

文章编号 1004-924X(2009)02-0439-06

结合全局信息的 SIFT 特征匹配算法

纪 华^{1,2}, 吴元昊¹, 孙宏海¹, 王延杰¹

- (1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033;
2. 中国科学院 研究生院, 北京 100039)

摘要:提出了结合全局信息的 SIFT(Scale Invariable Feature Transformation)特征匹配算法,解决了图像中存在多个相似区域时造成的误匹配问题。在尺度空间检测出特征点,生成包含两基于局部信息的 SIFT 向量和基于全局信息的全局向量;采用 BBF(Best Bin First)算法进行搜索,并采用欧氏距离作为度量函数进行特征向量的匹配。实验结果表明,由于在基于局部信息的 SIFT 向量中加入基于全局形状信息的全局向量,使得当特征点的尺度较小时,可以借助更大邻域范围内的信息对其进行描述,从而降低了由于局部信息相似而造成的误匹配的概率。所用图像误匹配的概率由 19% 下降到了 11%,极大地改善了匹配效果。

关键词:特征匹配;SIFT 算法;全局信息;BBF 算法

中图分类号:TP391.4 **文献标识码:**A

SIFT feature matching algorithm with global information

JI Hua^{1,2}, WU Yuan-hao¹, SUN Hong-hai¹, WANG Yan-jie¹

- (1. *Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;*
2. *Graduate University of Chinese Academy of Sciences, Beijing 100039, China*)

Abstract: An improved Scale Invariable Feature Transformation(SIFT) matching algorithm with global context vector is presented to solve the problems that SIFT descriptors result in a lot mismatches when an image has many similar regions. By detecting feature points in scale space, two kinds of feature vectors, a SIFT descriptor representing local properties and a global context vector, are computed. Then, according to BBF searching strategy, the feature vectors are matched by using Euclidean distance. The experimental results indicate that the improved algorithm can describe feature points in a larger region, and can reduce mismatch probability of experimental images from 19% to 11% because global context vectors based on global shape information are induced to the SIFT vectors based local information. These results reported above show proposed algorithm improves matching results greatly.

Key words: feature matching; SIFT algorithm; global information; BBF algorithm

收稿日期:2008-06-11;修订日期:2008-07-11.

基金项目:国家 863 高技术研究发展计划资助项目(No. 2006AA703405F)

1 引言

在基于特征的图像匹配技术中,其首要任务是提取稳定的特征,并进行描述。常用的方法有基于空间关系的匹配算法、基于不变量描述子的匹配算法^[1]、金字塔和小波算法等等。其中 SIFT (Scale Invariable Feature Transformation) 特征匹配算法^[2]是目前国内外特征匹配领域比较流行的一种算法,该算法能提取稳定的特征,匹配能力强,可以处理两幅图像之间发生平移、旋转、仿射变换、视角变换、光照变换情况下的匹配问题,甚至在某种程度上对任意角度拍摄的图像也具备较为稳定的特征匹配能力,从而可以实现差异较大的两幅图像之间的特征匹配。

不同特征点的局部邻域信息差别较大时, SIFT 特征匹配算法生成的描述子唯一,该算法可以得到满意的结果,但对于整幅画面中有多个相似区域的情况,该算法就会出现大量的误匹配点对。因此,本文对 SIFT 特征匹配算法进行了改进,加入全局纹理信息,使得 SIFT 描述子包含更大邻域范围内的曲线形状信息,从而可以降低误匹配的概率。

2 SIFT 算法简介

2.1 SIFT 特征点检测

David G. Lowe 在 2004 年总结了现有的基于不变量技术的特征检测方法,提出了一种基于尺度空间的、对图像缩放、旋转和仿射变换具有不变性的图像局部特征描述子——SIFT 算子。该算法首先在尺度空间进行特征检测,并确定关键点的位置和所处的尺度,然后使用关键点邻域的梯度的主方向作为该点的方向,以实现算子对尺度和方向的无关性,从而得到了一种对尺度和方向无关的算子。

2.1.1 检测尺度空间极值点

尺度空间理论最早出现于计算机视觉领域时其目的是模拟图像数据的多尺度特征。Koenderink^[3]证明高斯卷积核是实现尺度变换的唯一变换核,而 Lindeberg^[4]等人则进一步证明高斯核是唯一的线性核。一幅二维图像在不同尺度下的尺度空间表示可由图像与高斯核卷积得到:

$$L(x, y, z) = G(x, y, z) * I(x, y), \quad (1)$$

因此,要提取稳定的具有尺度无关性的特征点,就必须在图像二维平面空间和 DoG (Difference of Gaussian) 尺度空间中同时检测局部极值点。DoG 算子定义如下:

$$D(x, y, z) = (G(x, y, k) - G(x, y, z)) * I(x, y, z)$$

$$I(x, y, z) = L(x, y, k) - L(x, y, z) \quad (2)$$

在 DoG 空间检测极值时,需要把关键点与同一尺度的周围邻域 8 个像素和相邻尺度对应位置的周围邻域 9×2 个像素总共 26 个像素进行比较,以确保同时在尺度空间和二维图像空间检测局部极值。

