

基于深度学习的图像检索研究^{*}

杨馥溢 何 嘉

(成都信息工程大学计算机学院 成都 610225)

摘 要 首先深度学习是机器学习研究中的一个新的领域,广泛应用到图像识别中,已取得很好的成绩;其次如何从大规模的图像数据集中检索到目标图像已经逐渐成为多媒体研究领域的重点与难点。所以论文提出基于深度学习的图像检索技术,利用深度卷积神经网络模型进行特征抽取,将特征二值化,并修改网络损失函数,然后把二进制描述符应用于图像检索系统中,提高图像检索效率。即图像检索的关键步骤在于高准确的图像特征提取和高效率的特征匹配。

关键词 深度学习;深度卷积神经网络;图像检索;二进制描述符

中图分类号 TP301.6 **DOI:**10.3969/j.issn.1672-9722.2019.09.033

Research on Image Retrieval Based on Deep Learning

YANG Fuyi HE Jia

(College of Computer Science, Chengdu University of Information Technology, Chengdu 610225)

Abstract Firstly, deep learning is a new field in machine learning research. It is widely applied to image recognition and has achieved good results. Secondly, how to retrieve target image from large-scale image data is becoming the focus and difficulty of multimedia research. So in this paper, based on the depth study of image retrieval technology, the depth of the convolutional neural network model for feature extraction is used, the characteristics are binarized, and the network loss function is modified, and then the binary operator is applied to image retrieval system, the efficiency of image retrieval is improved. The key step of image retrieval is to make accurate image feature extraction and efficient feature matching.

Key Words deep learning, deep convolution neural network, image retrieval, binary descriptor

Class Number TP301.6

1 引言

随着计算机技术的飞速发展,数字图像的信息量越来越大,基于文本的图像检索已不能满足人们的需求。因此各种基于内容的图像检索算法也因此迅速兴起,图像检索的关键步骤是有效的图像特征提取和准确的特征匹配。

首先如何弥合底层像素信息和高层语义信息,早期通过手工的方式来提取特征描述符,在欧式空间内计算这些描述符的欧氏距离,然后根据欧氏距离来判定两张图片是否相似。手工提取特征不但费时费力而且提取的特征往往是不准确的,仍无法去表示图像的高层语义。研究表明卷积神经网络

能够学习丰富的中高层图像特征,更好地提取底层像素。其次随着图像数据量越来越大,如何高效地特征匹配,例如卷积神经网络第七层的输出包含丰富的图像信息,但是维度高达1000维,1000维的浮点数向量与1000维的浮点数向量之间求相似度,运算量较大。此时研究人员提出将高维特征向量压缩到低维空间,并且以01二进制的方式表达,这样在低维空间中计算两个二进制的汉明距离速度非常快,从而可以在一定程度上缓解时效问题。因此,如何构建出一个有效的图像二进制描述符成了关键。本文提出的一个基于深度学习的方法来学习二进制描述符,用于图像检索的研究中。

^{*} 收稿日期:2019年3月17日,修回日期:2019年4月20日

基金项目:四川省科技计划项目(编号:2012GZ0111)资助。

作者简介:杨馥溢,女,硕士研究生,研究方向:计算智能及数据挖掘。何嘉,女,博士研究生,教授,研究方向:人工智能。

2 国内外研究现状

特征描述符^[1]在计算机视觉中起着重要作用,已被广泛应用于众多的计算机视觉任务,如图像匹配、目标识别和图像检索^[2]等。

1) 实值描述符

1999年David G.Lowe提出尺度不变特征变换SIFT^[3]算法(Scale-invariant feature transform)来学习的丰富图像特征,该描述符都有很强的辨别能力,主要弥补了低级像素和高级语义信息之间的差距^[4]。但是它们是高维实值描述符,并且通常需要高计算成本。

