

一种基于扇形区域分割的 SIFT 特征描述符

曾 峦^{1,2} 顾大龙²

摘要 提出了一种在圆形区域内基于扇形区域分割的特征描述符构建方法. 首先, 针对 SIFT 描述符维数过高, 导致匹配速度慢的弱点, 提出在半径为 9 像素的圆形特征区域内划分为 8 个扇区, 在这些扇形特征邻域内统计 8 个方向的灰度梯度直方图, 形成 64 维描述符的方法, 降低了描述符的维数. 同时, 针对 SIFT 构建描述符的运算复杂性较高的事实, 提出在圆形区域内计算像素灰度梯度主方向, 以主方向为基准点把该区域划分为 8 个等面积扇区的方法, 取消了对特征区域的旋转变换, 降低了构建描述符的运算复杂性. 通过与 OpenCV SIFT 和 Lowe SIFT 进行多方面对比实验, 结果表明该方法的综合匹配速度具有显著提升, 在两幅图像存在一定程度的视点、模糊、旋转、比例、光照变化等情形下, 匹配性能有所增强.

关键词 机器视觉, 特征描述符, 图像匹配, SIFT

引用格式 曾峦, 顾大龙. 一种基于扇形区域分割的 SIFT 特征描述符. 自动化学报, 2012, 38(9): 1513–1519

DOI 10.3724/SP.J.1004.2012.01513

A SIFT Feature Descriptor Based on Sector Area Partitioning

ZENG Luan^{1,2} GU Da-Long²

Abstract This paper presents a new method of constructing feature descriptor based on sector area partitioning in a circular region. The large dimension of descriptor will decrease the matching speed of SIFT algorithm. In order to solve the problem, we partition a circular region whose radius is nine pixels into eight bins in angular direction. By computing the gradient orientation histogram in the eight directions for each bin, a descriptor with 64 dimensions is constituted. Therefore, the dimension of the descriptor is reduced. Meanwhile, based on the fact that computation of constructing SIFT descriptor is complex, we introduce a strategy that computes the dominate orientation of pixel gray gradients in the circular region, and then partitions the circular region into eight identical sector areas starting from the dominate orientation. Consequently, the computational complexity is reduced due to cancellation of rotation operation. By comprehensive comparison with the OpenCV SIFT and Lowe SIFT, the results indicate that the proposed method can increase the average matching speed significantly. Even if there exist affine distortion, defocusing, rotation, scaling or illumination variation, the matching performance can also be improved.

Key words Machine vision, feature descriptor, image matching, SIFT

Citation Zeng Luan, Gu Da-Long. A SIFT feature descriptor based on sector area partitioning. *Acta Automatica Sinica*, 2012, 38(9): 1513–1519

图像匹配在计算机视觉、模式识别和图像拼接等领域中起着决定性作用^[1–2], 基于特征的匹配方法具有计算量小、速度较快等特点, 是目前研究最多、应用最广的一种匹配方法. 但其匹配准确性和效率受到特征提取的稳定性和描述向量的仿射不变性的制约, 一直没有得到很好的解决.

SIFT 局部特征描述符^[3–4] 对图像方向、尺度、仿射变换、亮度、噪声、视角等变化具有良好的不变性, 已广泛应用于图像匹配中. 但 SIFT 算法用

于图像匹配时, 存在特征描述符构建过程复杂和描述向量维数比较高, 使得匹配运算时间过长等问题. 在匹配加速方面, Mikolajczyk 等^[5] 用线性规划的方法改进 SIFT 的匹配性能和速度; Klare 等^[6] 使用基于判别分析的局部特征 (Local feature-based discriminant analysis, LFDA) 方法减少图像特征区域后再建立 SIFT 描述符, 加速面部识别的过程.

在描述符改造方面, 刘立等^[7] 提出用一个圆形区域覆盖整个采样范围, 分 12 个梯度方向, 形成 12 维描述向量的方法, 但这种描述符在复杂场景下的匹配性能大幅下降; Mikolajczyk 等^[8] 认为圆形比矩形区域具有更好的定位性能, 采用环形扇区构建 272 维的梯度位置方向直方图 (Gradient location orientation histogram, GLOH) 描述符, 然后用主成分分析法 (Principal components analysis, PCA) 把它降为 128 维, 提高了匹配性能; Huang 等^[9] 仿照 GLOH 构建特征区域的方法, 采用正、负对比度

收稿日期 2011-05-03 录用日期 2011-09-14
Manuscript received May 3, 2011; accepted September 14, 2011
国家高技术研究发展计划 (863 计划) (2009AA7041007) 资助
Supported by National High Technology Research and Development Program of China (863 Program) (2009AA7041007)

本文责任编辑 戴琼海

Recommended by Associate Editor DAI Qiong-Hai

1. 哈尔滨工业大学超精密光电仪器工程研究所 哈尔滨 150080 2. 装备指挥技术学院 北京 101416

1. Institute of Ultra-precision Optical and Electronic Instrument Engineering, Harbin Institute of Technology, Harbin 150080 2. Academy of Equipment Command and Technology, Beijing 101416

