

魏鸿奥, 魏本昌. 残差自适应图像描述符 [J]. 智能计算机与应用, 2025, 15(9): 76–81. DOI: 10. 20169/j. issn. 2095–2163. 250912

残差自适应图像描述符

魏鸿奥, 魏本昌

(湖北汽车工业学院 电气与信息工程学院, 湖北 十堰 442002)

摘要: 针对局部特征聚合描述子向量 (Vector of Locally Aggregated Descriptors, VLAD) 在线生成阶段的量化残差计算过程中求得的残差均值较大进而影响 VLAD 检索精度的问题, 提出了一种残差自适应图像描述符 (Residual-Adaptive VLAD, RAVLAD)。预处理阶段, 通过 K-means 算法对特征空间进行划分, 以生成具有 2 层结构的视觉码书。RAVLAD 生成阶段, 量化残差基于 2 层码书上的最近邻视觉词自适应生成从而使残差均值减小, 量化残差的累积则基于第一层码书的各视觉词, RAVLAD 最终表示为第一层码书的各视觉词上残差累积向量的串联。为了抑制局部特征爆发现象, 采用三步归一化策略。在 Holidays 和 UKBench 数据集上的图像检索实验结果表明, RAVLAD 的检索性能要优于 VLAD 和其他诸多方法。

关键词: 图像检索; 图像描述符; 局部特征聚合描述子向量; 尺度不变特征变换描述子

中图分类号: TP391

文献标志码: A

文章编号: 2095–2163(2025)09–0076–06

Residual adaptive image descriptor

WEI Hong'ao, WEI Benchang

(School of Electrical and Information Engineering, Hubei University of Automotive Technology, Shiyan 442002, Hubei, China)

Abstract: In order to solve the problem that the retrieval accuracy of VLAD is affected by the large residual mean value obtained during the quantization residual calculation in the stage of VLAD online generation, RAVLAD is proposed. In the pre-processing stage, feature space is divided by K-means algorithm to generate visual code book with two-layer structure. In the stage of RAVLAD generation, quantized residuals are adaptively generated based on the nearest neighbor visual words in the two-layer code book to reduce the mean of residuals, and the accumulation of quantized residuals is based on the visual words in the first-layer code book. Therefore, RAVLAD is represented as the series of residuals cumulative vectors on the visual words in the first-layer code book. In order to suppress the phenomenon of local feature outburst, the three-step normalization strategy is adopted. The results of image retrieval experiments on Holidays and UKBench datasets show that RAVLAD has better retrieval performance than VLAD and many other methods.

Key words: image retrieval; image descriptor; VLAD; SIFT

0 引言

随着大数据、云计算时代的到来, 图像资源因其直观、易于理解的特点获得了人们的青睐。面对当今互联网中海量的图像资源^[1], 如何从中快速准确地找到用户需要的图像成为了亟待解决的重要问题, 而图像检索技术可以作为解决该问题的有效手段。图像检索分为基于文本的图像检索和基于内容的图像检索, 由于早期基于文本的图像检索依赖于人工对图像的文本标注, 不可避免地会产生检索速

度慢等问题, 因此图像检索已转向基于内容的检索机制。目前, 基于内容的图像检索已在刑事侦查^[2]、纺织业^[3]、版权保护^[4]等领域得到了广泛的应用。

在基于内容的图像检索过程中, 需要提取图像的特征, 并通过查询图像与目标图像特征之间的比对, 实现对相似图像的检索。因此, 图像特征提取的好坏会直接影响图像检索的效果。图像特征分为局部特征和全局特征。在图像的局部特征中, 最为经典的是尺度不变特征变换 (Scale Invariant Feature

