



引用格式:郭延辉,尹西杰,张宏.一种用于图像分类的局部二值改进算法[J].轻工学报,2017,32(3):73-77.

中图分类号:TP391.41 文献标识码:A

DOI:10.3969/j.issn.2096-1553.2017.3.012

文章编号:2096-1553(2017)03-0073-05

一种用于图像分类的局部二值改进算法

An improved local binary algorithm for image categorization

郭延辉,尹西杰,张宏

GUO Yan-hui, YIN Xi-jie, ZHANG Hong

关键词:

图像分类;局部二值描述符;汉明距离;特征提取;多尺度融合

山东女子学院 信息技术学院,山东 济南 250300

School of Information Technology, Shandong Women's University, Ji'nan 250300, China

Key words:

image categorization;
local binary descriptor;
Hamming distance;
feature extraction;
multiscale fusion

摘要:针对LBP特征提取过程中聚类耗时、占用内存较大的问题,提出了一种用于图像分类的局部二值改进算法。该算法用二进制描述符替代LBP中的十进制表示、用汉明距离替代欧氏距离进行特征聚类,同时将不同尺度的LBP特征进行融合,实现了多尺度图像局部二值描述。将提出的改进算法,尤其是多尺度融合特征,在PASCAL VOC 2007数据库与经典LBP方法进行对比分析,实验结果表明,本算法正确率更高,运行效率也有很大提高。

收稿日期:2016-11-01

基金项目:山东省高等学校科技计划项目(J14LN54)

作者简介:郭延辉(1982—),男,山东省临清市人,山东女子学院讲师,博士,主要研究方向为机器学习和模式识别。

Abstract: In the extraction process of LBP features, most consumption of time and memory were paid for clustering. In order to address these problems, an improved local binary algorithm for image categorization was proposed. The algorithm replaced decimal system encoding LBP with binary descriptor. Meanwhile, Hamming distance was employed rather than Euclidean metric for features clustering. The multi-scale LBP features was flued for a new local binary descriptor. The result of the experiment on the PASCAL VOC 2007 dataset showed that the adopted local binary descriptor was better than the classical LBP, specifically for time consumption.

0 引言

随着数字产品的不断普及,大量图片爆炸式地增长.这些图片的有效分类,能帮助人们快速获取高质量的信息.尽管图像分类在最近几年获得很大进展,但仍是计算机视觉领域很有挑战性的一个难题.图像分类方法一般由3步构成,依次为特征提取、特征选择、利用分类算法实现特征分类.在第一步中,比较流行的方法有尺度不变特性变换(SIFT)、方向梯度直方图(HOG)、局部二值模式(LBP)、卷积神经网络(CNN)等^[1-9];在第二步,最常用的方法是特征袋 BoF(Bag of Features);在第三步,将选取的特征通过分类器,如支持向量机(SVM),Soft-max 等^[10-11]实现图像分类.

LBP 被广泛应用在计算机视觉领域,与 SIFT 和 HOG 相比,由于局部二值法用二值表示,速度更快、内存空间更少,因此,在图像匹配中具有较好的性能.然而,LBP 在分类过程中,需要将二值数据转换成十进制数生成直方图,这严重影响了 LBP 的计算效率.为了解决以上问题,本文拟改进二值描述符,并使用汉明距离^[12-13]编码 LBP 特征,以得到高效的分类效果.

1 预备知识

通常,局部二值描述通过像素的邻接像素获得 0 或 1 值;然后,将加权之和转换成十进制数;最后,采用 BoF 模型编码这些二值描述.该方法分两个步骤:一是视觉词汇构建,二是直方图绘制.其中,距离测量在这两个步骤中起重要

作用,不但要真实地反映两元素之间的远近,还要方便计算.本文采用汉明距离进行距离测量.

1.1 汉明距离

汉明距离最早被应用在通信领域,主要用于固定长度的二进制字符串的校对工作. F_n 的汉明距离 $d(x, y)$ 是两个长度为 n 的字符串 x, y 对应位不同数字的个数.如在 F_4 中, $d(1001, 1100) = 3, d(0101, 0111) = 1$.

