一种新的局部不变特征检测和描述算法

杨恒王庆

(西北工业大学计算机学院 西安 710072)

摘 要 局部不变特征已经被成功地用来解决计算机视觉领域诸多实际问题. 文中提出一种新的局部不变特征检测和描述算法,提取出的特征能够对旋转、尺度缩放、光照等变化,甚至弱仿射变换保持不变. 一般说来,局部特征的提取分为特征检测和描述两个关键步骤. 在特征检测阶段,首先在每一层尺度图像上提取 Harris 角点,然后在以Harris 角点为中心的固定大小的搜索窗内搜索三维尺度空间的极值点作为局部特征点的位置和特征尺度,最后为每个特征点计算主方向. 文中的特征检测算法具有良好的可重复率性能. 在特征描述阶段,建立了梯度的距离和方向直方图来描述局部特征,文中的特征描述子不但具有良好的匹配性能,而且维数更低,十分有利于提高图像特征的匹配速度. 大量的图像匹配与图像检索实验结果验证了文中算法的有效性.

关键词 局部特征;特征检测;特征描述子;不变性;图像匹配 **POI** 号: 10,3724/SP.J. 1016,2010,00935

A Novel Local Invariant Feature Detection and Description Algorithm

YANG Heng WANG Qing

 $(School\ of\ Computer\ Science\ and\ Engineering\ ,\ Northwestern\ Polytechnical\ University\ ,\ Xi'an\quad 710072)$

Abstract Local invariant features have been successfully applied in many applications in computer vision. This paper proposes a novel local feature detection and description algorithm. The features are invariant to image rotation, scale and illumination changes, and even can be invariant to weak affine transformations. In general, the local feature extraction process can be divided into two key steps which are feature detection step and feature description step. In the detection step, firstly, the Harris corners are detected in every scale level image. Secondly, the local scale-space extrema is searched within a window which is center-localized on the multi-scale Harris corners. Finally, the predominant orientation is computed for each keypoint. The proposed feature detection algorithm has good repeatability performance. In the description step, a novel local descriptor is created based on the gradient distance and orientation histogram (GDOH). GDOH not only has good matching performance, but also has low dimensionality, which results in much faster feature matching speed. Extensive experimental results have demonstrated the effectiveness and efficiency of the proposed algorithm.

Keywords local feature; feature detection; feature descriptor; invariance; image matching

1 引 言

局部不变特征已经越来越广泛地应用在计算机

视觉的许多领域,如宽基线图像匹配、全景图拼接、目标识别、图像检索以及场景重建^[1-6].局部特征能够对图像间旋转、尺度缩放、视点变换、光照变化等图像变化因素保持不变性,对物体运动、遮挡、噪声

收稿日期:2008-01-03;最终修改稿收到日期:2009-11-06. 本课题得到国家自然科学基金(60873085)和国家"八六三"高技术研究发展计划项目基金(2007AA01Z314,2009AA01Z332)资助. 杨 恒,男,1980 年生,博士研究生,主要研究方向为图像处理与计算机视觉,研究兴趣包括局部不变特征提取、图像分类与检索等. 王 庆(通信作者),男,1969 年生,博士,教授,博士生导师,主要研究领域为图像处理和计算机视觉,研究兴趣包括基于图像的建模与渲染、图像分类与识别、动态光场信息处理与应用等. E-mail: qwang@nwpu. edu. cn.

等因素也可保持较好的可匹配性.局部特征建立在 尺度空间的理论上,一般说来,提取局部特征可以分 为两个关键步骤,即特征检测阶段和特征描述阶段. 特征检测阶段的任务是在尺度空间确定出关键点的 位置、特征尺度及其主方向,这些参数就可以描述出 一个局部图像区域.特征描述阶段的任务是为这个 局部图像区域建立描述子,该描述子应该能对图像 间的各种变化因素保持不变,以确保特征能够在这 些变化因素的影响下保持较好的可匹配性.近十几 年来,许多局部特征检测和描述算法被提出.

