

Sub-Window 尺度空间的 Attention-HardNet 特征匹配算法

齐向明, 冯一帆*

辽宁工程技术大学软件学院, 辽宁 葫芦岛 125105

摘要 为保护尺度空间边缘和角点信息, 提高特征匹配算法的可靠性, 提出一种 Sub-Window 尺度空间的 Attention-HardNet 特征匹配算法。该算法通过 Sub-window box filter 构建尺度空间来充分保留尺度空间图像边缘及角点信息; 使用 FAST 算法提取尺度空间特征点来提高特征点提取速度, 再利用圆形非极大值抑制算法对其进行优化, 提高准确率; 对 HardNet 特征提取网络添加 SENet 注意力机制, 构成 Attention-HardNet, 提取鲁棒性更强的 128 维浮点型特征描述符, 最后利用 L2 距离衡量不同描述符的相似性, 完成图像特征点匹配。在 Oxford 数据集上对匹配算法抗尺度、压缩、光照等性能进行测试, 由测试结果可以看出本文算法相较于常用匹配算法, 匹配正确率得到较大提升, 相较于 L2net、HardNet 等深度学习方法, 匹配正确率提高 3% 左右, 速度约提高 10%。

关键词 图像处理; sub-window 尺度空间; 圆形非极大值抑制算法; HardNet; SENet 注意力机制

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202158.2210006

Attention-HardNet Feature-Matching Algorithm in Sub-Window Scale Space

Qi Xiangming, Feng Yifan*

College of Software, Liaoning Technical University, Huludao, Liaoning 125105, China

Abstract This paper proposes an Attention-HardNet feature-matching algorithm in the sub-window scale space to protect the scale space's edge and corner information and improve the reliability of the feature-matching algorithm. First, a sub-window box filter was used to construct the scale space for fully retaining the scale-space image's edge and corner information. Second, the FAST algorithm was used to extract the scale-space feature points for increasing the speed of feature-point extraction, and the circular non-maximum suppression algorithm was used to suppress the scale-space feature points. It was optimized to improve the accuracy. Then, the SENet attention mechanism was added to the HardNet feature-extraction network to form the Attention-HardNet network, which extracted more robust 128-dimensional floating-point feature descriptors. Finally, L2 distance was used to measure the similarity of different descriptors. The image's feature-point matching was complete. On the Oxford dataset, tests on the matching algorithm's anti-scale, compression and illumination performances show that when compared with the commonly used matching algorithms, the matching accuracy rate of the proposed algorithm has been greatly improved. Compared with the deep learning methods such as L2net and HardNet, the matching accuracy rate of the proposed algorithm is increased by ~3%, and the speed is increased by ~10%.

Key words image processing; sub-window scale space; circular non-maximum suppression algorithm; HardNet; SENet attention mechanism

OCIS codes 100.2960; 100.3008; 100.3005

收稿日期: 2020-11-28; 修回日期: 2020-12-26; 录用日期: 2021-01-21

通信作者: *18342960860@163.com

1 引言

图像匹配是图像处理的重要研究方向,是三维重建^[1-2]、图像拼接^[3-4]、无人机航拍^[5-6]等图像处理领域的重要组成部分。基于局部特征点的图像匹配算法,因其根据尺度空间上提取的特征点和描述符来构建尺度空间,具有较好的尺度不变性,且与图像匹配算法的稳定性和效率密切相关,故成为目前图像匹配算法的研究热点。

国内外学者对尺度空间的构建方法做了许多研究工作。例如, Lowe^[7] 提出尺度不变特征变换(SIFT)算法,该算法通过高斯核函数构建高斯尺度空间,并通过对相邻高斯尺度空间作差构建高斯差分金字塔,用以提取特征点,该算法实时性较差。当特征点较少时,该算法对边缘光滑的目标无法准确提取特征点。在 SIFT 算法的基础上,张庆鹏等^[8] 提出一种融合彩色信息表征与光照补偿的 Harris-SIFT 算法,该方法首先在构建灰度尺度空间时引入颜色信息补偿量,增强了图像对比度,然后通过 Harris 算法提取角点,增加了匹配效率。但由于高斯卷积尺度空间浮点型运算量大,会影响尺度空间搭建的效率,故 Bay 等^[9] 提出 SURF 算法,即通过盒子滤波器快速构建尺度空间,该算法提升了匹配算法效率,但使尺度空间边缘信息丢失严重。无论是后续提出 ORB^[10]、DSP-SIFT^[11]、ASV-SIFT^[12] 等手工特征提取算法,还是基于深度学习 MatchNet^[13]、L2Net^[14]、HardNet^[15] 等方法,均是通过高斯核函数构建尺度空间。由高斯核函数构建的尺度空间效率较高,但易造成卷积图像边缘信息丢失,严重影响特征点和描述符的稳定性。

针对高斯核函数构建尺度空间不足的问题, Alcantarilla 等^[16] 提出 KAZE 算法,该算法通过 AOS 算法^[17] 求解非线性方程,构建非线性尺度空间,这一算法有效保留了尺度空间边缘信息,但解非线性方程时耗时较长,造成匹配算法效率较低。后又提出 AKAZE 算法^[18],使用 FED 算法^[19] 快速求解非线性方程,增加了匹配算法效率,但非线性尺度空间的构建效率仍然较低。在 KAZE 算法基础上,汪方斌等^[20] 提出 CKAZE 算法,该算法使用自适应扩散滤波函数保留了尺度空间图像的弱边缘,提高了算法中特征点描述向量的精度,但也降低了匹配算法的运行效率。

为充分保留尺度空间边缘信息,增加尺度空间的构建效率,以提升图像匹配算法的稳定性,提高算法的匹配速率。本文提出基于 Sub-window 尺度空间 HardNet 通道注意力网络特征匹配算法。首先通过 sub-window box filter^[21-22] 构建尺度空间;然后使用 FAST 算法^[10] 提取特征点,再通过圆形非极大值抑制算法^[23] 进行优化;最后对 HardNet 特征提取网络^[15] 引入 SENet 注意力机制^[24],命名为 Attention-HardNet,构建 128 维浮点型描述符,通过 L2 距离衡量不同描述符的相似性来完成图像匹配。

2 基本算法理论

2.1 Sub-window box filter

图像的滤波公式为

$$U(x, y) = \Phi(I(x, y)), \text{ where } (u, v) \in N(x, y), \quad (1)$$

式中 $I(x, y)$ 、 $U(x, y)$ 分别表示原始图像和滤波后的图像, $\Phi(\cdot)$ 表示滤波算法的权重函数, $N(x, y)$ 表示滤波窗口。传统滤波器的变量都集中在权重函数 $\Phi(\cdot)$ 上,权重函数 $\Phi(\cdot)$ 和滤波窗口 $N(x, y)$ 都独立于图像内容,因此不能保留边缘信息。