局部二值描述是从图像中抽取的维度相同的 0, 1 值.考虑到计算效率,汉明距离更适合 BoF 模型.在计算机中,通过异或操作能够快速实现距离计算.

1.2 汉明距离编码 BoF

使用汉明距离编码 BoF 主要是为了解决直方图维度快速增长的问题.汉明距离也可以通过 K -mean 算法等实现聚类.假设 $x \in \{0, 1\}^d$, d 为二进制字符串的长度.将 n 个观察值 $\{x_1, x_2, \dots, x_n\}$ 划分到 k 个集合 (S_1, S_2, \dots, S_k) 中,使分配到 k 个集合中的总误差最小,总误差公式为

$$\arg \min \sum_{i=1}^k \sum_{x_j \in S_i} |x_j - \mu_i|_{HD}$$

通过求解上式,聚类结果得到所有的质心,即视觉单词 $(\mu_1, \mu_2, \dots, \mu_k)$.假设从图片中抽取的一组描述为 (x_1, x_2, \dots, x_n) ,每个描述符 x_i 被分配到相应的视觉单词 μ_i 中:

$$\arg \min_{i=1 \dots k} |x_j - \mu_i|_{HD}$$

直方图是一个由这些局部描述符组成的 k 维非负向量 F .

BoF 模型的性能主要取决于 K -mean 聚类.在聚类过程中,距离测量方法对聚类结果至关

重要^[7],也会对存储空间和计算效率产生重要影响. 汉明距离和欧式距离聚类结果如图 1 所示,其中,网格点代表局部二值描述符,· 代表获得的质心.

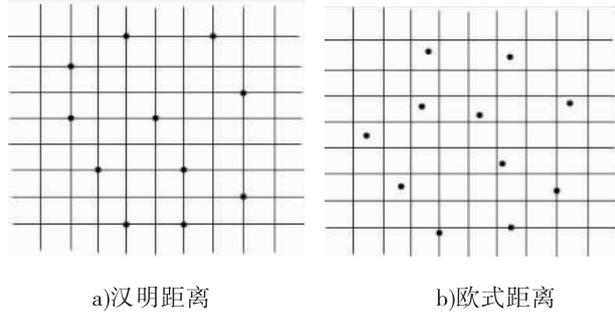


图 1 汉明距离和欧式距离聚类结果

Fig. 1 The clustering result of Hamming distance and Eudidean distance

由图 1 可见,汉明距离聚类结果都落在网格上,而欧式距离聚类结果大多落在网格内,因此汉明距离更适合局部二值描述.

2 图像分类框架

利用局部二值模式进行图像分类的主要步骤一般为二值特征抽取→利用聚类方法建模特征袋→根据特征袋计算图像的特征向量→用分类器实现图像分类. 具体过程如图 2 所示.

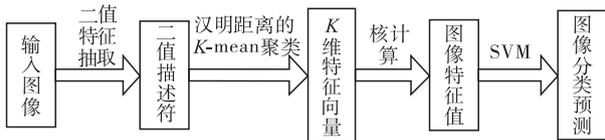


图 2 图像分类的过程与步骤

Fig. 2 The process and steps of image classification

2.1 局部二值特征提取

在进行像素特征值提取时,LBP 采用不同的半径,选取不同的邻居个数. 相应地,点 (x_c, y_c) 的 LBP 特征值定义如下:

$$\tau(g_c, g_p) = \begin{cases} 1 & g_c < g_p \\ 0 & \text{否则} \end{cases}$$

其中, g_p 是邻接像素的灰度值, g_c 是中心像素

值. 按照不同的邻接像素的半径 r , 一般会选择不同的像素点个数 p . 本文提取 1 个像素的 p 维 LBP 特征值为

$$f_p(g_c) = \sum_{1 \leq i \leq p} 2^{\tau(g_c, g_p)}$$

最后,求出所有像素的 p 维 LBP 特征值.

2.2 使用汉明距离的 BoF 模型

特征提取后,一幅图像被表示成一个 LBP 特征值的集合. 本文不需要计算十进制数的特征值,而是直接使用二值描述符描述特征值(0 或 1). 然后,用汉明距离对特征值进行聚类,得到每幅图片的 BoF 特征向量 F .

