

面向大规模 SFM 的快速鲁棒特征跟踪算法*

曹明伟,李书杰,贾伟,刘晓平

(合肥工业大学计算机与信息学院 合肥 230009)

摘要:运动恢复结构(SFM)是指通过分析二维图像序列恢复三维结构信息的过程,在计算机视觉的多种应用中起着重要的作用。特征跟踪是大规模 SFM 的核心组成部分,但现有的多视图特征跟踪算法在鲁棒性和效率上还存在不足,为解决这一问题,提出了一种快速和鲁棒的特征跟踪(FRFT)算法。首先,采用 AGAST 进行特征点检测,并使用图像矩为 AGAST 特征定义主方向,为构造旋转不变的描述子奠定基础;其次,在差分高斯金字塔空间内,根据中心点与邻域像素之间的差值构造特征描述子,避免光照和尺度变化对特征匹配的影响;再次,为了提高特征匹配效率,对特征集合进行聚类,采用 KD-Tree 加速特征匹配,提高算法的时间效率;最后,采用 4 种方式对 FRFT 算法进行验证,并与现有经典算法进行比较。实验结果表明,FRFT 算法在鲁棒性和时间效率方面均优于现有经典算法。

关键词:运动恢复结构;三维重建;摄像机跟踪;特征匹配;特征聚类

中图分类号: TP391.41 TH741.1 文献标识码: A 国家标准学科分类代码: 510.4050

Fast and robust feature tracking algorithm for large-scale SFM

Cao Mingwei, Li Shujie, Jia Wei, Liu Xiaoping

(School of Computer and Information, Hefei University of Technology, Hefei 230009, China)

Abstract: Structure from motion (SFM) refers to a process in which the 3D structure is created by analyzing 2D image sequences, which is very important in many applications of computer vision. Feature tracking is one of the core components of large-scale SFM. However, the robustness and time efficiency of the existing algorithms are needed to be improved. To address these issues, a fast and robust feature tracking (FRFT) algorithm is presented. Firstly, images moments are used to define a main direction for AGAST feature point, which can help to construct a rotation invariance descriptor. Secondly, in the space of the difference of Gaussian, the difference between the center point and its neighbor points is used to construct a descriptor for the OAGAST keypoint, which can avoid the influence of illumination and scale change on the feature matching. Thirdly, to improve the time efficiency of feature matching, the large feature set is clustered to some small ones, and KD-Tree is used to accelerate feature matching for improving the time efficiency of FRFT. Finally, the proposed method is evaluated with four ways, and compared with the state-of-the-art methods. Experimental results show that the proposed FRFT method outperforms the state-of-the-art ones on robustness and time efficiency.

Keywords: structure from motion; 3D reconstruction; camera tracking; feature matching; feature clustering

1 引言

运动恢复结构(structure from motion, SFM)^[1-2]是指通过二维图像序列恢复三维结构信息的过程,在计算机视觉的多种应用中起着重要的作用。例如,SFM 已经被广泛应用于三维重建^[3-5]、增强现实^[6]、三维地图构建^[7]、

3D 变化检测^[8]、图像填充^[9]、显著性检测^[10]和无人驾驶^[11]等领域。目前,已有的比较知名的 SFM 系统有 Bundler^[2]、ACTS^[12]、LS-ACTS^[13]和 VSFM (Visual SFM)^[14]等。在 SFM 系统中,特征跟踪是重要组成部分,是 SFM 研究的重点内容之一,在过去 10 年中,受到广泛关注,并取得了较大研究进展^[15-17]。

现有的 SFM 系统大多采用基于特征检测与匹配框

架(detection and matching framework, DMF)的特征跟踪方法,例如 Bundler 采用尺度不变特征变换(scale-invariant feature transform, SIFT)^[18]检测特征点和生成描述子,并采用强力匹配(brute force matching, BFM)算法进行特征匹配。ETH-3D^[19]采用加速的鲁棒特征(speed up robust feature, SURF)算法检测特征点和计算描述子,并采用 K -邻近(K -nearest neighbor, KNN)对特征进行匹配。VSFM 和 LS-ACTS 系统采用 GPU 加速版本的 SIFT 检测特征点,从而提高系统的时间效率。根据文献[20-21]所述,在大规模数据集上现有特征跟踪算法的鲁棒性和实时性还需进一步提高。为了弥补现有特征跟踪算法在鲁棒性和效率上的不足,在 DMF 框架下,本文提出一种快速鲁棒的特征跟踪(fast and robust feature tracking, FRFT)算法。FRFT 算法由以下部分组成。

1) 提出一种带有方向的特征检测算子。鉴于自适应通用的加速分割测试(adaptive and generic accelerated segmented test, AGAST)^[22]特征检测算法具有较高重复性,即图像经过变换后仍能够检测出较多的特征点;此外,相比于 SIFT 和带有方向的 FAST 特征点与旋转不变的二进制鲁棒独立元素特征(binary robust independent elementary features, BRIEF)^[23]等算法,AGAST 还具有较高

的时间效率;因此,采用 AGAST 算法检测特征点。但是 AGAST 特征点不具有方向性,难以构造旋转不变的描述子,本文采用图像矩^[24]为 AGAST 特征点定义一个主方向,提出具有方向的 AGAST(oriented AGAST, OAGAST)特征。

2) 提出一种特征描述子。为了避免光照和尺度变换对特征匹配的影响,在差分高斯空间内,根据中心点与邻域像素之间的差值(difference between center and neighbors, DCN)构造特征描述子,即 DCN 描述子。

3) 提出一种基于聚类的多视图特征匹配方法。采用 AP(affinity propagation)^[25]聚类算法对特征集合进行聚类,为每类特征集合构造一棵 KD-Tree 加速特征匹配。