累加值代替梯度直方图, 64 维的对比度上下文直方图 (Contrast context histogram, CCH) 描述符性能与 SIFT 相当, 但匹配模糊和视点变化图像时性能稍差一些; 与此相类似, 刘萍萍等^[10] 提出了一种 64 维的规范化强度对比描述子; Tola 等^[11] 认为圆形区域具有自然的抗旋转扰动能力, 提出用 25 个圆覆盖特征区域的方法构建了 200 维描述符。

吕冀等^[12] 用简单金字塔代替 DoG 图像来换取描述符构建速度, 但对灰度平滑且噪声影响较大的图像匹配性能较差; Ke 等^[13] 提出的 PCA-SIFT 描述符, 虽然用 PCA 的方法降低了描述向量的维数, 但采用了更大的特征邻域构建描述符, 又抵消了降维所带来的速度提高; Wong^[14] 提出了在采样区域与 SIFT 一样的条件下, 先统计 $0^\circ \sim 180^\circ$ 范围内的 4 个梯度方向直方图, 构成描述符的前 64 位, 来增加背景和颜色改变时描述符的稳定性, 然后统计 $0^\circ \sim 360^\circ$ 范围内的 4 个梯度方向直方图, 构成描述符的后 64 位, 来保证关键点的独特性, 匹配时实际使用的是 128 位的描述符, 且在求取主方向和统计梯度方向直方图时计算量较大; Cui 等^[15] 认为关键点中心区域的像素最为稳定, 采用可变区域大小的方法使采样区域的每一个子区域都与关键点相邻, 构建了不规则分块方向直方图 - SIFT (Irregular orientation histogram binning - SIFT, IOHB-SIFT) 描述符, 取消了高斯加权操作, 加快了描述符构建速度, 但维数并没有降低; Chen 等^[16] 把 SIFT 和生理学原理相结合, 模仿人眼视觉提取图像特征的过程, 构建了多尺度的韦伯局部描述符 (Weber local descriptor, WLD); Heikkilä 等^[17] 把中心对称的局部二元图 (Center symmetric local binary pattern, CS-LBP) 与 SIFT 相结合, 在描述符构建阶段取消了高斯加权和梯度方向等运算, 提高了描述符的构建速度, 描述符为 256 维; 曾慧等^[18] 在邻域像素点数为 8 时, 构建了 512 维的改进的加权中心对称的局部三元图 (Improved weighted center symmetric local trinary pattern, IWCS-LTP) 描述符。

SURF 方法^[19] 用积分图像求导、方框滤波近似代替二阶高斯滤波、简化主方向确定等办法对 SIFT 进行加速, 并用方形区域内像素的 Haar 小波响应 dx , dy , $|dx|$, $|dy|$ 的累积和作为特征向量, 形成了 64 维的描述符, 从减少梯度投影方向的角度对描述降维. 本文则试图用改变特征区域形状的方法对 SIFT 描述符进行降维, 综合文献 [8-11, 15] 的优点, 提出了一种在圆形区域内基于扇形区域分割的描述符构建方法, 实验结果表明在相同匹配阈值时本文算法和 OpenCV SIFT、Lowe SIFT 相比, 提取到的正确匹配对更多, 正确匹配率更高, 且运算速度更快、描述符的维数更低。

1 基于扇形区域分割的特征描述符

为了保证特征描述符的旋转不变性, SIFT 算法采用统计关键点周围 $16 \text{ 像素} \times 16 \text{ 像素}$ 采样区域中所有像素的灰度梯度方向投影值, 建立梯度直方图, 以直方图主峰值为该特征点的主方向, 并把该区域的像素坐标旋转到主方向上, 使得两幅待匹配图像对应的区域具有相同的方向指向, 消除了图像因发生了旋转使对应区域梯度方向的不一致性。

由于采用了方形区域, 使得在两幅图像中对应的区域发生了旋转变换后, 如果不对该区域的所有像素进行旋转, 直接取同样大小的方形区域构建描述符, 就会有一部分区域不重叠, 产生较大误差, 如图 1(a) 的灰色区域. 因此, 在 OpenCV SIFT 算法的实际操作中, 为了使旋转后能取得与基准图像同样大小的像素区域, 采用了比采样区域大一些的方形区域进行旋转, 取其落在采样区域的像素构建描述符, 如图 1(a) 的点线所示. 这种做法除了增加了旋转操作的像素数, 运算时间加长外, 还有如下缺点: 1) 由于 SIFT 在构建描述过程中, 主方向测量以 10° 步进, 不可避免地会有指向量化误差, 如图 1(b) 所示, 假设实际图像已旋转到 (x', y') 的位置, 而由于量化、测量误差的原因使旋转区域只达到 (x'', y'') 的位置, 这样, 在图 1(b) 的灰色区域就不是与原区域相同像素的内容, 即实线所围的区域与虚线所围的区域所包含的像素在整体上就会有差别, 从而降低了旋转后的区域相似性; 2) 对特征区域进行旋转变换过程中, 需要进行灰度插值和重采样, 在一定程度上改变了像素间的灰度梯度关系, 在构建描述符向量时, 就会带来更多的误差, 降低相似度, 增加匹配难度。

