

DOI:10.11918/202305044

仿人眼双目图像特征点提取与匹配方法

向浩鸣,夏晓华,葛兆凯,曹雨松

(道路施工技术与装备教育部重点实验室(长安大学),西安 710064)

摘要: 模仿人眼的视觉特性已成为机器迈向智能感知、智能认知的研究热点和难点。图像边缘蕴含着丰富的信息,因此人眼对场景中物体的边缘更加敏感。为在机器上实现这一视觉特性,提出了一种仿人眼双目图像特征点提取与匹配方法。首先选择边缘特征提取能力突出的 SUSAN (small univalue segment assimilating nucleus) 算子作为特征检测器;然后改进尺度不变特征变换 (scale-invariant feature transform, SIFT) 描述子的采样邻域,减少远离特征点的梯度信息因视点和视角差异带来的匹配误差,保留靠近特征点的主要梯度信息;随后对输入图像建立多尺度结构,在不同尺度上计算同一特征的主要梯度信息;最后利用平方根核比较梯度信息的相似性,生成多尺度描述子,增强描述向量的独特性。实验采用多种评价指标分别对提出的多尺度描述子和整体算法进行评估,并与经典的 SIFT、SURF (speeded up robust features)、Root-SIFT 等算法和先进的 BEBLID (boosted efficient binary local image descriptor)、SuperGlue、DFM 等算法进行对比。结果表明:提出的多尺度描述子能够提高边缘特征点的匹配准确率,并对光照变化具有更强的适应能力,体现出更好的匹配稳定性;与其他算法相比,本算法具有更高的匹配准确性。

关键词: 特征点提取;特征点匹配;仿人眼双目图像;多尺度结构;特征描述子

中图分类号: TP391 **文献标志码:** A **文章编号:** 0367-6234(2024)04-0092-09

Feature point extraction and matching method of humanoid-eye binocular images

XIANG Haoming, XIA Xiaohua, GE Zhaokai, CAO Yusong

(Key Laboratory of Road Construction Technology and Equipment (Chang'an University), Ministry of Education, Xi'an 710064, China)

Abstract: Imitating the human visual characteristics has become a research hot and challenging research topic for machines to move towards intelligent perception and intelligent cognition. Human eyes are more sensitive to edge of objects in the scene because edge contains abundant information. To realize this visual characteristic in machines, a feature point extraction and matching method of humanoid-eye binocular images is proposed. Firstly, smallest univalue segment assimilating nucleus (SUSAN) operator with outstanding edge feature extraction capability is selected as feature detector. Then, the sampling neighborhood of scale invariant feature transform (SIFT) descriptor is improved to reduce the matching error of gradient information far away from feature points due to viewpoint and view direction differences, and the main gradient information close to feature points is retained. Whereafter, a multi-scale structure is established for the input image, and the main gradient information of the same feature is computed at different scales. Finally, the square root kernel is used to compare the similarity of the gradient information, and the multi-scale descriptor is generated to enhance the uniqueness of the description vector. In the experiment, a variety of evaluation indexes are used to evaluate the proposed multi-scale descriptor and overall algorithm respectively, and compared with the classical SIFT, speeded up robust features (SURF), Root-SIFT and the advanced boosted efficient binary local image descriptor (BEBLID), SuperGlue, DFM algorithms. The results show that the proposed multi-scale descriptor improves the matching accuracy of edge feature points and has stronger adaptability to illumination changes, thereby demonstrating better matching stability. Compared with other algorithms, the proposed algorithm has higher matching accuracy.

Keywords: feature point extraction; feature point matching; humanoid-eye binocular images; multi-scale structure; feature descriptor

对人眼开展仿生学研究,是机器视觉领域的一项重要内容,将使机器具备更为有效、高级的智能感

知和认知能力。仿人眼双目图像能够模拟人左、右眼注视同一目标获取的两幅图像,在场景三维重建、

收稿日期: 2023-05-17; 录用日期: 2023-06-26; 网络首发日期: 2024-03-15

网络首发地址: <https://link.cnki.net/urlid/23.1235.T.20240315.1343.008>

基金项目: 国家自然科学基金(61901056); 中央高校基本科研业务费专项资金(300102251203); 机器人技术与系统国家重点实验室开放基金(SKLRs-2021-KF-10)

作者简介: 向浩鸣(1998—),男,硕士研究生;夏晓华(1987—),男,副教授,博士生导师

通信作者: 夏晓华, xhxia@chd.edu.cn

仿人机器人等领域具有广阔的应用前景。

在常规的双目模型中,两相机平行放置,如图1(a)所示。此时,处于重叠区域的某点在成像面上仅存在距离差。而人眼在观察物体时,双眼会形成会聚,使视线聚焦于同一点。图1(b)为简化的仿人眼双目模型,与平行式模型相比,具有更大的视野重叠区域。两相机间不仅存在距离差,还存在视向差,这使得仿人眼双目图像间的变换关系变得更为复杂,图像特征匹配的准确性受到极大挑战。