2.1.2 精确定位极值点

DoG 算子会产生较强的边缘响应,因此要通过拟合三维二次函数以精确确定关键点的位置和尺度,同时去除低对比度的关键点和不稳定的边缘响应点,以增强匹配稳定性,提高抗噪声能力。

空间尺度函数 $D(x, y, z)$ 在局部极值点 (x_0, y_0, z_0) 处的泰勒展开式如下:

$$D(x, y, z) = D(x_0, y_0, z_0) + \frac{\partial D}{\partial X_0} X + \frac{1}{2} X^T \frac{\partial^2 D}{\partial X_0^2} X, \quad (3)$$

对上式求导,并令其为 0,得到精确的位置 X_{\max} ,如公式(5)所示:

$$X_{\max} = - \left(\frac{\partial^2 D}{\partial X_0^2} \right)^{-1} \frac{\partial D}{\partial X_0} \quad (4)$$

在已经检测到的特征点中,要去掉低对比度的特征点和不稳定的边缘响应点。

去除低对比度的点:把公式(4)代入公式(3),只取前两项可得:

$$D(X_{\max}) = D + \frac{1}{2} \frac{\partial D}{\partial X_0} X_{\max}, \quad (5)$$

若 $|D(X_{\max})| < 0.03$,该特征点就保留下来,否则丢弃。

去除不稳定的边缘点:一个定义不好的高斯差分算子的极值在横跨边缘的地方有较大的曲率,而在垂直边缘的方向有较小的主曲率。主曲率可以通过一个 2×2 的 Hessian 矩阵求出。

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (6)$$

D 的主曲率和 H 的特征值成正比,令 λ_1 为较大的特征值, λ_2 为较小的特征值,则:

$$\begin{aligned} \text{Tr}(H) &= D_{xx} + D_{yy} = \lambda_1 + \lambda_2 \\ \text{Det}(H) &= D_{xx} D_{yy} - (D_{xy})^2 = \lambda_1 \lambda_2 \end{aligned} \quad (7)$$

令 $\lambda = \frac{r+1}{r}$, 则

$$\frac{\text{Tr}(H)^2}{\text{Del}(H)} = \frac{(r+1)^2}{r^2} = \frac{(r+1)^2}{r^2} = \frac{(r+1)^2}{r^2} \quad (8)$$

$(r+1)^2/r$ 的值在两个特征值相等时最小, 随着 r 的增大而增大, 因此, 为了检测主曲率是否小于某阈值 r , 只需检测 $\frac{\text{Tr}(H)^2}{\text{Det}(H)} < \frac{(r+1)^2}{r}$ 是否成立, 本

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (9)$$

$$(x, y) = \arctan(L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))$$

公式(9)为 (x, y) 处梯度的模值和方向公式, 其中所用的尺度为每个关键点所在的尺度。实际计算时, 可在以关键点为中心的邻域窗口内采样, 并用直方图统计邻域像素的梯度方向, 梯度直方图的范围是 $0^\circ \sim 360^\circ$, 每 10° 一个柱, 总共 36 个柱。直方图的峰值位置则代表了该关键点处邻域梯度的主方向, 即作为该关键点的方向。在梯度直方图中, 当存在另一个相当于主峰值 80% 能量的峰值时, 则将这个方向认为是该关键点的辅方向。一个关键点可能会被指定具有多个方向, 这可以增强匹配的鲁棒性。

2.2 SIFT 特征向量生成^[2,5]

首先将坐标轴旋转为关键点的方向, 以确保旋转不变性。接下来以关键点为中心取 16×16 窗口, 分割成 4×4 个子区域, 在每个子区域上计算 8 个方向的梯度直方图, 绘制每个梯度方向的累加值, 即可形成一个种子点, 因此, 一共可以生成 16 个种子点, 这样对于每个关键点就可以产生一个长度为 128 的数据, 即最终形成一个长度为 128 的 SIFT 特征向量。此时 SIFT 特征向量已经去除了尺度变化、旋转等几何变形因素的影响, 再继续将特征向量的长度归一化, 则可进一步去除光照变化的影响。