基金项目: 湖北省教育厅科技项目 (B2019077); 湖北汽车工业学院博士基金 (BK201603)。

作者简介: 魏鸿奥 (1999—), 男, 硕士研究生, 主要研究方向: 图像检索。

通信作者: 魏本昌 (1975—), 男, 博士, 讲师, 硕士生导师, 主要研究方向: 计算机视觉。Email: bc_david@163.com。

收稿日期: 2023–12–21

Transform, SIFT)^[5]描述子,由于其具有较强的鲁棒性,因此在图像处理领域有着较多的应用。Bay等学者^[6]提出的SURF(Speeded Up Robust Feature)描述子是对SIFT描述子的改进,虽然其维度只有SIFT描述子的一半,但检索精度能够和SIFT相当。此后,又有研究学者基于SIFT进行改进而提出PCA-SIFT^[7]、ICA-SIFT^[8]等描述子。除此之外,其它的局部描述子还有面向稠密特征提取的DAISY^[9]、更适用于移动场景的二值描述子ORB^[10]等。尽管图像的局部特征大多具有鲁棒性等特点,但如果将其直接用于图像检索,将产生较低的检索效率和较大的内存占用,从而降低图像的检索性能。针对这些问题,研究人员又提出了图像的全局特征,早期的全局特征有图像纹理特征^[11]和图像全局颜色直方图^[12]。之后,研究学者又提出了词袋模型(Bag of Words, BoW)^[13]、局部聚集描述子向量(Vector of Locally Aggregated Descriptors, VLAD)^[14-15]等通过聚集图像局部特征而生成的图像全局描述符。与BoW统计计数的方法相比,VLAD通过累积图像的局部特征与最近邻视觉词的残差来生成图像全局描述符的方式不仅能更好地描述图像,还能有效地减少信息损失,因而得到了广泛使用。

现已针对VLAD提出了多种改进方法。文献[16]提出了一种通过距离聚类的残差中心聚合方法来解决VLAD在残差累积过程中出现的过累积和欠累积问题。文献[17]针对VLAD编码过程出现的量化损失问题,提出一种基于内积加权的VLAD软编码方法。文献[18]通过拼接图像VLAD特征及图像空间分布熵向量的方法对VLAD进行改进,提出SEVLAD。文献[19]提出一种层次码书结构,使得SIFT描述子的量化更加精细,从而实现了VLAD的改进,提出EVLAD。文献[20]使用位置信息熵对SIFT描述子的位置信息进行描述,并对各视觉词上生成的128维残差和向量添加长度为2的位置信息熵值,以增强VLAD的检索能力。文献[21]利用每一个聚类内SIFT描述子的量化残差生成残差集,然后对残差集进行训练生成残差码本,最终将SIFT描述子量化产生的主残差和主残差量化产生的次残差求和得到FVLAD。文献[22]提出一种基于各聚类中局部特征计算各聚类分布熵向量的方法,并通过各聚类的分布熵向量与残差和向量的紧凑融合生成各聚类CEVLAD子向量,最终将子向量串联以生成CEVLAD描述符。

通过分析VLAD特征的编码过程,可以发现:

VLAD算法在离线预处理阶段聚类局部描述子生成视觉码书时,仅进行一次K-means聚类,即只生成一层粗码字的视觉码书。在线VLAD生成过程中,各图像的局部描述子也只基于这层粗码字的视觉码书计算量化残差,这种情况会导致局部描述子基于最近邻码字计算量化残差时生成的残差均值较大,进而影响VLAD对图像的描述能力。因此,针对该问题,提出了一种新的图像描述符生成算法。该算法在离线预处理阶段,局部特征空间会通过K-means算法生成具有2层结构的视觉码书。在线描述符生成阶段,量化残差基于2层码书共同生成从而使残差均值减小;为了抑制局部特征的爆发现象,算法中采用三步归一化策略,从而最终生成残差自适应图像描述符(Residual-Adaptive VLAD, RAVLAD)。

