

ORB 特征检测算法整理

一. 算法简介

由于 SIFT 特征的计算量过大，无法应用于实时系统或低功耗设备，因此有很多研究着力于如何加快其效率。其中一类研究以降低其计算量为思路，典型的如 SURF 特征；另一类以加快其计算速度为思路，例如 GPU 加速。ORB 特征作为 SIFT 特征和 SURF 特征的一个高效的替代方案，全称 Oriented FAST and Rotated BRIEF，由 Rublee 等人于 2011 年提出。其论文中指出，ORB 特征相比于 SIFT，有相似的匹配性能，更强的抗噪能力，同时具备实时性；其描述子效果与 SIFT 类似，优于 SURF，却比 SIFT 要快两个数量级。另一个很重要的方面是，使用 ORB 可以避开 SIFT 和 SURF 的专利限制。

ORB 特征基于 FAST 关键点和 BRIEF 描述子，是在此基础上改进而成。FAST 关键点中并没有用于描述特征方向的定向算子，ORB 在 FAST 的基础上使用质心算子确定方向，构成 **oFAST 关键点**。论文中提及，BRIEF 描述子在光照、模糊、视角失真等方面的鲁棒性不亚于 SIFT，但是对于平面旋转却相对敏感；ORB 在此基础上做出了改进，构成 **rBRIEF 描述子**。

总体来说，ORB 特征有着非常低的计算量和较好的性能，其特征有旋转不变性，但是并不具备尺度不变性。

二. oFAST 关键点

2.1. FAST 关键点

FAST 关键点由于其较高的计算效率而被广泛使用。但是 FAST 没有方向，ORB 特征为 FAST 添加了一个方向，就是 oFAST。

FAST (Features From Accelerated Segment Test) 关键点是由 Rosten 于 2006 年首次提出，稍作修改后在 2010 年发表。如图 1，FAST 通过检测中心像素 p 于围绕该像素的圆环（共 16 个点）的灰度差是否在阈值 t 之内，若有连续 n 个点的灰度差都在阈值范围之外，即，

$$\begin{aligned} I_{p \rightarrow x} &< I_p - t \\ I_{p \rightarrow x} &> I_p + t \end{aligned}$$

则可判定该点为关键点。

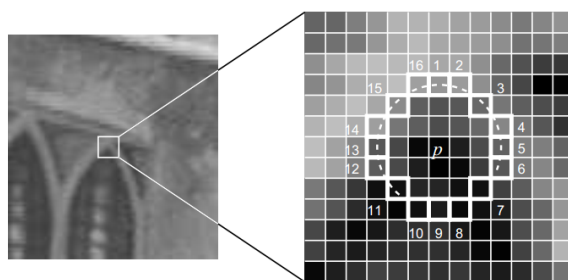


图 1 FAST 检测

一般取 $n = 9$ 可以获得最好的效果，也称 FAST-9。

为加快检测速度，可以用检测第 1、5、9、13 像素的方式对候选点进行初步筛选，即首先检测 1 和 9 像素，若两像素都在阈值范围内，则可判定其不是关键点，直接排除；之后检测合格的点的 5、13 像素，需要 4 个像素中至少有 3 个在阈值范围之外，才能将像素列为候选点。之后对候选点剩余的像素进行检测，进而检测其是否是关键点。在 Rosten 的论文中，采用一定区域内非极大值抑制的方式来解决临近区域内关键点过于集中的问题。即，在 3×3 （视实际情况， 5×5 或 7×7 均可）的范围内，计算每一关键点的 FAST 响应值 V ，只保 V 最

大的关键点， V 的值为关键点 p 和它周围 16 个像素的绝对灰度差之和。

在 ORB 中，由于 FAST 并不会提供多尺度特征，因此 Rublee 构建了一个图像的尺度金字塔，之后分别在每一层进行 FAST 检测，之后对得到的关键点计算 Harris 响应，按 Harris 响应值排序，选取前 N 个点。

2.2. 灰度质心法

由于 FAST 没有提供方向信息，ORB 使用了一种简捷有效的方式计算关键点的方向，即灰度质心法。假定关键点 O 周围的灰度值得出的质心与该点存在一定的偏移，由此得到的向量即可用于计算关键点方向。

角点周围图像块的矩被定义为

$$m_{pq} = \sum_{x,y} x^p y^q \cdot I(x, y)$$

进而得出图像块的灰度质心

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

关键点 O 的方向向量为 \overrightarrow{OC} ，方向角为

$$\theta = \text{atan2}(m_{01}, m_{10})$$

函数 $\text{atan2}()$ 是 $\arctan()$ 的象限相关的版本，按 \arctan 理解即可。 $\theta \in [-\pi, \pi]$ 。

为改进测量的旋转不变性，论文中将矩的计算限定在半径为 r 的圆形区域内， r 的选取是一个经验值。而 x 、 y 的取值范围为 $[-r, r]$ 。

三. rBRIEF 描述子

ORB 中所采用的 rBRIEF 描述子是 **BRIEF 的一种改进版本**，BRIEF 描述子是一种由图像块（patch）灰度检测集构成的二进制位序列描述子。在将 patch 进行 Gaussian 模糊后，随机抽取两个子窗口并比较其灰度，比较结果即为灰度检测值。