这种邻域方向性信息联合的思想增强了算法的抗噪声能力, 同时对于含有定位误差的特征匹配也提供了较好的容错性。

2.3 SIFT 算法存在的问题

在计算梯度直方图时, 为了使离关键点越近的梯度贡献越大, 每个梯度的幅值都要被高斯函数加权后累加到相应的位置。虽然高斯函数的定义域是无限的, 但能量比较集中, 在实际计算中, 为了使生成的特征向量具有尺度无关性, 取高斯

论文中取 $r=10$ 。

(3) 关键点主方向分配

利用关键点邻域像素的梯度方向分布特性为每个关键点制定方向参数, 使算子具备旋转不变性。

加权的半径是关键点尺度的函数, 使半径以外的高斯加权系数为 0。因此, 当关键点的尺度较小时, 生成特征向量的有效邻域范围就比较小, 如果图像中存在多个相似区域, SIFT 算法得到的特征向量就有很大的相似性, 从而造成误匹配。

3 改进的 SIFT 算法^[6]

针对上一节所提到的问题, 本文对 SIFT 算法进行了改进, 对于每一个被检测到的特征点, 建立一个由两部分组成的向量: 一部分是代表局部特性的 SIFT 描述子, 另一部分是用来区分相似局部特征的全局纹理向量。因此, 改进的 SIFT 算法特征向量定义如下:

$$F = \begin{bmatrix} L \\ (1 - \alpha)G \end{bmatrix} \quad (10)$$

其中, L 是长度为 128 的 SIFT 向量, G 是长度为 60 的全局向量, α 是相对权重因子。

与 SIFT 局部描述子生成相似, 全局纹理也生成一个直方图。我们计算每个像素点的最大曲率, 给定一个点 (x, y) , 曲率就定义为海森矩阵的较大特征值的绝对值, 公式如下:

$$C(x, y) = | \lambda_1(x, y) | \quad (11)$$

其中, $\lambda_1(x, y)$ 是海森矩阵较大的特征值。

对于每一个特征点, 以其为中心建立对数极坐标, 并把以整幅图像对角线长度为直径的圆域划分成 60 个区域, 如图 1 所示, 因此, 全局形状纹理信息是 5×12 的直方图, 在直方图的每一个对应位置累积曲率值。这里的方法并不是严格的对数极坐标, 因为沿极径方向的前两个增量相等, 增量分别为: $\frac{\pi}{16}, \frac{\pi}{16}, \frac{\pi}{8}, \frac{\pi}{4}, \frac{\pi}{2}$ 。

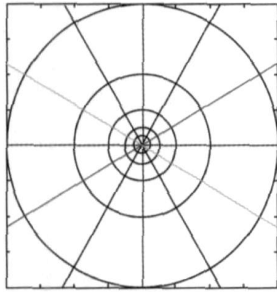


图 1 生成全局向量时的区域划分

Fig. 1 Region division in generating global vector

如果 $\tilde{X} = (\tilde{x}, \tilde{y})^T$ 是特征点的位置,主方向是 θ ,那么

$$a = \left[\frac{6}{\pi} \left(\arctan \left(\frac{\tilde{y} - \tilde{y}_0}{\tilde{x} - \tilde{x}_0} \right) - \theta \right) \right] \quad (12)$$

$$d = \max \left(1, \left\lceil \log_2 \left(\frac{\tilde{x} - \tilde{x}_0}{r} \right) + 6 \right\rceil \right)$$

a 和 d 分别为角度和径向距离离散值, $\hat{G}_{a,d} = \sum_{(x,y) \in N_{a,d}} C(x,y)$, 其中 $C(x,y)$ 是曲率图像。

由公式(12)可知,全局纹理向量也具有旋转不变性,但是全局纹理向量是图像尺寸的函数,而不是特征点尺度的函数,因此,它并不具有完全的尺度不变性。对于小尺度的特征点,局部邻域范围比较小,相似的可能性比较大,需要较大范围的全局形状信息,要求形状纹理区域的半径要远远大于特征点的尺度;而对于大尺度的特征点,局部描述子所描述的邻域范围已经足够大了,不需要全局的形状信息。因此,为了平衡固定尺寸的全局形状信息和可变尺寸的 SIFT 描述子,这里使用加权函数。