2 本文算法

为了满足大规模 SFM 系统的需求,在 DMF 框架基础之上提出一种快速鲁棒的特征跟踪算法-FRFT 算法,FRFT 算法的流程如图 1 所示,对于给定的图像序列,采用 OAGAST 算法检测特征点,然后使用 DCN 算法计算特征描述子;在特征匹配阶段首先采用 AP 算法对特征集合进行聚类,然后为每类特征集合创建一棵 KD-Tree 加速特征匹配,最终生成 Tracks。

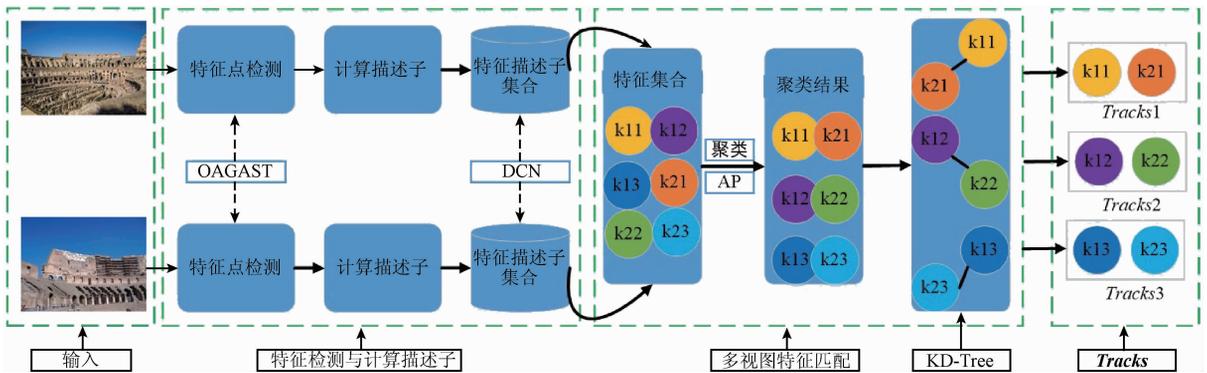


图 1 FRFT 算法流程

Fig. 1 The flowchart of the FRFT algorithm

2.1 特征检测与计算描述子

2.1.1 OAGAST 特征检测子

为了弥补 AGAST 特征不具有主方向性,容易造成特征的错误匹配^[26],采用图像矩为 AGAST 特征点定义一个主方向,提出 OAGAST 特征。

对于给定的特征点的 $O(x, y)$, 计算它的邻域 $N_o(x, y)$ 的图像矩 $m_{p,q}$, 计算方法如式(1)所示。

$$m_{p,q} = \sum_{x,y} x^p y^q N_o(x, y) \quad (1)$$

根据图像矩理论,采用一阶矩 m_{10} 和 m_{01} 以及零阶矩 m_{00} 计算 $N_o(x, y)$ 的质心 C 。

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (2)$$

由摄像机的成像原理可知,在图像区域内像素是非均匀分布,因此 $N_o(x, y)$ 的质心 C 与几何中心 O 不在同一位置,从 O 到 C 可以构造一条向量 \vec{OC} (如图 2 箭头),将向量 \vec{OC} 与 X 轴之间的夹角定义为特征点 O 的主方向,角度 θ 的计算方法如式(3)所示。

$$\theta = \tan^{-1}(m_{01}, m_{10}) = \tan^{-1} \left(\frac{\sum_{x,y} y N_o(x, y)}{\sum_{x,y} x N_o(x, y)} \right) \quad (3)$$

通过以上过程,可以计算图像的 OAGAST 特征点。

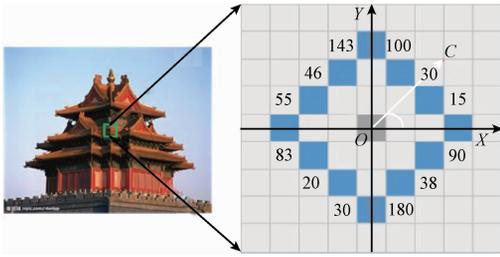


图2 OAGAST 特征
Fig.2 OAGAST Feature

2.1.2 DCN 描述子

提出 DCN 特征描述子,此特征描述子在差分高斯空间内,采用中心像素与 7×7 邻域内像素之间的差值作为描述子元素,从而构造 48 维的浮点型描述子,对光照和尺度变化具有较强的鲁棒性。DCN 描述子的计算过程如图 3 所示。

对于给定的特征点 $O(x,y)$,根据它的主方向 θ 可以计算出旋转矩阵 R ,由旋转矩阵 R 可以获得特征点 $O(x,y)$ 的旋转邻域 $R_o(x,y)$ (图 3(c) 所示为图 3(b) 中矩形区域扩大 10 倍后的结果),计算方法如式(4)所示。