Sub-window box filter 是一种在图像内容上执行自适应子窗口滤波的滤波器,它使用盒子滤波结合子窗口回归来保留边缘和角点信息。权重函数 $\Phi(\cdot)$ 与原始图像 $I(x, y)$ 的内容无关,但可支持子窗口自适应图像内容。

如图 1 所示,假设图像上角像素点或者边缘像素点为 $U'(x, y)$,子窗口中心像素点位于 8 个可能区域的角点或边缘,平滑区域中像素点可以被视为同时存在于 4 个子窗口或 2 个半子窗口上,可使用子窗口回归法统一处理图像中的角点、边缘和平坦区域。8 个子窗口可表示为

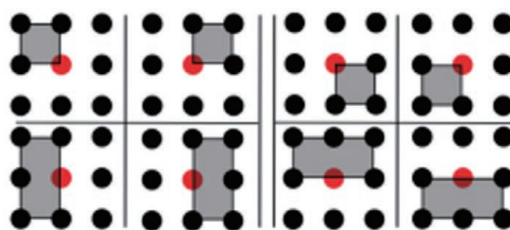


图 1 中心像素点所在位置的 8 种可能性

Fig. 1 Eight possibilities for the location of the central pixel

$$N_1(x, y) = \{(u, v) \mid x - r \leq u \leq x, y \leq v \leq y + r\}, \quad (2)$$

$$N_2(x, y) = \{(u, v) \mid x \leq u \leq x + r, y \leq v \leq y + r\}, \tag{3}$$

$$N_3(x, y) = \{(u, v) \mid x \leq u \leq x + r, y - r \leq v \leq y\}, \tag{4}$$

$$N_4(x, y) = \{(u, v) \mid x - r \leq u \leq x, y - r \leq v \leq y\}, \tag{5}$$

$$N_5(x, y) = \{(u, v) \mid x - r \leq u \leq x, |v - y| \leq r\}, \tag{6}$$

$$N_6(x, y) = \{(u, v) \mid x \leq u \leq x + r, |v - y| \leq r\}, \tag{7}$$

$$N_7(x, y) = \{(u, v) \mid |u - x| \leq r, y \leq v \leq y + r\}, \tag{8}$$

$$N_8(x, y) = \{(u, v) \mid |u - x| \leq r, y - r \leq v \leq y\}, \tag{9}$$

式中 N_1 至 N_4 可用于保护图像角点信息, N_5 至 N_8 可用于保护图像边缘信息。自适应形状只能是这 8 个子窗口中的一个, 因此不能像先前方法那样

任意改变, 这使得滤波器计算成本大幅降低。权重函数 $\Phi(\cdot)$ 在不同子窗口 $N_i(x, y)$ 是离散的, 可表示为

$$k_{1,2,3,4}(u, v) = \begin{cases} (r + 10)^{-2}, & (u, v) \in N_{1,2,3,4}, \\ 0, & \text{otherwise} \end{cases}, \tag{10}$$

$$k_{5,6,7,8}(u, v) = \begin{cases} (r + 10)^{-1}(2r + 1)^{-1}, & (u, v) \in N_{5,6,7,8}, \\ 0, & \text{otherwise} \end{cases}. \tag{11}$$

图 1 中 8 个区域可统一表示为

$$U^{(t+1)} = \operatorname{argmin}_U \iint \min_{i=1,2,\dots,8} |k_i U - U^{(t)}| dx dy, \tag{12}$$

式中 t 为迭代次数。

在 8 个核函数中, 运行并选择最接近当前值的过程为非线性各向异性扩散过程, 该过程计算像素

数时具有线性复杂度。Sub-window box filter 算法的伪代码见图 2。

```

Sub-window box filter pseudo code
Input:  $I(x, y)$ , size of filtering window  $r$ , iterations  $t$ ,
 $U^{(1)}(x, y) = I(x, y)$ 
while  $N \leq t$  do
  for  $(x, y) \in \Omega$  do
     $d_i = k_i \times U^{(t)} - U^{(t)}, i \in \{1, 2, \dots, 8\}$ 
     $m = \operatorname{argmin}_{i \in \{1, 2, \dots, 8\}} \{|d_i|\}$ 
     $U^{(t+1)}(x, y) = U^{(t)}(x, y) + d_m$ 
  end for
   $t++$ 
end while
Output: filtering the image  $U^{(t)}(x, y)$ 
    
```

图 2 Sub-window box filter 算法的伪代码

Fig. 2 Pseudocode of sub-window box filter algorithm

2.2 HardNet 特征提取网络

HardNet 共有 7 层全卷积网络结构, 每层卷积后加入批量标准化, 通过卷积步长实现降采样操作, 最后一层通过 dropout 正则化得到 128 维浮点型描

述符, 如图 3 所示。

为得到更稳定的成对匹配点, 构建 HardNet 损失函数来模仿 SIFT 算法, 此时要求成对匹配点的欧氏距离尽可能小, 与其他特征点的欧氏距离尽可

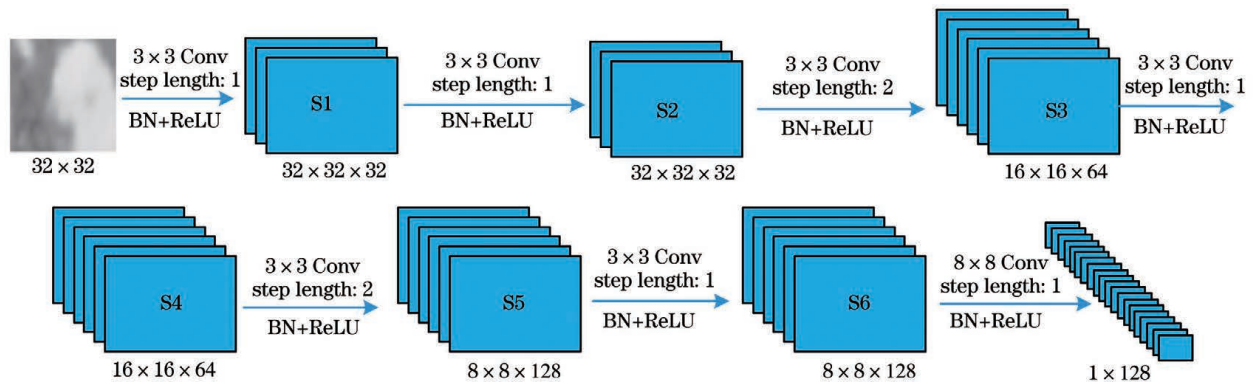


图 3 HardNet 网络结构

Fig. 3 HardNet network structure

能大。

若是训练集 χ 中共有 $2n$ 个匹配块, 则 $\chi = \{A_i, P_i\}_{i=1,2,\dots,n}$, 其中 A_i 表示锚点样本块, P_i 为正样本块。 A_i 和 P_i 样本块的 L2 距离可表示为

$$d_2(a_i, p_j) = \sqrt{\sum_{m=1}^o \sum_{n=1}^o [(a_i)_{mn} - (p_j)_{mn}]^2}, \quad (13)$$

