

SURF 算法的降维研究

卢清薇, 罗旌钰, 王云峰

(厦门大学 厦门大学电子工程系, 福建 厦门 361005)

摘要: SURF(Speed-up robust features)算法进行图像特征点匹配时需要循环遍历待匹配图像所有特征点, 计算特征点之间的 SURF64 描述距离, 耗时大。本文对 SURF 算法进行了 16 维与 4 维的降维研究。实验结果表明, 16 维 SURF 算法性能与 64 维 SURF 算法基本相当, 但能大幅度降低运算时间; 4 维运算性能降低较大, 不能用于特征点匹配, 但 4 维 SURF 描述算法可以扩展到图像的各个像素点, 用于 ICP 算法及图像的稠密匹配。

关键词: 图像匹配; 特征点; SURF 算法; 降维

中图分类号: TP391 文献标识码: A DOI: 10.3969/j.issn.1003-6970.2017.12.028

本文著录格式: 卢清薇, 罗旌钰, 王云峰. SURF 算法的降维研究[J]. 软件, 2017, 38 (12): 148-152

Research on Dimensional Reduction for SURF Algorithm

LU Qing-wei, LUO Jing-yu, Wang Yun-feng

(Department of electronic engineering, Xiamen University, Xiamen Fujian 361005, China)

【Abstract】: Matching image feature point by SURF(Speed-up robust features) algorithm needs to loop through all feature points on the image to be matched, and compute the sixty-four dimensional distance of SURF between feature points, which will take a long computation time. The sixteen dimensional SURF and four SURF are implemented. The experimental results show that the performance of the sixteen dimensional SURF is almost the same as that of the sixty-four dimensional SURF, moreover the sixteen dimensional SURF takes less time than sixty-four dimensional SURF. The performance of four dimensional SURF is much inferior to that of sixty-four dimensional SURF, so it cannot be used to match feature point. However the four dimensional description of SURF can be expanded to all feature points, which can be applied in ICP algorithm and dense matching algorithm.

【Key words】: Image matching; Feature points; SURF algorithm; Dimensional reduction

0 引言

即时定位与地图创建 (SLAM, Simultaneous Localization and Mapping) 系统在环境勘测、抢险救灾、物体追踪、终端移动自治等方面得到了广泛的应用。随着研究的深入以及计算机视觉的发展, SLAM 系统研究的重心已经从基于昂贵的激光和声纳等传感器转移到基于视觉 SLAM 上了。视觉 SLAM 的核心问题是多场景间的匹配问题, 选择有效的匹配算法是重构环境的重要前提。

最早单独使用视觉传感器来估计移动终端的运动状态是在 19 世纪 80 年代末, Moravec 等^[1]在他的研究工作中不但首次介绍了运动估计的流程, 还提

出了最早的图像角点检测算法。Matthies 等^[2]使用双目视觉系统, 结合 Moravec 的研究成果, 进行 SLAM 研究。Davison 等^[3]利用视觉匹配的方法实现了对一个自由运动摄像头的运动状态求解。Henry 等^[4]率先利用新一代视觉传感器 Kinect 采集室内场景数据, 联合形状和表面信息进行图像之间的配准, 可以完成 Kinect 的即时定位与场景地图的创建。Endres 等^[5]基于手持 Kinect, 利用 SIFT (Scale Invariant Feature Transform) 算法从彩色图像提取特征点, 进行匹配, 结合图像深度信息, 在三维空间中构建可视化的地图。上述文献都是通过图像特征点匹配, 计算图像之间的变换关系来完成即时定位与地图创建。因此, 特征点匹配算法的效率严重影响 SLAM 系统的性能。

作者简介: 卢清薇(1994-), 女, 硕士研究生, 主要研究方向: SLAM 系统; 罗旌钰(1994-), 男, 硕士研究生, 主要研究方向: SLAM 系统。