2.3 图像分类

获取输入图像的 BoF 描述之后,本文通过 SVM 对图像进行分类. SVM 作为基于直方图的图像分类器已经得到广泛应用. 本文用 χ^2 作为测量两个特征向量 F 和 F' 的距离(n 为特征向量的个数):

$$\text{dist}_{\chi^2}(F, F') = \sum_{i=1}^n \frac{(F_i - F'_i)^2}{F_i + F'_i}$$

将基于 χ^2 的核函数应用到 SVM 中,训练 SVM 分类器. 核函数公式如下:

$$K_{\chi^2}(F, F') = e^{-\frac{1}{D} \text{dist}_{\chi^2}(F, F')}$$

其中, D 为距离归一化参数. 本文中, D 为所有训练数据的平均距离. 最后,将训练好的 SVM 模型应用到测试图像集,并预测分类结果.

3 实验结果与分析

3.1 实验设置

为了验证本文提出的方法,采用比较经典的 LBP 算法作为对比对象,在 PASCAL VOC 2007 数据集上进行实验,对比了原始 LBP 特征和基于汉明距离的 BoF 特征的平均正确率 (mAP),以及 BoF 模型使用不同距离测量方法的耗时. 实验采用的二进制描述符与选取邻接像素的个数 p 相同,视觉词汇的大小用 c 表示. 文献[1] 提出,词汇越多,图像的分类效果越

好。但是,词汇量太多会导致直方图太稀疏,比较也更困难。为了控制好词汇量的大小,本文做了一系列的实验,最后的设置如下: $\{r = 1, p = 8\}$, $c = 220$; $\{r = 2, p = 16\}$, $c = 1300$; $\{r = 3, p = 24\}$, $c = 1400$; $\{r = 2, p = 16\}$, $c = 1300$; $\{r = 3, p = 24\}$, $c = 1400$ 。采用 LibSVM^[14] 进行分类器训练。

3.2 实验结果比较

两种方法不同距离测度的 mAP 结果如表 1 所示。由表 1 可见,本文方法的 mAP 优于经典的 LBP 算法。由于 LBP 在计算特征值的时候需要加权, LBP 的邻接像素值不能超过 32。而本文提出的方法能够有效地避免这个问题。

表 1 两种方法不同距离测度的 mAP 结果

Table 1 The mAP comparison of different distance measure of two methods %

不同距离测度	mAP
LBP $\{r = 1, p = 8\}$	28.30
BoF $\{r = 1, p = 8\}$	28.40
LBP $\{r = 2, p = 16\}$	32.11
BoF $\{r = 2, p = 16\}$	33.09
LBP $\{r = 3, p = 24\}$	36.13
BoF $\{r = 3, p = 24\}$	35.09
LBP $\{r = 4, p = 32\}$	—
BoF $\{r = 4, p = 32\}$	32.52
LBP $\{r = 5, p = 40\}$	—
BoF $\{r = 5, p = 40\}$	33.24
LBP $\{r = 6, p = 48\}$	—
BoF $\{r = 6, p = 48\}$	32.78

本文将不同尺度的 LBP 特征值进行组合来评估 LBP 的多尺度融合。选取 $\{r = 1, 2, 3, p = 8 + 16 + 24\}$ 作为每个像素的特征提取方案。组合后的二进制特征维度为 48。同时,笔者将传统的 LBP 直方图特征 $\{r = 1, p = 8\}$, $\{r = 2, p = 16\}$, $\{r = 3, p = 24\}$ 进行融合。本文方法的 mAP 为 35.17%, 而融合后的 LBP 获得的 mAP 仅为 32.49%。

本文采用汉明距离实现特征聚类,并对汉明距离和欧式距离做了时间消耗对比实验。该

实验条件为 Intel(R) Core(TM) i7 CPU 940@2.93 GHz, 9 GB RAM。实验结果表明,同样条件下,欧氏距离的时间耗费约是汉明距离的 200 倍(汉明距离所需时间为 98.93 s, 欧氏距离需要 19 046.43 s)。特征提取设置为 $\{r = 1, p = 8\}$ 时,利用汉明距离获得的 mAP 为 28.20%, 略微低于欧氏距离的 mAP 为 29.13%。但是,汉明距离可以节省更多的时间,而且可以在特征提取时选用更大的半径和更多的邻接像素。

