

中图分类号: TP391.4 文献标志码: A 文章编号: 1006-8961(2011)02-0141-11
论文索引信息: 孙浩, 王程, 王润生. 局部不变特征综述 [J]. 中国图象图形学报, 2011, 16(2): 141-151

局部不变特征综述

孙浩¹⁾, 王程^{1), 2)}, 王润生¹⁾

¹⁾(国防科学技术大学 ATR 国家重点实验室, 长沙 410073) ²⁾(厦门大学计算机科学系, 厦门 361005)

摘要: 局部不变特征是近年来计算机视觉领域的研究热点。局部不变特征在宽基线匹配、特定目标识别、目标类别识别、图像及视频检索、机器人导航、场景分类、纹理识别和数据挖掘等多个领域得到了广泛的应用。本文基于局部不变特征检测、局部不变特征描述和局部不变特征匹配 3 个基本问题, 综述了文献中现有的局部不变特征研究方法, 并比较了各类方法的优缺点。根据特征层次的不同, 局部不变特征检测方法可以分为角点不变特征、bbb 不变特征和区域不变特征检测方法 3 类。局部不变特征的描述方法可以分为基于分布的描述方法、基于滤波的描述方法、基于矩的描述方法和其他描述方法。局部不变特征匹配的研究主要集中在相似性度量、匹配策略和匹配验证 3 个方面。最后在分析各类研究方法的基础上, 总结了局部不变特征研究目前存在的一些问题及可能的发展方向。

关键词: 局部不变特征; 局部不变特征检测; 局部不变特征描述; 局部不变特征匹配

A review of local invariant features

Sun Hao¹⁾, Wang Cheng^{1), 2)}, Wang Runsheng¹⁾

¹⁾(State Key Laboratory of ATR, National University of Defense Technology, Changsha 410073 China)

²⁾(Department of Computer Science, Xiamen University, Xiamen 361005 China)

Abstract Local invariant features are receiving increasing attention from computer vision research community. Local invariant features have been widely utilized in a large number of applications e.g., wide baseline matching, object recognition, and categorization, image retrieval, visual search, robot localization, scene classification, texture recognition and data mining. This paper gives an overview of the various approaches and properties of local invariant features. We focus on three major areas: (1) local invariant feature detectors, (2) local invariant feature descriptors, and (3) local invariant feature matching. Most of the existing local invariant feature detectors can be categorized into corner detectors, bbb detectors or region detectors. Local descriptors can be categorized into distribution-based, filter-based, moment-based descriptors and others descriptors. Similarity measurement, matching strategy and matching verification are three key components of robust matching algorithms. Finally, some research challenges and future directions are discussed.

Keywords local invariant features; local invariant feature detector; local invariant feature descriptors; local invariant feature matching

0 引言

图像局部特征的研究已经有很长的历史, 早期

研究可以追溯到 20 世纪 70 年代的 Moravec 算子^[1]。文献中存在大量关于角点、边缘、bbb 和区域等局部特征的研究方法。近年来区分性强、对多种几何和光度变换具有不变性的局部不变特征在宽

收稿日期: 2009-05-18 修回日期: 2009-11-04

基金项目: 国家自然科学基金项目 (40971245)。

第一作者简介: 孙浩 (1984—), 男, 国防科学技术大学电子科学与技术专业博士研究生, 研究领域包括图像分析、理解与信息融合。E-mail: clhaosun@gmail.com

通讯作者: 王程, chwang_nud@263.net

基线匹配^[2-4]、特定目标识别^[5-8]、目标类别识别^[9-10]、图像及视频检索^[11-13]、机器人导航^[14]、场景分类^[15]、纹理识别^[16-17]和数据挖掘^[18]等多个领域内获得广泛的应用,是国内外的研究热点。

局部不变特征是指局部特征的检测或描述对图像的各种变化,例如几何变换、光度变换、卷积变换、视角变化等保持不变。局部不变特征的基本思想是提取图像内容的本质属性特征,这些特征与图像内容的具体表现形式无关或具有自适应性(即表现形式变化时特征提取自适应的变化以描述相同的图像内容)。局部不变特征通常存在一个局部支撑邻域,与经典的图像分割算法不同,局部支撑邻域可能是图像的任何子集,支撑区域的边界不一定对应图像外观(例如颜色或纹理)的变化。

局部不变特征不仅能够观测条件变化大、遮挡和杂乱干扰的情况下获得可靠的匹配,而且能够有效描述图像内容进行图像检索或场景、目标识别等。局部不变特征可以克服语义层次图像分割的需要。从复杂背景中分割出前景目标是十分困难的课题,基于低层特征的方法很难实现有意义的分割,把图像内容表示为局部不变区域的集合(多个区域可能存在重合,图像中一些部分也可能不存在局部不变区域),可以回避分割问题。基于局部不变特征的方法本质上是对图像内容进行隐式分割,局部不变特征既可能位于感兴趣的前景目标上也可能位于背景或目标边界上,后续的高层处理需要基于局部不变特征提取感兴趣的信息。

局部不变特征的研究包含 3 个基本问题:一是局部不变特征的检测,二是局部不变特征的描述,三是局部不变特征的匹配。根据不同的准则,局部不变特征的研究方法可以分为不同的类别,按照使用的色调空间的不同可以分为局部灰度不变特征和局部彩色不变特征;按照特征层次的不同可以分为角点不变特征、b bb 不变特征和区域不变特征;按照几何变换不变性的自由度可以分为平移不变特征、旋转不变特征、尺度不变特征、欧氏不变特征、相似不变特征、仿射不变特征和投影不变特征;按照处理思路的不同可以分为基于轮廓曲率的不变特征、基于灰度梯度、灰度变化和显著性的不变特征,基于生物视觉启发的不变特征,基于多尺度的不变特征和基于分割的不变特征。