通讯作者: 王云峰(1977-), 男, 副教授, 博士, 研究方向: 数字集成电路设计及 SLAM 系统软硬件设计。

SIFT^[6-8]和 SURF^[9]是应用最广的两种图像特征点匹配算法。SIFT 是由 Lowe 率先提出的一种鲁棒性强的图像局部特征提取与匹配算法,具有对旋转、光照、尺度变化等不变性。而 SURF 是由 Bay 在 SIFT 的基础上提出的一种特征点提取匹配算法, SURF 在各个方面均接近或超越了 SIFT 的性能,并且计算效率比 SIFT 算法提高了 2 倍左右,得到了广泛的应用^[10-12]。但在匹配时需要循环遍历待匹配图像所有特征点,计算各个特征点之间的 64 维 SURF 描述距离,耗时大。一些改进 SURF 算法已经被提出。BBF-SURF^[13]未对 SURF 算法做出任何改变,只是软件实现方法的改进;即在匹配过程中采用 K-D tree 进行遍历搜索,提高运算效率。该实现方法可应用到各类改进 SURF 特征点匹配算法中。Oriented BRIEF-SURF^[14]算法同 SURF 算法相比,不同点在于没有采用 64 维描述向量,而是采用了 BRIEF 描述算子对特征点进行描述,减弱了匹配算法的针对图像变换的不变性,降低模糊图像的正确匹配率,运算效率提升幅度不大。ROI-SURF^[15]采用 ROI 算子进行特征点提取,结合随机点半径剔除法减少了特征点提取的数目,然后采用与 SURF 相同的描述与匹配,但是实验结果显示该算法降低了正确匹配率且运算效率提升幅度小。FAST-SURF^[16]是采用 FAST 检测器提取特征点,不具备尺度不变性,因此即使运算效率提升大,但适用范围有限。本文将对 SURF 算法的描述进行降维,并基于 Mikolajczyk 图像数据库进行性能分析。

1 SURF 特征点匹配算法

图像特征点匹配算法可以分为特征点提取与描述、特征点匹配^[9]过程。SURF 算法是目前应用最多的特征点匹配算法,它采用了高斯二阶微分模板进行特征点提取,使用 64 维向量描述特征点,并利用线性遍历搜索的方法进行匹配。

1.1 特征点提取与描述

SURF 算法基于图像中各点的 Hessian 矩阵进行特征点提取,对于图像中任意一点 $P=(x,y)$,其灰度值为 $I(P)=I(x,y)$,则 Hessian 矩阵^[9] $H(P, \sigma_k)$ 如式 (1) 所示:

$$H(P, \sigma) = \begin{bmatrix} L_{xx}(P, \sigma_k) & L_{xy}(P, \sigma_k) \\ L_{xy}(P, \sigma_k) & L_{yy}(P, \sigma_k) \end{bmatrix} \quad (1)$$

设 $\partial^2 g(\sigma_k)/\partial x^2$ 为边长为 N 的高斯二阶偏导滤波模板, σ_k 为高斯函数的参数,则式 (1) 中 $L_{xx}(P, \sigma_k)$

为 $\partial^2 g(\sigma_k)/\partial x^2$ 与以 P 为中心的边长为 N 的正方形局部图像灰度的卷积。高斯函数二阶偏导滤波模板 $\partial^2 g(\sigma_k)/\partial x^2$ 可由式 (2) 表示:

$$\frac{\partial^2}{\partial x^2} g(\sigma_k) = \begin{bmatrix} w_{11} & \cdots & w_{1N} \\ \vdots & \vdots & \vdots \\ w_{N1} & \cdots & w_{NN} \end{bmatrix} \quad (2)$$

则:

$$L_{xx}(P, \sigma_k) = \sum_{i=1}^N \sum_{j=1}^N w_{ij} I(x - \frac{(N+1)}{2} + i, y - \frac{(N+1)}{2} + j) \quad (3)$$

