

基于一种快速鲁棒特征描述子的图像匹配

吴 铮, 孙 立, 夏一民, 汪亚明

WU Zheng, SUN Li, XIA Yi-min, WANG Ya-ming

浙江理工大学 信息与电子学院, 杭州 310018

College of Informatics and Electronics, Zhejiang Sci-Tech University, Hangzhou 310018, China

E-mail: sunrainhard@gmail.com

WU Zheng, SUN Li, XIA Yi-min, et al. Image correspondence based speed feature descriptor. Computer Engineering and Applications, 2008, 44(7): 86-88.

Abstract: This paper proposes a method about image correspondence based on a speed robust feature descriptor. This relies on integral images and Fast-Hessian detector to extract interest points. For each feature point, the dominant orientation is assigned by computing Haar-wavelet responses. Based on these orientations constructing a window around the feature point, the feature points are decrypted by the sum of Haar-wavelet responses on the domination orientation and the verticality. Image correspondences are made based on distance between these descriptors.

Key words: image correspondence; integral image; Fast-Hessian detector; speed up robust feature descriptor

摘 要: 基于一种快速鲁棒特征描述子进行图像特征点之间的匹配。此方法基于积分图像, 使用快速 Hessian 检测子来提取图像特征点。对每个特征点, 通过计算 Haar 小波, 来寻找它的主方向。然后构造窗口区域, 以主方向和与主方向垂直的方向上的小波和来描述特征点。最后以特征描述向量的距离为标准寻找不同图像之间的特征匹配。

关键词: 图像匹配; 积分图像; 快速 Hessian 检测子; 快速鲁棒特征描述子

文章编号: 1002-8331(2008)07-0086-03 **文献标识码:** A **中图分类号:** TP391

1 引言

图像匹配是计算机视觉的一个重要问题, 它是许多应用的提前, 诸如摄像机标定、三维重建、运动分析等。一般比较常用的是基于特征点的匹配。它可以分为三个部分: 首先从图像中检测出特征点; 然后根据特征点的邻域点来描述该特征点; 最后在不同的图像之间寻找这些特征点的对应关系。常用的特征点检测方法有 Harris^[1]角点检测与强角点^[2,3]检测, 但是它们不是尺度不变的, 也就是说, 对于不同尺度的图像, 往往无法建立相应的特征点对应关系。因此, 国内外提出了不少尺度不变的特征点检测子与描述子。Lindeberg^[4]介绍了自动尺度选择的特征检测子, 它使用 Hessian 矩阵和 Laplacian 算子提取特征。Mikolajczyk 和 Schmid^[5]改进了 Lindeberg 的方法, 提出一种鲁棒的、尺度不变的特征检测子, 使用 Harris 检测特征点的位置和 Laplacian 检测子选择尺度因子。Lowe^[6]从速度方面考虑使用 DoG (Difference of Gaussians) 来估计 LoG (Laplacian of Gaussian) 检测子。Lowe 提出一种 SIFT (Scale-Invariant Feature Transform) 描述子, 它通过计算特征点邻域的梯度直方图来描述特征点。接着, Ke 和 Sukthankar^[7]将 PCA 主量分析应用到 SIFT 中提出了 PCA-SIFT 描述子。目前实际运用中, SIFT 是一

种很好的特征描述子, 但是 SIFT 速度不是很快。

本文将基于一种快速鲁棒特征描述子^[8]进行图像的特征点匹配。它也是一种尺度不变的特征点检测子和描述子, 并且具有速度快, 鲁棒性好的特点。第 2 章介绍一种快速 Hessian 特征检测子; 第 3 章介绍 SURF (Speed Up Robust Features) 这种快速鲁棒特征描述子; 第 4 章是实验结果; 最后给出结论。

2 快速 Hessian 特征检测

首先给出积分图像 (Integral image) 的定义^[9], 它可以加快卷积等运算: 假设图像用 $I(x)$ 表示, 其中 $x=(x, y)$ 表示某像素点, 则积分图像 $I_{\Sigma}(x)$ 表示以图像原点和点 x 为顶点的矩阵区域内所有像素之和, 即:

$$I_{\Sigma}(x) = \sum_{i=0}^{i \leq x} \sum_{j=0}^{j \leq y} I(i, j) \quad (1)$$

Hessian 矩阵具有良好的计算时间和精度表现, 给出图像上 I 的某点 $x=(x, y)$, 尺度 σ 下的 Hessian 矩阵 $H(x, \sigma)$ 为:

$$H(x, \sigma) = \begin{bmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{bmatrix} \quad (2)$$

基金项目: 国家自然科学基金 (the National Natural Science Foundation of China under Grant No.60473038)。