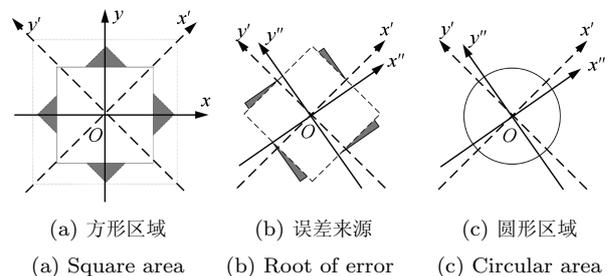


图 1 方形和圆形区域比较

Fig. 1 Compared the square with the circular area

圆形区域则不相同, 即使主方向有误差, 圆形区域旋转前、后所覆盖的像素是一致的, 如图 1(c) 所示. 实际上, 如果把圆的直径选为正方形的边长的话就相当于去除了方形区域外围受干扰的地方, 使得其旋转后的整体相似性加强. 孟繁杰等^[20] 认为圆形区域具有严格的几何不变性, 该文把 Harris-Laplace 提取到的角点集合划分到同心等面积的扇形区域中, 用于更有效地表示兴趣点从中心向四周各方向辐射

分布的情况, 然后, 在扇形区域内构建兴趣点的空间离散度, 用于表示每个子区域角点分布的疏密程度, 作为一个局部匹配特征。

当图像发生微小的比例变化时, 如从图 2(a) 缩小到图 2(b), 方形采样区域中 4×4 个子区域受到的影响是不一样的, 如图 2(b) 的中心区 (亮灰色区域) 最稳定, 其次是深灰色区域, 影响最大的是最外围的黑色区域。因此, 方形采样区域构建的描述符受外围子区域的影响较大, SIFT 采用高斯加权的方法来减小这种影响。而圆形采样区域, 则把影响最大的黑色区域尽可能地排除在外, 如图 2(c) 所示, 再加上高斯加权的作用, 外围子区域的影响被大大减小了, 从而增加了描述符的比例不变性。

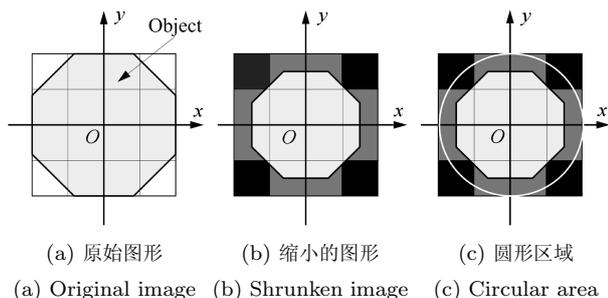


图 2 比例变化的影响

Fig. 2 Influence of scale variation

因此, 本文提出了一种在圆形邻域内用扇形区域分割构建描述符的方法, 如图 3 所示, 主要步骤如下。

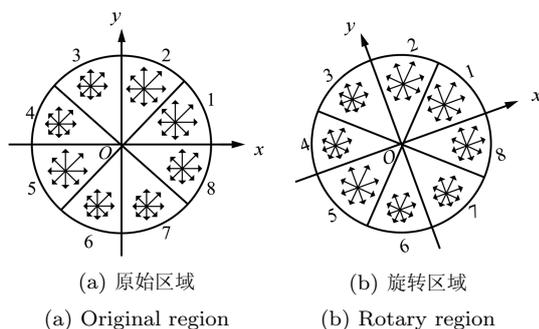


图 3 基于扇形区域的描述符

Fig. 3 Descriptor based on sector area

步骤 1. 以 SIFT 提取的关键点为中心, 取半径为 9 像素的圆形区域 (所含像素与原 SIFT 方法相同) 统计特征点的主方向。

步骤 2. 以主方向为起点, 将圆形邻域均匀划分为 8 个扇区, 每个扇形区域作为一个种子点, 分别统计扇区内所有像素按图像坐标 $0^\circ, 45^\circ, 90^\circ, 135^\circ, 180^\circ, 225^\circ, 270^\circ, 315^\circ$ 这 8 个方向的灰度梯度投影

值, 形成 8×8 个特征向量, 并按主方向进行归一化处理。

步骤 3. 对每个像素的梯度进行高斯加权运算, 以减少外围点的噪声对特征值的影响。

步骤 4. 以 1 号扇区形成的 8 个向量编码为 1~8 号向量, 2 号扇区形成 9~16 号向量, 以此类推, 8 号扇区形成 57~64 号向量, 形成 64 维向量描述符。

步骤 5. 对上述 64 维向量进行亮度归一化处理, 以降低光照变化带来的影响。

采用扇形区域分割构建特征描述符时, 关键是以主方向为起始点, 将圆形区域划分为 8 个扇形区域, 如图 3(a) 所示的 x 方向。当此圆形区域发生旋转变化时 (假定是沿区域中心旋转), 理想状态下其像素间的关系保持不变, 只是灰度梯度的主方向发生了变化。假定此时其主方向旋转至图 3(b) 所示的 x 方向, 我们仍然以这个方向为起始点把圆形区域划分为 8 个扇形区域, 那么, 图 3(b) 的 1~8 号扇区与图 3(a) 的 1~8 号扇区所包含的像素内容完全相对应, 每个种子点在统计扇区内像素点向 8 个方向进行灰度梯度投影时, 也以主方向为起始方向, 这样得到的特征向量就具有旋转不变性, 就可以取消对区域像素旋转的操作。旋转带来的大量的乘法操作变为少量的加法操作, 不但缩短了计算时间, 而且最大程度地保持两个特征区域像素间的位置和灰度关系, 进一步增强了描述符的旋转不变性。