$L_{xy}(P, \sigma_k)$, $L_{yy}(P, \sigma_k)$ 具有相似的含义,但模板不同。Hessian 矩阵对应的行列式^[9] $det(H)$ 为:

$$det(H) = L_{xx}L_{yy} - L_{xy}^2 \quad (4)$$

为了在有缩放关系的图像间找到相互匹配的特征点, SURF 算法通过改变 σ_k 获得不同尺度下的滤波器,分组对图像进行处理,得到相应尺度空间下的响应图;滤波器尺度 S_k 与 σ_k 的关系如式 (5) 所示:

$$S_k = 9 \times \sigma_k / N_k \quad (5)$$

式中: N_k 为滤波器边长。

特征点判断时,逐组进行,只要有一组满足式 (6) 即为特征点。

$$\begin{cases} det(H) > \varepsilon \\ det(H)(x, y, S) = \max[det(H)(i, j, S)] \\ i \in [x-1, x+1]; \\ j \in [y-1, y+1]; \\ S \in [S_{k-1}, S_{k+1}]; \end{cases} \quad (6)$$

式中 ε 值的大小影响特征点数目,通常是根据经验值选取,大多数文献选取 ε 为 0.650,本文也取此值。

特征点描述^[9]首先利用 Haar 小波滤波器确定描述的主方向,再以特征点为中心,以主方向为水平方向确定 4×4 个描述的子区域;每个子区域中用 Haar 小波滤波器计算水平方向的响应 d_x 和垂直方向的响应 d_y ,得到 4 维向量 $(\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|)$,把 4×4 个子块区域的向量连接起来就得到了一个 64 维的特征点描述向量。

1.2 特征点匹配

特征点匹配是为两幅图像共有的特征点建立一一对应关系的过程。SURF 算法采用的是遍历式线性搜索方法。设算法中 P_A 是图像 A 的特征点集 S_A 中任意一点, P_B 是图像 B 的特征点集 S_B 中任意一点,它们的特征点描述向量分别为 $Des(P_A)$ 和 $Des(P_B)$, $Des_i(P_A)$ 和 $Des_i(P_A)$ 分别是特征点描述向量的第 i 个

分量, P_A 和 P_B 之间的欧式距离 $Dist(P_A, P_B)$ 为:

$$Dist(P_A, P_B) = \left[\sum_{i=1}^{64} (Des_i(P_A) - Des_i(P_B))^2 \right]^{1/2} \quad (7)$$

把 A 图中的点 P_A 与 B 图 S_B 中的所有点的距离计算一遍, 得到最近距离 ND (Nearest Distance) 和次最近距离 NND (Next Nearest Distance), 则距离之比 RoD (Ratio of Distance) 如式 (8) 所示:

$$RoD = ND / NND \quad (8)$$

当 RoD 小于一个阈值 $threshold$ 时, 就认为这对点是匹配点。 $threshold \in [0.5, 0.7]$ 是个比较理想的范围^[9], 本文选取 0.65。SURF 算法匹配的具体过程如表 1 所示:

表 1 特征点匹配算法

Tab.1 The SURF feature point matching

输入: S_A : 图 A 的特征点集	
S_B : 图 B 的特征点集	
输出: $Pair$: 匹配点对	
1	for P_A in S_A
2	$ND = NND = 100.0$;
3	for P_B in S_B
4	$D = Dist(P_A, P_B)$;
5	if ($D < ND$)
6	$ND = D$;
7	$NND = ND$;
8	else ($D < NND$)
9	$NND = D$;
10	end if
11	end for
12	if ($ND/NND < threshold$)
13	(P_A, P_B) 压入 $Pair$;
14	end if
15	end for
16	return $Pair$;

2 SURF 算法的降维实现方法

2.1 SURF 算法的 64 维描述

SURF 特征点描述是利用特征点周围信息来描述这个点。首先以特征点为中心, 用尺寸为 $4S$ 的 Haar 小波滤波器在以 $6S$ 为半径的圆形邻域里, 求得每个像素点的 Haar 小波响应, 然后用以特征点为中心的高斯函数 ($\sigma = 2S$) 对这些响应进行加权。用一个圆心角为 $\pi/3$ 扇形 (如图 1) 以特征点为中心环绕一周, 计算该扇形处于每个角度时, 它所包括的图像点的 Haar 小波响应之和, 取和值的最大值为该