式中: $i=1, 2, \dots, n; j=1, 2, \dots, n; a_i, p_j$ 分别为 A_i 和 P_i 样本块的描述符。以 $d_2(a_i, p_j)$ 为矩阵元构建距离矩阵 $D, D = \text{pdist}(a, p)$, 其中 pdist 表示采用 $\text{distance.pdist}(\cdot)$ 方法。

为得到最小化匹配描述符和最近非匹配描述符的距离, 可构建三元损失函数为

$$L = \frac{1}{n} \sum_{i=1,2,\dots,n} \max\{0, 1 + d_2(a_i, p_i) - \min[d_2(a_i, p_{j_{\min}}), d_2(a_{k_{\min}}, p_i)]\}, \quad (14)$$

式中, $p_{j_{\min}}$ 为描述符 a_i 最近非匹配描述符, $a_{k_{\min}}$ 是描述符 p_i 最近非匹配描述符。三元组训练集构建

过程见图 4。

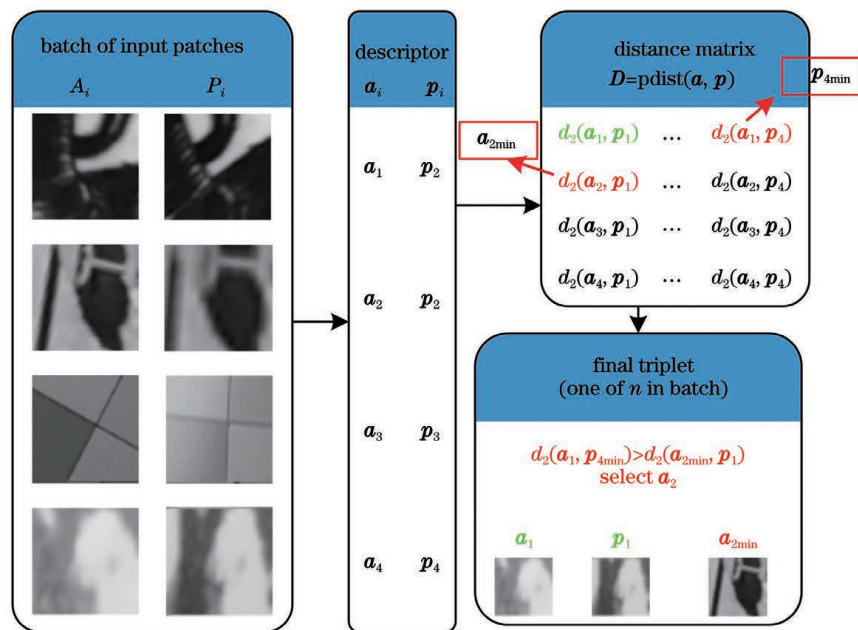


图 4 三元组训练集构建过程

Fig. 4 Construction process of triplet training set

3 特征匹配算法

为增加尺度空间信息量, 提升匹配算法鲁棒

性, 本研究提出了 Sub-window 尺度空间的 Attention-HardNet 特征匹配算法。具体步骤如下: 首先, 对两张待匹配图像构建 Sub-window box

filter 尺度空间;其次,通过 Fast 算法快速定位特征点并使用圆形非极大值抑制算法筛选特征点;然后,对 HardNet 添加 SENet 通道注意力结构,提取

局部特征描述符;最后,利用 L2 距离衡量不同描述符相似性,完成特征点匹配。具体过程如流程图 5 所示。

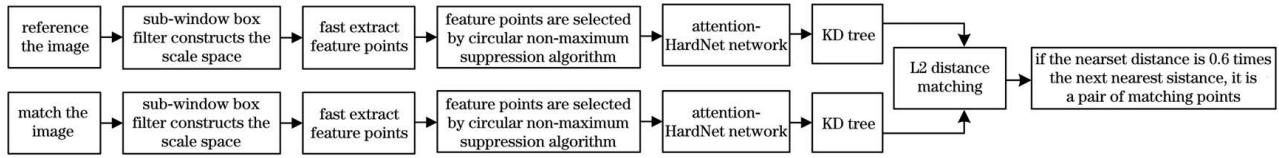


图 5 算法流程

Fig. 5 Algorithm process

3.1 尺度空间构造

为使图像匹配算法具有较好的尺度不变性,特征点提取和描述符构建需在尺度空间上进行,因此尺度空间的优劣将直接影响匹配算法的稳定性。对原图像降采样或者图像插值,可得到正金字塔或倒金字塔模型,再对金字塔进行平滑操作可得到尺度

空间。由第 2 节分析可知,Sub-window box filter 在确保高效率的同时,还具有较强的保边性。故本文提出通过 Sub-window box filter 构建尺度空间的方案。类似 SIFT 算法,本文尺度空间层数用 i' 表示,每层金字塔降采样因子用 ρ 表示,子窗口滤波器迭代次数为 t ,Sub-window box filter 尺度空间可表示为

$$L_i(x, y, t) = \operatorname{argmin}_U \iint [\min_{i=1,2,\dots,8} |k_i U_i(x, y) - U_i^{(t)}(x, y)|] dx dy, \quad (15)$$

式中, $U_i(x, y) = I(x, y) \times \rho$, $I(x, y)$ 为原始图像, ρ 为采样因子。迭代次数 t 与高斯滤波器尺度因子 σ 相同,迭代次数越大,滤波后图像的平滑程度越高。每层尺度空间降采样因子和迭代次数分别为

$$\rho_{i'+1} = 0.5 \times \rho_{i'}, \quad (16)$$

$$t_{i'} = 3 \times i', \quad (17)$$

式中 i' 代表第 i' 层金字塔。在本文中, i' 最大取值为 5,每层具体参数如表 1。

表 1 参数设置

Table 1 Parameter settings

Layer number	1	2	3	4	5
ρ	1	0.75	0.5	0.25	0.125
t	1	6	9	12	15

根据表 1 参数,构建尺度空间如图 6 所示,随着层数的增加,图像尺寸逐渐变小,图像非边缘和角点部分逐渐模糊,但边缘和角点信息保留得较好,可以为特征点检测和描述符提取保留更多细节信息。