2) 为了降低计算复杂度,提出了几个轻量级二进制描述符

2010年EPFL的Calonder提出BRIEF^[5](Binary Robust Independent Elementary Features)算法,该算法是对已检测到的特征点进行描述,它是一种二进制编码的描述子,事先利用特征点检测算法得到特征点的位置,再在特征点的中心取邻域,随机选取两个点建立特征描述符。但是BRIEF描述子不具有旋转不变性与尺度不变性。

2011年Ethan Rublee等人提出ORB^[6](Oriented Brife)算法,该算法首先检测特征点,利用计算中心矩的方法计算角点方向,将该方向作为主方向建立特征描述符。但是效果不明显。

同年Stefan Leutenegg提出BRISK^[7](Binary Robust Invariant Scalable Keypoints)算法,采样模式是均匀采样模式,即在同一圆上等间隔的进行采样,同时为了保证尺度不变性,考虑在找极值点时不单单要找原图像这一个尺度的,而是找多个尺度中的极值点,并把离散的尺度进行拟合以得到特征点真实的尺度。

2012年Alexandre Alahi提出FREAK^[8](Fast Retina Keypoint)算法,采样模式发生了改变,它采取了更为接近于人眼视网膜接收图像信息的采样模型。

上述二进制描述符对排序和匹配非常有效。给定的二进制描述符,可以通过经由XOR按位操作计算二进制描述符之间的汉明距离来快速测量图像的相似性^[9]。由于这些早期的二进制描述符是通过简单的紧密度比较计算的,它们通常对尺度,旋转和噪声敏感是不稳定的。

3) 哈希算法学习二进制描述符

局部敏感哈希LSH^[10](Locality-Sensitive Hashing)是把空间中的两个数据经过一种映射方式,如

果这两个数据在新的空间相邻的概率就会越大,而不相邻的数据映射到同一个桶的概率就会越小。

语义哈希SH^[11](semantic hashing)是为每个图像寻找一种二进制编码,使两个对象的相似度与其编码之间的海明距离相关,即相似度高的对象对应的编码海明距离小,相似度低的对象对应的编码海明距离大。

谱哈希SpeH^[12](spectral hashing)通过频谱图分割生成高效的二进制描述符。

迭代量化ITQ^[13](Iterative Quantization)使用迭代优化策略找到具有最小二值化损失的投影。

4) 深度学习哈希

CNNH(Convolution Neural Network Hashing)^[14]:首先通过对相似度矩阵(矩阵中的每个元素指示对应的两个样本是否相似)进行分解,得到样本的二值编码;然后利用CNN对得到的二值编码进行拟合。

DNNH(Deep Neural Network Hashing)^[15]:网络使用三张图像构成的三元组进行训练。在三元组中,其中的第一张图像和第二张图像是相似的,而第一张图像和第三张图像则是不相似的。基于三元组的损失函数的目标是:在得到的Hamming空间中,相似样本间的距离小于不相似样本间的距离。

DSH(Deep Supervised Hashing)^[16]:在该文中添加了一个正则项,使网络的输出值接近于二值编码。当网络的输出和期望得到的值偏差越大的时候,损失也越大,但是同时,梯度的值保持在-1或+1,来保证训练过程的稳定性。

3 基于深度学习的二进制描述符

3.1 深度卷积神经网络VGGNet16

VGGNet^[17]于2014年牛津大学Visual Geometry Group视觉几何组组员所提出,在ILSVRC取得了定位第一、分类第二的成绩。与AlexNet相比较,VGGNet共16层,其中13层的卷积层,3层的全连接层,网络结构如表1所示。VGGNet16主要改进后的好处主要有三点:一是3*3的尺寸可以捕获上下左右和中心的信息;二是两个3*3的卷积层可以代替5*5,可以减少参数,节省资源;三是卷积层后与非线性激活函数的结合,多个卷积层则有多个非线性激活函数,这样更能准确的提取出深层图像特征。当然VGGNET也有一定的缺点,因为网络层次过深,则参数过多需花费大量的训练时间,但是随着GPU并行计算的出现,该问题也能随之解决。