这种形状的描述符在其对应区域发生微小的比例变化时, 相对矩形来说每个扇形区域能够保持比例更多的像素是对应像素, 且变化区域的面积是相同的, 并分布在高斯加权区域的最外围, 其影响被减弱了, 就可进一步增加 SIFT 的缩放不变性 (大的尺度变化由图像的金字塔变换来抵御)。

2 实验结果与分析

在 OpenCV SIFT 提供的源程序基础上, 对在圆形邻域内进行扇形区域分割构建描述符的方法进行了编程实现 (特征检测、描述符构建使用 Visual C++ 2008 编程, 匹配算法采用 Matlab 7.4 编程), 构建了扇区梯度方向直方图 (Sector area gradients orientation histogram, SAGOH) 描述符, 并进行了相应的实验。

2.1 特征描述符性能测试

为了评价新特征描述符的匹配性能, 采用图 4 所示的流行测试图像¹, 对新的描述符进行了测试, 并与 OpenCV SIFT、Lowe 算法和 SURF 算法的匹配性能进行了对比。其中, 图 4(a) 的尺寸为 800 像

¹ 测试图像选自 <http://www.robots.ox.ac.uk/vgg/research/affine/>, July 20, 2011

素 × 600 像素, 图 4(b) 的尺寸为 850 像素 × 680 像素, 图 4(c) 的尺寸为 800 像素 × 640 像素, 图 4(d) 的尺寸为 1000 像素 × 700 像素, 图 4(e) 的尺寸为 900 像素 × 600 像素。

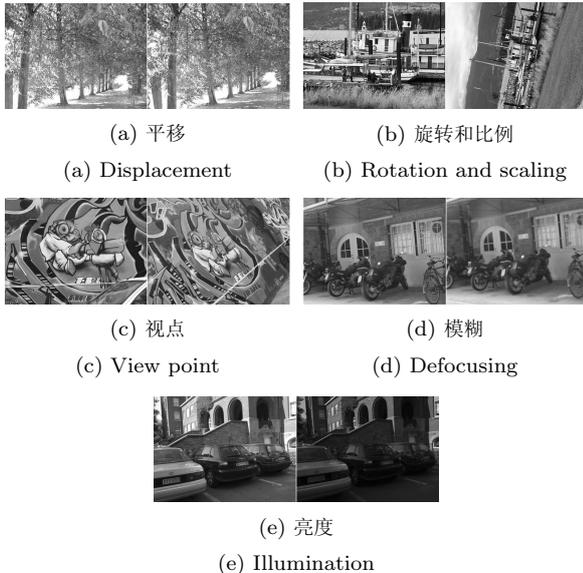


图 4 测试图像
Fig. 4 Test images

匹配过程中, 采用特征向量最邻近和次邻近欧氏距离的比值 (Ratio) 作为匹配对的判断阈值. 为了便于和 Lowe 算法相比较, 调整了 OpenCV 的特征提取参数, 使它们提取到的特征点数目相当.

根据 Lowe 的实验^[4], Ratio 取 0.8 时, 可以去除 90% 的误匹配, 但正确匹配也会丢失 5% 左右, 此值增大时, 误匹配数量急剧上升; 此值取 0.7 时, 可以去除 96% 的误匹配, 而正确匹配则丢失 8% 左右, 此值减少时, 误匹配数量降速减缓, 正确匹配数量降速增加. 随着 Ratio 值的降低, 误匹配数量会不断减少, 一般降到 0.3 时就可完全剔除误匹配. 实际上, Ratio 值的物理含义是匹配对的独特性, 当比率较大时, 表明参考图像中的某个特征点与待匹配图像中至少有两个特征点非常相似, 而当比率小时, 表明该特征对的相似性是非常独特的, 通过在不同的比率下考察正确匹配对数目和正确匹配率, 可以说明描述符的匹配性能. 因此, 本文在 Ratio 值为 0.3~0.8 的范围内匹配时对 SAGOH-SIFT, OpenCV-SIFT, Lowe-SIFT 和 SURF 进行对比实验, 图 5~图 9 分别是对图 4(a)~图 4(e) 中的图像对进行测试实验的结果.