1 相关概念

局部不变特征研究涉及很多概念,本节首先从数学形式上对局部不变特征进行了描述,然后阐述了局部不变特征的主要性质。由于文献中大量的研究是关于局部特征对平面几何变换和光度变换的不变性研究,本节介绍了几何变换^[19]的概念并列举了 4 种常见的平面几何变换及其性质,以 RGB 空间为例介绍了光度变换^[20]的概念。局部不变特征尺度不变特性的研究是基于图像的多尺度表示及自动尺度选择,最后给出了尺度空间的描述及其基本定义。

1.1 局部不变特征

图像函数表示为 $f(x, y)$, $g_{\Omega}(f)$ 为定义在图像局部邻域 Ω 上的特征函数, $H(f)$ 表示对图像进行的各种变换。特征函数 $g_{\Omega}(f)$ 对变换 H 具有不变性是指对任意的图像函数 f 满足:

$$g_{\Omega}(f) = g_{\Omega}(H(f)) \quad (1)$$

特征函数 $g_{\Omega}(f)$ 对变换 H 具有不变性时提取的特征为局部不变特征,其中特征不变性的自由度由 H 的自由度决定。

1.2 局部不变特征性质

局部不变特征应该具有以下特性:

1)重复性 相同场景或目标在不同成像条件下图像提取的局部不变特征应该是相同的;

2)区分性 局部不变特征应包含较大的灰度或色度模式变化,易于区分;

3)局部性 局部不变特征应具有局部性,减小遮挡的概率,同时可以采用简单的变换模型对图像间的变换进行近似建模;

4)精确性 局部不变特征应可以在空域、尺度域及形状域上精确定位;

5)不变性 局部不变特征的检测和描述对各种变换应具有不变性;

6)鲁棒性 局部不变特征的检测和描述应对图像噪声、量化误差、模糊等不敏感。

1.3 几何变换

几何变换作用于图像平面空间坐标^[19],假设变换矩阵用 H_g 表示,则几何变换可表示为

$$(x', y', 1)^T = H_g(x, y, 1)^T \quad (2)$$

表 1 列举了 4 种平面几何变换及其主要性质。

表 1 平面几何变换及其性质

Tab 1 Planar transformation properties

类别	变换矩阵	自由度	不变量
欧氏变换	$\begin{pmatrix} r_{11} & r_{12} & t_x \\ r_{21} & r_{22} & t_y \\ 0 & 0 & 1 \end{pmatrix}$	3	距离 面积
相似变换	$\begin{pmatrix} sr_{11} & sr_{12} & t_x \\ sr_{21} & sr_{22} & t_y \\ 0 & 0 & 1 \end{pmatrix}$	4	距离比值 角度
仿射变换	$\begin{pmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{pmatrix}$	6	平行 平行线距离比 面积比
投影变换	$\begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}$	8	共线性 距离比的比

1.4 光度变换

光度变换作用于像素的灰度或色度值^[20], 以 RGB 颜色空间为例, 光度变换可以表示为

$$\begin{pmatrix} r_2 \\ g_2 \\ b_2 \end{pmatrix} = \begin{pmatrix} \alpha_r & 0 & 0 \\ 0 & \alpha_g & 0 \\ 0 & 0 & \alpha_b \end{pmatrix} \begin{pmatrix} r_1 \\ g_1 \\ b_1 \end{pmatrix} + \begin{pmatrix} \beta_r \\ \beta_g \\ \beta_b \end{pmatrix} \quad (3)$$

式中 r, g, b 为 RGB 空间的 3 种彩色波段, α 和 β 为标量照度因子。

1.5 尺度空间

尺度空间理论是多尺度图像表示的框架^[21], 其基本思想是描述自然界目标的多尺度特性, 即观测尺度的不同导致对目标特性感知的不同。在无法获得感兴趣尺度的先验信息下, 面向未知场景的图像解译算法需要同时考虑多尺度的图像表示。2 维图像函数 $f(x, y)$ 的尺度空间表示 L 定义为

$$L(x, y; \sigma) = \iint (\xi x - \eta y - \eta) \varphi(\xi, \eta; \sigma^2) d\xi d\eta \quad (4)$$

式中 φ 表示尺度参数为 σ 的高斯核。在图像的多尺度表示基础上, 通过基于尺度空间归一化导数最大化的自动尺度选择, 可以实现图像特征提取的尺度不变特性。特征提取算法可以适应不同尺寸的目标、局部结构或目标与成像系统之间距离变化等因素导致的未知尺度变化。

2 局部不变特征检测

是在图像中定位感兴趣的点、blob 边缘或区域。按照特征层次的不同, 文献中的局部不变特征检测算法可分为角点不变特征、blob 不变特征和区域不变特征检测算法 3 类。

2.1 角点特征

2.1.1 Moravec 算子

Moravec 算子^[1]通过滑动二值的矩形窗口寻找最小灰度变化的局部最大值。Moravec 算子定义一个像素点为角点的条件是该像素点在各个方向上都具有较大的灰度变化。Moravec 算子的缺点是由于窗口的滑动只在每个 45° 方向故算子响应具有非等方性, 容易检测边缘上的点。Moravec 算子具有平移变换不变性。

2.1.2 Harris 算子

Harris 算子^[22], 也称为 Plessey 算子, 是由 Harris 和 Stephens 为了改善 Moravec 算子性能提出的。Harris 算子以二阶矩阵 (又称为自相关矩阵) 为基础, 二阶矩阵描述了像素点局部邻域内的梯度分布信息:

$$M = \sigma_D^2 \varphi(\sigma_I) * \begin{pmatrix} I_x^2(x, \sigma_D) & I_x I_y(x, \sigma_D) \\ I_x I_y(x, \sigma_D) & I_y^2(x, \sigma_D) \end{pmatrix} \quad (5)$$

二阶矩阵通过差分尺度为 σ_D 的高斯核进行局部图像导数的计算, 然后利用积分尺度为 σ_I 的高斯平滑窗对像素点局部邻域内的导数进行加权平均, Harris 算子采用角点响应函数作为检测角点特征的依据:

$$c = \det M - \lambda \text{tr}(M) \quad (6)$$

式中 λ 常取 0.04

Harris 算子具有平移和旋转不变性, 对光照条件的变化不敏感。在文献 [23] 的角点特征的比较实验中, Harris 角点特征的重复性和区分性被证明是最好的。通过在局部极值点的邻域内对角点响应函数进行二次逼近, Harris 算子可以达到亚像素的定位精度。