每个像素的曲率值被反转的高斯函数加权,权值函数是:

$$w(x,y) = 1 - e^{-((x-x_f)^2 + (y-y_f)^2)/(2^2)} \quad (13)$$

其中, (x_f, y_f) 是特征点的位置,取与 SIFT 局部特征邻域加权时相同的尺度。当尺度比较小时, $w(x,y)$ 比较大,使得较小的局部邻域所对应的全局信息比重较大;当尺度较大时, $w(x,y)$ 比较小,从而较少了全局信息的贡献。因此,加权函数使得 SIFT 局部描述子所能描述的邻域以外的区域信息起了更重要的作用,而且实现了从局部区域到全局区域的平滑过度。最后,对全局纹理向量进行归一化使算法对光照变化具有不变性。

$$\hat{G} = \frac{\tilde{G}}{\bar{G}} \quad (14)$$

4 结合全局信息的特征向量匹配

4.1 相似判定度量

当两幅图像的特征向量生成以后,可采用关键点特征向量的欧氏距离作为两幅图像中关键点的相似判定度量。由于特征向量包括两部分,这里分别计算各自的欧氏距离,公式如下:

$$d_L = |L_i - L_j| = \sqrt{\sum_k (L_{i,k} - L_{j,k})^2}$$

$$d_G = |G_i - G_j| = \sqrt{\sum_k (G_{i,k} - G_{j,k})^2} \quad (15)$$

最终的距离为:

$$d = d_L + (1 - \alpha) d_G \quad (16)$$

与公式(11)相同,它用来控制全局信息所占的比重。

4.2 匹配方法

为了减少一个特征点可能存在多个相似匹配点而造成的误匹配情况,这里采用最近邻和次近邻特征点距离之比来减少误匹配。如果最近的距离和次近的距离比值小于某个阈值 T_d ,则认为该点对为匹配点对,否则丢弃。降低阈值, SIFT 匹配点对数目会减少,但更加稳定。

如何找到最近邻和次近邻是本算法的关键问题。穷举法是最为有效的方法,但是如果特征点数目特别大时,计算量就会以指数级别增长。

4.3 BBF 算法^[7-8]

鉴于上述匹配算法中存在的问题,本文提出采用 BBF(Best Bin First) 来寻找最近邻和次近邻,它是对 k-d 树搜索算法的改进。实际上, k-d 树搜索算法大部分时间花费在检查节点上,而只有一部分节点满足最近邻条件,因此,可以采用近似的最近邻算法,通过限制 k-d 树中叶子节点数来缩短搜索时间。

一种简单的优化改进方法是以节点和被查询节点距离递增的顺序来搜索节点。当沿一个方向的分支搜索一节点时,优先级队列会被加入一个成员,该成员记录了该节点相关的信息,包括当前节点在树中的位置和该节点与被查询节点之间的距离。当一个叶节点被搜索到后,从队首删除一项,然后,再搜索包括最近节点的其他分支。

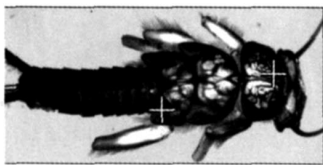
5 实验及其结果分析

以昆虫图像为例,如图 2(a)、(b) 所示,两幅

昆虫图片存在局部光照变化和整体的非刚性变化,但是基本上没有缩放和旋转。从图 2(c)、(d) 可以看出,检测出的特征点数目相当,且部分对应位置生成的特征向量相似。



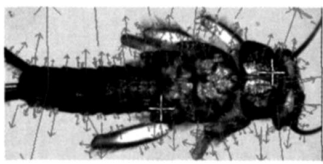
(a) 第一幅图像
(a) Image one



(b) 第二幅图像
(b) Image two



(c) 第一幅图像,检测出 536 个特征点
(c) Number of feature points is 536

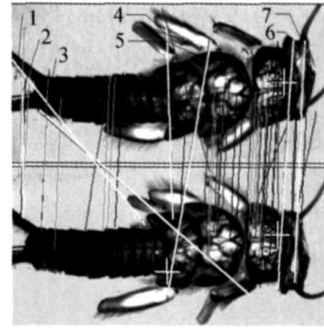


(d) 第二幅图像,检测出 534 个特征点
(d) Number of feature points is 534

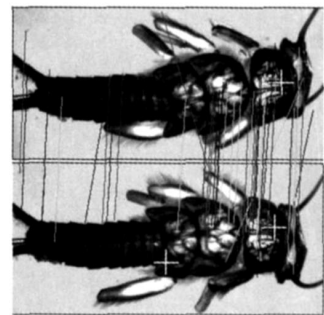
图 2 实验图片及其 SIFT 向量

Fig. 2 Experiment images and their SIFT vectors

图 3 是算法改进前后的实验结果,如图 3(a) 所示,加粗的直线 1~7 所连接的都是误匹配的点对,它们具有共同的特点,即两个特征点周围很小的邻域具有相似性,如果只考虑局部信息,必然会造成误匹配。加入全局信息之后的匹配结果如图 3(b) 所示,可以看到,直线 1~6 已经被剔除了,而直线 7 变成了 2 条正确的匹配直线,说明本文提出的算法,不仅可以减少误匹配的点对,还可以