图 5 是具有平移变化且稍微有一点旋转的图像对 (图 4(a)) 进行测试的结果. 从图中可见, SAGOH 在各种比率下的正确匹配对数目 (图 5(a)) 和正确匹配率 (图 5(b)) 方面, 都明显超过了 OpenCV SIFT 和 Lowe SIFT. 在平移方面, SAGOH 所围的

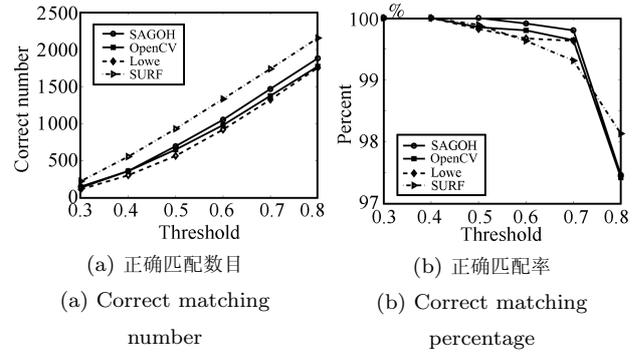


图 5 平移变化匹配结果
Fig. 5 Matching results for displacement

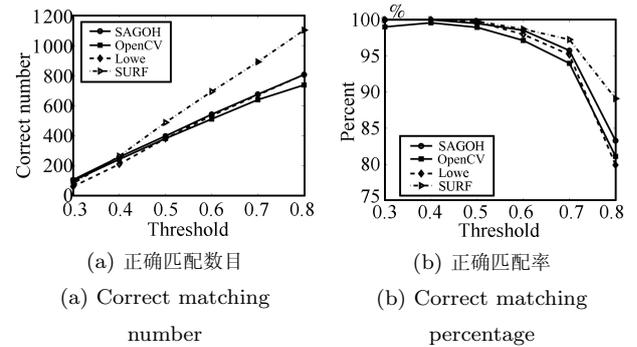


图 6 比例和旋转变化匹配结果
Fig. 6 Matching results for scaling and rotation

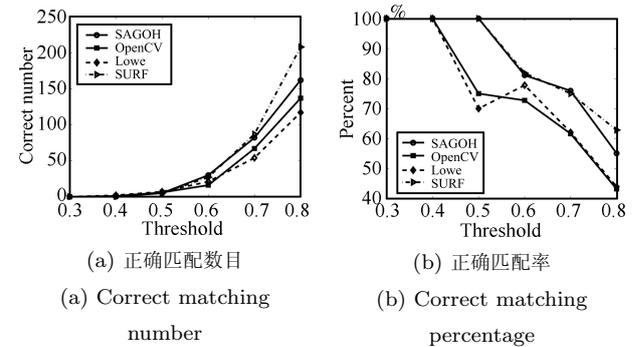


图 7 视点变化匹配结果
Fig. 7 Matching results for view point variation

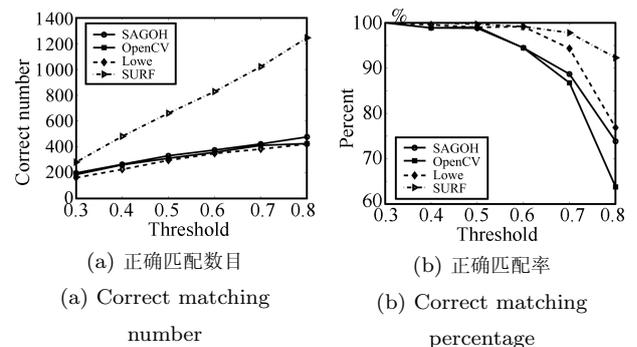


图 8 模糊变化匹配结果
Fig. 8 Matching results for defocusing

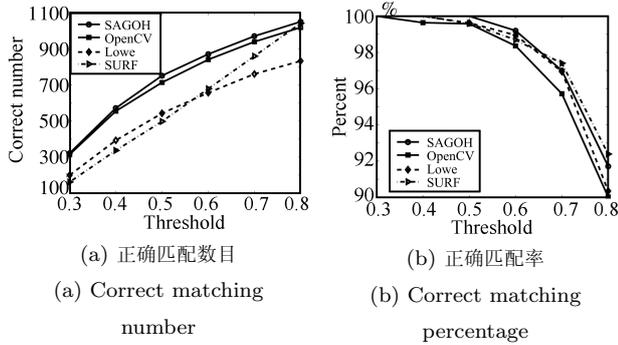


图9 亮度变化匹配结果

Fig. 9 Matching results for illumination variation

面积与 OpenCV SIFT 和 Lowe SIFT 相同, 匹配能力相当. 由于图像有微小的旋转变化, 此结果验证了圆形区域比方形区域抗旋转能力要强一些论断. 这说明扇区分割带来的抗平移和微小旋转变化的能力比描述符降维后带来的识别能力损失要大得多.

图 6 是对具有 2.5 倍比例变化且有 75° 旋转变化的图像对 (图 4(b)) 进行测试的结果. 从图中可见, SAGOH 在正确匹配对数目 (图 6(a)) 和正确匹配率 (图 6(b)) 方面, 都超过了 OpenCV SIFT 和 Lowe SIFT. 在旋转变化方面, 正如第 1 节分析的那样, 由于用扇区分割圆形区域, 取消了对采样区域的旋转操作, 减小了方形区域易受干扰的外围区域, 增强了描述符的抗旋转能力. 在比例变化方面, 除了上述原因外, 每个子区域都与最稳定的中心区域相连接, 增强了描述符的抗比例变化能力^[15]. 同样说明了扇区分割带来的抗旋转和比例变化的能力比描述符降维后带来的识别能力损失要大.