2.1.3 SUSAN 算子

考虑到基于局部梯度的方法对噪声影响比较敏感而且计算量大, Smith 和 Brady^[24] 提出了一种基于形态学的角点特征检测方法。如果多个像素属于同一目标, 那么在相对较小的局部邻域内像素的亮度应该是一致的。基于这一假设, SUSAN 算子通过在圆形模板区域内进行亮度比较检测角点特征。对于图像中的每一个像素, 考虑一个固定半径的圆形邻

域,以该像素作为中心参照,圆形邻域内的所有像素根据与参照像素的亮度关系,被分类成相似像素和不相似像素。通过这种方式为每个像素点生成一个关联的局部亮度相似性区域,区域的大小包含了该像素点处的图像结构信息,如图 1 所示,其中圆形邻域内的黑色区域代表相似区域,白色区域代表不相似区域。SUSAN 算子定义一个像素点为角点的条件为像素点的关联相似性区域内的像素数达到局部极小值并且小于预先设定的固定门限。SUSAN 算子具有平移和旋转不变性。

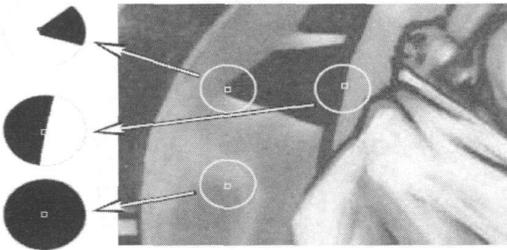


图 1 SUSAN 算子

Fig 1 SUSAN corner detector

2.1.4 FAST 算子

Rosten 等人^[25]在 SUSAN 角点特征检测方法基础上利用机器学习方法提出 FAST 角点算子。FAST 算法包含 3 个主要步骤:

- 1) 对固定半径圆上的像素进行分割测试,通过逻辑测试可以去处大量的非特征候选点;
- 2) 基于分类的角点特征检测,利用 ID3 tree 分类器^[26]根据 16 个特征判决候选点是否为角点特征,每个特征的状态为 $-1, 0, 1$;
- 3) 利用非极大值抑制进行角点特征的验证。

FAST 角点算子具有平移和旋转不变性、可靠性高、对噪声鲁棒性好、计算量小。

2.1.5 Harris-Laplace 算子

局部邻域泰勒展开可得到 Hessian 矩阵^[27]

$$\mathbf{H} = \begin{pmatrix} I_{xx}(x, \sigma_D) & I_{xy}(x, \sigma_D) \\ I_{xy}(x, \sigma_D) & I_{yy}(x, \sigma_D) \end{pmatrix} \quad (7)$$

式中二阶导数由图像与差分尺度为 σ_D 的高斯核进行卷积获得。基于 Hessian 矩阵行列式和迹的度量都有很好的性质, Hessian 矩阵的迹又称为 Laplacian 算子。

Lindeberg^[28]提出了一种特征尺度选择方法,其

思想是当给定函数在尺度上变化时,取得极值时对应的尺度为局部图像结构的特征尺度。在特征尺度上特征检测算子与局部图像结构达到最大的相似性,通过在特征尺度上进行特征的检测实现特征检测算子的尺度不变特性。Mokolajczyk 等人基于自动特征尺度选择的思想提出了 Harris-Laplace 算子^[29]。Harris-Laplace 算子首先在空间域上利用 Harris 算子的检测角点特征,然后以 Laplacian 算子为尺度度量在尺度空间上为 Harris 角点特征选择特征尺度,从而实现特征检测算子对平移、旋转和尺度变换的不变性。图 2 表示 Harris-Laplace 算子的特征尺度选择,其中 (a) (b) 为两幅不同焦距下获得的图像(存在尺度变换), (c) (d) 为两幅图像对应的 Laplacian 算子尺度响应曲线,特征尺度分别为 10.1 和 3.9。两幅图像之间的尺度变换因子为 2.5。图中所画圆形区域的半径为特征尺度的 3 倍。Harris-Laplace 角点特征检测算子对相似变换保持不变。

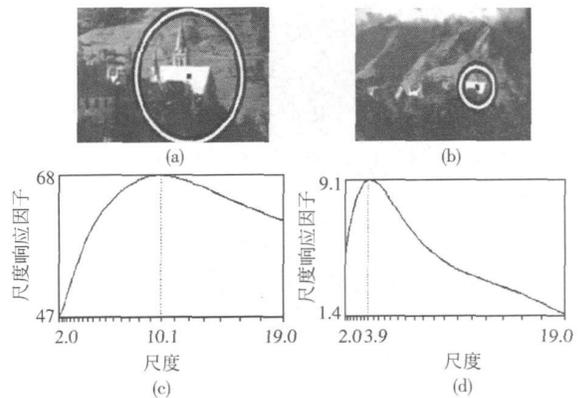


图 2 特征尺度

Fig 2 Characteristic scale

2.2 blob 特征

2.2.1 Hessian 算子

基于 Hessian 矩阵行列式和迹的度量都具有很好的性质,用于局部特征检测时两者都检测出图像中的 blob 结构^[30-31]。Laplacian 是可分离的线性滤波器,用于 blob 特征检测时存在一个缺点,即在信号变化主要为一个方向的轮廓或笔直边缘附近常常出现局部极值。由于这些局部极值处的定位对噪声和邻域的纹理变化比较敏感,所以是不稳定的。当 Laplacian 算子检测的 blob 结构用于寻找图像特征对,求解图像变换参数时会带来较大的误差。基于

Hessian 矩阵行列式的特征检测算法只能检测出和滤波器尺度对应的固定大小的 blob 特征。基于 Hessian 矩阵行列式和迹的 blob 检测算子对欧氏变换具有不变性。