3.2 特征点提取

SIFT 算法和 AKAZE 算法均需浮点型计算,特征点提取效率较低。为提高特征点提取速度,本研究选择无浮点型 FAST 算法来提取特征点。原 FAST 算法通过非极大值抑制算法抑制响应强度较低点,但非极大值抑制算法需多次计算相邻点的欧氏距离,存在大量的浮点型计算,降低了特征点提取

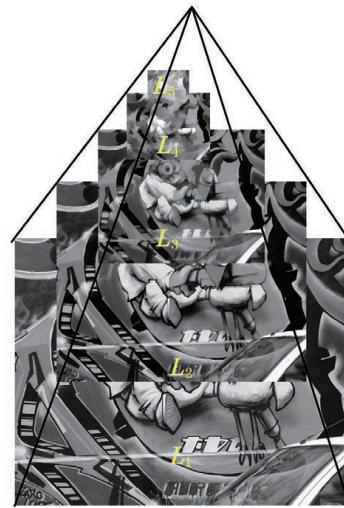


图 6 构建尺度空间

Fig. 6 Construct the scale space

效率。本节研究基于圆形特征点抑制算法,通过设置不同抑制半径和点响应强度筛选出稳定性较强的特征点。

任选一点,以 3 个像素为半径画圆,完全覆盖圆边 16 个像素点。取模板圆边 16 个像素点分别与中心点像素值 $I(p)$ 作差值运算,即

$$N = \sum_{x \in \text{circle}(p)} |I(x) - I(p)| > e_d, \quad (18)$$

其中 $\text{circle}(p)$ 表示以 p 像素点为圆心的圆, x 表示圆边某个像素点, p 为圆心像素点, e_d 表示相差阈值, $I(x)$ 表示圆边的像素值, $I(p)$ 表示中心点像素

值, N 为圆边连续大于阈值的像素点个数。根据 FAST 算法约定, 如果 N 大于 12, 则 $I(p)$ 被确定为特征点。FAST 特征点 Harris 矩阵可表示为

$$\mathbf{M}(x, y) = \begin{bmatrix} \frac{d^2 p_x}{dx^2} & p_x p_y \\ p_x p_y & \frac{d^2 p_y}{dy^2} \end{bmatrix}. \quad (19)$$

特征点 Harris 响应可写为

$$f(x, y) = \frac{\det \mathbf{M}(x, y)}{\text{tr} \mathbf{M}(x, y)} = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2}, \quad (20)$$

其中 λ_1, λ_2 为 Harris 矩阵特征值, $f(x, y)$ 为特征点响应强度。

得到关键点和响应强度后, 对关键点根据响应强度 $f(x, y)$ 进行排序, 得到集合 P , 然后对目标图像网格化, 每个网格的宽度为

$$c = \epsilon_r r / \sqrt{2}, \quad (21)$$

其中 ϵ_r 为阈值参数, r 为抑制半径。原 FAST 算法中, 非极大值抑制算法中不同点的欧氏距离, 可用点与点之间的网格数量代替, 故可实现加速。圆形非极大值抑制算法的伪代码见图 7。

Circular non-maximum suppression algorithm pseudo code

input: image size x, y ; feature point P ;

initialization: $h = x$; image grid is G_w ; width is c ; number of retained feature points is b

while $q > b$

$r = 1 + (h - 1) / 2$

for $p_i \in P$

$i++$;

if the feature point p_i in the grid is not included

preserve feature points: $P' = p_i \cup P'$

end if

$q++$

End for

$h = r - 1$

end while

output: set of feature points after screening P'

图 7 圆形非极大值抑制算法伪代码

Fig. 7 Pseudo code of circular non-maximum suppression algorithm

得到特征点后, 通过 ORB 算法构建主方向, 即提取特征采样区域的质心, 将特征点到质心的方向作为特征采样区域的主方向 θ 。

3.3 Attention-HardNet 特征提取网络

特征采样区域内的边缘和角点是区分不同特征的主要信息, 而较为平滑区域对特征采样区域信息的表达有限。为更有效地提取采样区域内重要程度较高的特征, 并抑制重要程度较低的特征, 本文提出一种具有注意力机制的特征采样网络。本文在 HardNet 中添加 SENet 结构^[24], 使 HardNet 具有通道注意力机制, 即通过学习获得每个通道的重要程度。采样区域特征是连续的, 为更好地得到特征点在特征采样区域内的上下文信息, 引入 CN 网络 (Context-Norm Net)^[25], 如图 8 所示。

先对 S1 卷积层作全局平均池化操作, 然后经过全连接层、ReLU 激活函数、全连接层和 Sigmoid 激活函数层, 得到挤压和激励 (SE) 通道注意力模块。如图 8 所示, SE 通道注意力模块输出维数为 $1 \times 1 \times 64$, 将其与 S3 层进行内积运算后, 嵌入到 HardNet 中。

在批归一化 (BN) 层前添加 CN 层, 可对特征分布进行归一化处理, 并捕捉采样区域的上下文信息。若 $o_i^{(l)} \in \mathbb{R}^{c^{(l)}}$ 是第 l 层卷积的第 i 层通道输出, $C^{(l)}$ 为第 l 层卷积, 则 CN 层归一化 $o_i^{(l)}$ 的表达式为

$$\text{CN}(o_i^{(l)}) = \frac{[o_i^{(l)} - u^{(l)}]}{\sigma^{(l)}}, \quad (22)$$