1 残差自适应图像描述符

1.1 VLAD描述符

VLAD是基于图像的SIFT特征和词袋模型生成,其生成过程如下。

在离线预处理阶段中,先提取出各个图像的SIFT特征,以形成局部特征空间。然后,对局部特征空间通过K-means算法聚成K类,并将K个类的类心作为K个视觉词生成大小为K的视觉码书,再将该码书标识为 $C = \{c_1, c_2, \dots, c_K\}$,其中, c_i 为第i个类的类心。在线VLAD生成阶段中,先将一幅图像的局部特征集合 $X = \{x_1, x_2, \dots, x_N\}$ 通过最近邻原则量化至视觉码书中与各自相距最近的视觉词上,量化过程如下:

$$i = \arg \min_{k: c_k \in C} \|x_p - c_k\| \quad (1)$$

其中, i 表示局部特征 x_p 按照最近邻原则量化至视觉词 c_i 上。并假设该图像量化至 c_i 上的所有局部特征为 $x_i = \{x_{i1}, x_{i2}, \dots, x_{ij}, \dots, x_{im}\}$,其中 x_{ij} 表示量化至 c_i 上的第j个局部特征。然后,基于视觉词 c_i ,计算并累加 x_i 中各特征与 c_i 的残差,生成 c_i 上的残差和向量 v_i :

$$v_i = \sum_{j=1}^m (x_{ij} - c_i) \quad (2)$$

其次,串联视觉码书C的各视觉词(c_1, c_2, \dots, c_K)上生成的残差和向量以生成高维向量 $v = [v_1^T, v_2^T, \dots, v_K^T]^T$ 。最后,为了抑制局部特征的爆发现象,需要对向量 v 进行L2归一化,生成图像的VLAD描述符。

在线VLAD生成过程的前3个阶段如图1所示。假设在预处理阶段,通过K-means算法将特征

空间聚成4类,并生成大小为4的视觉码书 $C = \{c_1, c_2, c_3, c_4\}$ 。在特征量化阶段,局部特征基于最近邻原则量化至相应的视觉词上,假设量化至视觉词 c_3 、 c_4 的局部特征数量各2个,量化至视觉词 c_1 的局部特征个数1个。在计算量化残差阶段,局部特征要基于所量化至的视觉词计算残差,并将求得的残差聚集至各视觉词。在量化残差累积阶段,则需要对各视觉词上聚集的残差进行累加以生成残差累积向量。VLAD 最终表示为各视觉词上残差累积向量的串联。

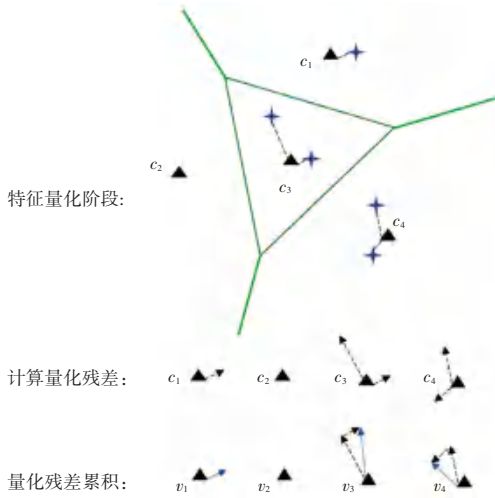


图1 在线 VLAD 生成过程示意图

Fig. 1 Schematic diagram of online VLAD generation process

1.2 RAVLAD

在编码 VLAD 的过程中,仅使用了一层粗码字的视觉码书,会导致局部描述子基于该码书上的最近邻码字计算量化残差时生成较大的残差均值,进而影响 VLAD 的检索精度。因此,针对该问题,提出了残差自适应图像描述符 RAVLAD。虽然 RAVLAD 的生成也分为离线预处理阶段和在线 RAVLAD 生成阶段,但对于各阶段的处理均和 VLAD 不同,对此可展开研究分述如下。

1.2.1 离线预处理阶段

在该阶段,要基于 K-means 算法对局部特征空间进行处理,以生成具有2层结构的视觉码书,具体方法如下:

(1) 提取图像库中部分图像的 SIFT 特征,以形成局部特征空间。

(2) 基于 K-means 算法将特征空间划分为 K 个 Voronoi 单元,并将 K 个类心作为 K 个视觉词生成大小为 K 的第一层视觉码书,并将第1层视觉码书标识为 $C = \{c_1, c_2, \dots, c_K\}$ 。其中, c_i 为第 i 个

Voronoi 单元的类心。

(3) 对每个 Voronoi 单元中的局部特征,再基于 K-means 算法聚成 L 类,并将 L 个类心作为 L 个视觉词生成大小为 L 的第2层视觉子码书。由此得到,第1层视觉码书的大小为 K ,第2层共有 K 个大小为 L 的视觉子码书,记为 $\{u_1, u_2, \dots, u_K\}$,其中 $u_i = \{u_{i1}, u_{i2}, \dots, u_{iL}\}$,这里 u_i 为第 i 个 Voronoi 单元所生成的第2层视觉子码书, u_{ij} 为 u_i 的第 j 个视觉词。

1.2.2 在线 RAVLAD 生成阶段

基于 1.2.1 节中层次视觉码书的 RAVLAD 生成过程如下:

(1) 设某幅图像的局部特征集合 $X = \{x_1, x_2, \dots, x_N\}$ 。首先,将 X 中各局部特征按照最近邻原则进行两步量化,针对局部特征 x_p ,两步量化过程如下:

$$i = \arg \min_{k: c_k \in C} \|x_p - c_k\| \quad (3)$$

其中, i 表示局部特征 x_p 第1步量化到第1层视觉码书的视觉词 c_i 上。进而推得:

$$j = \arg \min_{l: u_{ij} \in u_i} \|x_p - u_{il}\| \quad (4)$$

其中, j 表示已量化到 c_i 上的特征 x_p 第2步量化到第2层视觉子码书的视觉词 u_{ij} 上。

(2) 计算 x_p 基于第1层视觉码书的视觉词 c_i 的残差,计算公式为:

$$r^c = x_p - c_i \quad (5)$$

由此得到残差 r^c ;同时,也要计算 x_p 基于第2层视觉子码书的视觉词 u_{ij} 的残差,计算公式为:

$$r^u = x_p - u_{ij} \quad (6)$$

由式(6)求得残差 r^u 。

(3) 在残差 r^c 、 r^u 中,自适应地选择相对较小的残差。即:若 $\|r^u\| < \|r^c\|$,则选择残差 r^u ;反之选择 r^c 。经自适应选择后,所保留的残差向量记作 m ,并将 m 聚集至 c_i 。

(4) 累积 c_i 上聚集的所有残差向量,生成第 i 个 Voronoi 单元上的残差和向量 v_i 。

(5) 将 K 个 Voronoi 单元的残差和向量串联,形成高维向量 v 。对此可以表示为:

$$v = [v_1^T, v_2^T, \dots, v_K^T]^T \quad (7)$$

在线 RAVLAD 生成过程的前3个阶段如图2所示。假设在预处理阶段,特征空间先被划分为4个子空间,并生成第1层视觉码书 $C = \{c_1, c_2, c_3, c_4\}$,然后基于 K-means 算法再对各子空间划分,以生成第2层各视觉子码书 $u_1 = \{u_{11}, u_{12}, u_{13}\}$ 、 $u_2 = \{u_{21}, u_{22}, u_{23}\}$ 、 $u_3 = \{u_{31}, u_{32}, u_{33}\}$ 、 $u_4 = \{u_{41}, u_{42},$

u_{43} 。在特征量化阶段,局部特征基于最近邻原则进行两步量化,从而量化至第2层视觉子码书的视觉词上。在计算量化残差阶段,局部特征要分别基于两步量化过程中量化至的2个视觉词计算残差,并自适应地选择相对小的残差,图2中虚线表示自适应选择之后保留的残差,并将残差聚集至第1层视觉码书的视觉词上。在量化残差累积阶段,基于第1层视觉码书的视觉词,累积聚集至其上的残差。RAVLAD最终表示为第1层视觉码书各视觉词上残差累积向量的串联。