2.2.2 Hessian-Laplace 算子

Hessian-Laplace 算子^[29]的思想与 Harris-Laplace 算子的思想相似, 即首先在空间上检测 blob 结构, 然后通过 Laplacian 算子选择特征尺度以实现尺度变换的不变性。Hessian-Laplace 算子对相似变换保持不变性。

2.2.3 高斯差分算子

高斯差分算子 DoG (difference-of-Gaussian)^[32]通过近似 Laplacian 在图像中检测 blob 特征。Laplacian 在尺度空间理论中也称为扩散方程, 它是图像在尺度方向上的导数。尺度方向上相邻点差分是对尺度导数的简单近似, 相邻的不同尺度图像之间的差分是对尺度空间导数的近似。当采用高斯卷积来表示不同尺度上的图像时, 高斯差分图像通过近似 Laplacian-of-Gaussian 实现尺度空间导数, 从而避免了在 x 方向和 y 方向上的二阶导数的计算, 减小了计算量。高斯差分算子的处理流程如图 3 首先利用高斯卷积模板对图像进行平滑, 平滑后的相邻图像进行组合计算高斯差分图像; 然后在差分图像中寻找空间和尺度上的局部极值, 利用非极大值抑制和二次方程迭代对检测的特征位置进行筛选和精确定位; 最后由于 Laplacian 对边缘有强响应, 利用 Hessian 矩阵特征值的相对强弱滤除边缘点。高斯差分算子计算速度快, 对相似变换具有不变性。

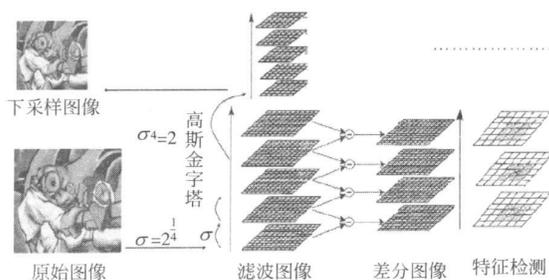


图 3 高斯差分算子

Fig 3 Difference of Gaussian detector

2.2.4 SURF 算子

Viola 和 Jones^[33]在实时人脸检测领域提出了积

分图像的概念, 积分图像可以用来快速地计算 Haar 小波或 box 卷积滤波器, SURF 算子^[34]利用积分图像快速计算近似的 Hessian 矩阵。SURF 算子与 Hessian-Laplace 算子一样基于 Hessian 矩阵, 但 Hessian-Laplace 算子分别采用矩阵行列式和迹检测空间及尺度上的局部极值点, SURF 算子利用 Hessian 矩阵的行列式同时检测空间和尺度上的极值点。SURF 算子通过在积分图像基础上引入 box 滤波器对高斯核进行近似, 从而实现 Hessian 矩阵行列式的快速计算。SURF 算子对相似变换具有不变性。

2.3 区域特征

2.3.1 Harris/Hessian-Affine 区域特征

Mikołajczyk^[29]提出了 Harris/Hessian 仿射不变区域特征, 其算法的具体流程为

- 1) 利用 Harris 角点响应函数或 Hessian 矩阵的行列式进行空间域上感兴趣点的提取;
- 2) 利用 Laplacian 算子寻找感兴趣点在尺度空间上的特征尺度;
- 3) 通过二阶矩阵的特征值和特征向量为感兴趣点估计仿射区域;
- 4) 归一化仿射区域为圆形区域;
- 5) 提取归一化后的感兴趣点的空间位置和特征尺度;
- 6) 如果归一化后感兴趣点的二阶矩阵的特征值不相等, 则返回步骤 3);

2.3.2 基于边缘的区域特征

基于边缘的区域特征 EBR^[2]利用 Harris 角点局部邻域内的边缘几何信息来构建仿射不变特性, 其理由是 1) 边缘在仿射变换下稳定, 对视角、尺度和光照的变化有很好的适应性; 2) 利用边缘几何可以减小处理问题的维数, 6D 的仿射问题变成 1D 的边缘几何问题。图 4 表示基于边缘的区域特征提取, 其算法具体流程如下:

- 1) 检测 Harris 角点 $p(x, y)$, 通过 Canny 边缘算子提取相邻的边缘 $p \rightarrow p_b, p \rightarrow p_s$;
- 2) $p_1(x, y), p_2(x, y)$ 偏移 $p(x, y)$ 的速度与相对仿射参数 l_1, l_2 的关系为

$$l_i = \int \text{abs}(|p_i^{(1)}(s_i) \cdot p - p_i(s_i)|) ds_i \quad (8)$$

$i = 1, 2$

式中 s_i 为任意的曲线参数, 对于每一个 $l(l_1 = l_2)$, $p(x, y), p_1(x, y), p_2(x, y)$ 定义一个区域 $\Omega(l)$, $\Omega(l)$ 是由 $p \rightarrow p_1, p \rightarrow p_2$ 扩展的平行四边形;

3) 当区域 $\Omega(l)$ 覆盖的局部图像的光度量达到极值时, 停止 $p_1(x, y), p_2(x, y)$ 的偏移;

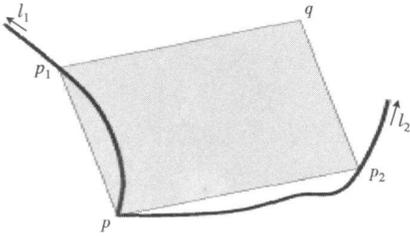


图 4 基于边缘的区域特征
Fig 4 EBR feature

2.3.3 基于灰度的区域特征

Tuytelaar 等人^[2]提出了灰度区域特征 IBR, 其思想是基于多尺度灰度极值检测进行区域提取, 算法流程为首先利用非极大值抑制进行局部灰度极值的检测, 以检测到的极值点为中心定义一个放射性灰度函数

$$f_i(t) = |I(t) - I_0| (\max(t^{-1} \int_0^t |I(t) - I_0| dt, d))^{-1} \quad (9)$$