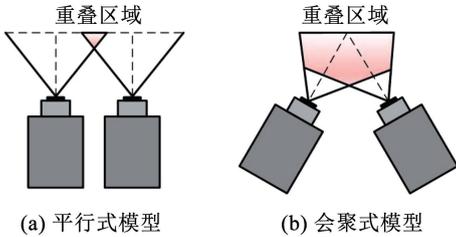


图1 双目相机模型

Fig. 1 Binocular camera model

国内外研究学者针对图像特征点提取与匹配准确性问题作了很多研究,提出了一系列算法。例如, Lowe^[1]提出尺度不变特征变换(scale-invariant feature transform, SIFT)算法,通过提取特征点邻域内的梯度信息构建描述向量,对多种图像变换都具有良好的稳定性。在此基础上, Bay等^[2]提出快速鲁棒特征(speeded up robust features, SURF)算法,对扇形滑动窗口内的图像计算 Haar 小波,并对小波模板的响应进行统计,生成 64 维描述子,极大提高了计算效率。但该方法在视角变化下的匹配准确性低于 SIFT^[3]。近年来, Dong等^[4]对特征点所属的尺度空间再次进行采样,通过池化生成 DSP-SIFT 描述子,提高了对多种图像变换的稳定性。Aguilera等^[5]在 Gabor 滤波器的基础上改进传递函数,得到的 LGHD 描述子对图像光照变化具有良好的适应性。Cao等^[6]利用递增的同心方形窗口描述特征点,得到的描述子维数较小,提高了匹配效率,但也丢失了部分可判别性。何显辉等^[7]结合颜色和纹理信息,改进 SURF 描述子,能够在视角变化的图像上取得较高的匹配精度,但得到的匹配点对数量较少。

上述方法提取的匹配特征点对多位于图像的非边缘处,而人眼更加关注图像边缘信息^[8],这意味着边缘匹配对的准确与否将极大影响人眼对图像匹配准确性的整体感观,因此需要提取更多的边缘特征点并对其进行准确匹配,以满足在机器上实现仿人视觉特性的要求。此外,由于仿人眼双目图像存在显著的视点和视角差异,匹配点在两幅图像上

的邻域像素信息会发生较大改变,增加特征点准确匹配的难度,而人眼的中心视力和周边视力能够以多种分辨率感知特征点邻域信息^[9],提高对准确匹配点对的识别能力。因此,人眼对边缘信息敏感和多分辨率感知的特性能准确对应两幅差异较大图像的公共特征,本文根据该原理提出了一种仿人眼双目图像特征点提取与匹配方法,该方法能产生较多的边缘特征点对,提高匹配准确性,对光照变化也具有有良好的适应性。

1 方法

1.1 边缘特征点提取

获得边缘匹配点对的前提是提取到足够的边缘特征点。作为一种边缘特征点检测算法, SUSAN 算子提取的特征能较为完整地描述物体轮廓信息^[10],如图 2 所示。因此,本文选择 SUSAN 算子作为特征检测器,模拟人眼注重边缘信息的视觉特性。



图2 SUSAN 检测器的特征点提取效果

Fig. 2 Feature point extraction effect of SUSAN detector

SUSAN 算子采用一个圆形模板对输入图像的像素依次进行处理,通过计算模板内像素点的灰度值与模板中心灰度值的差异,将与模板中心灰度值相似的像素点归类为同一区域,该区域称为核值相似区。判断候选点是否为角点的过程如图 3 所示,其中灰色区域为图像表面, $a \sim d$ 分别为圆形模板在图像中的不同位置。如图 3 中的 a , 当圆形模板位于角点位置时,核值相似区面积最小;当圆形模板靠近图像表面时,如图 3 中的 b 或 c , 核值相似区面积逐渐增大;当圆形模板完全处于图像内部时,如图 3 中的 d , 核值相似区面积最大。

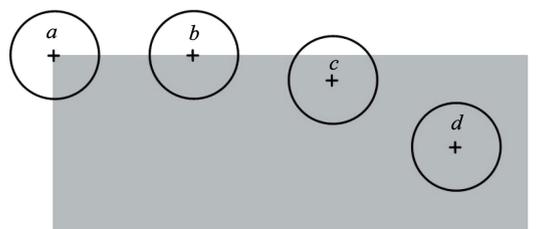


图3 核值相似区面积示意

Fig. 3 Diagram of SUSAN area

判断像素点是否属于核值相似区的公式为

$$C(r, r_0) = \begin{cases} 1, & |I(r) - I(r_0)| \leq t \\ 0, & |I(r) - I(r_0)| > t \end{cases} \quad (1)$$

式中: $I(r)$ 为圆形模板中非核像素的灰度值, $I(r_0)$ 为核的像素值, t 为相似度阈值。对模板中每个像素都进行比较后, $I(r_0)$ 处的核值相似区面积为

$$n(r_0) = \sum_{i=1}^{\max} C(r_i, r_0) \quad (2)$$

在得到每个像素点的核值相似区面积后, 将其与几何阈值 g 进行比较, 若小于 g , 则该点被认为是一个角点。角点响应可表示为

$$R(r_0) = \begin{cases} g - n(r_0), & n(r_0) < g \\ 0, & n(r_0) \geq g \end{cases} \quad (3)$$