特征点所对应的主方向。

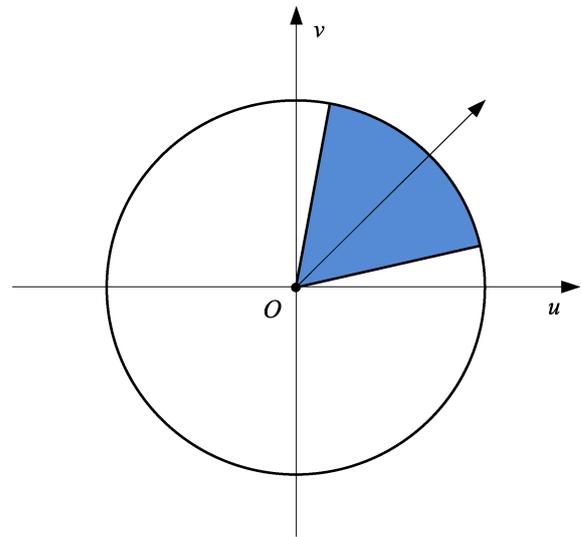


图 1 滑动的扇形窗口

Fig.1 The fan widow for calculating direction

以特征点为中心、上述得到的主方向为水平方向的确定一个正方形领域, 其边长为 $20S$, 将该正方形区域分成 4×4 个子块区域, 每个子块区域中用 Haar 小波滤波器处理。 d_x 表示水平方向的 Haar 小波响应, d_y 表示垂直方向的 Haar 小波响应, 在构建描述子向量之前, 对于所有的 d_x 、 d_y 都要用一个以特征点为中心的高斯函数加权, 该高斯函数的 $\sigma = 3.3S$ 。在每个子块区域中对 d_x 、 d_y 、 $|d_x|$ 、 $|d_y|$ 求和, 从而得到一个 4 维向量 ($\Sigma d_x, \Sigma d_y, \Sigma |d_x|, \Sigma |d_y|$), 把 4×4 个子块区域的向量连接起来就得到了一个 64 维的特征点描述向量。

2.2 降维实现方法

生成 SURF 描述符时首先要以特征点为中心确定的边长为 $20S$ 的正方形邻域, 该邻域被称为 $20S$ 方邻域。如图 2 所示, 则 $20s$ 方邻域内共有 400 个采样像素, 以 25 个采样像素为单元区域, 则 $20s$ 方邻域被划分为 $4 \times 4 = 16$ 个子区域。对每个子区域进行 Haar 小波滤波, 并且在每个子块区域中对 d_x 、 d_y 、 $|d_x|$ 、 $|d_y|$ 求和就得到一个 4 维向量 ($\Sigma d_x, \Sigma d_y, \Sigma |d_x|, \Sigma |d_y|$)。由于每个子区域生成 4 维描述向量, 所以 16 个子区域共生成 64 维描述向量。

SURF 描述的核心思想是以一个正方形的区域特征来代表其中心特征点。显而易见这种描述方式比以单个点的特征参数代表特征点的方式更加具备鲁棒性。这种描述方式完全可以扩展到图像的所有