式中 t 为沿射线的距离参数, $I(t)$ 为 t 处的灰度, I_0 为局部灰度极值, d 为非零常数。

如图 5 所示, 通过连接放射线上所有放射性灰度函数的最值点形成一个局部仿射不变区域。

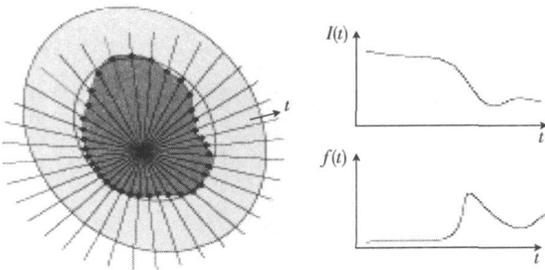


图 5 仿射不变区域
Fig 5 A affine-invariant region

通常由最值点连接成的局部仿射不变区域是不规则的需要利用矩特征进行椭圆拟和。图 6 为基于灰度的区域特征提取算法的流程图。

2.3.4 基于显著性的区域特征 (Salient Regions)

Kadir 和 Brady^[35] 基于椭圆区域 Ω 的灰度概率

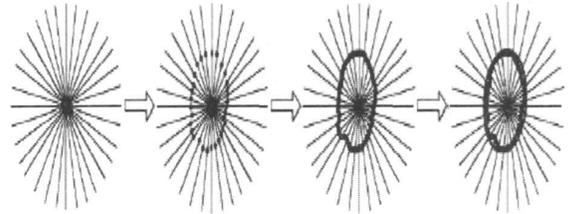


图 6 基于灰度的区域特征
Fig 6 BR feature

密度函数提取显著性区域, 算法的流程为

- 1) 在每个像素 $p(x, y)$ 处, 计算以 $p(x, y)$ 为中心, 尺度为 s 方向为 θ 主轴比为 λ 的椭圆区域的灰度概率密度函数 $p(I)$ 的熵;
- 2) 寻找尺度空间上的熵极值, 记录 (s, θ, λ) 为候选的显著性区域, 熵定义为

$$\Theta = - \sum_I p(I) \log p(I) \quad (10)$$

- 3) 每个极值处概率密度函数 $p(I, s, \theta, \lambda)$ 对尺度求偏导

$$w = \frac{s^2}{2s-1} \sum_I \left| \frac{\partial p(I, s, \theta, \lambda)}{\partial s} \right| \quad (11)$$

- 4) 计算椭圆区域的显著性 $y = Hw$ 并根据显著性排序, 保留前 P 个区域为显著性区域;

2.3.5 MSER 区域特征

MSER (maximally stable extremal region) 区域特征由 Matas 等人^[36] 提出, 其实现思想类似于分水岭图像分割算法。通过不断的改变门限对图像进行二值化分割, 算法提取那些在一系列门限下面积稳定的区域作为最稳定极值区域特征。

MSER 特征提取的结果是任意形状的区域特征, 区域由包含它的边界像素点来定义, 区域内的像素灰度值一致的低于或高于其区域外的灰度值。MSER 的一个主要优点就是它对连续或非线性的空间变换都有很好的鲁棒性。

以上分别介绍了角点不变特征、blob 不变特征、区域不变特征的检测方法及其主要性质, 由上文对局部不变特征检测方法的讨论可以得出以下两点结论:

- 1) 角点特征与 blob 特征具有良好的互补特性, 实际应用中应结合使用;
- 2) 特征不变性自由度的增加通常是以算法的复杂度增加、计算量增加、特征重复性下降等为代价的。

表 2 总结了以上不变特征检测方法及其主要性质。

表 2 不变特征检测算子及其性质

Tab 2 Invariant Feature detectors properties

特征算子	类别	自由度	重复性	计算时间	定位精度	鲁棒性
Moravec	角点	2	中	短	中	中
Harris	角点	3	高	短	高	高
SUSAN	角点	3	中	短	中	中
FAST	角点	3	中	短	中	高
Harris-Laplace	角点	4	高	中	高	中
Hessian	blob	3	中	中	中	中
Hessian-Laplace	blob	4	中	中	中	高
DoG	blob	4	中	短	中	中
SURF	blob	4	中	短	低	中
Harris-Affine	区域	6	高	中	高	中
Hessian-Affine	区域	6	中	中	中	高
EBR	区域	6	高	长	高	低
IBR	区域	6	中	中	中	中
Salient Regions	区域	6	低	长	低	中
M SER	区域	6	高	短	高	中

3 局部不变特征描述

特征描述是特征匹配的前提, 特征描述的目的是量化特征的属性以表述特征信息。文献中存在大量的局部不变特征描述方法, 大致可以分为基于分布的特征描述方法、基于滤波的特征描述方法、基于矩的特征描述方法和其他描述方法。

3.1 基于分布的描述方法

Lowe 提出的 SIFT (scale invariant feature transform) 特征^[31], SIFT 特征描述符在文献 [37] 的特征描述符性能比较实验中多数情况下优于其他特征描述符。图 7 为 SIFT 特征描述符的示意图, 其中, size 为空间采样的大小以像素为单位, SIFT 描述符是梯度位置和方向的 3D 直方图。在 SIFT 特征点 (DoG 算子提取) 的特征尺度上以梯度主方向为参照的 SIFT 描述符, 对相似变换具有不变性。

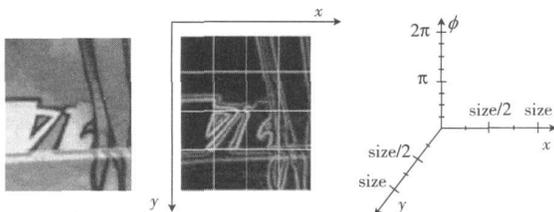


图 7 SIFT 描述符

Fig 7 SIFT descriptors

Ke 和 Sukthankar^[38] 在 SIFT 描述符基础上提出了 PCA-SIFT 描述符, PCA-SIFT 描述符比 SIFT 描述符简单, 计算量小, 但在特征向量的降维过程中特征描述符的区分性下降。在特征点的特征尺度上以梯度主方向为参照提取的 PCA-SIFT 描述符对相似变换具有不变性。