其中 $\text{CN}(\cdot)$ 为 CN 层归一化, $u^{(l)}$ 和 $\sigma^{(l)}$ 分别为小批量均值和小批量方差。

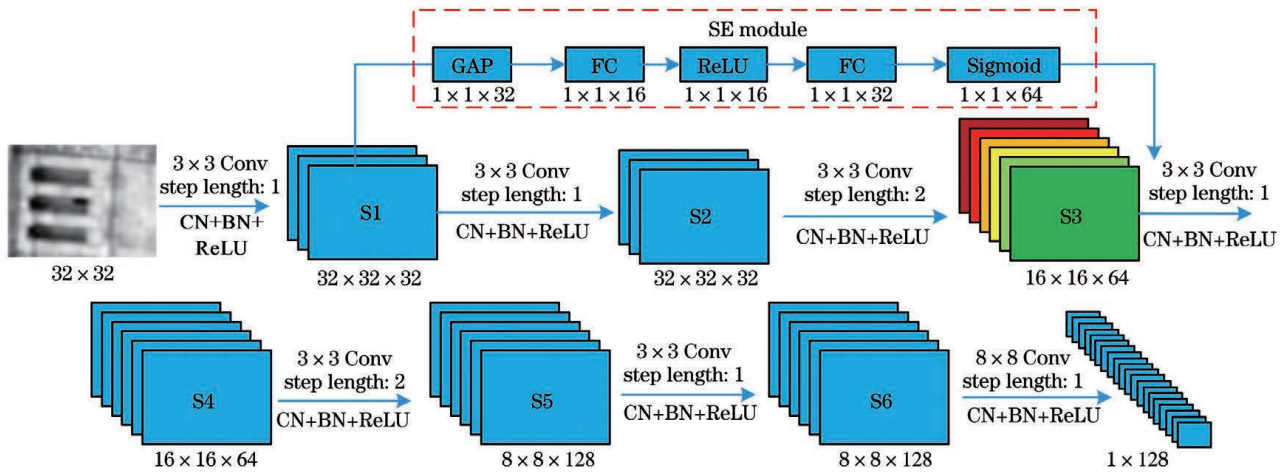


图 8 Attention-HardNet

Fig. 8 Attention-HardNet

$$u^{(l)} = \frac{1}{I} \sum_{i''=1}^I o_{i''}^{(l)}, \quad (23)$$

$$\sigma^{(l)} = \sqrt{\frac{1}{I} \sum_{i''=1}^I [o_{i''}^{(l)} - u^{(l)}]^2}, \quad (24)$$

式中 I 为小批量大小。

Attention-HardNet 损失函数, 通过 HardNet 三元损失函数构建, 正向传播得到 128 维浮点型描述符, 通过 L2 距离衡量不同描述符的相似性并通过 KD 树算法加速寻找同一对特征点, 若最近 L2

距离是次最近 L2 距离的 0.6 倍, 则该组描述符对应的特征点为同一对特征点。

4 实验设计

为验证所提算法的优越性, 选择使用 SIFT 算法、KAZE 算法、ORB 算法、L2Net 算法、HardNet 算法与本文所提算法进行对比实验, 在 Oxford 数据集对上述算法的匹配正确率和匹配时间作统计分析。图 9 为 Oxford 数据集部分图像, 可衡量匹配算



图 9 实验参考图像(左)及待匹配图像(右)。(a)Bike 图像;(b)UBC 图像;(c)Wall 图像;(d)Leuven 图像;(e)Bark 图像;(f)Boat 图像;(g)Graffiti 图像;(h)Trees 图像

Fig. 9 Experimental reference images (left) and images to be matched (right). (a) Bike image; (b) UBC image; (c) Wall image; (d) Leuven image; (e) Bark image; (f) Boat image; (g) Graffiti image; (h) Trees image

法的抗模糊变化、抗 JPEG 压缩变化、仿射不变性、光照鲁棒性、旋转和尺度变化不变性。

4.1 实验环境及模型训练

操作系统为 Ubuntu18.04, 处理器 Intel(R) Core(TM)i7-7820x, 显卡为 1080Ti, 使用 Pytorch1.5 深度学习框架训练和布置模型。

KAZE 算法、SIFT 算法和 ORB 算法均使用 Opencv3.1 库函数, 为更好地判断不同算法的鲁棒性, 上述算法提取图像特征点后, 根据 KeyPoint 的 response 值排序, 并选择前 1000 个特征点进行匹配; 本文采用的圆形非极大值算法中保留的特征点数量也为 1000。L2Net 算法和 HardNet 算法, 均通过 SIFT 算法提取特征点, 然后以特征点为中心选择 32×32 的窗口作为特征采样区域。

本文通过 HPatches 数据集^[26]训练 Attention-HardNet, 权值初始化增益为 0.6, 偏差为 0.01 正交项, 通过随机梯度下降法训练模型, 初始学习率为

0.1, 衰减系数为 0.0001, 训练样本批量大小为 512。

4.2 算法鲁棒性分析

图 9 中; Bike 和 Trees 组图像具有模糊变换, 可判断匹配算法的抗模糊变化能力; UBC 图像具有 JPEG 压缩变化, 可衡量不同算法的压缩比不变性; Wall 和 Graffito 组图像具有仿射变化, 可判断不同匹配算法的抗仿射变换能力; Leuven 组图像具有亮度变化, 可衡量算法的抗光照鲁棒性; Bark 和 Boat 组图像同时具有旋转和尺度变化, 可综合判断匹配算法抗旋转和尺度变化能力。上述实验图像每组共有 6 张, 第 1 张为参考图像, 其余 5 张为待匹配图像, 5 张待匹配图像随着编号的增加图像变化程度逐渐增加, 图中均为待匹配图像的第 5 张图像。

图 10 为 SIFT 算法、ORB 算法、KAZE 算法、L2Net 算法、HardNet 算法和本文算法用于图 9 实验中图像的模拟数据。图 10 的纵坐标为正确匹配率, 正确匹配率越大表明算法的鲁棒性越好; 横坐标