不同的 g 值可以检测出不同的特征, 如角点、边缘等。本文将 g 值设定为 24, 以便同时提取角点和边缘特征点, 使这些特征点能较为完整地描述物体轮廓信息, 模拟人眼对边缘信息更加敏感的视觉特性。

1.2 多尺度描述子构建

SIFT 描述子对仿射变换、光照变化和噪声变化等具有良好的适应性^[11]。由于仿人眼双目图像存在显著的视点和视角差异, 因此本文对 SIFT 进行改进, 与人眼中心视力与周边视力的多分辨率感知特性相结合, 通过图像金字塔结构产生多尺度信息, 并采用平方根核度量梯度信息之间的相似性, 构建一种更为可靠的多尺度特征描述子。

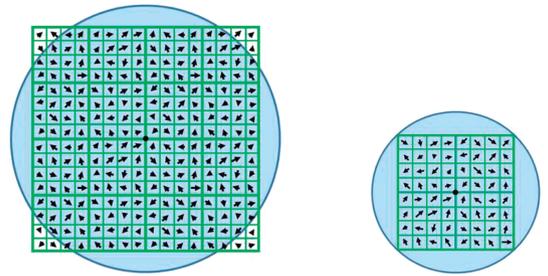
SIFT 算法在构建特征描述子时以特征点为中心, 取 16×16 的邻域并将其划分为 4×4 的子区域, 在每个子区域统计 8 方向的梯度直方图, 形成 $4 \times 4 \times 8 = 128$ 维的特征向量, 如图 4(a) 所示。由于仿人眼双目图像对之间存在较大的视点和视角差异, 同一特征点在左、右目图像中的相对位置会发生改变, 导致其周围场景发生变化。考虑到这一问题, 本文首先缩小特征点的邻域大小, 以特征点为中心取 8×8 的邻域并将其划分为 2×2 的子区域, 然后将每个子区域划分为 8 个方向并进行梯度幅值统计, 利用圆形模板对梯度幅值进行加权处理, 进一步提高靠近关键点的梯度信息对匹配的贡献度。所使用的高斯加权函数为

$$w_{i,j} = e^{-\frac{(i^2+j^2)}{2 \times 1.5^2}} \quad (4)$$

式中 i, j 为样本点与关键点之间的相对位置。最后形成 $2 \times 2 \times 8 = 32$ 维的特征向量, 如图 4(b) 所示。

改进的 32 维描述子摒弃远离特征点的梯度信息, 减少因视点和视角差异带来的梯度信息变化, 能够提高后续准确匹配的抗干扰能力。SIFT 描述子在单一尺度上计算梯度信息、生成特征向量, 对靠近

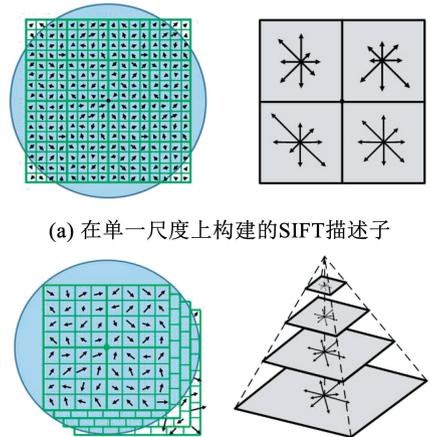
关键点的梯度信息利用有限。为进一步提高特征点的信息描述能力, 增强同名特征的匹配稳定性, 本文结合人眼中心视力的高分辨率与周边视力的低分辨率感知特性, 在 4 层图像金字塔的每一层图像上对同名特征点计算改进的 32 维描述子, 并按照尺度大小进行拼接, 生成 $4 \times 32 = 128$ 维, 具有多尺度信息的特征描述子, 获取的特征点描述信息更加全面, 能够提高对主要梯度信息的利用率。两种描述子的构建方式如图 5 所示。



(a) SIFT 描述子的 16×16 邻域 (b) 改进描述子的 8×8 邻域

图 4 邻域改进示意

Fig. 4 Diagram of improved neighborhood area



(a) 在单一尺度上构建的 SIFT 描述子

(b) 在不同尺度上构建的多尺度描述子

图 5 构建多尺度描述子

Fig. 5 Multi-scale descriptor construction

此时得到的描述子以欧式距离作为相似性度量, 而用平方根核比较梯度相似性时得到的匹配效果更好^[12]。因此本文使用平方根核衡量多尺度特征向量之间的相似性。其基本思想是对特征向量的每个元素进行 L1 范式归一化, 并开平方根, 以达到用平方根核替代原先的特征向量的目的。

设 \mathbf{x} 和 \mathbf{y} 是具有欧式距离规范的 n 维特征向量, 则 \mathbf{x}, \mathbf{y} 之间的欧式距离为

$$d_E(\mathbf{x}, \mathbf{y})^2 = 2 - 2S_E(\mathbf{x}, \mathbf{y}) \quad (5)$$

式中 $S_E(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y}$, 表示 \mathbf{x} 和 \mathbf{y} 之间的欧拉距离与它们的相似性。 \mathbf{x} 和 \mathbf{y} 的平方根核定义为

$$H(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \sqrt{x_i y_i} \quad (6)$$

