征的视觉物体表示方法[博士学位论文]. 北京: 北京理工大学, 2010.



艾浩军(1972—),男,湖北汉川人,博士,副教授,CCF 高级会员,主要研究领域为多媒体信号处理与分析,室内定位。  
E-mail: aihj@whu.edu.cn



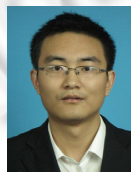
张敏(1985—),男,硕士,主要研究领域为模式识别。  
E-mail: 109484986@qq.com



方禹(1988—),男,硕士,主要研究领域为室内定位。  
E-mail: 378407802@qq.com



赵梦蕾(1988—),女,硕士,主要研究领域为室内定位。  
E-mail: 529430398@qq.com



李泰舟(1990—),男,硕士生,主要研究领域为室内定位。  
E-mail: litaizhou30@gmail.com



王红霞(1977—),女,博士,讲师,主要研究领域为图像处理,模式识别。  
E-mail: whx\_green@163.com