(1)特征检测

Schmid 等[7]利用 Harris 角点检测算法提取局 部特征,该局部特征能够抗旋转变换. Harris 特征 检测器对噪声和光照变化具有很强的鲁棒性,但是 它对图像间的尺度缩放不具备不变性[8]. Dufournaud 等[9]提出了多尺度 Harris 特征检测算法,即在 各个尺度图像上用 Harris 算子进行特征提取,该方 法可以解决不同尺度图像间的特征匹配问题. Brown 等[2]使用了改进的多尺度 Harris 特征提取 算法,用 Harris 矩阵特征值的调和平均值作为极值 的判别函数. 多尺度的 Harris 特征检测算法虽然可 以解决存在尺度缩放的图像间匹配特征的问题,但 是对于同一个图像局部结构,该算法会提取出许多 重复的特征,而这些特征仅仅在位置和尺度上略有 不同,这种大量的重复性会增加误配率和匹配计算 的复杂度. 为了克服这个缺点, Lindeberg[10] 在深入 研究尺度空间理论的基础上,提出 LoG(Laplacianof-Gaussian)尺度空间的三维局部极值点可以作 为稳定的局部特征. Lowe[11] 提出了著名的尺度不 变特征变换算法 SIFT (Scale Invariant Feature Transform), SIFT 在特征检测阶段采用 DoG (Difference-of-Gaussian)算子,搜索尺度空间上的 局部三维极值点,并以此来确定局部特征的位置和 尺度,再通过拟和三维二次函数进行位置和尺度的 精确定位,同时去除低对比度的特征点和不稳定的 边缘响应点,以提高抗噪声能力. DoG 算子是 LoG 算子的近似,但是其计算速度更快. Mikolajczyk^[12] 分析并指出 DoG 和 LoG 算子的缺点是提取出来的 特征未必都是稳定的. 因为这些极值点有可能会位 于亮度只在一个方向变化的边缘上,这样的点容易 受到噪声或者细小纹理变化的干扰. 同时, Mikolajczyk 还提出 Harris-Laplace 尺度不变特征检测算 法. 该算法首先在每一层尺度图像上检测 Harris 角点,然后用 Laplace 函数去除掉不在尺度轴上达 到极值的角点. Bay 等人^[13]于 2006 年提出 Fast-Hessian 特征检测算法,该算法基于汉森矩阵定位局部特征.

(2) 特征描述

Johnson 等[14]引入了一种可以描述物体形状的 纺线图(spin Image)作为描述子,该描述子能够对 特征点的刚性变换保持不变, Brown[2] 对关键点周 围的点进行采样组成一个 8×8 大小的图像片,再对 形成的 64 维描述向量进行规范化以消除光照影响, 最后对特征向量进行 Harr 小波变换,用前三个非 零小波系数作为索引以提高特征匹配速度. Gool[15] 引入了广义颜色矩反映信号的多谱特征,这些矩特 征描述了区域的形状和颜色特征. Florack 等[16] 用 一组不同阶的微分算子组成描述子,该描述子可以 保持旋转不变性. Schaffalitzky 等[1] 提出了一组复 滤波器,这种方法避免了特征主方向的估计问题,此 外复滤波器是正交的,因此可以用欧氏距离作为相 似性度量准则. Lowe[11] 提出了著名的 SIFT 描述 子,该描述子对特征周围一定区域内像素点梯度的 模值和方向用三维直方图进行统计,统计时对关键 点附近 4×4 个子区域都计算 8 个方向直方图,最终 形成 128 维特征向量. SIFT 描述子具有很强的抗噪 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. 但是由于 SIFT 提出的特征描述 子维数较高,使得进行图像匹配时速度较慢.为了降 低 SIFT 特征维数, Ke 等[17] 在 SIFT 算法基础上提 出了 PCA-SIFT 算法, 先对每个关键点周围 39×39 区域内像素统计梯度的两个垂直分量组成 3042 维的 描述向量,然后使用主成分分析(PCA)将高维特征 向量投影到前20维特征向量空间,由于PCA-SIFT 的维数低,使得其匹配速度大大增加. Mikolajczyk 等[18] 比较评价了各种建立描述子的算法性能,并通 过实验表明 SIFT 算法的性能是最优异的,同时 Mikolajczyk 还提出了一种扩展 SIFT 的描述子 GLOH (Gradient Location and Orientation Histogram),GLOH 采用对数极坐标的形式对关键点周 围划分出17个位置区域,对每一个位置区域内的梯 度统计 16 个方向的直方图,这样共形成 272 维特征 描述向量,最后再用主成分分析的方法将特征维数 降至 128 维.

本文提出一种新的尺度不变局部特征检测和描述算法,提取出来的特征能够对图像间旋转、尺度缩放、线性光照变化以及部分视点变换保持不变性.在特征检测阶段,我们在多尺度 Harris 角点附近一定

区域内搜索 DoG 尺度空间的局部极值点,以确定特征的位置和尺度,再采用与 SIFT 一样的方法统计每个特征点的主方向. 在特征描述阶段,我们在特征点周围的圆形区域内统计梯度的距离和方向直方图,并以此建立描述子. 实验表明,本文的特征检测算法在图像间具有尺度缩放、旋转变换、光照变化、噪声干扰以及小视点变换的情况下都具有良好的可重复率. 另外,本文提出的特征描述子不但与 SIFT, PCA-SIFT 以及 GLOH 描述子具有相当的性能,而且特征向量维数更低,使得在图像特征匹配的应用中具有明显的速度优势.

2 关键点检测

我们首先简要分析 Harris-Laplace 检测算法^[12], 并在此基础上提出本文的 Harris-Difference 特征检 测算法.