灰度检测值 τ 为

$$\tau(p; x, y) \triangleq \begin{cases} 1 & : p(x) < p(y) \\ 0 & : p(x) \geq p(y) \end{cases}$$

其中 p 为所在 patch， x 、 y 均为 patch 中的像素点（称为一组 test）， $p(x)$ 、 $p(y)$ 为像素所在子窗口的灰度。每个关键点的描述子向量由 n 个检测值构成，即

$$f_n(p) \triangleq \sum_{1 \leq i \leq n} 2^{i-1} \tau(p; x_i, y_i)$$

在抽取 x 、 y 时，可以有 Gaussian 分布、均匀分布等多种分布形式。文中表示 Gaussian 分布， $n = 256$ 时能够得到相对较好的效果。在实现时，使用积分图计算 Gaussian 模糊，采用 5×5 子窗口及 31×31 的 patch。

BRIEF 虽然鲁棒性较强、计算简便，但是并不具备旋转不变性。为了**旋转不变性**，在此基础上将其改进为 **steered BRIEF 描述子**。根据之前计算出的关键点方向角 θ ，可得到平面旋转矩阵 R_θ 。对于灰度检测中的 n 组 test，可以构建 $2 \times n$ 矩阵 S

$$S = \begin{bmatrix} x_1, \dots, x_n \\ y_1, \dots, y_n \end{bmatrix}$$

对其进行旋转

$$S_\theta = R_\theta S$$

则 steered BRIEF 向量为

$$g_n(p, \theta) = f_n(p) | (x_i, y_i) \in S_\theta$$

实现时将 θ 离散化，以 $2\pi/30$ 为增量，构建旋转角查找表，以加快计算效率。

虽然 steered BRIEF 解决了旋转不变性，但是由于其方差较小，不同组 test 间相关性较高，因而可区分性较差，会影响到最终的匹配结果。为解决这一问题，Rublee 采用基于学习的方法，确定提取 test 的最佳位置，以得到更合适的描述子向量，即 **rBRIEF 描述子**。学习的思路为，在所有可能的二进制 test 中，寻求一个均值接近 0.5，同时有着高方差、低相关性的 test。

Rublee 从 PASCAL 2006 图像集中建立一个由 300k 个关键点组成的训练集，列举 31×31 的 patch 中所有可能的二进制 test，每组 test 都是从中提取出的一对 5×5 子窗口。取 $\omega_p = 31$ ，

$\omega_t = 5$ ，则子窗口共有 $N = (\omega_p - \omega_t)^2$ 种选择，考虑到其中可能会有重叠，最终 Rublee 选取了 $M = 205590$ 组 test。

最终的学习算法为

1. 在每个 patch 中计算每组 test 的检测值
2. 以每组 test 的均值到 0.5 的距离排序，形成向量 \mathbf{T}
3. 进行贪婪搜索：
 - A. 将 \mathbf{T} 中首个 test 放入结果集 \mathbf{R} ，并将其从 \mathbf{T} 中删除
 - B. 将 \mathbf{T} 中下一个 test 与 \mathbf{R} 中所有的 test 对比，若阈值相关性达到阈值，则抛弃之；反之，则令其加入 \mathbf{R}
 - C. 重复 A、B 步，直到 \mathbf{R} 中 256 个位置填满；若 \mathbf{R} 中 test 数目不足 256，则提高阈值，再次尝试。

最终提取出的 256 组 test 在均值接近 0.5 的同时，有着较低的相关性，较大的方差，及较好的差异性。之后在这一 test 提取方式的基础上，通过 \mathbf{R}_θ 计算旋转，再计算描述子向量。

也就是说，**rBRIEF 将 test 的提取位置固定了下来。**

在 OpenCV 的实现中，若将 patch 设定为 31×31 ，子窗口设定为 5×5 ，则在灰度检测中使用 rBRIEF 的固定位置提取 test；若参数有变，则采用随机提取的方式。

参考

- [1] Rublee E, Rabaud V, Konolige K, et al. ORB: An efficient alternative to SIFT or SURF[C]//2011 International Conference on Computer Vision, Barcelona, Spain, 2011: 2564-2571.
- [2] Rosten E, Drummond T. Machine learning for high-speed corner detection[C]//Proceedings of the European Conference on Computer Vision. Heidelberg: Springer, 2006, 3951: 430-443.
- [3] Calonder M, Lepetit V, Strecha C. BRIEF: Binary robust independent elementary features [G]. LNCS 6314: Computer Vision-ECCV. Berlin: Springer Berlin Heidelberg, 2010: 778-792.
- [4] Kaehler A, Bradski G. 学习 OpenCV3[M]. 阿丘科技团队，译. 北京:清华大学出版社, 2018.