GLOH (gradient location orientation histogram) 描述符^[37] 是 SIFT 描述符的扩展, 其区分性和鲁棒性都优于 SIFT 描述符。GLOH 描述符对相似变换具有不变性。

Belongie 等人^[6] 提出了 Shape context 描述符, 其思想与 SIFT 描述符类似。Shape context 描述符可以视为边缘点位置和方向的 3D 直方图统计, 可以有效的对轮廓信息进行描述, 它在形状匹配和目标识别领域得到广泛应用。

Johnson 和 Hebert^[39] 面向 3 维目标识别中提出了 spin image 描述符。灰度域的 spin image 描述了区域内像素灰度值 i 和像素到区域中心距离 d 的分布信息, 是 2D 的灰度直方图。图 8 为 spin image 描述符的计算, 其中左图为图像块, 右图为对应的 spin image 描述符。

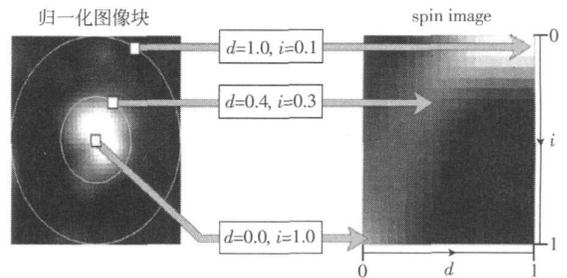


图 8 Spin image 描述符

Fig 8 Spin image descriptors

3.2 基于滤波的描述方法

Freeman 和 Adelson^[40] 提出了不同类别的 steerable 滤波器, 通过组合 steerable 滤波器获得旋转不变的区域特征描述。Baumberg^[41] 基于复数滤波器进行宽基线下的特征匹配。Schaffalitzky 和 Zisserman^[3] 利用复数滤波器进行局部特征的描述。图 9 表示 16 种复数滤波器的响应, 滤波器的响应与高斯导数相似, 具有旋转不变性。

3.3 基于矩的描述方法

矩不变量描述方法的研究最早可以追溯到 Hu

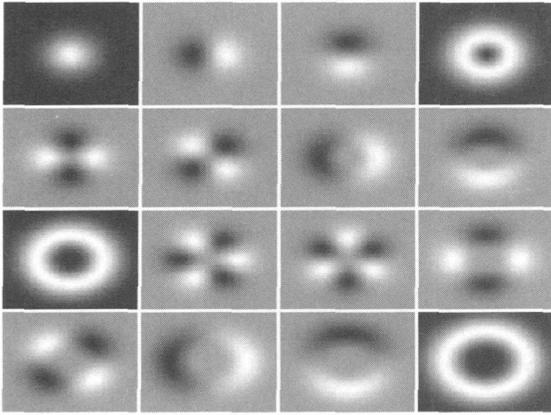


图 9 复数滤波器描述符

Fig 9 Complex filter descriptors

矩^[42]。矩可以定义在任意的 2D、3D 和高维空间,直接作用于二值图像、灰度图像、彩色图像或图像局部特征上。2D 矩的一般定义式可写为

$$M_{pq} = \iint_{\Omega} \varphi_{pq}(x, y) g(f(x, y)) dx dy \quad (12)$$

$$p, q = 0, 1, 2, \dots$$

式中 $f(x, y)$ 为图像函数, $g(f(x, y))$ 为特征函数, $\varphi_{pq}(x, y)$ 为核函数, Ω 为定义区域。

矩不变量可以分为非正交矩和正交矩两类^[43]。常用的非正交矩有几何矩、旋转矩和复合矩等。正交矩又可以分为连续正交矩和离散正交矩,其中连续正交矩有 Ledengre 矩 (LM)、Zem ke 矩 (ZM) 和 Pseudo-Zemike 矩 (PZM) 等,离散正交矩有 Tchebiche 矩 (TM)、Kraw tchouk 矩 (KM)、Racah 矩 (RAM) 和 Dual Hahn 矩 (DHM) 等。表 3 列举了各种矩不变量描述符及其主要性质^[43-44]。

表 3 矩不变量及其性质

Tab 3 Moment invariants properties

名称	平移	旋转	尺度	卷积	鲁棒性
几何矩	不变	变化	不变	不变	低
复合矩	不变	不变	不变	不变	低
旋转矩	变化	不变	不变	不变	低
LM 矩	不变	变化	不变	变化	低
ZM 矩	不变	不变	不变	变化	中
PZM 矩	变化	不变	不变	变化	中
TM 矩	不变	变化	不变	变化	中
KM 矩	不变	变化	变化	变化	高
RAM 矩	不变	变化	变化	变化	高
DHM 矩	不变	变化	变化	变化	高

4 局部不变特征匹配

特征匹配是图像配准、目标识别等多种应用的前提,特征匹配的目的在于量化特征向量之间的差异以识别相似或区分不同的特征。关于局部不变特征匹配的研究主要集中在相似性度量,匹配策略和匹配验证 3 个方面。

针对不同的特征描述方法,应选择适合的相似性度量。相似性度量不但要能有效的区分不同特征向量之间的差别而且要对相同特征向量之间的类内变化具有鲁棒性。相似性度量^[45-46]的选择反应了度量特征对之间相似性的不变特性。常用的相似性度量有相关系数,欧氏距离,直方图相交,EMD (earth moving distance) 距离^[47-48],二次式距离和马氏距离等。直方图相交、二次式距离和 EMD 距离常用来比较直方图特征向量之间的差异性,欧氏距离适用于特征向量的各分量之间是正交无关的且各维度的重要程度相同的情况,马氏距离适用于特征向量的各个分量间具有相关性或者具有不同权重的情况。