表 1 VGGNet16 网络结构

Input	conv3-256	conv3-512
conv3-64	conv1-256	conv1-512
conv3-64	maxpool	maxpool
maxpool	conv3-512	FC-4096
conv3-128	conv3-512	FC-4096
conv3-128	conv1-512	FC-1000
maxpool	maxpool	Softmax
conv3-256	conv3-512	

3.2 改进损失函数

首先使用 VGGNet16 网络用来提取图像特征,然后用新的全连接层代替深度卷积神经网络的 Softmax 分类层,并对该层进行二值化(用 0、1 表示),得到相应二进制描述符。此时二进制与二进制计算汉明距离可以提高图像检索的速度;最后,本文优化了损失函数。第一优化在二值化过程中产生的量化损失,减少量化损失,使量化前和量化后结果最为接近;第二考虑不同的旋转角度对图像检索的影响,减少原图像和旋转图像之间的二进制描述符误差,捕获更多不相关的信息,更能实现有效的图像检索。深度卷积神经网络结构图如图 1 所示。

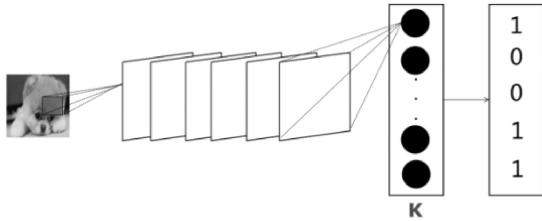


图 1 深度卷积神经网络结构图

根据本文描述,所学习到的二进制特征描述符应该满足两个基本条件:一是高质量提取图像特征,二是减少计算花费。因此首先假如共有 n 个样本,其中 $X=(x_1, x_2, x_3, \dots, x_n)$ 、 $Y=(y_1, y_2, y_3, \dots, y_n)$, 图像样本通过 VGGNet16 网络的学习,学习到一组非线性投影参数 $W=(w_1, w_2, w_3, \dots, w_k)$;其次将输入图像量化为二进制向量 B_n ;最后为了学习一串二进制描述符并应用于图像检索中,所学习到的 W 应满足如下要求:

- 1)减少量化损失;
- 2)减少旋转前和旋转后误差损失。

根据卷积神经网络前向传播算法求得 $f_k(X; W)$ 如公式(1)所示。

$$f(X, W) = f_k \left(\sum_{i=1}^n w_k x_i \right) \quad (1)$$

其中 $k=12\text{bit}, 24\text{bit}, 32\text{bit}$;

以某样本 X 为例,为了得到 0、1 的二进制字符串,量化过程求得 B_n 如公式(2)所示。

$$B_n = (-1/2)(\text{sign}f(X; W) - 1) \quad (2)$$

其中符号函数 $\text{sign}f(X; W)$ 如公式(3)所示。

$$\text{sign}f(X; W) = \begin{cases} 1, & f(X; W) > 0 \\ -1, & f(X; W) < 0 \end{cases} \quad (3)$$

3.2.1 减少量化损失

本文的方法是将输入图像映射成一串二进制串。在量化过程中,为了能更多地保留图像的原始信息,因此需要减少量化过程中的损失,这样留下的原始图像信息就越多。根据公式(2)设计减少量化损失 $L_1(W)$ 如公式(4)所示。

$$\min L_1(W) = \sum_{n=1}^N \| \text{sign}f(X; W) - f(X; W) \|^2 = \sum_{n=1}^N \| (2B_n - 1) - f(X; W) \|^2 \quad (4)$$

3.2.2 减少旋转前和旋转后误差损失

由于 VGGNet16 网络仅仅只考虑了多尺度的研究,但是图像旋转后会给检索带来一定的困难,因此参考选择了不同的旋转角度来提高提取图像特征的准确率,主要是减少图像旋转前后的二进制误差^[18],从而提高检索效率。随着增加图像旋转度时,估计误差可能会变大,根据估计误差,首先使用惩罚因子参数 ε 来表示对损失的重视程度。惩罚因子最初用于 SVM^[19] 分类中,表示分类与损失之间的权值。因此通过惩罚参数将旋转后样本 x_i 分类为 y_i 类。当惩罚参数较小时,则起不到对旋转样本的惩罚作用,随着增大惩罚因子的值,虽然能实现将旋转后的样本点完全正确的分类,但是这样将会导致过拟合,泛化能力不够,故选取合适的惩罚参数 ε 对本实验很重要。然后根据上述量化过程减小前后旋转误差如公式(5)所示。