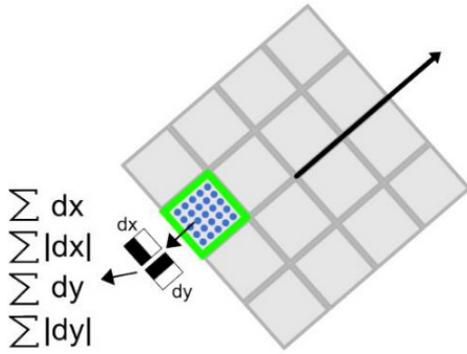


图 2 SURF 描述生成示意图

Fig.2 The generation of SURF description

点，影响扩展应用的原因在于其维数。应用时对图像全部像素点进行 64 维描述向量运算需要大量的运算时间，不具备可行性。基于此，本文实现了对 SURF 描述向量实现了 16 维与 4 维的降维。16 维的降维方法是固定单位区域为 25 个像素点；然后选取描述区域时是以特征点为中心，以边长为 10S 构造正方形邻域。则正方形邻域可以划分成 4 个子区域，每个区域采用原 SURF 的处理方式，构造 4 维描述向量。4 个区域则共 16 维描述向量。4 维的降维方法同 16 维类似，以边长为 5S 构造正方形邻域。整个邻域只包含一个子区域，则只有 4 维描述向量。

3 降维 SURF 算法性能分析

Mikolajczyk 图像数据集是用来测试图像匹配算法的性能的图像集合，其中包含了 8 组图像，每组图像集合中都有 6 幅图像，其中每组集合中的图像都是相机在不同参数情况下拍摄得到的，分别具有尺度、压缩、旋转、光照、视角等图像变换。其中 Boat 图像集和 Bark 图像集是旋转图集，可以用来检测匹配算法的旋转鲁棒性。匹配算法对模糊变换的鲁棒性可以使用 Bikes 图像集合 Trees 图像集来检测，因为这两个图像集是由原图像和不同尺度的高斯核卷积得到的。Graf 图像集和 Wall 图像集中的图像是由视角变换得到的，可以用来匹配算法在视角变换下的稳定性。匹配算法在光照变换的稳定性可以使用 Leuven 图像集来检测，Leuven 图像集为不同亮度下的同一物体的图像。最后一组图像集 Ubc 图像集内的图像具有不同的压缩程度，可以用来检测匹配算法在压缩变换下的稳定性。

本文选用具有代表性的 ubc 图集、boat 图集、bikes 图集、leuven 图集和 graf 图集进行降维性能分析。根据这几组图像的特征，将每组图像的第一幅图像和剩余的图像一一匹配，各个维度的性能比较如表 2-表 6 所示。

表 2 ubc 图集上性能比较

Tab.2 The performance comparison on ubc

图像对	维数/时间					
	64 维匹配率	64 维匹配时间	16 维匹配率	16 维匹配时间	4 维匹配率	4 维匹配时间
ubc1 与 ubc2	4408/5974	2.99528	4198/6016	1.6852	1242/6091	0.80775
ubc1 与 ubc3	3547/5813	2.96914	3447/5689	1.86963	43/6095	0.742386
ubc1 与 ubc4	2254/5528	2.78198	2310/5845	1.59463	10/6096	0.699977
ubc1 与 ubc5	1361/5697	2.4533	1247/5292	1.51382	12/6093	0.700123
ubc1 与 ubc6	503/5542	1.9505	437/5175	1.23753	29/6069	0.514954

表 3 boat 图集上正确匹配数目比较

Tab.3 The performance comparison on boat

图像对	维数		
	64 维	16 维	4 维
boat1 与 boat2	2025	923	10
boat1 与 boat3	409	832	11
boat1 与 boat4	679	154	13
boat1 与 boat5	38	47	20
boat1 与 boat6	33	97	19

表 4 bikes 图集上正确匹配数目比较

Tab.4 The performance comparison on bikes

图像对	维数		
	64 维	16 维	4 维
bikes1 与 bikes2	2806	2663	18
bikes1 与 bikes3	2398	2030	14
bikes1 与 bikes4	1343	913	17
bikes1 与 bikes5	881	272	26
bikes1 与 bikes6	84	279	35

表 5 Leuven 图集上正确匹配数目比较
Tab.5 The performance comparison on Leuven

图像对	维数		
	64 维	16 维	4 维
Leuven1 与 leuven2	2871	2606	25
Leuven1 与 Leuven3	1978	2030	19
Leuven1 与 Leuven4	1374	1681	21
Leuven1 与 Leuven5	1409	991	17
Leuven1 与 Leuven6	1004	833	31