设 \mathbf{m} 和 \mathbf{n} 为多尺度描述子的特征向量,首先对 \mathbf{m} 、 \mathbf{n} 进行 L1 规范化,再对每个元素求平方根,即

$$S_E(\sqrt{\mathbf{m}}, \sqrt{\mathbf{n}}) = \sqrt{\mathbf{m}^T \mathbf{n}} \quad (7)$$

故有 $S_E(\sqrt{\mathbf{m}}, \sqrt{\mathbf{n}}) = H(\mathbf{m}, \mathbf{n})$, 而 $S_E(\sqrt{\mathbf{m}}, \sqrt{\mathbf{m}}) = 1$, 因此可得到 L2 范式的多尺度描述向量。此时,用欧式距离比较多尺度描述子就相当于用平方根核替代原先的特征向量,即

$$d_E(\sqrt{\mathbf{m}}, \sqrt{\mathbf{n}})^2 = 2 - 2H(\mathbf{m}, \mathbf{n}) \quad (8)$$

1.3 特征点匹配

采用双向匹配的方式对得到的多尺度描述子进行初步匹配,具体步骤为:

Step1 将仿人眼左、右目图像中提取的特征点集合分别记为 P 、 Q 。

Step2 首先将 P 作为基准集合,遍历其中所有的点,以欧氏距离作为相似性准则,找到 Q 中与其对应的匹配点,记为集合 $Z\{P \rightarrow Q\}$ 。

Step3 然后采用相同的策略,以 Q 作为基准集合,遍历其中所有的点,搜索其在 P 中的对应匹配点,记为集合 $W\{Q \rightarrow P\}$ 。

Step4 比较集合 Z 和 W 中的所有匹配点对,若某匹配点对在两个集合中存在相同的对应关系,则这两点被认为是一对正确匹配点对。

此时得到的初步匹配点集中仍会不可避免地存在部分误匹配对,因此本文采用随机抽样一致算法(RANdom SAmple consensus, RANSAC)^[13]对匹配点对再次进行筛选。RANSAC 算法通过迭代计算的方式在初步匹配点集中寻找满足某一正确模型对应的最大内点集,能够有效地剔除误匹配点对。

2 实验与分析

为充分说明本文算法在提高匹配准确率和匹配精度方面的优势,对提出的多尺度描述子和整体算法分别进行评价。首先将多尺度描述子与 SIFT、SURF、Root-SIFT^[12]、LGHD 以及 BEBLID^[14] 描述子进行对比分析;然后在不同光照变化的仿人眼双目图像上对多尺度描述子的适应性进行验证;最后对整体算法进行单应性估计实验(homography estimation),并与 GMS^[15]、RIFT^[16]、SuperGlue^[17] 和 DFM^[18] 算法进行对比。所有实验均在 Windows 10, 64 位操作系统,基于 Intel Core i5-6300HQ、16 G 内存的测试环境下进行。

2.1 描述子对比实验

为体现多尺度描述子的匹配性能,将其与 SIFT、SURF、Root-SIFT、LGHD 以及 BEBLID 描述子

进行性能比较。SIFT 算法通过统计邻域梯度信息生成描述向量。SURF 算法通过计算水平和垂直方向的 Haar 小波响应获取描述信息,这两种描述子因对多种图像变换都具有良好的鲁棒性而得到广泛应用。Root-SIFT 算法使用海林格距离替代欧式距离以比较 SIFT 描述子之间的相似性,能够产生更多的准确匹配点对。LGHD 算法通过构建多尺度和多方向的 Log-Gabor 滤波器计算频率和空间信息来描述特征点邻域信息,对光照变化具有良好的适应性。BEBLID 算法作为近年来先进的机器学习算法之一^[19],利用积分图像计算平均灰度值之差,能够显著提高图像匹配精度。

2.1.1 评价指标

为进行定量比较,本文采用文献[20-21]中的评价指标对多尺度描述子的性能进行量化比较,具体如下。

1) 假定匹配率(putative match ratio, PMR)为初步匹配点对数量与左、右目图像中提取较少特征点数量的比值,表示算法的选择性,假定匹配率越低说明算法的限制性越强,即

$$P_p = \frac{N_m}{\min(N_{\text{left}}, N_{\text{right}})} \quad (9)$$

式中: P_p 为假定匹配率, N_m 为初步匹配点对数量, N_{left} 、 N_{right} 分别为左、右目图像中提取的特征点数量。

2) 正确匹配数(number of correct matches, NCM)表示最终得到的匹配点对数量。本文以经过 RANSAC 算法筛选后得到的匹配点对数量作为正确匹配点对数量,即