$$\min L_2(w) = \sum_{n=1}^N \sum_{\theta=-R}^R \varepsilon \| B_{(n, \theta)} - Y \|^2 + \sum_{n=1}^N \sum_{\theta=-R}^R \| B_{(n, \theta)} - B_n \|^2 \quad (5)$$

3.3 总体损失函数

结合公式(4)和公式(5),提出本文的总体损失函数如公式(6)所示,根据反向传播和随机梯度下降算法更新 W 。

$$L(W) = \alpha L_1(W) + \beta L_2(W) \quad (6)$$

3.4 算法流程

根据本文设计思想,大致算法步骤如表 2 所示。

表 2 算法步骤	
Input	$X=[x_1, x_2, \dots, x_n]$
Initialize	W (Pre-training) Fix W Calculate B_n using 公式(2) Fix b_n Calculate W using 公式(4)
Optimization	Fix W Calculate B_n using 公式(2) Fix W Calculate B_n using 公式(5)
Output	W

基于内容的图像检索流程如图 2 所示。

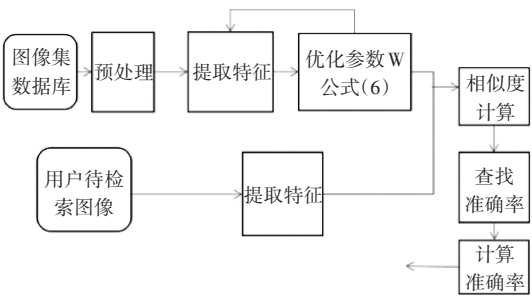


图 2 图像检索流程图

4 实验与分析

4.1 数据集

CIFAR10 由 60000 张 32*32 的 RGB 彩色图片构成,共 10 个分类。50000 张训练,10000 张测试。这个数据集最大的特点在于识别迁移到了普适物体,部分数据集如图 3 所示。



图 3 CIFAR10 数据集图

4.2 深度学习框架

本文采用开源框架 CAFFE,CAFFE 的全称是 Convolutional Architecture for Fast Feature Embedding,它是一个开源的深度学习框架,核心语言是 C++,支持 Python 和 Matlab 接口,既可以在 CPU 上

运行,也可以在 GPU 上运行。

4.3 图像检索度量算法

对图像提取相应的特征后,特征与特征之间选择合适的度量算法需要进行相似度匹配。因为本文比较的特征是一串 0、1 的字符,所以可以选取用汉明距离来计算相似度。对两个字符 x,y 进行异或运算,统计结果为 1 的个数,那么这个数就是汉明距离。汉明距离越大表明差异越大,越不相似。

4.4 图像检索算法的评价标准

1)查准率

查准率(Precision)指一次检索结果中,返回的图像中的正确图像占有所有检出的比例,公式如(7)所示。

$$Precision = P(R/S) = \frac{P(S \cap R)}{P(R)} = \frac{s}{s + u} \tag{7}$$

2)查全率

查全率(Recall)指在一次检索结果中,返回结果里的正确图像在检索系统中所有相关图像中所占的比例,公式如(8)所示。

$$Recall = P(R/S) = \frac{P(S \cap R)}{P(R)} = \frac{s}{s + v} \tag{8}$$

3)平均精度均值 (mean Average Precision-mAP)

由于查准率和查全率这两种评价标准在一定程度上是相互限制的,通常查全率高,则查准率低;查准率高,则查全率低。因此引入了精度均值(Average Precision- AP),用来表示查准率和查全率两条曲线下的面积,而 mAP 是多个类别 AP 的平均值。因此本文选择平均精度均值来作为衡量标准。