表 6 graf 图集上正确匹配数目比较
Tab.6 The performance comparison on graf

图像对	维数		
	64 维	16 维	4 维
graf1 与 graf2	1123	1055	14
graf1 与 graf3	196	20	14
graf1 与 graf4	17	19	10
graf1 与 graf5	21	18	16
graf1 与 graf6	25	21	14

SURF 描述在各个图像集上的所体现的匹配性能均有大幅下降, 不适合用于图像特征点匹配。但由于其维数较少, 并且具备一定的各种变换的鲁棒性, 可以扩展到图像的每个像素点, 用于 ICP 算法及稠密匹配。

4 结论

本文通过改变 SURF 算法特征点描述的正方形邻域的大小, 实现了 16 维与 4 维的特征点 SURF 描述。实验结果表明, 16 维 SURF 描述用于图像特征点匹配可以大幅降低匹配时间, 并且大多数环境下匹配性能与原 64 维 SURF 相当; 仅不适合用于存在大量旋转的环境下的图像特征点匹配。4 维 SURF 描述的匹配效果大幅下降, 不宜用于特征点匹配。但可以扩展到图像的每个像素点, 用于 ICP 算法及稠密匹配。

参考文献

[1] MORAVEC H P. Obstacle Avoidance and Navigation in the Real World by a Seeing Robot Rover [D]. California: Stanford University, 1980.

[2] MATTHIES L, SHAFER S. Error modeling in stereo navigation[J]. Robotics and Automation, IEEE, 1987, 3(3): 239-248.

[3] DAVISON A J. Real-time simultaneous localization and mapping with a single Camera[C], Computer Vision. Nice, France, 2003:1403-1410.

[4] HENRY P, REN X, FOX D, et al. RGB-D mapping:Using depth cameras for dense 3D modeling of indoor environments[J]. International Journal of Robotics Research, 2010, 31(5): 647-663.

[5] ENDRES F, HESS J, ENGELHARD N, et al. An evaluation of the RGB-D system[C], Robotics and Automation. St. Paul, Minnesota, 2012: 1691-1696.

[6] LOWE DG. Distinctive image features from scaleinvariant Keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110.

[7] TAKEISHI N, TANIMOTO A, YAIRI T. Evaluation of interest-region detectors and descriptors for automatic landmark tracking on asteroids[J]. Transactions of the Japan Society for Aeronautical and Space Sciences, 2015, 58(1): 45-53.

[8] WU Hao, TIAN Guo-hui, LI Yan. Building of cognizing semantic map in large-scale semi-unknown environment[J]. Journal of Central South University. 2014, 21(5): 1804-1815.

[9] HERBERT B, TINNE T, LUC V G. SURF: Speeded Up Robust Features [J]. Computer Vision and Image Understanding, 2008, 110(3): 346-359.

[10] 刘博文, 童立靖. 基于多视角三维扫描数据的图像配准[J]. 软件, 2016, 37(9): 29-32.

[11] 李守宪, 魏芳. 移动终端Logo识别系统[J]. 软件, 2016, 37(02): 98-102.

[12] 翁秀玲, 王云峰, 余小亮. 一种图像立体匹配的误匹配点剔除与校正方法[J]. 软件, 2016, 37(9): 20-24.

[13] Jia X, Wang X, Dong Z. Image matching method based on improved SURF algorithm[C], IEEE International Conference on Computer and Communications. IEEE, 2015: 142-145.

[14] Ye Y. The Research of SLAM Monocular Vision Based on The Improved SURF Feather[C], International Conference on Computational Intelligence and Communication Networks. IEEE, 2014: 344-348.

[15] Ma Y L S. Research on Image Based on Improved SURF Feature Matching[C], Seventh International Symposium on Computational Intelligence and Design. IEEE, 2015: 581-584.

[16] Zhang H, Hu Q. Fast image matching based-on improved SURF algorithm[C], International Conference on Electronics, Communications and Control. IEEE, 2011: 1460-1463.