2.1 Harris-Laplace 特征检测算法

Harris-Laplace 算法首先在每一层尺度图像上使用 Harris 函数(如式(1))来定位图像中的角点,如果 Harris 函数在周围 8 个点中取得极值,则该点就为 Harris 角点. 其中 σ_I 表示积分尺度, σ_D 是差分尺度, $L(x,\sigma_n)=G(\sigma_n)*I(x)$ 表示尺度空间函数,由标准高斯核 $G(\sigma_n)$ 与图像 I(x) 卷积得到. μ 是Harris 矩阵,det() 表示矩阵的行列式,trace() 是矩阵的迹, α 是常数,取值为 0.04, L_x 和 L_y 分别是尺度函数对 x 和 y 的偏导数.

$$\mu(\boldsymbol{x}, \sigma_{l}, \sigma_{D}) = \sigma_{D}^{2}G(\sigma_{l}) * \begin{bmatrix} L_{x}^{2}(\boldsymbol{x}, \sigma_{D}) & L_{x}L_{y}(\boldsymbol{x}, \sigma_{D}) \\ L_{x}L_{y}(\boldsymbol{x}, \sigma_{D}) & L_{y}^{2}(\boldsymbol{x}, \sigma_{D}) \end{bmatrix}.$$

然后,对于每个 Harris 角点,用 Laplace 函数选择特征尺度,即如果 Laplace 函数值在尺度轴上与附近两个点比较达到极值并且大于给定阈值(如式(2)),则该角点作为稳定的特征点予以保留;否则就排除这个 Harris 角点.

 $LoG(\mathbf{x},\sigma_n) = \sigma_n^2 |L_{xx}(\mathbf{x},\sigma_n) + L_{yy}(\mathbf{x},\sigma_n)| > threshold,$ $LoG(\mathbf{x},\sigma_n) > LoG(\mathbf{x},\sigma_p), p \in \{n-1,n+1\}$ (2)
其中 LoG 表示 Laplace 函数, L_{xx} 与 L_{yy} 分别是尺度 函数对 x 和 y 的二阶偏导数,threshold 是给定的 阈值.

然而在实验中我们发现,如果用 Laplace 函数 选择特征尺度,仅有不到 10%的初始角点能够得到 保存,绝大多数的 Harris 角点将会被滤除掉. 另外, 选择的 Harris 角点也只是在一维尺度轴上的极值点,并不是三维尺度空间的局部极值点,这可能会导致不稳定的特征提取.

2.2 Harris-Difference 特征检测算法

针对 Harris-Laplace 算法的不足,本文提出 Harris-Difference 特征检测算法,该算法可以保留 相对较多的和相对稳定的 Harris 角点或者其附近 点,提高特征检测的性能. Harris-Difference 特征检 测算法分为以下 3 个步骤.

- 1. 先用文献[11]的方法构造金字塔尺度空间,然后在不同尺度图像上检测 Harris 角点.
- 2. 使用多尺度 Harris 角点作为初始点,在其附近一个固定大小的窗口内搜索 DoG 函数在三维局部空间(3×3×3)上的极值点,如式(3),其中 k 是常数,与金字塔每一层内不同尺度图像的个数有关.

$$DoG(\mathbf{x}, \sigma_n) = |L(\mathbf{x}, k\sigma_n) - L(\mathbf{x}, \sigma_n)|$$
 (3)

- 2.1. 选择一幅尺度图像上的一个 Harris 角点,验证其DoG 函数值是否是三维局部空间(3×3×3)的极值点,如果是,保留该 Harris 角点并作为特征点,跳转至步 2.3;否则,去除该 Harris 角点,跳转至步 2.2.
- 2.2. 在以当前角点为中心的 5×5 大小的窗口内,搜索DoG函数在局部三维空间(3×3×3)的极值点. 如果找到这样的点,只保留距离搜索窗中心最近的点并作为特征点,然后跳转至步 2.3;否则,直接跳转至步 2.3.
- 2.3. 如果每一层尺度图像上的所有 Harris 角点都已经 计算过,则保存所有特征点的位置和尺度信息,结束步 2;否则,跳转至步 2.1.

图 1 显示了两组不同变换条件下图像对随着搜索窗尺寸变化的可重复率曲线(见 5.1 节),可以看出,在搜索窗大小为 5 的时候,可重复率最高,当搜索窗大于 5 以后,检测算法的性能反而会略有下降.原因是当搜索窗过大后,虽然可以找到更多的特征点,但这些点未必是稳定的,它们并不能保证可重复率会提高.因此,在本文的后续实验中,搜索窗大小都设置成 5×5.

3. 使用文献[11]中的方法,为每个特征点计算主方向. 在后面建立描述子的过程中,通过旋转坐标至主方向上,可以使得描述子具有旋转不变性.