作者简介: 吴铮 (1981-), 女, 硕士, 主要研究领域为计算机视觉与图像处理; 孙立 (1981-), 男, 硕士, 主要研究领域为计算机视觉与图像处理; 夏一民 (1979-), 男, 硕士, 主要研究领域为计算机视觉与图像处理; 汪亚明 (1972-), 男, 博士, 教授, 硕士生导师, 常务副院长, 主要研究领域为计算机视觉与模式识别、医学成像与医学图像处理。

收稿日期: 2007-06-21

修回日期: 2007-09-10

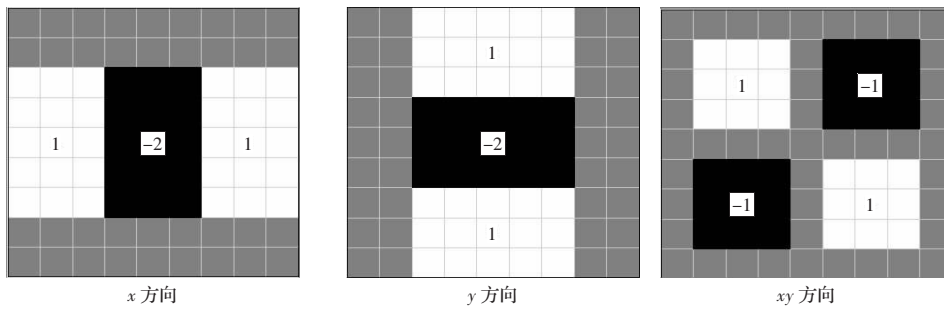


图1 方框滤波

其中, $L_{xx}(x, \sigma)$ 是图像点 $I(x)$ 与二阶高斯滤波 $\frac{\partial^2}{\partial x^2} g(\sigma)$ 的卷积; $L_{yy}(x, \sigma)$ 和 $L_{xy}(x, \sigma)$ 类似。

实际计算中, 高斯滤波需要离散化, 而且高斯滤波可能产生一些走样。本文使用一种方框滤波(box filters)来近似二阶高斯滤波, 它在积分图像的基础上计算迅速, 并与图像大小无关。

如图 1 所示的 9×9 的方框滤波是尺度因子 $\sigma=1.2$ 的二阶高斯滤波的估计, 设该估计为 D_{xx}, D_{xy}, D_{yy} 。为了平衡与 Hessian 矩阵的行列式值, 考虑有:

$$\frac{|L_{xy}(1.2)|_F |D_{xx}(9)|_F}{|L_{xx}(1.2)|_F |D_{xy}(9)|_F} = 0.912 \approx 0.9 \quad (3)$$

其中, $|A|_F = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2 \right)^{\frac{1}{2}}$ 是矩阵的 Frobenius 范数。因此:

$$\det(H_{approx}) = D_{xx} D_{yy} - (0.9 D_{xy})^2 \quad (4)$$

最后, 滤波结果按掩膜大小进行归一化。

尺度空间通常可以用金字塔图像来实现。因为之前使用积分图像和方框滤波, 所以并不需要通过迭代同样的滤波来重复计算, 取而代之通过放大滤波的大小来计算。上述 9×9 的滤波为初始层, 它对应尺度因子为 $\sigma=1.2$ (即二次高斯滤波的 $\sigma=1.2$)。接下去的层依次放大方框滤波的大小来计算, 滤波大小依次为 $9 \times 9, 15 \times 15, 21 \times 21, 27 \times 27$ 等等。当尺度比较大时, 滤波大小的步长也需要相应的增加。以每四个滤波为一组, 每个组之间的步长依次按 6、12、24 增长, 即接下去的滤波大小为 $39 \times 39, 51 \times 51, 63 \times 63, 75 \times 75$ 。

对于某一点, 它处于某一个尺度图像下, 其上下各对应上下尺度图像上的点, 加上该点的邻域, 因此可以构造一个 $3 \times 3 \times 3$ 的立体邻域。为了确定图像的特征点, 按照文献[10], 对于某一尺度图像下的点 $x=(x, y, \sigma)$, 通过尺度空间图像的插值得到 Hessian 矩阵行列式的局部最大值。

$$H(x) = H + \frac{\partial H^T}{\partial x} + \frac{1}{2} x^T \frac{\partial^2 H}{\partial x^2} x \quad (5)$$

特征点的位置由式(5)的极值决定, 即:

$$\hat{x} = - \frac{\partial^2 H^{-1}}{\partial x^2} \frac{\partial H}{\partial x} \quad (6)$$