$$N_c = N_r \quad (10)$$

式中: N_c 为正确匹配点对数量, N_r 为经过 RANSAC 算法筛选后得到的匹配点对数量。

3) 可重复率(repeatability, REP)为正确匹配点对数量与左、右目图像中提取较少特征点数量的比值,可重复率越高说明算法的准确率越高,即

$$P_r = \frac{N_c}{\min(N_{\text{left}}, N_{\text{right}})} \quad (11)$$

式中 P_r 为可重复率。

2.1.2 实验过程

实验选取 6 组仿人眼双目图像进行特征提取与匹配,测试图像如图 6 所示,图 6 中的圆圈表示仿人眼双目视线会聚点。其中:第 1、2 组实验图像分别表示平面场景和三维场景,第 3、4 组实验图像分别表示近距离场景和远距离场景,第 5、6 组实验图像分别表示简单场景和复杂场景。为充分体现描述子的匹配性能,在特征提取阶段均采用 SUSAN 检测器,匹配方式和筛选过程均保持一致,图 7 展示了不同描述子对比的实验结果。



图 6 仿人眼双目图像

Fig. 6 Humanoid-eye binocular image

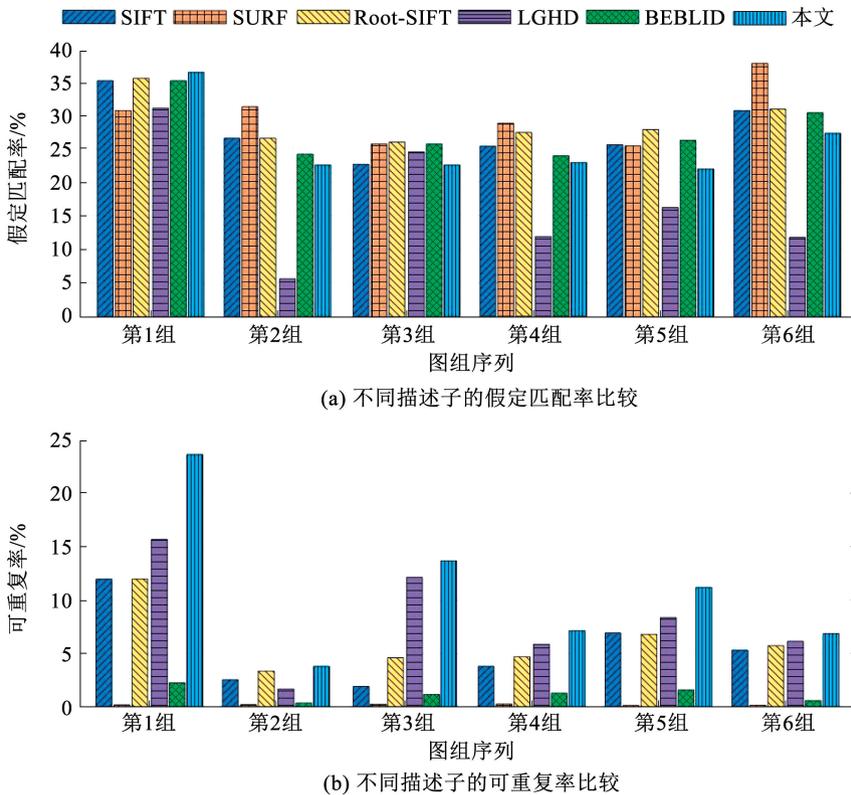


图 7 不同描述子的性能比较结果

Fig. 7 Performance comparison results of different descriptors

从图 7 可以看出,本文算法的假定匹配率略低,却能取得最高的可重复率,说明本文提出的多尺度描述子能够有效增强对特征点描述的独特性,降低不相似特征点间的误匹配率,从而减少误匹配点对数量。这主要是由于在构建描述子时结合人眼中心视力与周边视力的多分辨率感知特性,采用多尺度

结构计算特征点邻域梯度信息,得到的信息量更加全面,并能充分利用特征点周围的主要梯度信息,提高特征点初始匹配时的抗干扰能力,限制不可靠特征点之间的匹配能力,以达到有效剔除误匹配点对的目的。从图 7(b)可以看出,本文提出的多尺度描述子在 6 组实验图像上取得的平均可重复率为

11.31%,比 SIFT 描述子高出 5.62%,比第 2 位的 LGHD 描述子和第 3 位的 Root-SIFT 描述子分别高出 2.88%、5.03%,表明多尺度描述子在 SIFT 描述子的基础上能够显著提高匹配准确率,较其他算法相比同样具有明显优势。

为验证本文提出的多尺度描述子在光照变化下具有一定的适应能力,选取了 5 组存在 3 种光照强度的仿人眼双目图像进行实验,如图 8 所示,图 8 中的圆圈表示仿人眼双目视线会聚点。图 9 展示了其中一组图像的正确匹配和误匹配结果。



图 8 3 种光照强度变化下的仿人眼双目图像

Fig. 8 Humanoid-eye binocular images under three kinds of illumination intensity changes



图 9 3 种光照强度变化下不同描述子的匹配效果

Fig. 9 Matching effect of different descriptors under three kinds of illumination intensity variations

由图 9 可知, SURF 算法和 BEBLID 算法在图像存在较大的视点和视角差异时易出现误匹配。SIFT 算法和 Root-SIFT 算法对视点和视角变化具有一定