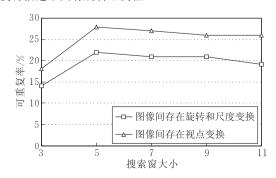
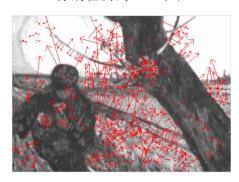


图 1 随搜索窗口大小变化的可重复率曲线

图 2 是本文算法在两幅标准图像上提取出来的局部特征点的示例,左右两幅图像大小分别是512×348 和 850×684,分别检测到 896 个和 6762





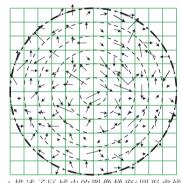
个局部特征点.特征点用箭头表示,箭头末端表示特征点位置,箭头长度表示特征尺度大小,箭头指向表示特征点主方向.



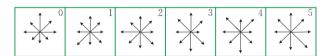
3 局部特征描述子

图像的局部特征描述子应该能够对图像间的旋转、尺度缩放、光照变化甚至小视点变换保持不变性,而且具有很高的区分力和鲁棒性.受 SIFT 描述子的启发,本文的描述子也是基于特征点邻域梯度统计的思想,以增强描述子的抗噪声能力以及对定位误差的容错性.建立新描述子的目标是保持与SIFT、PCA-SIFT 以及 GLOH 这些描述子相当的性能,同时降低描述向量的维数,以减小图像特征匹配的时间复杂度.

本文提出的描述子基于梯度的距离和方向直方 图(Gradient Distance and Orientation Histogram, GDOH),图 3 解释了建立描述子的过程.首先,计算 特征点周围圆形区域内的梯度的模值和方向(如 图 3(a)中的箭头),圆形区域的大小与特征点的尺 度成正比. 然后,用一个高斯窗函数对每一个梯度模 值加权,高斯窗大小与圆形区域大小一致(如 图 3(a)中的最外层圆形点画线). 高斯窗加权的目 的是为了减少距离中心远的点的影响. 接着,将每个 梯度方向旋转至主方向以保证旋转不变性,同时计 算每个梯度点到中心点的距离. 根据距离和旋转后 的方向,建立具有48个单元(距离划分6个单元,方 向划分8个单元)的直方图,直方图每个单元的统计 值是由加权的梯度模值累加得到的,描述子如图 3 (b) 所示. 一般来讲, 计算出来的距离单元和方向单 元都不是整数,为了使得相邻距离单元以及相邻方 向单元间过渡平滑,我们采用双线性插值算法来分 配梯度模值对相邻单元间的贡献, 最终形成的 48 维 描述子向量就对应着直方图的每个单元的值,如 图 3(b)中的箭头长度. 最后,将特征向量归一化以进一步去除光照的影响.



(a) 描述子区域内的图像梯度(圆形虚线表示不同距离的划分,圆形点画线表示高斯加权窗,黑色箭头表示梯度的模值和方向)



(b) GDOH描述子(6个格子表示距离的划分, 箭头指向表示8个 方向划分, 箭头长度表示直方图每个单元的值)

图 3 GDOH 描述子建立过程示意图

在实验中,我们发现距离划分6个单元且方向划分8个单元或者距离划分8个单元且方向划分8个单元和可以取得最优的性能,这两种方案分别可以产生48维和64维的特征描述向量,考虑到向量维数对后面匹配速度的影响,我们在后续实验中都选择前一种划分方法,即为每一个特征建立一个48维的描述向量.

4 特征匹配

特征描述子建立好以后,使用 BBF (Best-Bin-First)^[11]算法进行特征点间的粗匹配. BBF 是一种

近似的最近邻搜索算法,能够以很高的概率找到待匹配点的最近邻点和次近邻,如果最近邻和待匹配点的距离与次近邻和匹配点的距离之比小于匹配门限阈值(实验中取 0.4),则认为该最近邻点是待匹配点的候选匹配.在对所有点进行粗配之后,使用RANSAC(RANdom Sample Consensus)算法估计两个图像对之间的单应变换矩阵并作为几何约束,进而去除误配,完成特征点之间的精确匹配.

5 实验结果和分析

实验中选用的测试图像均来自标准图像数据 集^①,该图像数据库中图像包含了各种变换条件下 的场景图像.实验采用的评价准则包括:

- (1)可重复率(repeatability)^[14],评价特征检测算法.
- (2) 查全率-查错率(recall vs. 1-precision)曲线^[17],评价描述子的性能.
 - (3) 匹配的时间代价,比较特征匹配的速度.
 - (4)特征匹配实例,主观比较匹配性能.
- (5)图像检索应用,验证本文算法在实际应用中的性能.

所有的实验都运行在 $2.80 \mathrm{GHz}$ CPU 和 $512 \mathrm{MB}$ 内存的 PC 机环境上,采用 VC6. 0 编程语言实现.