图 7 是对具有较大视点变化的图像对 (图 4(c)) 进行测试的结果. 从图中可见, SAGOH 在正确匹配对数目 (图 7(a)) 和正确匹配率 (图 7(b)) 方面, 都明显超过了 OpenCV SIFT 和 Lowe SIFT. Li 等^[21]认为, 用规格化的椭圆区域代替矩形区域, 可以增加描述符的仿射不变性, 该文采用与 GLOH 相似的方式构建描述符区域. 而圆形区域与椭圆区域比较接近, 因此, SAGOH 抗仿射变换能力要比 OpenCV SIFT 和 Lowe SIFT 强一些. 但上述三种描述符的匹配对都不是很多, 说明了其抗仿射变化的能力是有限的, 这可以借鉴 Li 等^[21]的方法作进一步的改进.

图 8 是对具有模糊变化的图像对 (图 4(d)) 进行测试的结果. 从图中可见, SAGOH 的正确匹配对数目 (图 8(a)), 明显超过了 OpenCV SIFT 和 Lowe SIFT. 正确匹配率 (图 8(b)) 比 OpenCV SIFT 高, 但比 Lowe SIFT 低. 在正确匹配对数目方面, 由于图像发生了模糊变化, 当特征区域随主方向进行旋转时, OpenCV SIFT 和 Lowe SIFT 要对特征区域进行插值和重采样, 使该区域更加模糊, 破坏了像素

间的梯度关系, 描述符识别能力下降. 而 SAGOH 不进行插值和重采样操作, 保持了像素间的梯度关系, 描述符识别能力就比较强, 正确匹配对数目就要多一些. 在正确匹配率方面, SIFT 在描述符构建过程中要去除一些低对比度的关键点, 图像模糊后, 低对比度的关键点增加, 由于 Lowe SIFT 取的阈值与 SAGOH 和 OpenCV SIFT 不同 (后面两种相同), 提取到的关键点数目要少得多 (两幅图像分别为 SAGOH: 5097 和 1075; OpenCV: 5061 和 1060; Lowe: 3825 和 800), 降低了正确匹配对数目的同时, 也大大降低了错误匹配对数目. 而在实际使用过程中, 往往更加关注正确匹配对的数目, 误匹配可以用其他优化的方法来消除 (如 Zeng 等^[22]提出的方法).

图 9 是对具有亮度变化的图像对 (图 4(e)) 进行测试的结果. 从图中可见, SAGOH 在正确匹配对数目 (图 9(a)) 和正确匹配率 (图 9(b)) 方面, 都超过了 OpenCV SIFT 和 Lowe SIFT. 图像亮度变化改变了像素间的灰度梯度关系, 在测量关键点主方向时就会带来误差, 使 OpenCV SIFT 和 Lowe SIFT 实际上会对某些子区域进行旋转操作, 插值和重采样对这些子区域的模糊会降低描述符的识别能力.

在和 SURF 相比较时, 使用 SURFmex 程序², 并调整了 SURF 算法中提取特征的控制参数, 使它在五组实验图像中提取到的特征点数目基本与 SAGOH 相当 (图 4(a): 7122 和 7192; 图 4(b): 8564 和 8475; 图 4(c): 4304 和 4969; 图 4(d): 5122 和 1853; 图 4(e): 3283 和 1890). 对于图 4(a) 所示的纹理特性较强的图像, SURF 方法的正确匹配对数目比 SAGOH 多, 但其正确匹配率在比率为 0.7、0.6、0.5 时都比 SAGOH 低 (在比率为 0.4、0.3 时则相等); 对于图 4(e) 所示的亮度有大幅度变化的图像, SURF 的正确匹配对数目明显低于 SAGOH, 正确匹配率则基本相当; 对于图 4(c) 所示的视点变化的图像, 除在比率为 0.8 时性能好于 SAGOH 外, 其他比率下性能基本相当; 对于图 4(b) 和图 4(d) 则正确匹配对数目和正确匹配率都明显好于 SAGOH.

2.2 运行时间测试

首先进行运算复杂性分析. SAGOH 与 OpenCV SIFT 在关键点提取阶段的做法相同, 因此本文只比较两者在描述符构建和匹配阶段运算复杂性上的差别. 两种方法的特征邻域的半径都与尺度有关, 在基础尺度层 (假定尺度因子为 1 时) SAGOH 特征邻域的半径为 9 像素, 邻域中像素数目

² <http://www.lwtbb.com/d-3bmh.html/>, July 20, 2011

表 1 运时间测试结果
Table 1 Test results of runtime

Serial	Key point (left)	Key point (right)	Runtime (Open CV) (s)	Runtime (SAGOH) (s)	Increment (%)
图 4(a)	9 753	9 007	317.46	169.71	46.54
图 4(b)	9 563	6 297	217.14	84.11	61.27
图 4(c)	3 745	4 994	75.50	33.19	56.04
图 4(d)	5 061	1 060	20.85	18.73	10.17
图 4(e)	3 286	2 181	22.62	17.72	22.66