的鲁棒性,但随着光照强度减弱,其重复率也会随之降低,说明其对图像光照变化的适应能力较差。LGHD 作为一种适用于非线性强度变化图像之间的

特征点匹配算法,能够在图像存在光照变化时保持较为稳定的重复率,但得到的正确匹配点对的数量有限。相较于上述算法,本文算法能够取得数量最多的正确匹配点对,并且在光照强度降低时仍能保持较为稳定的重复率,说明对图像光照变化具有良好的适应性。

为了在图像存在光照强度变化时对算法进行定量比较,对正确匹配数和可重复率这两种直接反映算法匹配性能的指标在不同光照强度下的实验结果进行平均化,计算结果见表 1。

表 1 光照强度变化下 6 种算法的平均计算结果

Tab.1 Average calculation results of six algorithms under illumination intensity variations

算法	正确匹配数/对	可重复率/%
SIFT	286	8.93
SURF	11	0.36
Root-SIFT	302	9.41
LGHD	416	13.31
BEBLID	85	2.55
本文	458	14.59

由表 1 可知,在不同光照强度变化下,本文算法可取得数量最多的正确匹配点对和最高的可重复率。与 SIFT 算法相比,本文算法得到的正确匹配点对数量增加 60.14%,可重复率提高 5.66%,说明本文提出的多尺度描述子与 SIFT 描述子相比具有更强的光照变化适应能力。相较于第 2 位的 LGHD 算法和第 3 位的 Root-SIFT 算法,本文算法得到的正确匹配点对数量分别增加 10.10%、51.66%,可重复率分别提高 1.28%、5.18%,说明本文算法不仅能提高存在视点和视角差异图像的匹配准确率,而且对图像光照变化具有良好的适应性。

最后,本文对不同描述子的运行效率进行了对比分析,表 2 为 6 种描述子在图 6 所示的仿人眼双目图像上的平均运行时间。

表 2 6 种算法的平均运行时间

Tab.2 Average running time of six algorithms

算法	大小	运行时间/s
SIFT	128f	265.08
SURF	64f	145.18
Root-SIFT	128f	264.25
LGHD	384f	130.23
BEBLID	512b	49.13
本文	128f	265.24

注:f为浮点型描述子,b为二进制描述子。

由表 2 可知, SURF、LGHD 和 BEBLID 算法的

运行时间较短,但其产生的正确匹配点对数量较少、可重复率也较低。而与本文描述子维数相同的 SIFT 和 Root-SIFT 算法的匹配效果远低于本文算法。与这两种算法相比,本文算法的运行时间花费略高,这是由于 SIFT 和 Root-SIFT 算法只在单一尺度上构建描述子,而本文则在 4 个不同尺度上构建描述子,这会不可避免地增加匹配耗时。

2.2 总体算法对比实验

为进一步说明本文算法在图像存在视点和视角差异下的匹配准确性,将所提出的总体算法与 GMS、RIFT、SuperGlue 和 DFM 算法进行对比。GMS 算法以运动估计为基础,是一种具有代表性的特征匹配过滤算法,在实验中采用其默认的 ORB 算法进行特征点提取与匹配。RIFT 算法使用相位一致性替代图像强度来检测特征点,并通过 Log-Gabor 滤波器构建用于特征描述的最大索引图,不仅提高了特征检测的稳定性,也克服了梯度信息对特征描述的限制性。SuperGlue 和 DFM 是两种先进的基于深度学习的特征匹配算法,前者利用图神经网络对特征点进行匹配,在实验中采用其默认的 SuperPoint 算法提取特征点及其描述子;后者使用预训练的 VGG 网络提取特征点,对待匹配图像应用由粗到细的策略实现特征点对的准确匹配。

2.2.1 评价指标

本文对上述算法进行了单应性估计实验以比较各自的匹配性能。采用文献[21]中的精确率作为评价指标,表示准确估计值与所有估计值的比值。具体来说,本文利用 RANSAC 算法推导出仿人眼双目图像对之间的变换模型,并将其与真实投影模型进行比较,若误差小于 3 像素,则认为该单应性估计准确。

2.2.2 实验过程

选取 5 组仿人眼双目图像进行单应性估计实验,实验结果见表 3。

表 3 仿人眼双目图像单应性估计的定量比较

Tab.3 Quantitative comparison of homography estimation in humanoid-eye binocular images

图像	精确率/%				
	GMS	RIFT	SuperGlue	DFM	本文
图组 1	86.40	91.50	59.81	82.39	85.14
图组 2	80.59	95.76	48.19	76.95	98.15
图组 3	78.95	85.71	48.57	82.12	99.18
图组 4	76.68	94.95	76.13	85.78	98.03
图组 5	71.27	95.70	67.13	84.69	97.93
平均值	78.78	92.72	59.97	82.39	95.69

由表3可知,本文算法在除图组1外的4个图组中均取得了97.00%以上的精确率。与其他4种算法相比,本文算法的平均精确率最高,比第2位的RIFT算法和第3位的DFM算法分别高出2.97%、13.30%,说明本文算法相较于其他4种常规算法和基于深度学习的特征点匹配算法在匹配准确性方面具有明显的优势。