5.1 可重复率

可重复率是指在两幅图像上检测到的对应特征数的比率,它体现了特征检测算法的性能,是衡量检测算法的重要指标. 如果检测算法在两幅待匹配图像上的相同区域内分别检测出 m_1 和 m_2 个特征点,则可重复率 r_R 计算如下:

$$r_{R} = \frac{C(m_{1}, m_{2})}{\min(m_{1}, m_{2})} \tag{4}$$

其中 $\min(m_1, m_2)$ 为两幅图像中最少的特征点数. $C(m_1, m_2)$ 指的是对应特征点数,对应的特征点需要满足如下两个条件[12]:

- (1) 特征点的位置误差 $\varepsilon_L = |x_a H \cdot x_b| <$ 误差限阈值,其中 H 是两幅图像间的单应变换矩阵,误差限阈值是设定好的常数(3 像素).
- (2) 特征区域的表面误差 $\varepsilon_S = \left| 1 s^2 \frac{\sigma_a^2}{\sigma_b^2} \right| < 0.2$, 其中 s 是图像间实际的尺度缩放因子, σ_a 和 σ_b 分别是两个特征点的特征尺度.

图 4~图 8 分别给出了随着图像间的尺度变化、旋转变化、光照变化、噪声干扰以及视点变换的

条件下, DoG、Harris-Laplace、Fast-Hessian 以及本 文的 Harris-Difference 特征检测算法的可重复率变 化曲线,图 4(a)1 为参考图像,2~5 分别是尺度放 大 2 倍~5 倍的变换图像,从图 4(b)可以看出在有 尺度缩放变化的条件下,DoG 算法可以获得最好的 可重复率性能,而本文的 Harris-Difference 算法要 略微低于 DoG 算法,但是要高于 Fast-Hessian 以及 Harris-Laplace 算法. 图 5(a)1 为参考图像,2~7 分 别是旋转角度为 30°,60°,90°,120°,150°以及 180°的 变换图像,从图 5(b)可以看出,4 种算法的可重复率 曲线都具有周期性, 当旋转角度为 90°和 180°时出 现峰值,本文 Harris-Difference 算法的可重复率略 微低于 DoG 算法,但仍然高于 Fast-Hessian 以及 Harris-Laplace 算法. DoG 算法在尺度和旋转变换 中都获得了最好的检测性能,原因是 DoG 算法检测 了尺度空间中所有的局部极值点, 但是 DoG 检测出 来的特征点可能是不稳定的,因为它检测的这些极 值点有可能会落在亮度信号只在一个方向变化的边 缘附近,这就导致这些检测出来的特征点对于噪声 干扰或者细微的纹理变化很敏感, 而基于 Harris 的 检测算法以及基于 Hessian 矩阵的算法检测的是图 像角点特征,即亮度信号在两个正交方向同时变化 比较明显的图像特征点,这种特性使得这类算法具 有较高的鲁棒性和抗干扰能力,这种特性在图 6、 图 7 和图 8 的结果中得到了体现. 图 6(a)1 为参考 图像,2~6分别是亮度增强2倍~6倍的变换图像, 从图 6(b)可以看出在有亮度变化的条件下,4 种算 法的可重复率性能相当,而 Harris-Difference 算法 要略微优于其它3种算法.图7(a)1为参考图像, 2~6 是加入了高斯噪声的模糊图像,其高斯半径 σ 分别为 1,1.5,2,2.5,3,从图 7(b)可以看出在有噪 声污染的条件下, Harris-Difference、Harris-Laplace 和 Fast-Hessian 3 种算法的可重复率都要明显高于 DoG 算法,而 Harris-Difference 算法的性能是最优 的.图 8 给出了在视点变换条件下的实验结果,(a)1 为参考图像,2~6的视点变换角度分别为20°,30°, 40°,50°以及60°,从图8(b)中可以看出在视点变换 不大的情况下(不超过 40°), Harris-Difference 算法 的可重复率要高于其它3种算法,而当视点变换超 过 40°时,由于图像间发生较大的畸变,4 种算法的 可重复率均降为 0. 综合以上实验结果可以得出结

① 实验图像数据集来源. http://lear. inrialpes. fr/people/mikolajczyk/

论,Harris-Difference 算法在图像间具有几何变换 (旋转和尺度变换)的条件下,可重复率要略低于 DoG 算法;而在图像间具有光照变化、高斯污染以 及视点变换的条件下,Harris-Difference 算法性能 要优于 DoG 算法.此外,在任何条件下,本文算法都 优于 Harris-Laplace 和 Fast-Hessian 算法.这说明 本文提出的 Harris-Difference 算法具有良好的特征 检测性能.