4.5 实验结果

本实验采用 GPU 运行,节省运行时间,使用 CIFAR10 彩色图像数据集进行实验,CIFAR10 数据集像素为 32*32,因为本文网络考虑到使用 VGG-NET16,所以将数据进行预处理,首先归一化为 256*256,然后中心剪裁为 224*224。采用在 ImageNet 大规模数据集上训练得到的预训练权重来初始化网络参数,随后使用随机梯度下降(Stochastic Gradient Descent,SGD)方法和反向传播算法根据总体损失函数公式(6)优化 W。本实验网络中的批处理大小设置为 32,旋转图像角度如公式(5)中分别为 0、±10、±20、±30。根据网络 loss 层的参数 loss_weight,该参数的主要作用表现为对该层损失的重视程度,越重视则取 1,故设置α为 1,惩罚参数ε取 100。当β=0.1 时,返回检索前 top 1000 的实验结果并与其他算法进行比较,各经典算法精确度比

较如表 3 所示。

表 3 各算法精确度比较(mAP %)

Method	12bit	24bit	32bit
LSH	0.1277	0.1367	0.1407
SH	0.1311	0.1278	0.1364
ITQ	0.1080	0.1088	0.1117
CNNH	0.5425	0.5604	0.5640
DNNH	0.5708	0.5875	0.5574
DSH	0.6157	0.6512	0.6755
DH	0.6924	0.7219	0.7387
Ours	0.7312	0.7497	0.7628

在不同的旋转角度下,当旋转角度 $\theta=[-10, 10]$ 时,当旋转角度 $\theta=[-20, 20]$ 时以及当旋转角度 $\theta=[-30, 30]$ 时,各算法精确度比较如图 4 所示。

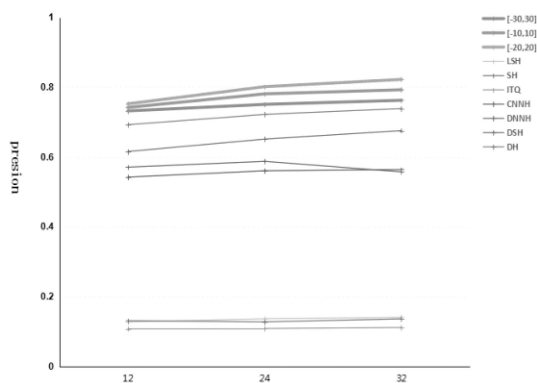


图 4 不同旋转角度的精确度图

其中当二进制长度为 12bit,不同旋转角度的分类错误率如表 4 所示。

表 4 分类错误率比较(%)

Method	Test Error
$\theta=0$	0.4877
$\theta=[-10, 10]$	0.5011
$\theta=[-20, 20]$	0.4780
$\theta=[-30, 30]$	0.5325

4.6 实验分析

首先根据表 3 可以发现二进制位数越长,精确度越高,并且在 VGGNet16 网络下改进损失函数检索效率有明显得提高;其次根据图 4 可以得到,当图像旋转角度在 $[-20, 20]$ 时,图像检索效率较好。

5 结语

在本文中,提出了基于 VGGNet16 的深度学习网络模型来学习一串二进制描述符用于图像检索中。主要采用改进损失函数的方式来学习“好的”二进制描述符,并在 CIFAR10 数据集上验证该实验是否有效。实验表明减少量化损失以及考虑旋转因素可以提高图像检索的准确率。同时惩罚参数