3 SURF 特征描述子

SURF 特征描述子可以分为两步: 第一步根据特征点周围的一个圆形区域, 找到一个方向; 第二步在这个选定方向上构造一个矩形区域, 并提取出所需要的描述信息。

对以每一个特征点为圆心, 半径为 $6s$ 的邻域中的点分别计算 x 和 y 方向上的 Haar 小波响应, 如图 2 所示。

s 表示该特征点所在的尺度因子。Haar 小波的尺寸则选为 $4s$ 。这样得到一系列的向量, 以 60° 为一个窗口, 绕着原点旋转, 计算 Haar 小波响应的向量和, 选择最长的向量为该特征点的主方向, 如图 3 所示。

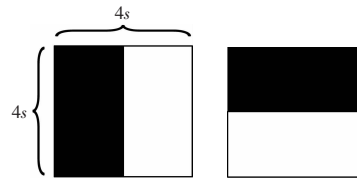


图2 Haar 小波

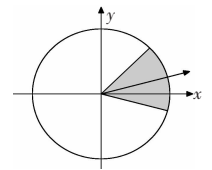


图3 主方向的选择

选定特征点主方向之后, 按主要方向构造一个大小为 $20s$ 的窗口, 并将该窗口区域分成 4×4 的子区域。对每一个子区域, 按照 5×5 的大小进行取样, 分别计算相对于主方向的水平和垂直方向上的 Haar 小波响应, 记为 d_x 和 d_y 。为了增加鲁棒性和局部误差, 结果上乘以一个高斯滤波, 然后计算它们的和与绝对和, 可以得到一个四维向量 $v = (\sum d_x, \sum d_y, \sum |d_x|, \sum |d_y|)$ 。因此 4×4 个子区域可以得到 64 个值, 归一化后的这 64 个值就构成特征点的 SURF 描述子。

基于快速 Hessian 检测子和 SURF 描述子, 每个特征点可以用一个 64 维的向量表示。对于一幅图像中的某一个特征点, 在另一幅图像中的众多特征点的 SURF 描述子向量中寻找与其距离(欧式距离或马式距离)最近的某点作为其对应点, 在实际计算中, 如果最近点与次近点相距不远, 则不把该最近点作为相应的对应点。

4 实验

本文采用快速 Hessian 检测子和 SURF 特征描述子进行图像匹配: 首先基于快速 Hessian 检测子进行特征点的选取, 然后寻找到特征点的主方向, 最后能过 SURF 描述子得到一组描述特征点特征的向量, 基于这些向量来寻找不同图像之间的特征点匹配。本文实验所用图片均来自 <http://www.robots.ox.ac.uk/~vgg/research/affine/>。如图 4 所示为 bark 的原始图, 图 5 为检测到的特征点及主方向和所在尺度, 图 6 为最后匹配结果图。表 1 显示八组不同情况下使用 SURF 特征描述子进行图像匹配, 对每一组图, 均选择第 1、2 幅图进行实验。

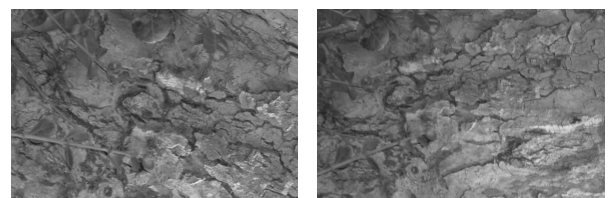


图4 bark 原始图

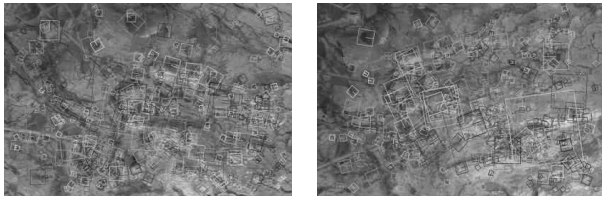


图5 bark 特征点及其主方向和所在尺度

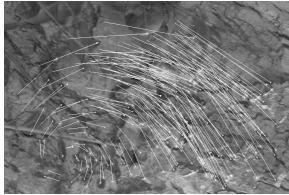


图6 匹配结果图

表1 八组图像匹配结果

图像变化类型	图像组名	图像大小	图1	图1 计算时间/ms	图2	图2 计算时间/ms	匹配点数
模糊	bikes	1 000×700	1 013	1 266	776	1 063	583
模糊	trees	1 000×700	3 318	3 461	3 230	3 314	970
不同视角	graf	800×640	1 543	1 656	1 768	1 882	569
不同视角	wall	1 000×700	2 227	2 365	2 116	2 218	1 012
缩放旋转	bark	765×512	811	935	739	871	179
缩放旋转	boat	850×680	2 336	2 474	2 223	2 319	852
光亮变化	leuven	900×600	1 064	1 207	846	1 009	555
JPEG 压缩	ubc	800×640	1 298	1 425	1 291	1 431	1 110