为了对具体的投影误差进行量化比较,实验过

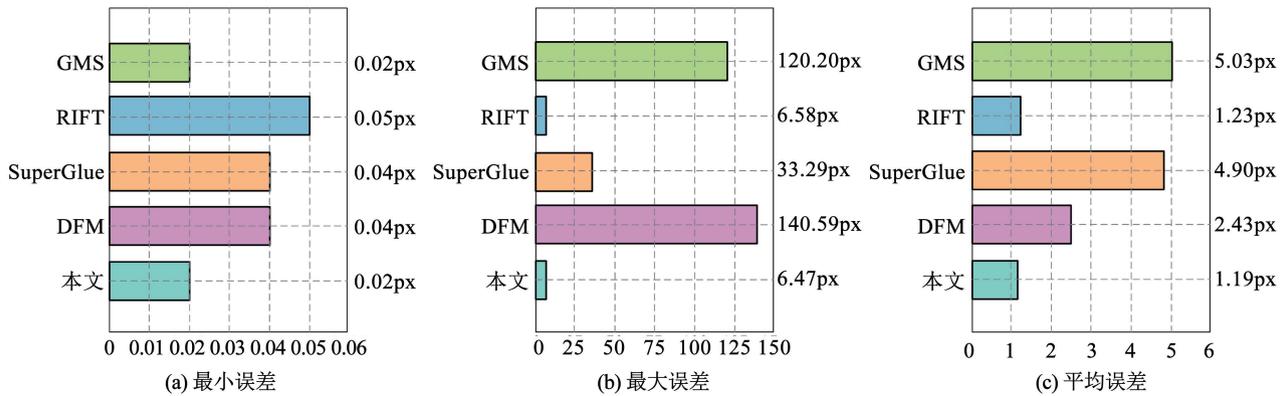


图10 5种算法的平均单应性估计误差

Fig. 10 Average estimated homography error of five algorithms

3 结论

1)结合人眼对边缘信息敏感和多分辨率感知的特性,以平方根核作为相似性度量,构建了一种多尺度描述子。

2)实验结果表明:与相关描述子相比,多尺度描述子的可重复率至少提高2.88%,对光照强度变化也具有较好的适应性;与相关总体算法相比,本文算法的匹配精确率至少提高2.97%,并能使距离误差达到最小。

3)本文算法能够提取较多的边缘特征点对,增强描述子的独特性,限制不可靠特征之间的匹配能力,实现视点和视角差异较大的仿人眼双目图像的准确匹配。这对于推动仿生视觉技术在智能机器上的应用具有重要的意义。

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. DOI: 10.1023/B:VISI.0000029664.99615.94

[2]BAY H, ESS A, TUYTELAARS T, et al. Speeded-up robust features (SURF)[J]. Computer Vision and Image Understanding, 2008, 110(3): 346. DOI: 10.1016/j.cviu.2007.09.014

[3]LUO Juan, GWUN O. A comparison of SIFT, PCA-SIFT and SURF [J]. International Journal of Image Processing, 2009, 3(4): 143

[4]DONG Jingming, SOATTO S. Domain-size pooling in local descriptors: DSP-SIFT [C]//2015 IEEE Conference on Computer

程中对相关算法的单应性估计误差的最小值、最大值以及平均值进行比较,图10展示了这3种距离误差在5个图组中的平均计算结果。可以看出,本文算法的最小误差、最大误差和平均误差均低于其他4种算法,这表明本文算法能够取得较为准确的匹配结果,提高对存在视点和视角变化图像的匹配稳定性。

Vision and Pattern Recognition (CVPR). Boston, MA: IEEE, 2015: 5097. DOI: 10.1109/CVPR.2015.7299145

[5]AGUILERA C A, SAPPA A D, TOLEDO R. LGHD: A feature descriptor for matching across non-linear intensity variations [C]//2015 IEEE International Conference on Image Processing (ICIP). Quebec City, QC, Canada: IEEE, 2015: 178. DOI: 10.1109/ICIP.2015.7350783

[6]CAO Yu, PANG Bo, LIU Xin, et al. An improved Harris-SIFT algorithm for image matching [C]//International Conference on Advanced Hybrid Information Processing. Cham: Springer, 2018: 56. DOI:10.1007/978-3-319-73317-3_8

[7]何显辉,王凯,张平,等.融合颜色和纹理的多特征匹配算法[J].激光杂志,2022,43(3):87

HE Xianhui, WANG Kai, ZHANG Ping, et al. Multi feature matching algorithm integrating color and texture [J]. Laser Journal, 2022, 43(3): 87. DOI: 10.14016/j.cnki.jgzz.2022.03.087

[8]徐翰文,张闯,陈苏婷.基于视觉显著性的图像特征提取算法[J].计算机应用,2022,42(Sup.2):72

XU Hanwen, ZHANG Chuang, CHEN Suting. Image feature extraction algorithm based on visual saliency [J]. Journal of Computer Applications, 2022, 42(Sup.2): 72. DOI:10.11772/j.issn.1001-9081.2022040574