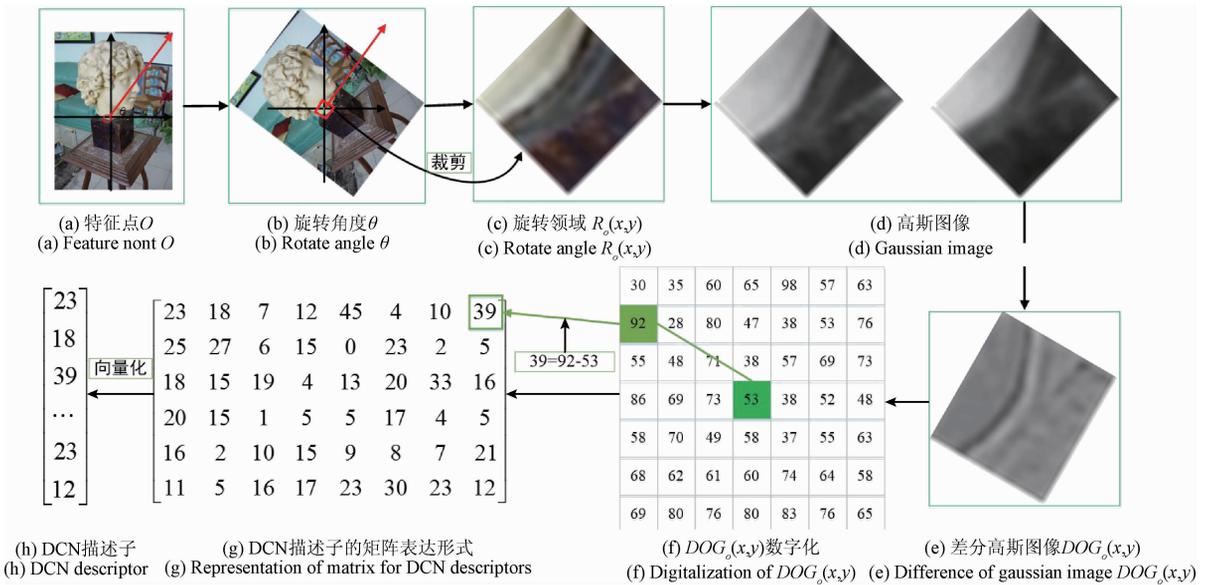


图3 DCN 描述子的计算过程

Fig.3 The computing pipeline of the DCN descriptor

$$R_o(x,y) = N_o(x,y) * R \tag{4}$$

由于差分高斯(difference of Gaussians, DOG)对光照和尺度变化具有较强的鲁棒性,因此,计算 $R_o(x,y)$ 的差分高斯图像 $DOG_o(x,y)$ (图 3(e) 是图 3(d) 中两个高斯图像的差值),计算方法如式(5)所示。

$$DOG_o(x,y) = R_o(x,y) \cdot (G(x,y,k\sigma) - G(x,y,\sigma)) \tag{5}$$

式中: $G(x,y,\sigma) = \frac{1}{2\pi\sigma} \exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$ 表示方差为 σ 的高斯密度函数; k 为常量,用于控制 DOG 尺度空间变化。

在差分高斯空间内,采用 $O(x,y)$ 与 $DOG_o(x,y)$ 内像素之间的差值作为 DCN 描述子;图 3(f) 为 $DOG_o(x,y)$ 数字化表达形式,以行为单位计算每个元素与中心点之间的差值,获得差值矩阵 M ,计算方法如式(6)所示。

$$M_i(x,y) = |O(x,y) - DOG_i(x,y)| \tag{6}$$

式中: $DOG_i(x,y) \in DOG_o(x,y)$ 。

以行为单位将差值矩阵 M 串联为 48 位的浮点向量,即为 DCN 描述子,计算方法如式(7)所示。

$$DCN(i) = hstack(M_i(x,y))^T \tag{7}$$

式中: $hstack(*)$ 表示相邻元素水平首尾连接。

由于 DCN 描述子是在 DOG 空间内,采用中心像素与邻域像素之间的差值所构造,因此,能够有效避免光照和尺度变化的影响,同时有效解决 BRIEF 等描述子的奇异性问题。

2.1.3 描述子匹配

由于 DCN 描述子元素是浮点型数值,采用与 SIFT 类似的度量方法,即采用 L2 距离计算描述子之间的相似性,如式(8)所示。

$$Dist_{l2}(DCN_1, DCN_2) = \|DCN_1 - DCN_2\|_2 \tag{8}$$

2.2 多视图特征跟踪

多视图特征跟踪的目的是生成 *Tracks*。从 *Tracks* 的定义和计算过程对本文的多视图特征跟踪方法进行详

细分析。

2.2.1 Tracks 的定义

Tracks 表示不同图像上相互匹配的特征点组成的特征跟踪链,如图 4(a)所示,对于给定的 5 幅图像,每幅图像上的不同颜色的圆点表示特征点,目的是生成图 4(e)所示的 **Tracks**。记 I_U 表示图像序列集合, I_i 表示单幅图

像,即 $I_i \in I_U$; $F_{i,U}$ 表示第 i 幅图像的特征点集合, $f_{i,j}$ 表示第 i 幅图像上的第 j 个特征点,即 $f_{i,j} \in F_{i,U}$ 。**Tracks** 的第 j 条特征跟踪链的定义如式(9)所示。