注: wall 的图1大小为1 000×700,图2大小为880×680

5 结论

本文基于 SURF 特征描述子进行图像匹配,充分表现了快速 Hessian 检测子和 SURF 描述子的速度优势。相比较有名的

SIFT 特征描述子,SURF 比 SIFT 快速,而且 SURF 同 SIFT 一样具有很好的尺度不变性。虽然在不同视角下的图像匹配中,SURF 比 SIFT 效果比较差,但在运动中 SURF 具有很好的效果,速度也比 SIFT 快。总的来说,SURF 具有良好的速度和精度。

参考文献:

- [1] Harris C, Stephens M J. A combined corner and edge detector[C]//Alvey Conference, 1988: 147-152.
- [2] Shi Jianbo, Tomasi C. Good features to track[C]//IEEE Conference on CVPR'94, Seattle, 1994: 593-600.
- [3] Tommasini T, Fusiello A, Trucco E, et al. Making good features track better[C]//Proceedings of the IEEE Computer Society Conference on CVPR'98, Washington, DC, USA: IEEE Computer Society, 1998: 178.
- [4] Lindeberg T. Feature detection with automatic scale selection[J]. International Journal of Computer Vision, 1998, 30(2): 79-116.
- [5] Mikolajczyk K, Schmid C. Indexing based on scale invariant interest points[C]//International Conference on Computer Vision, 2001, I: 525-531.
- [6] Lowe D G. Object recognition from local scale-invariant features[C]//International Conference on Computer Vision, 1999: 1150-1157.
- [7] Ke Y, Sukthankar R. PCA-SIFT: a more distinctive representation for local image descriptors[J]. IEEE CVPR, 2004, II: 506-513.
- [8] Bay H, Tuytelaars T, van Gool L J. SURF: Speeded Up Robust Features[C]//European Conference on Computer Vision, 2006, I: 404-417.
- [9] Viola P, Jones M J. Rapid object detection using a boosted cascade of simple features[J]. IEEE CVPR, 2001, I: 511-518.
- [10] Brown M, Lowe D. Invariant features from interest point groups[C]//British Machine Vision Conference, 2002.
- [11] Lee D S, Erol B, Huu J J. Segmenting people in meeting videos using mixture background and object models[C]//Advances in Multimedia Information Processing, PCM, 2002.
- [12] Haußecker H, Jahne B. A tensor approach for local structure analysis in multi dimensional images[C]//Proceedings 3D Image Analysis and Synthesis'1996, 1996.
- [13] Jahne B, Haußecker H, Scharr H. Study of dynamical processes with tensor-based spatio temporal image processing techniques[C]//LNCS 1407: Computer Vision-ECCV, 1998: 322-326.
- [14] Kothe U. Integrated edge and junction detection with the boundary tensor[C]//the 9th IEEE International Conference on Computer Vision 2003, 2003, 1: 424-431.
- [15] Ngo C W, Pong T C, Zhang H J. Motion analysis and segmentation through spatio-temporal slices[C]//International Conference on Image processing, 2003: 341-355.
- [16] Kuhne G, Weickert J, Schuster O, et al. A tensor-driven active contour model for moving object segmentation[C]//Proc IEEE Image Processing, 2001.
- [17] Video analysis [EB/OL]. [2005-06-01]. <http://knight.cis.temple.edu/~video/VA>.
- [18] Segmentation of videos using color, motion and spatial information [EB/OL]. [2007-04]. <http://www.cs.ucf.edu/~vision/projects/videoSegmentation/videoSegmentation.html>.

(上接 56 页)

6 结论

本文对视频的局部灰度特征进行了分析,建立了用于视频运动物体分割的时空结构张量。实验效果令人满意。但是还存在很多不足,由于视频数据的复杂性,距同时满足精度和实时的要求还有一定的距离。实验中的阈值是经过实验获得,阈值的自适应也是一个需要解决的问题。

参考文献:

- [1] Adiv G. Determining three-dimensional motion and structure from optical flow generated by several moving objects[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1985, 7(4): 384-401.
- [2] Oliver N M, Rosario B, Pentland A. Bayesian computer vision system for modeling human interactions[J]. IEEE Trans on PAMI, 2000, 22(8): 831-843.
- [3] Murray D W, Buxton B F. Scene segmentation from visual motion using global optimization[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1987, 9(2): 220-228.
- [4] Harville M, Gordon G, Woodfill J. Foreground segmentation using adaptive mixture models in color and depth[C]//Proceedings of Detection and Recognition of Events in Video, 2001.