[9]STEWART E E M, VALSECCHI M, SCHÜTZ A C. A review of interactions between peripheral and foveal vision [J]. Journal of Vision, 2020, 20(12): 2. DOI: 10.1167/jov.20.12.2

[10]SMITH S M, BRADY J M. SUSAN: A new approach to low level image processing [J]. International Journal of Computer Vision, 1997, 23(1): 45. DOI: 10.1023/A:1007963824710

[11]丁国绅,乔延利,易维宁,等.基于光谱图像空间的改进SIFT特征提取与匹配[J].北京理工大学学报,2022,42(2):192

DING Guoshen, QIAO Yanli, YI Weining, et al. Improved SIFT feature extraction and matching based on spectral image space [J].

- Transactions of Beijing Institute of Technology, 2022, 42(2): 192. DOI: 10.15918/j.tbit1001-0645.2020.024
- [12] ARANDJEIOVIĆ R, ZISSERMAN A. Three things everyone should know to improve object retrieval [C]//2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Providence, RI: IEEE, 2012: 2911. DOI: 10.1109/CVPR.2012.6248018
- [13] FISCHLER M A, BOLLES R C. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography [M]//Readings in Computer Vision. Amsterdam; Elsevier, 1987: 726. DOI: 10.1016/b978-0-08-051581-6.50070-2
- [14] SUÁREZ I, SFEIR G, BUENAPOSADA J M, et al. BEBLID: Boosted efficient binary local image descriptor [J]. Pattern Recognition Letters, 2020, 133: 366. DOI: 10.1016/j.patrec.2020.04.005
- [15] BIAN Jiawang, LIN Wenyan, LIU Yun, et al. GMS: Grid-based motion statistics for fast, ultra-robust feature correspondence [J]. International Journal of Computer Vision, 2020, 128(6): 1580. DOI: 10.1007/s11263-019-01280-3
- [16] LI Jiayuan, HU Qingwu, AI Mingyao. RIFT: Multi-modal image matching based on radiation-variation insensitive feature transform [J]. IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society, 2019, 29: 3296. DOI: 10.1109/TIP.2019.2959244
- [17] SARLIN P E, DETONE D, MALISIEWICZ T, et al. SuperGlue: Learning feature matching with graph neural networks [C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA: IEEE, 2020: 4937. DOI: 10.1109/CVPR42600.2020.00499
- [18] EFE U, INCE K G, AYDIN ALATAN A. DFM: A performance baseline for deep feature matching [C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Nashville, TN: IEEE, 2021: 4279. DOI: 10.1109/CVPRW53098.2021.00484
- [19] MATALOV D P, LIMONOVA E E, SKORYUKINA N S, et al. Memory efficient local features descriptor for identity document detection on mobile and embedded devices [J]. IEEE Access, 2022, 11: 1104. DOI: 10.1109/ACCESS.2022.3233463
- [20] WANG Zhengbing, FENG Xugang, XU Guili, et al. A robust visible and infrared image matching algorithm for power equipment based on phase congruency and scale-invariant feature [J]. Optics and Lasers in Engineering, 2023, 164: 107517. DOI: 10.1016/j.optlaseng.2023.107517
- [21] XIA Yifan, JIANG Jie, LU Yifan, et al. Robust feature matching via progressive smoothness consensus [J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2023, 196: 502. DOI: 10.1016/j.isprsjprs.2023.01.016

(编辑 张 红)

封面图片说明

封面图片来自本期论文“分形粗糙面的主轴承弹流脂润滑分析”，是西安建筑科技大学机电工程学院李玲教授课题组为探究弹性流体动压润滑(弹流润滑)状态下RV减速器主轴承接触表面之间的润滑特性，将复杂RV减速器主轴承结构模型进行简化的过程。本研究基于润滑脂的非牛顿特性以及粗糙表面分形理论，提出一种主轴承的点接触弹流脂润滑数值模型。首先，对该模型进行数值求解，得到脂膜压力、脂膜厚度的分布规律；然后，将其分别与其他点接触弹流润滑模型的数值结果和实验结果进行对比，验证了所建模型的正确性；最后，分析了主轴承表面光滑和粗糙状态下流变指数、分形维数、卷吸速度、载荷和润滑脂黏度对润滑性能的影响。研究表明：润滑脂的非牛顿性越明显，脂膜厚度越小且颈缩现象愈加不明显，接触区附近脂膜压力越符合赫兹压力分布，二次压力峰逐渐消失；考虑主轴承分形粗糙表面的弹流润滑特征更切合实际，增大分形维数，接触区真实接触面积增大，有利于降低脂膜压力，增加脂膜厚度；卷吸速度、载荷和润滑脂黏度对脂膜厚度分布的影响显著，对脂膜压力分布的影响较小；脂膜厚度以及最小膜厚越大，主轴承接触区脂膜不易破坏且越容易形成动压润滑。

(图文提供:李玲,魏高帅,王晶晶,林红,李丽霞,解妙霞.西安建筑科技大学机电工程学院)