$$Tracks_j = \{f_{1,j}, f_{2,j}, \dots, f_{n-1,j}, f_{n,j}\} \quad (9)$$

由 **Tracks** 的定义和图 4(a)所示的匹配实例可知,对于规模较大的图像序列集合 I_U ,多视图特征跟踪极其耗时。

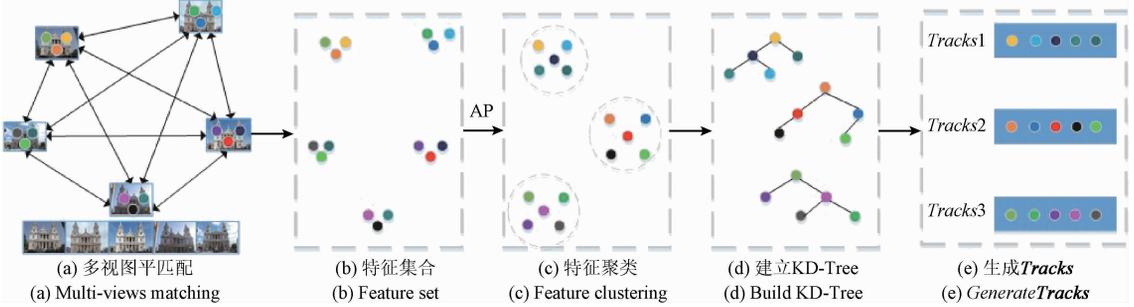


图 4 Tracks 的计算过程

Fig. 4 The procedure of Tracks

2.2.2 Tracks 的计算方法

对于给定的特征集合 F_U ,采用 AP 算法将 F_U 分为若干类,计算法方法如式(10)所示。

$$\{F_{i,U}\} = AP(F_U) \quad (10)$$

式中: AP 表示聚类算法; $\{F_{i,U}\}$ 表示被分类的特征集合, i 的个数由 AP 算法自动确定。

根据描述子的模长为每类特征集合建立一棵 KD-Tree,所需步骤如下。

1) 对于给定的 $F_{i,U}$,计算每个 DCN 描述子的模长 L_{den} 。

$$L_{den}^i = \|DCN_j\|_2 \quad (11)$$

2) 以模长最大的描述子 DCN_{max} 作为根节点,建立一个节点的 KD-Tree,如式(12)所示。

$$KDTree_i = \{DCN_{max}\} \quad (12)$$

3) 插入新的特征描述子 DCN_{j+1} ,如果 $L_{den}^{j+1} > L_{den}^j$,则将 DCN_{j+1} 作为 DCN_{max} 的右子树,插入方法如式(13)所示。

$$KDTree_i = \{DCN_{max} \{right\} DCN_{j+1}\} \quad (13)$$

反之,则将 DCN_{j+1} 作为 DCN_j 的左子树,插入方法如式(14)所示。

$$KDTree_i = \{DCN_{max} \{left\} DCN_{j+1}\} \quad (14)$$

4) 重复步骤 3),直到 $F_{i,U}$ 中每个描述子均被处理完毕上述步骤,即可为 $F_{i,U}$ 创建一棵 $KDTree_i$ 。

对于给定的特征描述子 DCN_j ,从 $KDTree_i$ 中寻找与之相匹配的特征描述子 DCN_{j+1} ,构造特征跟踪链 $Tracks_i$,计算方法如式(15)所示。

$$Tracks_i = \begin{cases} input, & (find(KDTree_i, DCN_j)) \\ donothing, & 其他 \end{cases} \quad (15)$$

重复特征跟踪链 $Tracks_i$ 的创建和检索过程,即可为输入特征集合 F_U 构造完整的 **Tracks**。

$$Tracks = \{Tracks_1, \dots, Tracks_n\} \quad (16)$$

本文提出的特征跟踪算法实现步骤如下。

输入:图像序列 I_1, I_2, \dots, I_N

输出:特征跟踪链 $Tracks_i, i \in [1, M]$

步骤 1:采用带有方向的差分高斯特征算法计算特征点 $OAGAST_i$ 和计算描述子 DCN_i ,创建特征集合 F_U ;

步骤 2:采用 AP 算法对特征集合 F_U 进行聚类,构造分类的特征集合 $\{F_{i,U}\}$,计算方法如式(14)所示;

步骤 3:为每类特征集合 $F_{i,U}$ 创建一棵 KD-Tree,计算方法如式(13)和(14)所示;

步骤 4:对于给定的特征 DCN_j ,从 $KDTree_i$ 中查找相匹配的描述子 DCN_{j+1} ,构造特征跟踪链 $Tracks_i$,计算方法如式(15)所示;

步骤 5:重复步骤 3)~4),直到 $\{F_{i,U}\}$ 内的每类特征集合均被处理完毕,即可生成 **Tracks** 如式(16)所示。

3 运动推断结构系统

将 FRFT 算法应用于全局 SFM (global structure from motion, GSFM) 系统中,主界面如图 5 所示。主要步骤如下。

1) 对于给定的图像序列,采用本文的 ODCN 算法检测特征点和计算描述子。

2) 采用 FRFT 算法对步骤 1) 生成的特征集合进行跟踪,并采用交叉验证 (cross checking, CC)^[26] 算法去除错误匹配,生成 *Tracks*。

3) 以 *Tracks* 作为输入,采用 5 点^[27] 算法计算摄像机的相对运动姿态,即旋转矩阵 R 和平移向量 t 。

4) 根据 *Tracks* 和摄像机姿态 $[R | t]$,采用三角化^[28] 算法生成初始点云模型。

4) 采用并行化的集束优化 (parallelized bundle adjustment, PBA) 算法^[29] 对初始点云模型进行优化,最终生成精确的三维点云模型。



图5 运动恢复结构系统的结果界面

Fig. 5 The result interface of the SFM system

4 实验结果与分析

FRFT 算法的开发环境为 Visual C++ 2013、Nvidia CUDA SDK 7.5 和 OpenCV SDK 2.4.11,所用 PC 配置为

Intel i5 CPU, 8.0GB RAM。实验分别在 Oxford benchmark^[30]、Building^[31] 数据集、OTB^[32] 数据集上对 FRET 算法进行验证。