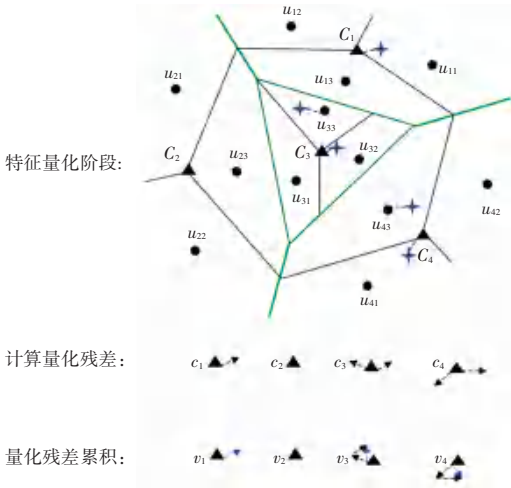


图2 在线RAVLAD生成过程示意图

Fig. 2 Schematic diagram of online RAVLAD generation process

1.2.3 归一化策略

为了抑制局部特征爆发现象,在RAVLAD的生成过程中,采用3步归一化策略:

(1) 对自适应选择后保留的残差向量 m 进行归一化。即:

$$m = m / \|m\| \quad (8)$$

(2) 对第 i 个 Voronoi 单元上生成的残差和向量 v_i 进行归一化,即:

$$v_i = v_i / \|v_i\| \quad (9)$$

(3) 对最终串联形成的高维向量进行 $L2$ 归一化。即:

$$v = v / \|v\| \quad (10)$$

1.2.4 RAVLAD 生成算法

算法1 图像的RAVLAD生成算法

输入 第1层视觉码书 $C = \{c_1, c_2, \dots, c_K\}$, 第2层视觉码书 $u_i (1 \leq i \leq K)$, 图像的局部特征集合 $X = \{x_1, x_2, \dots, x_N\}$

输出 RAVLAD v

// 初始化

```

1. for  $i = 1$  to  $K$ 
2.    $v_i = 0_d$ ;
3. end for
4. for  $p = 1$  to  $N$ 
   // 局部特征量化
5.    $i = \arg \min_{k: c_k \in C} \|x_p - c_k\|$ ;
6.    $j = \arg \min_{l: u_{ij} \in u_i} \|x_p - u_{ij}\|$ ;
7.    $r^c = x_p - c_i$ ; // 基于第1层码书的最近邻视觉词计算残差
8.    $r^u = x_p - u_{ij}$ ; // 基于第2层子码书的最近邻视觉词计算残差
   // 自适应选择残差
9.   if  $\|r^u\| < \|r^c\|$ 
10.     $m = r^u$ ;
11.   end if
12.   else
13.     $m = r^c$ ;
14.   end else
15.    $v_i = v_i + m / \|m\|$ ; // 残差累积
16. end for
17. for  $i = 1$  to  $K$ 
18.    $v_i = v_i / \|v_i\|$ ;
19. end for
20.  $v = [v_1^T, v_2^T, \dots, v_K^T]^T$ ; // 残差和串联
21.  $v = v / \|v\|$ 

```

1.2.5 复杂度分析

假设局部特征向量的维度是 D , 第1层视觉码书包含 K 个 D 维的视觉词, 第2层视觉子码书包含 L 个 D 维的视觉词, 一幅图像中局部特征的个数为 N 。第1层视觉码书的空间开销为 KD , 第2层视觉码书的空间开销共为 KLD , 则总的空间开销为 $(KD + KLD)$, 空间复杂度为 $O(KLD)$ 。生成RAVLAD描述符的时间主要分成2次量化过程。在第一次量化过程中, 每个局部特征要和 K 个类心比较一次, 则 N 个特征共需比较 NK 次, 所产生的时间开销为 NKD 。在第二次量化过程中, 每个局部特征要和 L 个类心再进行比较, 则 N 个特征共需比较 NL 次, 产生的时间开销为 NLD 。则时间复杂度为 $O(N(K + L)D)$ 。