(a) 未加全局向量时,匹配上 52 个特征点,误匹了 10 个,误匹配概率为 19%
(a) Without global vectors, the number of match points is 52, of which 10 points mismatch, the probability of mismatching is 19%



(b) 加入全局向量时,匹配上了 44 个特征点,误匹了 5 个,误匹配概率为 11%
(b) With global vectors, the number of match points is 44, of which 5 points mismatch, the probability of mismatching is 19%

图 3 改进前后实验结果比较

Fig. 3 Comparison before and after improvement

对误匹配的点对进行校正,因此可以大大降低误匹配的概率。

本文的实验取 $T_d = 0.5$, $\epsilon = 0.5$ 。未加入全局向量时,匹配上了 52 个特征点,其中误匹配了 10 个,误匹配的概率为 19%;加入全局向量后,匹配上了 44 个特征点,其中误匹配了 5 个,误匹配的概率为 11%,可见本文提出的算法将误匹配率降低了 8%,明显改善了匹配效果。

6 结 论

本文针对图像中存在多个相似区域,使得 SIFT 算法生成的特征向量比较接近,从而造成误

匹配的问题,提出了结合全局信息的 SIFT 特征匹配算法。在基于局部信息的 SIFT 向量中加入基于全局形状信息的全局向量,使得当特征点的尺度较小时,可以借助更大邻域范围内的信息来对其进行描述,从而降低了由于局部信息相似而造成的误匹配的概率。实验结果表明:改进后的算法区分能力更强,匹配正确性更高,所用图像误匹配的概率降低了 8%。

参考文献:

- [1] 丁雪梅,王维雅,黄向东. 基于差分 and 特征不变量的运动目标检测和跟踪[J]. 光学精密工程, 2007, 15(4): 570-576.
DING X M, WANG W Y, HUANG X D. New method for detecting and tracking of moving target based on difference and invariant[J]. *Opt. Precision Eng.*, 2007, 15(4): 570-576. (in Chinese)
- [2] LOWE D G. Distinctive image features from scale-invariant keypoints[J]. *International Journal of Computer Vision*, 2004, 60(2): 91-110.
- [3] KOENDERINK J. The structure of images[J]. *Biological Cybernetics*, 1984, 50: 363-396.
- [4] LINDBERGER T. Sscale-space for discrete signals[J]. *IEEE Transactions PAMI*, 1980, 207: 187-217.
- [5] 赵辉. SIFT 特征匹配技术讲义[J]. 盐城工学院学报

该算法还需要在以下几个方面进行改进:(1)对于每个特征点,都需要在整图像中计算全局描述子,计算量比较大,因此,需要在不影响算法效果的情况下对算法进行简化。(2)全局向量描述的范围大小是固定的,与图像尺寸有关,因此不具有尺度不变性,下一步的工作要使全局向量所描述的范围是特征点尺度的函数,使本文所提出的算法具有尺度不变性。

(自然科学版), 2007, 20(2): 1-5.

- ZHAO H. SIFT features matching technology[J]. *Journal of Yancheng Institute of Technology (Natural Science Edition)*, 2007, 20(2): 1-5. (in Chinese)
- [6] ERIC N M, HONG L I D, LINDA S. A SIFT descriptor with global context[J]. *Computer Vision and Pattern Recognition*, 2005(1): 20-25.
- [7] 周颜军. 数据结构[M]. 长春:吉林科学技术出版社, 2003.
ZHOU Y J. *Data Structures*[M]. Changchun: Jilin Science and Technology Publishing House, 2003. (in Chinese)
- [8] 戚世贵. 基于图像特征点的提取匹配及应用[D]. 长春:吉林大学硕士学位论文, 2006.
QI SH G. *The research of features extraction registration and application*[D]. Changchun: Jilin University Master Degree Thesis, 2006. (in Chinese)

作者简介:



纪华(1983-),女,河北唐山人,博士研究生,主要从事景象匹配、视频跟踪方面的研究。E-mail: jh12@163.com



吴元昊(1977-),男,吉林长春人,博士,助理研究员,主要从事数字图像处理方面的研究。E-mail: kennth07@163.com



孙宏海(1980-),男,吉林省吉林市人,博士,助理研究员,主要从事数字图像处理方面的研究。E-mail: shh426@gmail.com

导师简介:



王延杰(1963-),男,吉林长春人,研究员,博士生导师,主要研究方向为数字图像处理。E-mail: wangyj@ciomp.ac.cn