4 结语

本文提出了一种改进的用于图像分类的局部二值描述方法。该方法能够有效解决局部二值方法对高维度的限制,同时,由于汉明距离更适合计算机操作,大大提高了运行效率。本文还将提出的方法与经典 LBP 方法在 PASCAL VOC 2007 数据库上做了对比,计算时间有很大优势。今后,会将随机场理论加入特征提取过程中,实现邻接像素的选择。

参考文献:

- [1] LOWE D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91.
- [2] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C] // IEEE Conference on Computer Vision and Pattern Recognition (CVPR '05). Piscataway: IEEE, 2005, 1: 886.
- [3] OJALA T, PIETIKAINEN M, MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971.
- [4] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. Imagenet classification with deep convolution-

- al neural networks [C] // NIPS2012. Nevada: ACM Press,2012:1097.
- [5] CALONDER M, LEPETIT V, STRECHA C, et al. BRIEF: binary Robust independent elementary features [C] // ECCV2010. Crete: Springer Press,2010,6314:778.
- [6] KUMAR D A, ANNIE M L C. Clustering dichotomous data for health care [J]. International Journal of Information Sciences and Techniques,2012,2(2):23.
- [7] PANDIT S, GUPTA S. A comparative study on distance measuring approaches for clustering [J]. International Journal of Research in Computer Science,2011,2(1):29.
- [8] CAPUTO B, HAYMAN E, MALLIKARJUNA P. Class-specific material categorisation [C] // Tenth IEEE International Conference on Computer Vision (ICCV'2005). Piscataway: IEEE, 2005:1597.
- [9] ZHU C, BICHOT C E, CHEN L. Visual object recognition using daisy descriptor [C] // 2011 IEEE International Conference on Multimedia and Expo (ICME). Piscataway: IEEE,2011:1.
- [10] CORTES C, VAPNIK V. Support-vector networks [J]. Machine Learning, 1995, 20 (3) : 273.
- [11] GOLD S, RANGARAJAN A. Softmax to soft-sign: Neural network algorithms for combinatorial optimization [J]. Journal of Artificial Neural Networks,1996,2(4):381.
- [12] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection [C] // IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE,2005:886.
- [13] OJALA T, PIETIKÄINEN M, MÄENPÄÄ T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2002,24(7):971.
- [14] CHANG C C, LIN C J. LIBSVM: A library for support vector machines [J]. ACM Transactions on Intelligent Systems and Technology,2011,2(3):27.

(上接第26页)

- [6] 孟宏. 中医药理论与创意在化妆品中的应用 [J]. 北京日化,2016(2):2.
- [7] 董银卯,孟宏,何聪芬. 中医药理论与技术在化妆品中的应用 [J]. 日用化学品科学,2009,32(9):14.
- [8] 徐灵胎. 徐灵胎医书全集 [M]. 太原:山西科学技术出版社,2001.
- [9] 崔玲. 神农本草经 [M]. 天津:天津古籍出版社,2009.
- [10] 邓小锋,孟宏,李丽,等. 炮制技术在化妆品植物原料开发中的应用 [J]. 日用化学工业,2015,45(4):226.
- [11] 赵二劳,郭青枝,白建华. 几种中草药防晒成分的提取及其性能测定 [J]. 北方园艺,2010(11):197.
- [12] 李楚忠,高红军,丛琳. 天然植物保湿成分在护肤品中的应用概况 [J]. 日用化学品科学,2014,37(7):24.
- [13] SÜNTAR I, AKKOL E K, KELES H, et al. Exploration of the wound healing potential of *Helichrysum graveolens* (Bieb.) Sweet: isolation of apigenin as an active component [J]. J Ethnopharmacol,2013,149(1):103.
- [14] 谢艳君,孔维军,杨美华,等. 化妆品中常用中草药原料研究进展 [J]. 中国中药杂志,2015,40(20):3925.
- [15] 中华人民共和国国务院新闻办公室. 中国的中医药 [M]. 北京:人民出版社,2016.