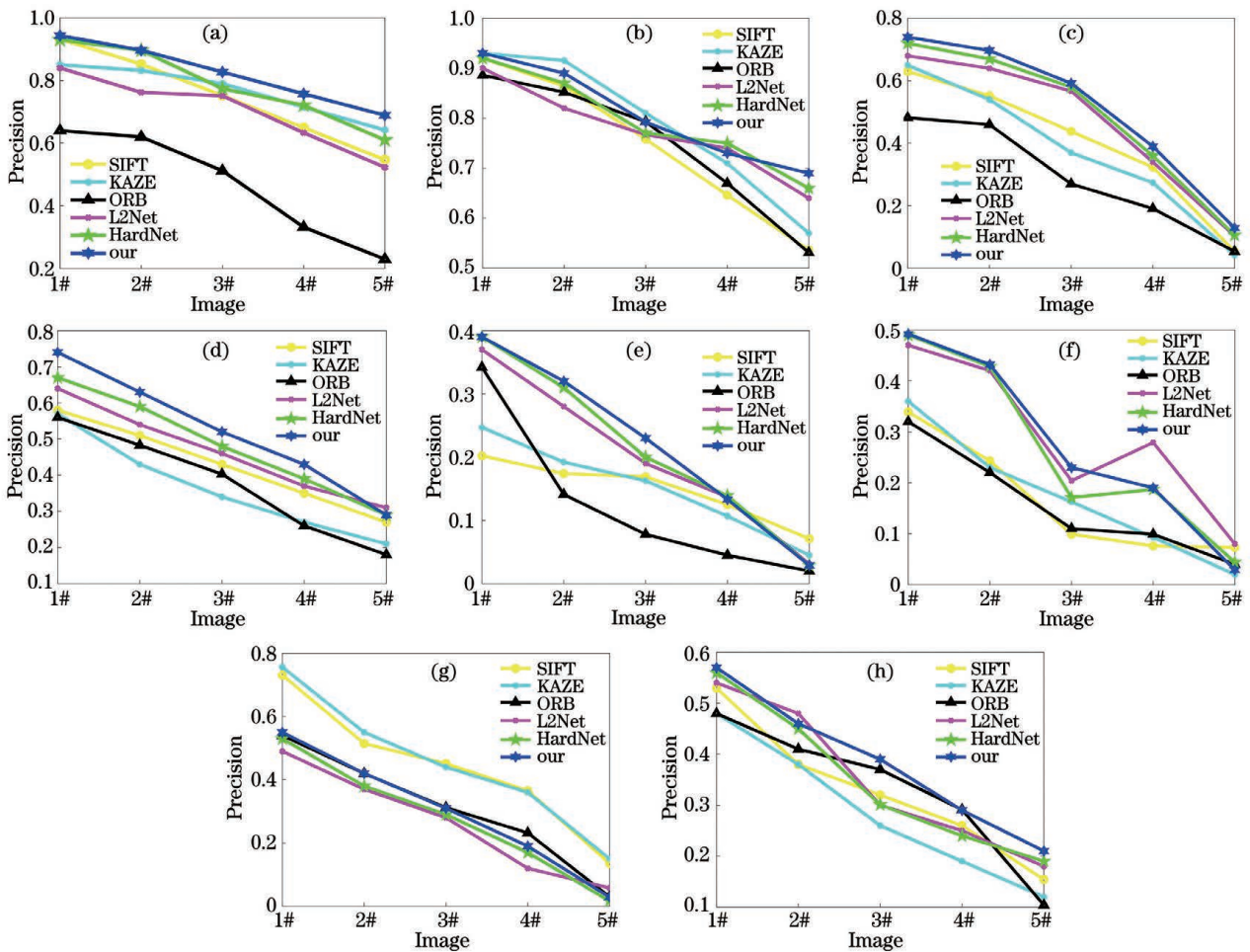


图 10 匹配正确率。(a) Bike 图像; (b) UBC 图像; (c) Wall 图像; (d) Leuven 图像; (e) Bark 图像; (f) Boat 图像; (g) Graffito 图像; (h) Trees 图像

Fig. 10 Matching accuracy. (a) Bike image; (b) UBC image; (c) Wall image; (d) Leuven image; (e) Bark image; (f) Boat image; (g) Graffito image; (h) Trees image

为待匹配图像编号,编号越大表明待匹配图像与目标图像差别越大。由图 8 可知,所提算法无论是抗模糊变化、JPEG 压缩变换,还是仿射不变性、光照鲁棒性、旋转和尺度变化,匹配正确率均远超 ORB 算法,所提算法的稳定性有大幅提升。

由图 10(a)和图 10(h)可知,对于具有模糊变化图像,所提算法匹配正确率与 HardNet、KAZE 算法相差较小,优于 L2Net 和 SIFT 算法。

从图 10(b)中 JPEG 压缩图像可知,匹配图像的压缩比较小时,上述几种算法相差较小;随着图像压缩比增加,所提算法的匹配正确率下降速率最低,最终匹配正确率仍高于 70%。

由图 10(c)和图 10(g)可知,对于具有视角变化图像,SIFT、ORB、KAZE 等算法匹配正确率下降迅速。对于 Wall 组实验图像,所提算法的匹配正确率高于其他几种算法;对于 Graffito 组实验图像,SIFT 和 KAZE 算法匹配正确率最高,所提算法的稳定性欠佳。

对于具有旋转和尺度变化的匹配图像,上述几种算法的匹配正确率均低于 0.5,但所提算法相较 ORB 和 SIFT 算法的匹配正确率高 0.1 左右;匹配图像尺度和旋转变最大时,所提算法的匹配正确率仍超过 0.1。由此分析可知,所提算法稳定性远优于 ORB 算法;相较 L2Net 和 HardNet 算法,匹配正确率有显著提升,提高了 HardNet 提取特征的能力。

图 11 为所提算法在图 9 数据集特征点的匹配效果图。对于图 9(a)、图 9(b)和图 9(d)匹配效果图,成对特征点连接线较为规整,能清晰分辨出成对特征点的具体位置;对于图 9(c)、图 9(e)、图 9(f)、图 9(g)和图 9(h)匹配效果图,连接线较为杂乱,错误匹配点数量较多。由此可知:所提算法对具有仿射变化、角度和尺度变换的图像,匹配效果较差;对具有模糊变化、JPEG 压缩变化、光照变化图像的匹配效果较好;对图 9(a)组具有模糊变化图像的匹配效果较好,但对图 9(h)组模糊变化图像的匹配效果欠佳。