2 实验结果及分析

本节将在公开的 INRIA Holidays 数据集^[23] 和 UKBench 数据集^[24] 上对 RAVLAD 的检索性能进行评价, 并比较其与其它描述符的检索性能。

2.1 实验数据集与评价标准

Holidays 数据集共包含 1 491 幅图像,所包含的图像主要以自然风光为主。Holidays 数据集样例如图 3 所示。该数据集被划分为 500 组,每组的第一张图像为查询图像(共 500 幅),正确的检索结果是该组的其他图像(共 991 幅),并且所有图像共提取出 4 455 091 个 128 维的 SIFT 特征,该数据集使用平均查准率(mean Average Precision, mAP)作为图像检索的性能评价指标。



图 3 Holidays 数据集样例图
Fig. 3 Examples of Holidays dataset

UKBench 数据集共包含 10 200 幅彩色图像,每 4 张图像取自同一物体的不同视点和光照,该数据集使用 N-S 评分作为其性能评价指标。UKBench 数据集样例如图 4 所示。

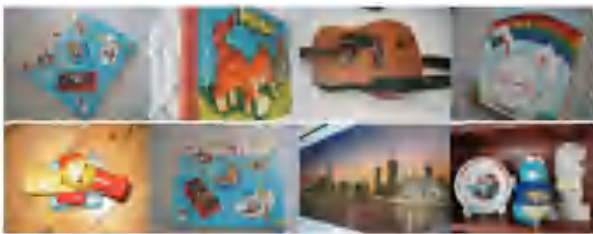


图 4 UKBench 数据集样例图
Fig. 4 Examples of UKBench dataset

2.2 参数 K 、 L 对 RAVLAD 检索性能的影响

为了探究参数 K 和 L 对 RAVLAD 检索性能的影响,在 Holidays 和 UKBench 数据集上进行相关实验,实验结果如图 5、图 6 所示。由图 5、图 6 可以看出,在 Holidays 和 UKBench 数据集上,若固定 L ,增加 K ,RAVLAD 的检索性能呈上升趋势,主要原因是随着 K 的增加,RAVLAD 描述符的维度会随之增大,其对图像描述的程度会更加精细,因此检索性能随之上升。由图 5 可知,在 Holidays 数据集上,若固定 K ,增大 L ,RAVLAD 的检索性能在整体上呈现出先增后减的趋势,并都在 $L = 15$ 时达到最大值。由图 6 可知,在 UKBench 数据集上,若固定 K ,增大 L ,RAVLAD 的检索性能会逐渐升高,但当 L 增加到 7 之后,该描述符检索性能的变化就较为平稳,这是由

于当第 2 层视觉子码书的大小 L 不断增大时,每个 Voronoi 单元会被划分得更加精细,所产生的残差均值以及残差的判别性相较于之前改变的程度都会减小,因此导致该现象的产生。综上所述,在后续比较 RAVLAD 和其它描述符的检索性能时,仅考虑 $L = 15$ 的情况。

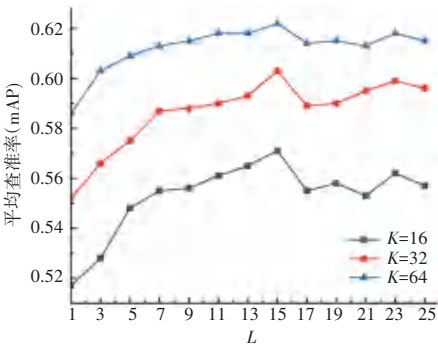


图 5 参数 K 、 L 对 RAVLAD 检索性能的影响 (Holidays)
Fig. 5 Impact of parameters K and L on RAVLAD retrieval performance (Holidays dataset)