为 256, 判断像素在圆形邻域的哪个扇区内需要进行 256 次反正切运算, 648 次乘法运算, 求取邻域内像素梯度和幅值需要进行 256 次反正切运算, 256 次开平方运算, 在进行梯度幅值加权时进行 256 次指数运算, 其余为少量代数运算. OpenCV SIFT 在基础尺度层 (尺度因子为 1) 为 22×22 的正方形区域, 其中核心区域大小为 16×16 , 其余区域为旋转操作的需要而增加的. 在求取像素梯度方向和幅值时需要进行 256 次反正切运算, 256 次开平方运算, 在进行梯度幅值加权时进行 256 次指数运算, 在进行旋转操作时需要用到特征邻域核心区域外的像素信息进行旋转操作, 共进行 2904 次乘法运算和 2 次正弦运算, 其余为少量代数运算. 因此, SAGOH 比 OpenCV SIFT 减少了 2836 次乘法运算和 2 次正弦运算, 增加了 256 次反正切运算, 总运算复杂性要低一些; 以 ratio 为参数进行匹配时, SAGOH 为 64 维, SIFT 描述符为 128 维, SAGOH 少进行 64 次减法、加法和乘法运算, 运算复杂性比 SIFT 要低.

然后进行实际测试. 在 DELL 550 硬件和 Matlab 7.4 软件平台上, 使用 Ratio 作为匹配对的判断阈值, 对图 4(a)~图 4(e) 所示的图像进行了匹配运算时间测试, 并与 OpenCV SIFT 进行对比, 具体情况如表 1 所示 (由于 Lowe SIFT 没有开源程序代码, 其描述符的构建过程和区域大小不清楚, 运行时间的比较失去了基准. 对图 4(a)~图 4(e) 所示的图像匹配测试时间分别为: 291.06s, 186.10s, 41.15s, 10.12s, 18.39s. 对图 4(d) 进行匹配时, 其特征点总数比 SAGOH 少了 1547 个, 因此匹配总时间反而比 SAGOH 短).

由于 SAGOH 在特征点提取阶段采用了与 OpenCV SIFT 相同的方法, 在描述符的构建上有所改变, 但其所围的面积是相同的, 只是取消了特征区域的旋转操作, 这样可以加快运行速度; 另外, 把 128 维的特征向量降到了 64 维, 在特征匹配阶段也会得到加速. 从表中可以看出, 特征点数目越多程序运行时间就缩短得越多, 运行效率就提升得越高. 同时, 图像旋转变化越多, 本文方法节省的时间也就越多, 如表中图 4(b) 对应的图像就是有 75° 左右的旋

转变化, 运行时间节省最多; 而对于图 4(d) 所对应的图像只有平移变化, 且提取到的特征点较少, 节省的时间就不多. 就上述 5 组实验图像而言, 匹配速度平均提升了约 39.3%.

相比之下, SURF 算法执行效率最高, 上述五组图像总匹配时间分别为: 59.59s, 83.09s, 25.27s, 13.34s 和 8.48s. 但匹配运算阶段比建立描述符阶段所用的时间要多得多, 因此, 研究描述符的降维方法非常重要.

3 结论

本文通过分析 SIFT 使用方形区域构建特征描述符抗旋转和比例变化性能的局限性, 针对其描述符维数过高导致匹配时间过长的弱点, 提出了一种新的描述符构建算法. 该算法用等面积的扇形区域分割圆形邻域构成描述符的子区域, 减少了方形区域最易受旋转和比例变化干扰的区域, 取消了对特征区域的旋转操作, 减少了构建描述符的运算复杂性, 降低了描述符的维数, 提高了运算速度. 实验结果表明, 本文的方法在比例、旋转、平移、模糊、亮度和视点变化等方面的匹配性能 (正确匹配对数和正确匹配率), 超过 OpenCV SIFT 和 Lowe SIFT 算法的基础上, 所用实验图像的平均匹配速度提高了约 39.3%; 为进一步提高 SIFT 算法的匹配速度提供了一种有效的方法. 下一步将针对本文算法中对视点变换的匹配性能提高得不够多的问题, 借鉴 Li 等^[21]的方法, 与本文算法相结合, 使用等面积的椭圆区域作为描述符的子区域, 对该描述符作进一步的改进.

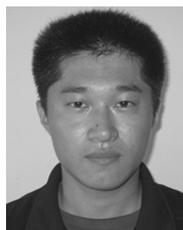
References

- Zhuang Yan, Chen Dong, Wang Wei, Han Jian-Da, Wang Yue-Chao. Status and development of natural scene understanding for vision-based outdoor mobile robot. *Acta Automatica Sinica*, 2010, **36**(1): 1–11
(庄严, 陈东, 王伟, 韩建达, 王越超. 移动机器人基于视觉室外自然场景理解的研究与进展. *自动化学报*, 2010, **36**(1): 1–11)
- Lin Hai-Feng, Ma Yu-Feng, Song Tao. Research on object tracking algorithm based on SIFT. *Acta Automatica Sinica*, 2010, **36**(8): 1204–1208