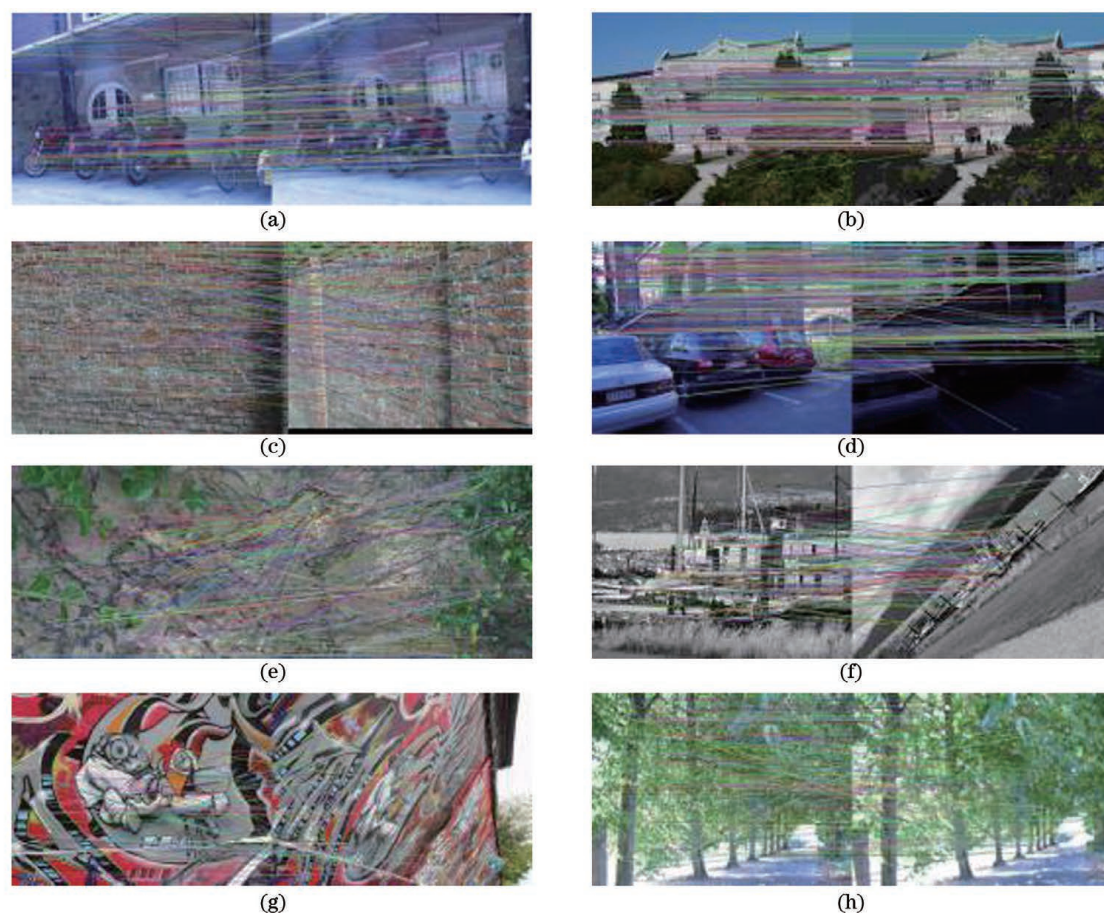


图 11 匹配效果图。(a)Bike 图像;(b)UBC 图像;(c)Wall 图像;(d)Leuven 图像;(e)Bark 图像;(f)Boat 图像;(g)Graffito 图像;(h)Trees 图像

Fig. 11 Matching renderings. (a) Bike image; (b) UBC image; (c) Wall image; (d) Leuven image; (e) Bark image; (f) Boat image; (g) Graffito image; (h) Trees image

4.3 算法效率分析

为更好地对比多种算法的效率,基于深度学习的 L2Net、HardNet 和本文所提算法均在搭载 i7-7820x 的同一设备上执行。基于局部特征点的图像匹配算法主要分为 4 个环节:金字塔构建、特征点检

测、描述符构建和特征匹配,为更好地衡量不同算法的效率,本文对图 8(b)图像匹配的各个环节耗时进行比较,如表 2。由表 2 可知:ORB 算法的匹配总耗时最短,效率最高;L2Net 算法的匹配效率最低;所提算法匹配速度优于 HardNet 和 L2Net 算法。

表 2 各个算法时间效率统计

Stage	SIFT	ORB	KAZE	L2Net	HardNet	Our
	algorithm	algorithm	algorithm	algorithm	algorithm	algorithm
Pyramid structure modeling	297.82	27.46	1671.09	297.81	297.84	150.74
Feature point extraction	187.81	5.51	138.71	187.81	187.81	5.05
Descriptor structure modeling	375.93	66.46	220.18	1972.60	1929.55	2013.27
Matching	123.83	31.64	75.10	124.87	123.94	123.83
Total	985.39	131.07	2105.08	2583.10	2539.14	2292.09

KAZE 算法通过 AOS 算法迭代求解非线性方程来构建尺度空间,耗时最长;SIFT 算法不仅需要构建高斯金字塔,还需要构建高斯差分金字塔,效率较低;ORB 算法利用不同尺度高斯核函数卷积得到尺度空间,效率最高;本文通过 Sub-window box filter 不同迭代次数构建不同尺度层,本文算法的效率仅次于 ORB 算法。

SIFT 算法不仅要和高斯差分金字塔上像素值作比较,还需使泰勒展开式精确到亚像素,耗时最长。KAZE 算法使用 Hessian 矩阵检测特征点,有大量浮点型运算,效率较低。ORB 算法和本文算法均使用 FAST 算法提取特征点,ORB 算法使用非极大值抑制算法筛选特征点,本文使用圆形非极大值抑制算法提取特征点。非极大值抑制算法需计算点与点的欧氏距离,存在大量浮点型运算,耗时较长;圆形非极大值抑制算法仅计算点与点的网格数量,均为整数加法运算,效率较高。因此本文算法提取特征点速度最快。

ORB 算法对特征采样区域内两两像素值进行编码,效率最高;SIFT 算法和 KAZE 算法均对特征采样区域提取浮点型描述符,KAZE 算法通过小波算子构建 64 维描述符,构建描述符的速度优于 SIFT 算法。L2Net 算法、HardNet 算法和本文算法均通过深度学习网络提取描述符,L2Net 和 HardNet 算法的预测模型相同,L2Net 和 HardNet 算法描述符的提取效率相差较小;本文算法在 HardNet 中引入 SE 模块和 Context-Norm 模块,特征提取效率较低。

ORB 算法的描述符为二进制浮点型,其使用汉

明距离衡量不同描述符的相似性,匹配速度最快。KAZE 算法的描述符为 64 维浮点型,匹配速度优于 SIFT、L2Net 等算法。SIFT 算法、L2Net 算法、HardNet 算法和所提算法,描述符均为 128 维浮点型,匹配速度较慢。

综上所述,ORB 算法的匹配效率最高,但鲁棒性最差;SIFT 和 KAZE 算法相似,但匹配正确率远不及本文所提算法。本文所提算法的匹配速度相较于 L2Net 和 HardNet 算法有较大幅度的提升,匹配正确率优于 L2Net 和 HardNet 算法,因此本文所提算法的匹配速度和稳定性均优于 L2Net、HardNet 算法。

5 结 论

为得到保边性较好、构建效率较高的尺度空间,本文通过 Sub-window box filter 构建尺度空间,完成了特征点匹配。通过 Sub-window box 滤波器构建尺度空间,能够保留更多的图像边缘信息,提升特征点和描述符的稳定性;通过圆形非极大值抑制算法可更高效地筛选出极值点,以提升特征点稳定性;对 HardNet 添加 SENet 注意力机制,可提取更多有效的特征信息,增加描述符的唯一性。经实验证明,所提算法鲁棒性较强,匹配速度较快,综合性能最强。再进一步的研究中,将对特征点提取方式作深入研究,旨在通过深度学习的方式得到稳定性更强的特征点。