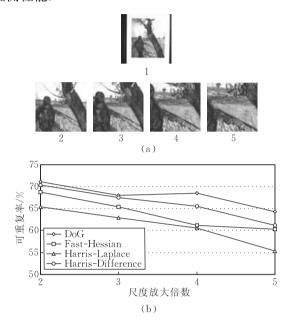


图 4 4 种算法随着尺度变化的可重复率曲线

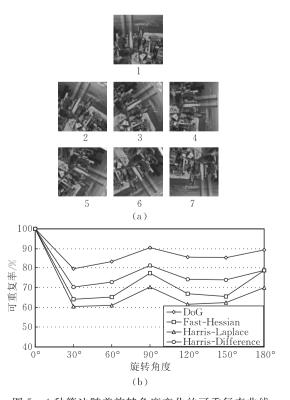


图 5 4 种算法随着旋转角度变化的可重复率曲线

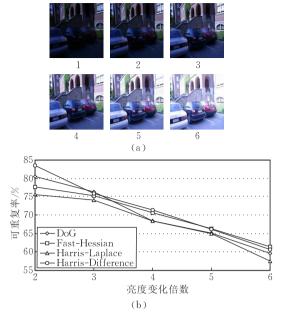


图 6 4 种算法随着亮度变化的可重复率曲线

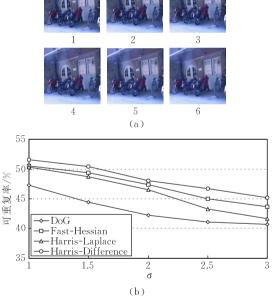


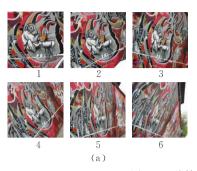
图 7 4 种算法随着不同程度高斯噪声干扰情况 下的可重复率曲线

5.2 查全率-查错率曲线

查全率-查错率曲线是对描述子性能的评价,查全率(recall)是指正确匹配数占对应(correspondence)特征数目的比率,查错率是错误匹配数目占整个匹配数目(正确匹配与错误匹配之和)的比率.查全率-查错率曲线评价准则的思想就是增加正确匹配数目的同时减少查错率.图 9 是评价查全率-查错率性能的 4 组实验数据,(a)中图像对间存在 55°的旋转变换和 1.6 倍的尺度缩放;(b)中图像间存在着 20°的视点变换;(c)中图像对间存在着 1.2 倍的

亮度变化;(d)中图像存在高斯噪声污染,高斯半径 σ=3. 图 10 分别给出了对应图 9 中的 4 组实验数据 的 SIFT、PCA-SIFT、GLOH 以及本文的 GDOH 算 法随着不同的匹配门限做出的查全率-查错率曲线

图,从图 10(a)中可以看出,当图像间存在几何变换 (旋转和尺度缩放)时,4 种描述子的性能是接近的, 当查错率小于 0.3 时,GLOH 的性能最优,而当查 错率大于 0.3 时,GDOH 的性能最优.当图像间存



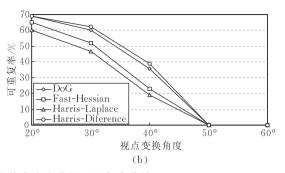


图 8 4 种算法随着视点变换角度变化的可重复率曲线

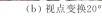








(a) 旋转变换55°和尺度变换1.6倍







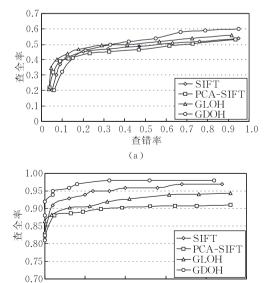




(c) 亮度变化1.2倍

(d) 高斯噪声, σ=3

图 9 评价查全率-查错率性能的 4 组实验数据



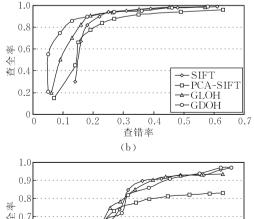
0.05

0.10

查错率 (c)

0.15

0.20



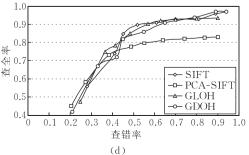


图 10 分别对应图 9 的 4 组不同变换类型下的 4 种描述子的查全率-查错率曲线比较

0.25

在较小的视点变换时(见图 10(b)),GDOH 的性能最优,尤其是在查错率低于 0.2 的情况下,GDOH 的优势较为明显,在查错率高于 0.2 时,4 种描述子的性能十分接近.图 10(c)的结果表明,在光照变化的情况下,4 种描述子都可以达到很好的性能(查全率都在 0.8 以上),而 GDOH 获得了最优的性能.在图像间存在噪声干扰的情况下(见图 10(d)),当查错率低于 0.5 时,PCA-SIFT 描述子的性能要明显低于其它 3 种描述子.综合这些实验结果可以得出结论,本文的 GDOH 描述子的性能和目前优秀的 SIFT、PCA-SIFT 以及 GLOH 描述子性能相当,甚至在某些变换条件下 GDOH 性能更优,比如小视点变换和光照变化.