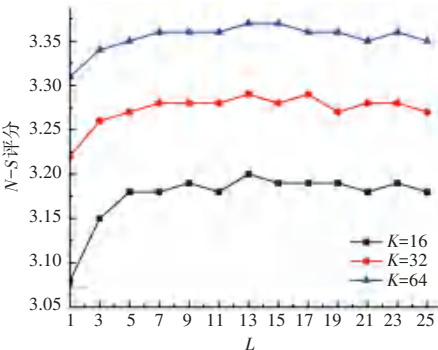


图 6 参数 K 、 L 对 RAVLAD 检索性能的影响 (UKBench)
Fig. 6 Impact of parameters K and L on RAVLAD retrieval performance (UKBench dataset)

2.3 RAVLAD 与其它描述符的比较

在 Holidays 和 UKBench 数据集上,比较 RAVLAD 与 BoW^[13]、VLAD^[14-15]、SEVLAD^[18]、EVLAD^[19]、FVLAD^[21]和 CEVLAD^[22]这些描述符在原始状态下的检索性能。其中,BoW、SEVLAD、FVLAD、CEVLAD 描述符的检索性能分别引自文献[14]、文献[18]、文献[21]、文献[22]、VLAD、EVLAD 描述符的检索性能为运行原文献提供的源码得出的结果。在 Holidays 和 UKBench 数据集上 RAVLAD 与其它描述符的检索性能比较见表 1、表 2。

由表 1 可以看出,在 Holidays 数据集上,RAVLAD 的检索精度高于其它描述符,当 $K = 64$ 时,RAVLAD 相比于 BoW、VLAD、EVLAD、FVLAD、CEVLAD 描述符的检索精度分别提高 21.8%、7.0%、1.5%、2.5%、5.0%,相比于 SEVLAD 的检索精度提高了 2.1%。

表 1 Holidays 数据集上检索性能比较

Table 1 Comparison of retrieval performance in the Holidays dataset

描述符	<i>K</i>	<i>D</i>	mAP
BoW ^[13]	20 000	20 000	0.404
VLAD ^[14-15]	64	8 192	0.552
SEVLAD ^[18]	64	8 256	0.601
EVLAD ^[19]	64	8 192	0.607
FVLAD ^[21]	64	8192	0.597
CEVLAD ^[22]	64	8 192	0.572
RAVLAD	64	8 192	0.622

表 2 UKBench 数据集上检索性能比较

Table 2 Comparison of retrieval performance in the UKBench dataset

描述符	<i>K</i>	<i>D</i>	mAP
BoW ^[13]	20 000	20 000	2.87
VLAD ^[14-15]	64	8 192	3.12
EVLAD ^[19]	64	8 192	3.31
FVLAD ^[21]	64	8 192	3.36
CEVLAD ^[22]	64	8 192	3.28
RAVLAD	64	8 192	3.37

由表 2 可以看出,在 UKBench 数据集上,RAVLAD 的检索精度也高于其它描述符,由于 SEVLAD 没有使用 UKBench 数据集进行实验,因此,只对比 RAVLAD 和其余 5 个描述符。 $K=64$ 时,RAVLAD 相比于 BoW、VLAD、FVLAD、EVLAD 描述符的检索精度分别提高 0.50、0.25、0.01、0.06,相比于 CEVLAD 的检索精度提高了 0.09。

3 结束语

针对 VLAD 在量化残差计算过程中求得的残差均值较大进而影响 VLAD 检索精度的问题,提出了残差自适应图像描述符 RAVLAD。在 RAVLAD 的生成过程中,残差基于两层码书共同生成从而使残差均值减小,针对局部特征的爆发现象,在生成 RAVLAD 的过程中采用三步归一化策略进行抑制。实验结果表明,RAVLAD 相比于诸多描述符在检索精度方面均有提升。

参考文献

[1] 杨慧,施水才. 基于内容的图像检索技术研究综述[J]. 软件导刊,2023,22(4):229-244.

[2] 兰蓉,马威. GIST 结合图像结构细节特征的现勘图像检索算法[J]. 西安邮电大学学报,2020,25(6):45-51.