参 考 文 献

- [1] Xiao H, Tian C, Zhang Y, et al. Stereo matching

- algorithm based on improved Census transform and gradient fusion [J]. *Laser & Optoelectronics Progress*, 2021, 58(2): 0215008.
- 萧红, 田川, 张毅, 等. 基于改进 Census 变换与梯度融合的立体匹配算法 [J]. *激光与光电子学进展*, 2021, 58(2): 0215008.
- [2] Zhang Z J, Cheng X J, Cao Y J, et al. Application of 3D reconstruction of relic sites combined with laser and vision point cloud[J]. *Chinese Journal of Lasers*, 2020, 47(11): 1110001.
- 张子健, 程效军, 曹宇杰, 等. 结合激光与视觉点云的古遗迹三维重建应用 [J]. *中国激光*, 2020, 47(11): 1110001.
- [3] Han J F, Zhang Y. Research on stitching algorithm based on UAV based on aerial photography [J]. *Laser & Optoelectronics Progress*, 2020, 57(20): 201003.
- 韩建峰, 张妍. 基于无人机航拍路面的拼接算法研究 [J]. *激光与光电子学进展*, 2020, 57(20): 201003.
- [4] Zhao X S, Chen X J, Ban Y, et al. Power function-weighted image stitching method involving improved SURF and Cell acceleration[J]. *Laser & Optoelectronics Progress*, 2020, 57(24): 241018.
- 赵潇洒, 陈西江, 班亚, 等. 融合改进 SURF 和 Cell 加速的幂函数加权图像拼接方法 [J]. *激光与光电子学进展*, 2020, 57(24): 241018.
- [5] Niu G C, Wang Y. Unmanned vehicle positioning and mapping method based on multi-constraint factor graph optimization[J]. *Journal of Beijing University of Aeronautics and Astronautics*, 2021, 47(2): 306-314.
- 牛国臣, 王瑜. 基于多约束因子图优化的无人车定位与建图方法 [J]. *北京航空航天大学学报*, 2021, 47(2): 306-314.
- [6] Li L, Gong C L, Huang H W, et al. Unmanned aerial vehicle glint image restoration algorithm based on discrete cosine transform[J]. *Acta Optica Sinica*, 2020, 40(19): 1928001.
- 李澜, 巩彩兰, 黄华文, 等. 基于离散余弦变换的无人机耀斑图像恢复算法 [J]. *光学学报*, 2020, 40(19): 1928001.
- [7] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [8] Zhang Q P, Cao Y. Image matching algorithm based on exposure and color information [J]. *Laser & Optoelectronics Progress*, 2019, 56(19): 191004.
- 张庆鹏, 曹宇. 一种融合光照和彩色信息的图像匹配算法 [J]. *激光与光电子学进展*, 2019, 56(19): 191004.
- [9] Bay H, Ess A, Tuytelaars T, et al. Speeded-up robust features (SURF) [J]. *Computer Vision and Image Understanding*, 2008, 110(3): 346-359.
- [10] Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alternative to SIFT or SURF [C] // 2011 International Conference on Computer Vision, November 6-13, 2011, Barcelona, Spain. New York: IEEE Press, 2011: 2564-2571.
- [11] Dong J M, Soatto S. Domain-size pooling in local descriptors: DSP-SIFT [C] // 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 7-12, 2015, Boston, MA, USA. New York: IEEE Press, 2015: 5097-5106.
- [12] Yang T Y, Lin Y, Chuang Y Y. Accumulated stability voting: a robust descriptor from descriptors of multiple scales [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 327-335.
- [13] Liu Z S, Li Z X, Zhang J Y, et al. Euclidean and hamming embedding for image patch description with convolutional networks [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), June 26-July 1, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 1145-1151.
- [14] Tian Y R, Fan B, Wu F C. L2-net: deep learning of discriminative patch descriptor in euclidean space [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 6128-6136.
- [15] Mishchuk A, Mishkin D, Radenovic F, et al. Working hard to know your neighbor's margins: local descriptor learning loss [C] // Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA. Trier: DBLP, 2017: 4826-4837.
- [16] Alcantarilla P F, Bartoli A, Davison A J. KAZE features [M] // Fitzgibbon A, Lazebnik S, Perona P, et al. Computer vision-ECCV 2012. Lecture notes in computer science. Heidelberg: Springer, 2012, 7577: 214-227.
- [17] Weickert J, Romeny B M T H, Viergever M A. Efficient and reliable schemes for nonlinear diffusion filtering [J]. *IEEE Transactions on Image Processing*, 1998, 7(3): 398-410.
- [18] Alcantarilla P, Nuevo J, Bartoli A. Fast explicit diffusion for accelerated features in nonlinear scale spaces [C] // Proceedings of the British Machine Vision Conference 2013, September 9-13, Bristol, UK.

- British: British Machine Vision Association, 2013: 131-142.
- [19] Grewenig S, Weickert J, Bruhn A. From box filtering to fast explicit diffusion [M] // Goesele M, Roth S, Kuijper A, et al. DAGM 2010: pattern recognition. Lecture notes in computer science. Heidelberg: Springer, 2010, 6379: 533-542.
- [20] Wang F B, Chu Z T, Zhu D R, et al. An improved KAZE feature detection and description algorithm [J]. Laser & Optoelectronics Progress, 2018, 55(9): 091007.
汪方斌, 储朱涛, 朱达荣, 等. 一种改进的 KAZE 特征检测描述算法 [J]. 激光与光电子学进展, 2018, 55(9): 091007.
- [21] Gong Y H, Liu B Z, Hou X X, et al. Sub-window box filter [C] // 2018 IEEE Visual Communications and Image Processing (VCIP), December 9-12, 2018, Taichung, Taiwan, China. New York: IEEE Press, 2018: 1-4.
- [22] Yin H, Gong Y H, Qiu G P. Side window filtering [C] // 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 15-20, 2019, Long Beach, CA, USA. New York: IEEE Press, 2019: 8750-8758.
- [23] Bailo O, Rameau F, Joo K, et al. Efficient adaptive non-maximal suppression algorithms for homogeneous spatial keypoint distribution [J]. Pattern Recognition Letters, 2018, 106: 53-60.
- [24] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7132-7141.
- [25] Yi K M, Trulls E, Ono Y, et al. Learning to find good correspondences [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 2666-2674.
- [26] Balntas V, Lenc K, Vedaldi A, et al. HPatches: a benchmark and evaluation of handcrafted and learned local descriptors [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 3852-3861.