5.3 特征匹配的时间复杂度比较

我们从图像数据集中选取 79 对图像,包括各种不同的图像变换类型(见表 1),比较 SIFT 算法和本文算法的特征检测和特征匹配的时间复杂度.表 1 给出实验结果,可以看出本文算法的平均特征检测时间与 SIFT 几乎一样,没有很明显的速度增益,但是总的平均匹配时间仅为 SIFT 的 40.9%(1419/3468),特征匹配的速度增益非常显著,这是由于本文提出的描述子向量维数更低.

表 1 本文提出的算法与 SIFT 算法的时间代价比较

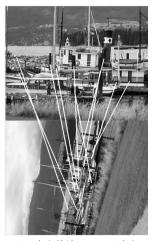
	SIFT/ms		本文算法/ms		
	平均检测 时间	平均匹配 时间	平均检测 时间	平均匹配 时间	
旋转 & 尺度 缩放(29 对)	35636	5841	33020	2303	
视点变换 (25 对)	29053	3694	20530	1586	
光照变化 (25 对图像)	7120	489	6911	226	
总平均	24529	3468	20805	1419	

5.4 特征匹配实例

图 11 给出了一组匹配结果的实例,左图和右图分别显示了 SIFT 算法和本文算法的匹配结果,设定的误差限是 1.5 个像素,这里只显示前 10 个特征匹配作为比较. 匹配的图像对之间的旋转变化和尺度缩放都比较大,而且图像细节信息比较丰富. 可以看出,在这个例子当中本文算法的正确匹配率明显高于 SIFT 算法,SIFT 算法的正确匹配率明显高于 SIFT 算法,SIFT 算法的正确匹配率仅有50%,而本文的算法对于所有的匹配都正确. 同时可以发现,SIFT 算法在这种具有丰富细节信息的图像间进行特征匹配的表现并不理想,比较容易受到细

小纹理的干扰,导致香错率比较高.





(a) SIFT: 5/10正确率

(b) 本文算法: 10/10正确率

图 11 SIFT 算法和本文算法的匹配实例比较

5.5 图像检索应用

我们在一个小的图像数据库上^①进行图像检索 实验以验证本文提出算法的有效性. 该数据库一共 有 30 幅图像,包含 10 组不同的物体场景,每组是由 3幅包含同一物体场景的不同变换的图像组成.我 们分别用SIFT算法和本文提出的算法分别提取图 像局部特征,然后为每个局部特征计算描述子向量, 接着对每两幅图像进行特征匹配,并将匹配的特征 个数作为图像间相似度测量准则. 最后按照文献[17] 的计分规则,对于每一幅图像返回2幅检索到的最 相似图像,如果这两幅图像和待检索图像在同一组, 加2分;如果有一幅在同一组,加1分;否则,没有加 分. 在这样的规则下满分可以达到 60 分. 用两种算 法得到的分数分别除以60就可以得到图像检索的 正确率. 表 2 分别列出了 SIFT 算法和本文算法的 图像检索的正确率,以及两种算法的平均特征检测 时间和平均特征匹配时间. 可以看出,在这个实验中 本文算法的图像检索性能要优于 SIFT 算法,比 SIFT 算法的检索正确率高出 6.7%; 在特征检测速 度方面,两种算法基本持平;而由于本文描述子维数 更低,使得本文算法的图像特征匹配速度优势明显, 比 SIFT 算法的速度提高 62.4%.

表 2 两种算法在图像检索实验中的性能比较

	图像检索 正确率/%	平均特征检测 时间/ms	平均特征匹配 时间/ms
SIFT	48.3	4712	682
本文算法	55.0	4582	420

① 图像检索数据库, http://www.cs.cmu.edu/~yke/pcasift/

6 结 论

本文提出了一种新的局部不变特征检测和描述算法,并应用于宽基线的图像匹配与检索.在局部特征检测方面,本文提出 Harris-Difference 算法,首先以多尺度的 Harris 角点作为初始点,然后在其附近一定范围内搜索 DoG 函数的局部空间极值点作为特征点的位置和尺度信息,最后为每个特征点计算主方向. Harris-Difference 具有比 Harris-Laplace和 Fast-Hessian 更高的可重复率,并且比 DoG 算法具有更好的抗干扰能力.在特征描述方面,本文提出的 GDOH 描述子不但与目前优秀的 SIFT,PCA-SIFT 以及 GLOH 描述子性能相当,甚至在某些变换条件下 GDOH 性能更优,比如小视点变换和光照变化,而更为重要的是 GDOH 特征向量的维数更低,十分有利于提高图像特征匹配的速度.大量图像匹配和检索的实验结果验证了本文算法的有效性.