4.1 在 Oxford 数据集上测试

在 Oxford 数据集上采用正确率 P 对每类特征的性能进行评估,计算方法如式(17)所示。

$$P = \frac{N_c}{N} \quad (17)$$

式中: N 表示原始匹配数, N_c 表示正确匹配数。

在 Oxford 数据集上对 ODCN 等算法的匹配精度进行测试,结果如图 6 所示(每个算法结果的第 1 列表示 barks 序列上的精度,其他依次按照 boat、graf、leven、trees、ubc 和 wall 排列)。在 barks 序列上 AB 算法、FB 算法和 MO 算的精度最低;ODCN 算法的精度最高,为 88%。在 bikes 序列上 BRISK 算法的精度最低,为 37%;ODCN 算法的精度最高,为 82%。在 boat 序列上,MO 算法的精度最低,为 9%;精度最高的是 ODCN 算法,为 78%。在 graf 序列上,精度最低的是 AB 算法,为 7%;精度最高的是 ODCN 算法,为 55%。在 leven 序列上,MO 算法的精度最低,为 72%;ODCN 算法的精度最高,为 91%。在 trees 序列上,BRISK 算法的精度最低,为 45%;精度最高的是 ODCN 算法,为 75%。在 ubc 序列上,MO 算法的精度最低,为 80%;ODCN 算法的精度最高,为 95%。在 wall 序列上,BRISK 算法的精度最低,为 52%;精度最高的是 ODCN 算法,为 72%。从以上分析可知,在 Oxford 数据集的每类测试序列上,ODCN 算法的精度最高,由此可知本文提出的 ODCN 算法具有较强的鲁棒性。

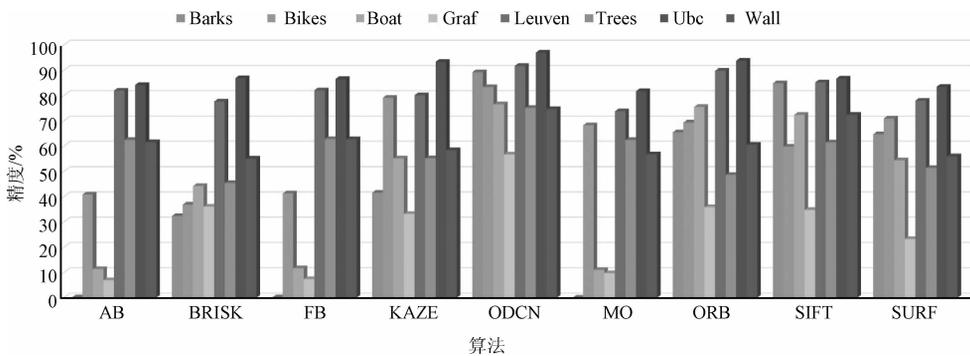


图6 在 Oxford 数据集上测试特征的精度

Fig. 6 The evaluated precision of each algorithm on the Oxford dataset

为了评估每类特征的时间效率,统计出在 Oxford 数据集上每类特征的时间性能(包含特征点检测、描述子生成和特征匹配),结果如表 1 所示。在 Oxford 数据集上,除 ubc 序列外,ODCN 算法的效率均为最

高;经过分析发现 MB 算法在 ubc 序列上检测出的特征点较少,在计算描述子和匹配阶段节省了大量时间,但是,这也从侧面反映 MB 算法的不足,特征的重复性太低。

表 1 在 Oxford 数据集上测试特征的时间效率

Table 1 The evaluated computational cost of each algorithm on the Oxford dataset

序列名	算法								
	AB	BRISK	FB	KAZE	ODCN	MB	ORB	SIFT	SURF
barks	12.306	0.701 8	3.684 9	2.262 2	0.554 1	1.107 1	1.506 8	7.077 4	2.477 7
bikes	3.855 1	0.707 5	1.179 3	3.693 4	0.602 4	0.615 1	0.782 7	6.144 1	3.279 1
boat	29.840 6	1.062 6	7.237 6	7.791 0	0.752 3	0.790 3	2.106	12.450 3	4.262 3
graf	6.007	0.786 9	2.009	4.778 1	0.753 4	1.135 6	1.899 3	6.755 5	3.644 6
leuven	5.872 6	0.727 5	2.072 8	3.104	0.689 3	0.701 6	1.221 1	5.178 7	2.815
trees	78.694 4	1.725 1	14.638 7	14.661 2	1.158 7	1.383 4	2.282 1	20.751 3	6.463 6
ubc	24.680 3	0.947 5	6.036	4.414 3	1.825 4	1.580 9	2.044 8	9.809	3.919
wall	85.174 8	1.430 4	19.852 2	7.571 8	1.385 2	1.924 3	2.243 2	19.127 5	7.200 9

4.2 在 Building 数据集上测试

在 Building 数据集上对提出的 FRFT 特征跟踪算法进行测试,并与 KLT (Kanade Lucas Tomasi)、Bundler 和 ETH-3D 进行比较,结果如图 7 所示。



图 7 在 Building 数据集上测试结果

Fig. 7 The evaluated results on the Building dataset

从测试结果可以看出, KLT 算法检测出的特征点数较少。Bundler 和 ETH-3D 的结果存在大量的错误匹配。FRFT 不但能够检测出较多的特征点,而且相邻图像之间的特征匹配较为精确,不存在 Bundler 和 ETH-3D 那样较多的错误匹配。统计出在 Building 数据集上每个特征跟踪算法的成功率 (ratio of success, ROS), 计算方法如式 (18) 所示, 结果如图 8 所示。

$$ROS = \frac{\#N}{N_{1,2}} \quad (18)$$

式中: #N 表示每个 Tracks 的正确匹配数, $N_{1,2}$ 表示第 1 幅图像与第 2 幅图像之间的正确匹配数。

从如图 8 可以看出 FRFT 算法的成功率最高, 为 93.68%。此外, 统计出每个特征跟踪算法的时间效率, 结果如图 9 所示, 可以看出 FRFT 算法的效率最高, 平均处理时间为 16.37 s; 在 Building 数据集上测试结果表明, 本文提出的 FRFT 算法具有较高的实时性和鲁棒性。