高效的匹配策略可以大大节省高维特征向量之间匹配时间。高维特征空间中最近邻或 k 最近邻一般被认为是待匹配特征向量的候选匹配。高维空间中如何有效地搜寻最近邻或 k 最近邻目前仍是一个开放的课题^[49],目前文献中常用的搜索方法有基于 kd-trees 的搜索方法^[49],基于 balltrees 的搜索方法^[49],基于 LSH (locality sensitive hashing)^[50] 的搜索方法和 ANN (approximate nearest neighbour)^[51-52] 快速搜索算法等。高维特征空间中相似性度量计算量大,可以将高维特征向量投影变换到线性空间^[53] 或者进行降维处理^[54] 以提高匹配效率。

由于图像间存在各种几何及光度变换、噪声、量化误差及图像中可能存在相似的局部结构等多种因素的影响,基于相似性度量的特征匹配结果中可能存在错误的匹配,需要引入其他的约束对匹配的结果进行验证。文献中常用的约束有局部平面结构约束^[27]、极线几何约束^[36]、全局约束和几何光度约束^[2]等。

在图像、视频检索及目标识别等领域应用较多的局部不变特征的匹配方法有基于门限的匹配方

法、基于最近邻的匹配方法和基于最近邻距离比率的匹配方法。3种方法各有优缺点,其中基于门限的匹配方法简单计算量小,基于最近邻距离比率的匹配方法准确性高,实际应用中应根据具体的应用背景进行匹配策略的选择。

5 结 论

局部不变特征是描述图像内容的重要工具,在宽基线匹配、目标识别、图像及视频检索等多个领域内得到了广泛的应用。基于局部不变特征检测、局部不变特征描述和局部不变特征匹配 3个基本问题对文献中的局部不变特征方法进行了回顾和分析,并比较了各类算法的优缺点。尽管目前局部不变特征的研究取得了很大的进步,但仍存在很多挑战。

1) 针对不同的应用场景,文献中定义了多种局部不变特征,其中许多方法都是从工程上引出的,算法关键参数的选取很多都是经验值,缺乏理论证明,扩展性差,如何在理论上对局部不变特征进行有效的描述及扩展面临很大挑战。

2) 现有的局部不变特征研究方法缺乏语义层次的解释,提取的局部不变特征既可能位于感兴趣目标上也可能位于复杂背景结构上或目标与背景的边界处。如何对局部不变特征进行语义层次的解译面临很大挑战。

3) 文献中存在大量的局部不变特征检测和描述方法,其各有优缺点。针对具体的应用,如何根据图像的统计特性自动选择具有互补特性的特征检测及描述方法还面临着很大挑战。

4) 局部不变特征的评价体系仍需完善。文献中大量使用的评价准则为重复性和匹配虚警率,对应为局部不变特征重复性和区分性的量化评价,如何对不变特征的局部性、精确性、鲁棒性和不同层次不变特征之间的互补性等性质进行有效的量化评价还面临很大的挑战。

参考文献 (References)

- [1] Moravec H. Towards automatic visual obstacle avoidance [C] // Proceedings of International Joint Conference on Artificial Intelligence. New York, USA: ACM Press, 1977: 584.
- [2] Tuytelaars T, Gool L V. Matching widely separated views based on affine invariant regions [J]. International Journal of Computer Vision, 2004, 59(1): 61-85.
- [3] Schaffalitzky F, Zisserman A. Multi-view matching for unordered image sets [C] // Proceedings of the 7th European Conference on Computer Vision. Cambridge, MA, USA: MIT Press, 2002: 414-431.
- [4] Pritchett P, Zisserman A. Wide baseline stereo matching [C] // Proceedings of the 6th International Conference on Computer Vision. New York, USA: ACM Press, 1998: 754-760.
- [5] Lowe D G. Object recognition from local scale-invariant features [C] // Proceedings of the 7th International Conference on Computer Vision. New York, USA: ACM Press, 1999: 1150-1157.
- [6] Belongie S, Malik J, Puzicha J. Shape matching and object recognition using shape contexts [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(4): 509-522.
- [7] Johnson A, Hebert M. Using spin images for efficient object recognition in cluttered 3D scenes [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(5): 433-449.
- [8] Obdrzalek S, Matas J. Object recognition using local affine frames on distinguished regions [C] // Proceedings of British Machine Vision Conference. Oxford, England: BMVA Press, 2002: 113-122.
- [9] Mikolajczyk K, Leibe B, Schiele B. Local features for object class recognition [C] // Proceedings of International Conference on Computer Vision. New York, USA: ACM Press, 2005: 1792-1799.
- [10] Dorko G, Schmid C. Selection of scale invariant neighborhoods for object class recognition [C] // Proceedings of International Conference on Computer Vision. New York, USA: ACM Press, 2003: 634-640.
- [11] Schmid C, Mohr R. Local grayvalue invariants for image retrieval [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1997, 19(5): 530-534.
- [12] Sivic J, Zisserman A. Video google: A text retrieval approach to object matching in videos [C] // Proceedings of International Conference on Computer Vision. New York, USA: ACM Press, 2003: 1470-1478.
- [13] Sivic J, Schaffalitzky F, Zisserman A. Object level grouping for video shots [J]. International Journal of Computer Vision, 2006, 67(2): 189-210.
- [14] Se S, Lowe D G, Little J. Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks [J]. International Journal of Robotics Research, 2002, 21(8): 735-758.
- [15] Oliva A, Torralba A. Modeling the shape of the scene: a holistic representation of the spatial envelope [J]. International Journal