[3] 谷迁,袁理,杨亚莉,等. 融合高级语义特征的色纺织物图像快速检索[J]. 纺织学报,2022,43(3):89-94.

[4] 余松森,陈晓升,苏海,等. 基于特征融合的商标检索方法[J]. 计算机工程与设计,2022,43(5):1288-1294.

[5] LOWE D G. Distinctive image features from scale-invariant

keypoints[J]. International Journal of Computer Vision, 2004, 60(2):91-110.

[6] BAY H, TUYTELAARS T, GOOL V L. SURF: speeded up robust features [C]// Proceedings of the European Conference on Computer Vision. Cham: Springer, 2008: 404-417.

[7] 于林飞,李云红,穆兴,等. 改进 PCA-SIFT 的电子标签图像匹配算法[J]. 西安工程大学学报,2020,34(2):67-72.

[8] BHAT V S, PUJARI J D. Neuro-fuzzy fusion in a multimodal face recognition using PCA, ICA and SIFT [J]. International Journal of Computational Vision and Robotics, 2016, 6(4):414.

[9] NGOBENI T, AJOODHA R. Satellite image classification using HOG and DAISY Feature Descriptors [C]//Proceedings of 2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME). Piscataway,NJ:IEEE, 2022: 1-6.

[10] XIE Y, WANG Q, CHANG Y, et al. Fast target recognition based on improved ORB feature[J]. Applied Sciences, 2022, 12(2):786.

[11] 史泽鹏,张元,杨晓文,等. 融合形状与纹理的点云局部特征描述算法[J]. 计算机工程与设计,2023,44(3):822-828.

[12] 骆俊锋. 结合高层语义与底层颜色特征的图像检索方法[J]. 电子元器件与信息技术,2021,5(3):188-190.

[13] SIVIC J, ZISSERMAN A. Video Google: A text retrieval approach to object matching in videos [C]// Proceedings of the 9th IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 2003: 1470-1477.

[14] JÉGOU H, DOUZE M, SCHMID C, et al. Aggregating local descriptors into a compact image representation [C]// Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Piscataway,NJ: IEEE, 2010: 3304-3311.

[15] JÉGOU H, PERRONNIN F, DOUZE M, et al. Aggregating local image descriptors into compact codes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(9):1704-1716.

[16] 张琳娜,梁列全,郑心炜,等. 基于最近邻量化距离聚类的残差中心聚合图像表示[J]. 数据采集与处理,2020,35(1):79-88.

[17] 龙显忠,熊健. 基于内积加权局部聚合描述子向量的图像分类[J]. 中国科技论文,2021,16(3):259-265.

[18] 孙明思,赵宏伟,赵浩宇,等. 利用空间分布熵的改进 VLAD 图像检索[J]. 光学精密工程,2021,29(1):152-159.

[19] 魏本昌,郑丽,管涛. 残差增强的图像描述符[J]. 计算机辅助设计与图形学学报,2019,31(6):1039-1045.

[20] 赵宏伟,王也然,刘萍萍,等. 利用位置信息熵改进 VLAD 的图像检索方法[J]. 哈尔滨工程大学学报,2018,39(8):1376-1381.

[21] LIU Ziqiong, WANG Shengjin, TIAN Qi. Fine-residual VLAD for image retrieval[J]. Neurocomputing, 2016, 173(P3):1183-1191.

[22] ZHOU Qiuzhan, WANG Cheng, LIU Pingping, et al. Distribution entropy boosted VLAD for image retrieval [J]. Entropy, 2016, 18(8):311-329.

[23] JÉGOU H, DOUZE M, SCHMID C. Hamming embedding and weak geometric consistency for large scale image search [C]// Proceedings of European Conference on Computer Vision. Cham: Springer, 2008: 304-307.

[24] NISTER D, STEWENIUS H. Scalable recognition with a vocabulary tree [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway,NJ: IEEE, 2006: 2161-2168.