参考文献

- [1] Schaffalitzky F, Zisserman A. Multi-view matching for unordered image sets, or "how do I organize my holiday snaps?"// Proceedings of the 7th European Conference on Computer Vision, Copenhagen, 2002; 414-431
- [2] Brown M, Szeliski R, Winder S. Multi-image matching using multi-scale oriented patches//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. San Diego, 2005: 510-517
- [3] Snavely N, Seitz S M, Szeliski R. Photo tourism: Exploring photo collections in 3D. ACM Transactions on Graphics, 2006, 25(3): 835-846
- [4] Brown M, Lowe DG. Automatic panoramic image stitching using invariant features. International Journal of Computer Vision, 2007, 74(1): 59-73
- [5] Mikolajczyk K, Leibe B, Schiele B. Local feature for object class recognition//Proceedings of the 10th IEEE International Conference on Computer Vision. Beijing, 2005: 1792-1799
- [6] Ming An-Long, Ma Hua-dong. Region-SIFT descriptor based correspondence between multiple cameras. Chinese

- Journal of Computers, 2008, 31(4): 650-661(in Chinese) (明安龙, 马华东. 多摄像机之间基于区域 SIFT 描述子的目标匹配. 计算机学报, 2008, 31(4): 650-661)
- [7] Schmid C, Mohr R. Local grayvalue invariants for image retrieval. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(5): 530-534
- [8] Schmid C, Mohr R, Bauckhage C. Evaluation of interest point detectors. International Journal of Computer Vision, 2000, 37(2): 151-172
- [9] Dufournaud Y, Schmid C, Horaud R. Matching images with different resolutions//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Hilton Head Island, South Carolina, 2000: 612-618
- [10] Lindeberg T. Feature detection with automatic scale selection. International Journal of Computer Vision, 1998, 30 (2): 79-116
- [11] Lowe D G. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 2004, 60(2): 91-110
- [12] Mikolajczyk K, Schmid C. Scale & affine invariant interest point detectors. International Journal of Computer Vision, 2004, 60(1): 63-86
- [13] Bay H, Tuytelaars T, Gool L V. SURF: Speeded up robust features//Proceedings of the 9th European Conference on Computer Vision. Graz, Austria, 2006: 404-417
- [14] Johnson A E, Hebert M. Recognizing objects by matching oriented points//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Puerto Rico, 1997; 684-689
- [15] Gool L V, Moons T, Ungureanu D. Affine/photometric invariants for planar intensity patterns//Proceedings of the 4th European Conference on Computer Vision. Cambridge, UK, 1996: 642-651
- [16] Florack L, Romeny B H, Koenderink J J, Viergever M A. General intensity transformations and differential invariants. Journal of Mathematical Imaging and Vision, 1994, 4(2): 171-187
- [17] Ke Y, Sukthankar R. PCA-sift: A more distinctive representation for local image descriptors//Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Washington, DC, 2004; 506-513
- [18] Mikolajczyk K, Schmid C. A performance evaluation of local descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1615-1630



YANG Heng, born in 1980, Ph. D. candidate. His research interests are image processing and computer vision, including local invariant feature extraction, image classification and retrieval.

WANG Qing, born in 1969, Ph. D., professor, Ph. D. supervisor. His research interests are image processing and computer vision, including image based modeling and rendering, image categorization and recognition and dynamic light field photograph.

Background

This paper focuses on the issues of local feature detection and description which can be widely applied in computer vision applications, such as wide-baseline image matching and image retrieval. In feature detection aspect, recent literatures are all based on the scale space theory. The proposed Harris-Difference algorithm outperforms other traditional algorithms, such as Harris-Laplace and Fast-Hessian, in terms of repeatability criterion. Furthermore, it obtains higher performance in noise resistance than DoG algorithm. In feature description aspect, the proposed GDOH (Gradient Distance and Orientation Histogram) algorithm can achieve similar and even better feature matching performance than the stateof-the-art descriptors, such as SIFT, PCA-SIFT and GLOH, which are widely used in vision applications. Moreover, the dimensionality of the GDOH descriptor is much lower than these widely used descriptors which results in higher efficiency in feature based image matching and retrieval systems. The research work and the result of this paper is a part of the project supported by National Natural Science Foundation of China (No. 60873085) and National High Technology Research and Development Program (863 Program) of China (grant No. 2007 AA01Z314). The NSFC project is to solve the issues of feature correspondence and optimization from large scale unordered wide baseline images. The '863' project aims at 4D scene reconstruction based on multiple webbaseline images. The images are automatically grouped according to the time sequence and the correspondent reconstructed 3D scene can be viewed on time line. The system provides users an efficient way to organize large amount unordered images in the temporal and spatial orders. Some systems on the reconstruction from images have been developed in recent years. The most representative one is the Photo Tourism, which can reconstruct 3D scene from hundreds of images in a few days. The project aims to improving the reconstruction accuracy and speed based on larger scale image set. The first task of the projects is to improve the feature matching performance, which is the key and important step of the whole system. This paper is to address such issue. The proposed method can improve the performance of feature matching in both the detection and description aspects.