- of Computer Vision 2001, 42(3): 145-175.
- [16] Vama M, Zissem A. A statistical approach to texture classification from single images [J]. International Journal of Computer Vision, 2005, 62(1/2): 61-81
- [17] Randen T, Husoy J H. Filtering for texture classification: a comparative study [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1999, 21(4): 291-310.
- [18] Sivic J, Zissem A. Video data mining using configurations of viewpoint invariant regions [C] // Proceedings of International Conference on Computer Vision and Pattern Recognition. Cambridge, MA, USA: MIT Press 2004 488-495
- [19] Hartley R I, Zissem A. Multiple View Geometry in Computer Vision [M]. New York: Cambridge University Press 2004 37-44
- [20] Li J, Allinson N M. A comprehensive review of current local features for computer vision [J]. Neurocomputing 2008, 71(1): 1771-1787
- [21] Lindeberg T. Scale-space [J]. Encyclopedia of Computer Science and Engineering 2009 IV (1): 2495-2504.
- [22] Harris C, Stephens M. A combined corner and edge detector [C] // Proceedings of Avey Vision Conference New York, USA: ACM Press 1988: 147-151.
- [23] Schmid C, Mohr R, Bauckhage C. Evaluation of interest point detectors [J]. International Journal of Computer Vision 2000, 37(2): 151-172
- [24] Smith S M, Brady J M. SUSAN—a new approach to low level image processing [J]. International Journal of Computer Vision 1997, 23(1): 45-78.
- [25] Rosten E, Drummond T. Machine learning for high-speed corner detection [C] // Proceedings of European Conference on Computer Vision New York, USA: ACM Press 2006: 430-443
- [26] Quinlan J R. Induction of decision trees [J]. Machine Learning 1986 1(1): 81-106
- [27] Mikołajczyk K, Tuytelaars T, Schmid C, et al A comparison of affine region detectors [J]. International Journal of Computer Vision 2005, 65(1/2): 43-72.
- [28] Lindeberg T. Feature detection with automatic scale selection [J]. International Journal of Computer Vision 1998 30(2): 79-116.
- [29] Mikołajczyk K, Schmid C. Scale & affine invariant interest point detectors [J]. International Journal of Computer Vision 2004, 60(1): 63-86
- [30] Mikołajczyk K. Interest Point Detection Invariant to Affine Transformations [D]. Nice France Institut National Polytechnique de Grenoble 2002.
- [31] Tuytelaars T, Mikołajczyk K. Local invariant feature detectors: a survey [J]. Foundations and Trends in Computer Graphics and Vision 2007, 3(3): 177-280
- [32] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision 2004 60(2): 91-110.
- [33] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features [C] // Proceedings of International Conference on Computer Vision and Pattern Recognition Cambridge, MA, USA: MIT Press 2001 511-518.
- [34] Bay H, Tuytelaars T, Gool L V. SURF: Speeded up robust features [C] // Proceedings of European Conference on Computer Vision New York, USA: ACM Press 2006 404-417.
- [35] Kadir T, Zissem A, Brady M. An affine invariant salient region detector [C] // Proceedings of the 8th European Conference on Computer Vision New York, USA: ACM Press 2004: 345-457.
- [36] Matas J, Chum O, Urban M, et al Robust wide baseline stereo from maximally stable extremal regions [C] // Proceedings of the 13th British Machine Vision Conference Oxford, England BMVA Press 2002 384-393
- [37] Mikołajczyk K, Schmid C. A performance evaluation of local descriptors [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence 2005 27(10): 1615-1630
- [38] Ke Y, Sukthankar PCA-SIFT: A more distinctive representation for local image descriptors [C] // Proceedings of International Conference on Computer Vision and Pattern Recognition New York, USA: ACM Press 2004: 506-513
- [39] Johnson A E, Hebert M. Using spin-images for efficient multiple model recognition in cluttered 3D scenes [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence 1999 21(5): 433-449
- [40] Freeman W, Adelson E. The design and use of steerable filters [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence 1991 13(9): 891-906
- [41] Baumberg A. Reliable feature matching across widely separated views [C] // Proceedings of International Conference on Computer Vision and Pattern Recognition New York, USA: ACM Press 2000: 774-781.
- [42] Hu M K. Visual pattern recognition by moment invariants [J]. IRE Transactions on Information Theory, 1962 8(1): 179-187
- [43] Shu H, Luo L, Coatrieux J L. Moment-based approaches in image [J]. IEEE Engineering in Medicine and Biology Magazine, 2007 26(5): 70-74.
- [44] Belkasin S O, Shridhar M, Ahmad I M. Pattern recognition with moment invariants: a comparative study and new results [J]. Pattern Recognition 1991, 24(1): 1117-1138.
- [45] Ramanan D, Baker S. Local distance functions: a taxonomy, new algorithm, and an evaluation [C] // Proceedings of the IEEE International Conference on Computer Vision New York, USA: ACM Press 2009 49-54.
- [46] Babenko B, Branson S, Belongie S. Similarity metrics for

- categorization from monolithic to category specific [C] // Proceedings of the IEEE International Conference on Computer Vision New York, USA: ACM Press, 2009: 1-7
- [47] Ling H, Okada K. An efficient earth mover's distance algorithm for robust histogram comparison [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(5): 840-853
- [48] Pele O, Werman M. Fast and robust earth mover's distances [C] // Proceedings of the IEEE International Conference on Computer Vision New York, USA: ACM Press, 2009: 121-126
- [49] Darrell T, Indyk P, Shakhnarovich G. Nearest Neighbor Methods in Learning and Vision: Theory and Practice [M]. Cambridge: MIT Press, 2006: 12-44
- [50] Andoni A, Indyk P. Near-optimal hashing algorithms for approximate nearest neighbor in high dimensions [J]. Communications of the ACM, 2008, 51(1): 117-122
- [51] Arya S, Mount D. ANN: Library for approximate nearest neighbor searching [EB/OL] [2009-11-05]. <http://www.cs.umd.edu/~mount/ANN/>
- [52] Muja M, Lowe D G. Fast approximate nearest neighbors with automatic algorithm configuration [C] // Proceedings of the Fourth International Conference on Computer Vision Theory and Applications New York, USA: ACM Press, 2009: 331-340
- [53] Mikolajczyk K, Matas J. Improving descriptors for fast tree matching by optimal linear projection [C] // Proceedings of the IEEE International Conference on Computer Vision New York, USA: ACM Press, 2007: 1-7
- [54] Cabender M, Lepetit V, Fua P, et al. Compact signatures for high-speed interest point detection and matching [C] // Proceedings of the IEEE International Conference on Computer Vision New York, USA: ACM Press, 2009: 357-364