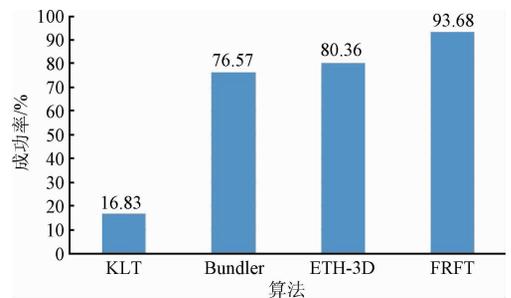


图 8 在 building 数据集上测试特征跟踪算法的成功率

Fig. 8 The ROS of each feature tracking algorithm on the building dataset

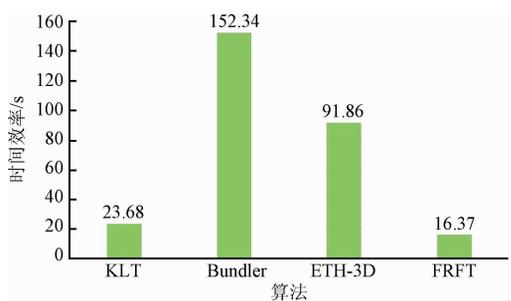


图 9 在 building 数据集上测试特征跟踪算法的时间效率

Fig. 9 The time efficiency of each feature tracking algorithm on the building dataset

4.3 在 OTB 数据集上测试

为了验证 FRFT 算法在室外环境下的鲁棒性, 在 OTB 数据集上选取具有光照变化的 david 序列、具有噪声的 walking 序列和具有部分遮挡的 women 序列对 FRFT 等算法进行测试, 结果如图 10 所示。可以看出, FRFT 算法能够有效避免遮挡、光照变化等影响, 跟踪到的特征点

数较多,且不存在交叉匹配。因此,从实验结果可知,FRFT算法在室外环境下具有较强的鲁棒性。

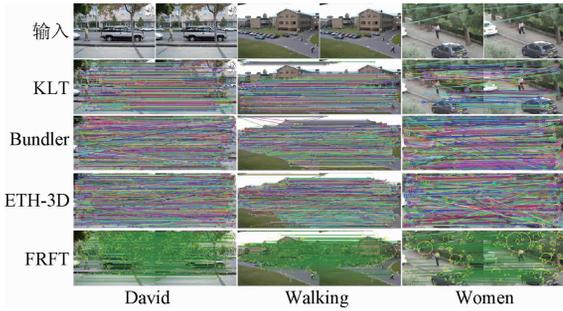


图10 在 OTB 数据集上测试结果

Fig. 10 The evaluated results on the OTB dataset

此外,统计了在 OTB 数据集上每个特征跟踪算法的成功率和平均时间效率,结果如图 11 和 12 所示。由图 11 可知,KLT 算法的成功率最低,为 57.38%;Bundler 的成功率为 72.64%;ETH-3D 的成功率为 78.23%;FRFT 算法的成功率最高,为 91.85%。在时间效率方面,FRFT 算法效率最高,平均处理时间为 11.68 s;KLT 算法的平均处理时间为 18.73 s;ETH-3D 的平均处理时间为 78.65 s;Bundler 的时间效率最低,平均处理时间为 120.47 s。因此,在 OTB 数据上测试结果表明,FRFT 算法在成功率和时间效率方面优越于 KLT 等经典算法。

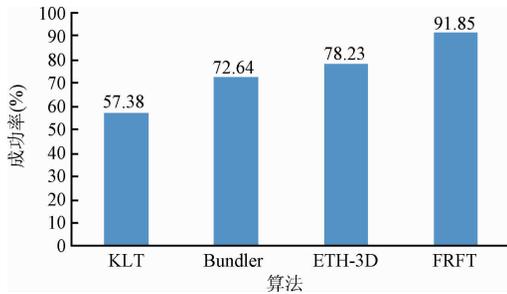


图 11 在 OTB 数据集上测试特征跟踪算法的成功率

Fig. 11 The ROS of each feature tracking algorithm on the OTB dataset

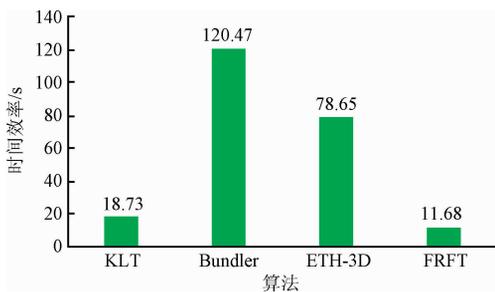


图 12 在 OTB 数据集上测试特征跟踪算法的时间效率

Fig. 12 The time efficiency of each feature tracking algorithm on the OTB dataset

4.4 SFM 的重建结果对比

为了验证 FRFT 算法在大规模 SFM 系统中的有效性,将其应用于图 5 所示的 GSFM 系统。大规模主要体现在两个方面:1)图像的分辨率高;2)数据量较大。根据文献[20]所述,在大规模 SFM 系统下,特征跟踪算法有效性的验证主要体现在高分辨率图像下是否能够快速有效地检测和匹配特征点,以便 SFM 系统重建出精确的点云模型。