的取值难以确定,太小起不到惩罚作用;太大则由于误差的影响会导致错误,因此应该再反复得做实验确定最优的惩罚因子值。

参考文献

- [1] K. He, F. Wen, J. Sun. K-means hashing: An affinity-preserving quantization method for learning binary compact codes[J]. CVPR, 2013, 9(4): 2938–2939.
- [2] 周明全, 耿国华, 韦娜. 基于内容图像检索技术[M]. 北京: 清华大学出版社, 2007: 15–17.
- ZHOU Mingquan, GENG Guohua, WEI Na. Based oncontent image retrieval technology [M]. Beijing: tsinghua university press, 2007: 15–17.
- [3] Lowe DG. Object Recognition from Local Scale-Invariant Features [C]//International Conference on Computer Vision IEEE, 1999: 54–64.
- [4] Girshick R, Donahue J, Darrell T, et al. RichFeature Hierarchies for Accurate Object Detection and Semantic Segmentation [C]//Computer Vision and Pattern Recognition IEEE, 2014: 582–586.
- [5] Calonder M, Lepetit V, Strecha C, et al. BRIEF: binary robust independent elementary features [C]//European Conference on Computer Vision: Springer-Verlag, 2010: 781–784.
- [6] Rublee E, Rabaud V, Konolige K, et al. ORB: An efficient alternative to SIFT or SURF [J]. International Conference on Computer Vision IEEE, 2012, 58(11): 2567–2569.
- [7] Stefan Leutenegger, Margarita Chli, Roland Y. Siegwart, Binary Robust Invariant Scalable Keypoints [J]. ICCV, 2011, 58 (11): 2550–2551.
- [8] Alexandre Alahi, Raphael Ortiz, Pierre Vanderghenst, Fast Retina Keypoint [J]. CVPR, 2012, 157 (10) : 514–515.
- [9] B.Fan, Q.Kong, T.Trzcinski, Z.Wang, C.Pan, and P.Fua. Receptive fields selection for binary feature description [J]. IEEE Trans.Image Proc, 2014, 23(6): 2587–2589.
- [10] A.Andoni, P.Indyk. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions [J]. In Proc.FOCS, 2008, 51(1): 461–462.
- [11] R.Salakhutdinov, G.E.Hinton. Semantic hashing [J]. Approx.Reasoning, 2009, 50(7): 972–973.
- [12] Y.Weiss, A.Torralba, R.Fergus. Spectral hashing [J]. In Proc.NIPS, 2008, 282(3): 1755–1756.
- [13] Y.Gong, S.Lazebnik, A.Gordo, and F.Perronnin. Iterative quantization: A approach to learning binary codes for large-scale image retrieval [J]. IEEE Trans.Pattern Anal. Mach. Intell, 2013, 35(12): 2918–2920.

(下转第 2285 页)