- (蒯海峰, 马宇峰, 宋涛. 基于 SIFT 特征目标跟踪算法研究. 自动化学报, 2010, **36**(8): 1204–1208)
- 3 Lowe D G. Object recognition from local scale-invariant features. In: Proceedings of the 7th IEEE International Conference on Computer Vision. Kerkyra, Greece: IEEE, 1999. 1150–1157
 - 4 Lowe D G. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004, **60**(2): 91–110
 - 5 Mikolajczyk K, Matas J. Improving descriptors for fast tree matching by optimal linear projection. In: Proceedings of the 11th IEEE International Conference on Computer Vision. Rio de Janeiro, Brazil: IEEE, 2007. 1–8
 - 6 Klare B, Li Z F, Jain A K. Matching forensic sketches to mug shot photos. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2011, **33**(3): 639–646
 - 7 Liu Li, Peng Fu-Yuan, Zhao Kun, Wan Ya-Ping. Simplified SIFT algorithm for fast image matching. *Infrared and Laser Engineering*, 2008, **37**(1): 181–184
(刘立, 彭复元, 赵坤, 万亚平. 采用简化 SIFT 算法实现快速图像匹配. 红外与激光工程, 2008, **37**(1): 181–184)
 - 8 Mikolajczyk K, Schmid C. A performance evaluation of local descriptors. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2005, **27**(10): 1615–1630
 - 9 Huang C R, Chen C S, Chung P C. Contrast context histogram — an efficient discriminating local descriptor for object recognition and image matching. *Pattern Recognition*, 2008, **41**(10): 3071–3077
 - 10 Liu Ping-Ping, Zhao Hong-Wei, Zang Xue-Bai, Dai Jin-Bo. A fast local feature description algorithm. *Acta Automatica Sinica*, 2010, **36**(1): 40–45
(刘萍萍, 赵宏伟, 臧雪柏, 戴金波. 一种快速局部特征描述算法. 自动化学报, 2010, **36**(1): 40–45)
 - 11 Tola E, Lepetit V, Fua P. DAISY: an efficient dense descriptor applied to wide-baseline stereo. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(5): 815–830
 - 12 Lv Ji, Gao Hong-Min, Wang Bo, Zhou Zhi-Qiang. Fast SIFT algorithm for autonomous image guidance. *Systems Engineering and Electronics*, 2009, **31**(5): 1147–1151
(吕冀, 高洪民, 汪渤, 周志强. 图像制导中的 SIFT 快速算法. 系统工程与电子技术, 2009, **31**(5): 1147–1151)
 - 13 Ke Y, Sukthankar R. PCA-SIFT: a more distinctive representation for local image descriptors. In: Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington D. C., USA: IEEE, 2004. 506–513
 - 14 Wong Y M. Invariant Local Feature for Image Matching [Master dissertation], Chinese University of Hong Kong, China, 2006
 - 15 Cui Y, Hasler N, Thormahlen T, Seidel H P. Scale invariant feature transform with irregular orientation histogram binning. In: Proceedings of the 6th International Conference on Image Analysis and Recognition. Halifax, Canada: Springer, 2009. 258–267
 - 16 Chen J, Shan S G, He C, Zhao G Y, Pietikainen M, Chen X L, Gao W. WLD: a robust local image descriptor. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2010, **32**(9): 1705–1720
 - 17 Heikkilä M, Pietikainen M, Schmid C. Description of interest regions with local binary patterns. *Pattern Recognition*, 2009, **42**(3): 425–436
 - 18 Zeng Hui, Mu Zhi-Chun, Wang Xiu-Qing. A robust method for local image feature region description. *Acta Automatica Sinica*, 2011, **37**(6): 658–664
(曾慧, 穆志纯, 王秀青. 一种鲁棒的图像局部特征区域的描述方法. 自动化学报, 2011, **37**(6): 658–664)
 - 19 Bay H, Ess A, Tuytelaars T, Gool L V. SURF: speeded up robust features. *Computer Vision and Image Understanding*, 2008, **110**(3): 346–359
 - 20 Meng Fan-Jie, Guo Bao-Long. Image retrieval by using local distribution features of interest points and multiple-instance learning. *Journal of Xidian University (Natural Science)*, 2011, **38**(2): 47–53
(孟繁杰, 郭宝龙. 使用兴趣点局部分布特征及多示例学习的图像检索方法. 西安电子科技大学学报(自然科学版), 2011, **38**(2): 47–53)
 - 21 Li C L, Ma L Z. A new framework for feature descriptor based on SIFT. *Pattern Recognition Letters*, 2009, **30**(5): 544–557
 - 22 Zeng L, Tan J B. Robust matching algorithm for image mosaic. In: Proceedings of the 6th International Symposium on Precision Engineering Measurements and Instrumentation. Hangzhou, China: SPIE, 2010. 166–175



曾 峦 装备指挥技术学院教授, 哈尔滨工业大学博士研究生. 主要研究方向为数字图像处理和计算机视觉. 本文通信作者. E-mail: zengluan@sina.com
(ZENG Luan Professor at the Academy of Equipment Command and Technology, Ph. D. candidate at Harbin Institute of Technology. His research interest covers digital image processing and computer vision. Corresponding author of this paper.)



顾大龙 装备指挥技术学院硕士研究生. 主要研究方向为数字图像处理和计算机视觉. E-mail: gudalong3568@sina.com
(GU Da-Long Master student at the Academy of Equipment Command and Technology. His research interest covers digital image processing and computer vision.)