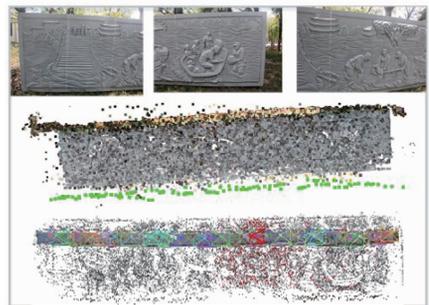
采用高分辨率的 Brieuc、Diboan 和 Wall 数据集对 GSFM 系统进行验证,同时与 VSFM^[14] 系统进行比较,结果如图 13 所示。在 Brieuc 和 Diboan 序列的结果中,第 1 列为输入样本,第 2 列为 GSFM 系统结果,第 3 列为 VSFM 系统结果。在 Wall 序列中,第 1 行为输入样本,第 2 行为 GSFM 系统结果,第 3 行为 VSFM 系统结果。从图 13 可以看出,GSFM 系统能够产生稠密的点云模型,且没有出现残缺;VSFM 系统重建出的点云模型相对稀疏。



(a) Brieuc



(b) Diboan



(c) Wall

图 13 SFM 系统重建结果

Fig. 13 The reconstructed results of the SFM system

4.5 关于多视图特征跟踪的讨论

根据本文对特征跟踪算法的评估和开发 GSFM 系统的实践经验,对影响特征跟踪算法效率和精度的因素进行总结,便于读者根据具体问题选择合适算法。

1) 影响特征跟踪精度的因素

(1) 选择合适的特征检点检测算法。选择重复性较高的特征检测子,能够在一定程度上缓解跟踪丢失问题。

(2) 选择合适的特征描述子。二进制描述子的精度低于浮点型描述子。

(3) 特征匹配策略。强力匹配策略的精度小于 KNN^[33] 和 FRFT 等,此外还可以采用 RANSAC 算法去除错误匹配,但是会降低时间效率。

2) 影响特征跟踪效率的因素

(1) 选择效率较高的特征检测算法。效率较高的特征检测算法有 FAST、AGAST、ORB 和 OAGAST。

(2) 选择效率较高的描述子生成算法。

(3) 选择效率较高的特征匹配策略。可供选择的有 Cascade Hashing 和本文提出的 FRFT 算法。

(4) 充分利用现有的硬件资源。随着 GPU 的普及,可以采用 GPU 技术对算法进行加速。例如采用 GPU 加速的 SIFTGPU^[34] 算法和本文提出的 FRFT 算法。

5 结 论

特征跟踪在 SFM 体系结构中具有重要的理论和实践研究价值。为了克服现有特征跟踪算法在鲁棒性和时间效率上的不足,在 DMF 特征跟踪框架下提出一种快速鲁棒的特征跟踪算法-FRFT 算法,主要贡献如下:1) 提出带有方向的 OAGAST 特征检测子,为计算具有旋转不变性的特征描述子奠定基础;2) 提出一种具有较强鲁棒性的 DCN 特征描述子,解决特征的奇异性对跟踪算法的影响,提高了特征跟踪算法的准确率;3) 提出一种基于聚类的多视图特征跟踪方法,为提高大规模三维重建的效率奠定基础;4) 根据实验结果,对可能影响特征跟踪算法性能的因素进行分类总结,便于读者根据具体问题选择合适的特征跟踪算法。

此外,在 YFCC 100M 这样的较大数据集上,FRFT 算法的时间效率还有待提高,因此,下一步将探索在多 GPU 上实现 FRFT 算法,进一步提高大规模 SFM 系统的时间效率。

参考文献

[1] 王宇希,刘越. 一种基于多摄像机的鲁棒运动结构重建方法[J]. 仪器仪表学报, 2015, 36(1): 230-238.
WANG Y X, LIU Y. Robust structure from motion approach based on multi cameras[J]. Chinese Journal of Scientific Instrument, 2015, 36(1):230-238.

[2] SNAVELY N, SEITZ S, SZELISKI R. Photo tourism: Exploring photo collections in 3D[J]. ACM Transaction

on Graphics (TOG), 2006, 25(3):835-846.

- [3] FRAHM J M, FITE-GEORGEL P, GALLOP D, et al. Building Rome on a cloudless day[C]. Computer Vision-ECCV 2010, 2010: 368-381.
- [4] 陈辉,马世伟, NUECHTER A, 等. 基于激光扫描和 SFM 的非同步点云三维重构方法[J]. 仪器仪表学报, 2016, 37(5):1148-1157.
CHEN H, MA SH W, NUECHTER A, et al. Non-synchronous point cloud algorithm for 3D reconstruction based on laser scanning and SFM[J]. Chinese Journal of Scientific Instrument, 2016, 37(5):1148-1157.
- [5] 史利民,郭复胜,胡占义. 利用空间几何信息的改进 PMVS 算法[J]. 自动化学报, 2011, 37(5): 560-568.
SHI L M, GUO F SH, HU ZH Y. An improved pmvs through scene geometric information [J]. Acta Automatica Sinica, 2011, 37(5):560-568.
- [6] KLEIN G, MURRAY D. Parallel tracking and mapping for small AR workspaces [C]. IEEE and ACM International Symposium on Mixed and Augmented Reality, 2007:225-234.
- [7] COLBERT M, BOUGUET J Y, BEIS J, et al. Building indoor multi-panorama experiences at scale [J]. ACM Siggraph, 2012, 70(2): 101-102.
- [8] SAKURADA K, OKATANI T, et al. Detecting changes in 3D structure of a scene from multi-view images captured by a vehicle-mounted camera [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013:137-144.
- [9] ARRIGONI F, ROSSI B, MALAPELLE F, et al. Robust global motion estimation with matrix completion [J]. International Archives of the Photogrammetry Remote Sensing & S, 2014: 40(5):63-70.
- [10] PARK H S, SHI J. Social saliency prediction[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015: 4777-4785.
- [11] SONG S, CHANDRAKER M. Robust scale estimation in real-time monocular SFM for autonomous driving [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014:1566-1573.
- [12] DONG Z L, ZHANG G F, JIA J Y, et al. Keyframe-based real-time camera tracking[C]. IEEE International Conference on Computer Vision (ICCV), 2009: 1538-1545.
- [13] ZHANG G F, DONG Z L, JIA J Y, et al. Efficient non-consecutive feature tracking for structure-from-motion [C]. Computer Vision-ECCV, 2010: 422-435.
- [14] WU CH CH. Towards linear-time incremental structure from motion [C]. IEEE International Conference on 3D Vision (3DV), 2013:127-134.
- [15] 朱奇光,张朋珍,李昊立,等. 基于全局和局部特征融合的图像匹配算法研究[J]. 仪器仪表学报, 2016,