参考文献

- [1] 吴扬. 基于监控视频的人体异常行为检测研究[D]. 哈尔滨: 哈尔滨工业大学, 2015.
WU Yang. Research on Detection of Abnormal Human Behavior Based on Surveillance Video [D]. Harbin: Harbin Institute of Technology, 2015.
- [2] MOON H M, CHAE S H, MOON D, et al. Intelligent video surveillance system using two-factor human information [J]. Telecommunication Systems, 2013, 52 (4) : 2249-2257.
- [3] YingZhang, Huchuan Lu, Lihe Zhang and Xiang Ruan. Combining motion and appearance cues for anomaly detection [J]. Pattern Recognition, 2016, 51(C) : 443-452.
- [4] Dheeraj Kumar and James C, Bezdek etc. A visual-numeric approach to clustering and anomaly detection for trajectory data [J]. The Visual Computer, 2017, 33 (3) : 265-281.
- [5] 梁荣华, 郑莉莉, 黄鲜萍. 基于SVM的人体异常行为检测算法研究[D]. 杭州: 浙江工业大学, 2012.
LIANG Ronghua, ZHENG Lili, HUANG Xianping. Research on Algorithm of Human Anomaly Detection Based on SVM [D]. Hangzhou: Zhejiang University of Technology, 2012
- [6] L.Kratz, K.Nishino. Anomaly Detection in Extremely Crowded Scenes Using Spatio-Temporal Motion Pattern Models [C]. Computer Society Conference on Computer Vision and Pattern Recognition: IEEE, 2009: 1446-1453.
- [7] CHEN Yu-feng, LIANG Guo-yuan, LEE Ka-keung, et al. Abnormal behavior detection by multi-SVM-based Bayesian network [C]. International Conference on Information Acquisition. Seogwipo-si: IEEE, 2007: 298-303.
- [8] 朱旭东, 刘志镜. 基于主题隐马尔科夫模型的人体异常行为识别[J]. 计算机科学, 2012, 39(3): 251-255.
ZHU Xudong, LIU Zhijing. Human Abnormal Behavior Recognition Based on the Theme Hidden Markov Model [J]. Computer Science, 2012, 39(3) : 251-255.
- [9] 张国印. 基于贝叶斯网络的 Android 恶意行为检测方法 [J]. 计算机工程与应用, 2016, 52(17) : 16-23.
ZHANG Guoyin. Android malicious behavior detection method based on Bayesian network [J]. Computer Engineering and Applications, 2016, 52(17) : 16-23.
- [10] 徐超. 基于视频图像中人的异常行为识别[D]. 西安: 西安电子科技大学, 2014.
XU Chao. Abnormal Behavior Recognition Based on Human Images in Video [D]. Xi'an: Xi'an University of Electronic Science and Technology, 2014.
- [11] 周义, 费树岷, 沈捷. 基于特征点轨迹的实时异常行为检测[J]. 工业控制计算机, 2016, 29(2): 80-82.
ZHOU Yi, FEI Shumin, SHEN Jie. Real Time Abnormal Behavior Detection Based on Feature Point Trajectory [J]. Industrial Control Computer, 2016, 29(2) : 80-82.
- [12] Barnich O, Van Droogenbroeck M. ViBe: a powerful random technique to estimate the background in video sequences [C]. International and Signal Processing (ICASSP) 2009 : 945-948.
- [13] 王丹, 樊永生, 王秀川. 一种基于梯度和Vibe算法的运动检测算法[J]. 电视技术, 2017, 41(3): 100-104.
WANG Dan, FAN Yongsheng, WANG Xiuchuan. A Motion Detection Algorithm Based on Gradient and Vibe Algorithm [J]. TV technology, 2017, 41(3) : 100-104.
- [14] 李全栋, 陈树越, 张微. 一种改进的无监督聚类的关键帧提取算法[J]. 应用光学, 2010, 31(5): 741-744.
LI Quandong, CHEN Shuyue, ZHANG Wei. An Improved Keyless Frame Extraction Algorithm for Unsupervised Clustering [J]. Applied Optics, 2010, 31 (5) : 741-744.
- [15] 杜鉴豪, 许力. 基于区域光流特征的异常行为检测[J]. 浙江大学学报, 2011, 45(7): 1161-1166.
DU Jianhao, XU Li. Abnormal Behavior Detection Based on Optical Flow Features [J]. Journal of Zhejiang University, 2011, 45(7) : 1161-1166.
- [14] Rongkai Xia, Yan Pan, Hanjiang Lai, Cong Liu, Shuicheng Yan. Supervised Hashing for Image Retrieval via Image Representation Learning [C]//AAAI, 2014: 4-6.
- [15] Hanjiang Lai, Yan Pan, Ye Liu, Shuicheng Yan. Simultaneous Feature Learning and Hash Coding with Deep Neural Networks [C]//CVPR, 2015: 3272-3273.
- [16] Xia R, Pan Y, Lai H, Liu C, and Yan S. Supervised hashing for image retrieval via image representation learning [C]//AAAI, 2014: 3-4.
- [17] Karen Simonyan, Andrew Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition [J]. CVPR, 2014, 9(4): 2-3.
- [18] Kevin Liny, Jiwen Luz, Chu-Song Chen, Jie Zhou. Learning Compact Binary Descriptors with Unsupervised Deep Neural Networks [J]. CVPR, 2016, 10 (5) : 1184-1185.
- [19] Nello Cristianini, John Shawe-Yaylor. 支持向量机导论 [M]. 李国正, 等译. 北京: 电子工业出版社, 2004: 27-29.
- [20] Nello cristini, John Shawe-Yaylor. Support vector machine introduction [M]. Li guozheng, et al. Beijing: Electronics industry press, 2004: 27-29.

(上接第 2280 页)