- 37(1):170-176.
- ZHU Q G, ZHANG P ZH, LI H L, et al. Investigation on the image stitching algorithm based on global and local feature fusion [J]. Chinese Journal of Scientific Instrument, 2016, 37(1):170-176.
- [16] 兰太寿,李炜荣,刘向军. 基于虚拟双目视觉的电器电磁机构三维动态测试研究[J]. 电子测量与仪器学报, 2014, 28(1):29-35.
- LAN T SH, LI W R, LIU X J. Three-dimensional dynamic test research for electromagnetism mechanism of electrical apparatus based on virtual binocular vision[J]. Journal of Electronic Measurement and Instrumentation, 2014, 28(1):29-35.
- [17] 许佳佳,张叶,张赫. 基于改进 Harris-SIFT 算子的快速图像配准算法[J]. 电子测量与仪器学报, 2005, 29(1):48-54.
- XU J J, ZHANG Y, ZHANG H. Fast image registration algorithm based on improved Harris-SIFT descriptor[J]. Journal of Electronic Measurement and Instrumentation, 2015, 29(1):48-54.
- [18] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2):91-110.
- [19] ZACH C. Robust bundle adjustment revisited [C]. Computer Vision-ECCV, 2014: 772-787.
- [20] HEINLY J, SCHONBERGER J L, DUNN E, et al. Reconstructing the world * in six days * (as captured by the Yahoo 100 million image dataset) [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015:3287-3295.
- [21] ZHANG G, VELA P A. Good features to track for visual SLAM [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015:1373-1382.
- [22] MAIR E, HAGER G D, BURSCHKA D, et al. Adaptive and generic corner detection based on the accelerated segment test [C]. Computer Vision-ECCV, 2010: 183-196.
- [23] RUBLEE E, RABAU D V, KONOLIGE K, et al. ORB: an efficient alternative to SIFT or SURF [C]. IEEE International Conference on Computer Vision, 2011: 2564-2571.
- [24] ROSIN P L. Measuring corner properties[J]. Computer Vision and Image Understanding, 1999, 73 (2): 291-307.
- [25] FREY B J, DUECK D. Clustering by passing messages between data points [J]. Science, 2007, 315 (5814): 972-976.
- [26] HEINLY J, DUNN E. Comparative evaluation of binary features[C]. Computer Vision-ECCV, 2012:759-773.
- [27] NISTER D. An efficient solution to the five-point relative pose problem[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(6): 756-770.
- [28] HARTLEY R, ZISSERMAN A. Multiple view geometry in computer vision [M]. Cambridge: Cambridge University Press, 2003.
- [29] WU CH CH, AGARWAL S, CURLESS B, et al. Multicore bundle adjustment [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011: 3057-3064.
- [30] MIKOLAJCZYK K, TUYTELAARS T, SCHMID C, et al. A comparison of affine region detectors [J]. International Journal of Computer Vision, 2005, 65(1-2): 43-72.
- [31] JAMES M R, ROBSON S. Straightforward reconstruction of 3D surfaces and topography with a camera: Accuracy and geoscience application [J]. Journal of Geophysical Research: Earth Surface, 2012, 117(F3):94-96.
- [32] WU Y, LIM J, YANG M H. Online object tracking: A benchmark [C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2013:2411-2418.
- [33] GARCIA V, DEBREUVE E, NIELSEN F, et al. K-nearest neighbor search: Fast GPU-based implementations and application to high-dimensional feature matching [C]. IEEE International Conference on Image Processing, 2010:3757-3760.
- [34] WU CH CH. SiftGPU: A GPU implementation of scale invariant feature transform (SIFT) (2007), URL <http://cs.unc.edu/~ccwu/siftgpu>, 2011.

作者简介



曹明伟, 2010年于宿州学院获得学士学位, 2014年获得合肥工业大学硕士学位, 现为合肥工业大学博士研究生, 主要研究方向为三维重建、3D目标检测、计算机视觉。
E-mail: cmwqq2008@163.com

Cao Mingwei received his B. Sc. degree from Suzhou University in 2010, and M. Sc. degree from Hefei University of Technology in 2014. Now, he is a Ph. D. student in Hefei University of Technology. His main research direction is 3D reconstruction, 3D object detection and computer vision.



刘晓平(通讯作者), 1985年于山东大学获得学士学位, 分别在1987年和1998年于合肥工业大学硕士学位和博士学位, 现为合肥工业大学教授、博士生导师, 主要研究方向为三维重建、计算机动画和仿真。
E-mail: lxp@hfut.edu.cn

Liu Xiaoping (Corresponding author) received his B. Sc. degree in 1985 from Shandong University, M. Sc. degree in 1987 and Ph. D. degree in 1998 both from Hefei University of Technology. Now, he is a professor and PhD supervisor in Hefei University of Technology. His main research direction includes 3D reconstruction, computer animation and simulation.