

一种快速的空间变换模型计算方法

路晓静 黄向生

(中国科学院自动化研究所 北京 100090)

摘要 提出一种基于 Sobel 算子特征点检测的快速空间变换模型计算方法。利用 Sobel 算子对边缘的敏感特性生成 Sobel 描述器,进行特征点的描述和匹配,并运用 PROSAC 算法快速消除误匹配,从而计算出空间变换模型。实验结果表明,算法采用简单的 Sobel 算子组合成描述器,快速提取出了图像中的特征点,并且根据相似性描述器准确地进行了特征点匹配,提高了空间变换模型的计算速度,为增强现实和图像拼接等视觉领域应用提供了前提^[1,2]。

关键词 Sobel 算子,描述器,特征点,特征点匹配,PROSAC,增强现实,图像拼接

中图分类号 TP391.4 **文献标识码** A

Fast Calculation Method of Space Transform Model

LU Xiao-jing HUANG Xiang-sheng

(Institute of Automation, Chinese Academy of Sciences, Beijing 100090, China)

Abstract A novel fast spatial transformation model calculation approach, based on Sobel operator feature point detection, was proposed. The Sobel descriptor is constructed by taking advantages of the sensitive characteristics of Sobel operator on edge. The features are extracted and matched. Then the wrong matching points are eliminated by PROSAC algorithm, and the spatial transformation model between two images is calculated. The experimental results demonstrate that the simple structure Sobel descriptor can extract the feature points with high speed and match the feature points between two images exactly, thus improve the calculation speed, provide a method for augmented reality and image mosaic etc.

Keywords Sobel operator, Descriptor, Feature point, Feature point matching, PROSAC, Augmented reality, Image mosaic

空间变换模型的计算是计算机图形图像、计算机视觉领域的基础计算环节,主要用于增强现实^[3,4]和图像拼接等领域。空间变换模型表征的是空间中具有重叠场景的图像之间的空间变换关系,包括尺度和旋转变量、水平和垂直方向位移以及水平和垂直方向的变形量等信息。

目前普遍求解空间变换模型的方法是将具有一定空间变换关系的图像进行特征区域提取与匹配,之后采用数据集优化方法对变换模型进行计算。其中,特征区域提取与匹配算法中较常使用的算子有 SIFT 算子、Harris 角点等。但对于增强现实等领域而言,通常需要较快的计算速度以及较高的计算精度。SIFT 算子虽然在进行特征点的提取与匹配的过程中精度高,但是计算速度慢^[5]; Harris 角点计算速度快但对于特征点的提取与匹配精度不高^[6]。本文在特征点提取与匹配中,使用了一种基于 Sobel 算子的新方法,利用 Sobel 算法对边缘的敏感及便于提取等特点,在保证特征点提取与匹配精度的同时,有效地减少了特征区域提取与匹配的时间。在数据集优化过程中采用 PROSAC 算法,优先选取匹配度较高的特征点对进行模型估计,速度优于传统的 RANSAC 算法。

本文以静态图像为基础,分为图像获取、特征点提取与匹配、变换模型计算 3 个方面,着重探讨和解决了空间变换模型计算的快速性问题,为今后的计算机图形图像、计算机视觉等

领域提供理论和实验前提。

1 图像获取

由于摄像头存在一定程度上的失真,为了避免后续的拼接出现明显的变形,需要标定摄像头,从而对图像进行校正预处理。本文采用张正友标定法^[7,8]来得到摄像头的内部参数和畸变系数,对每一次视频序列传来的新帧进行校正,之后将图像转换为灰度图。校正前与校正后的示意图如图 1 所示。

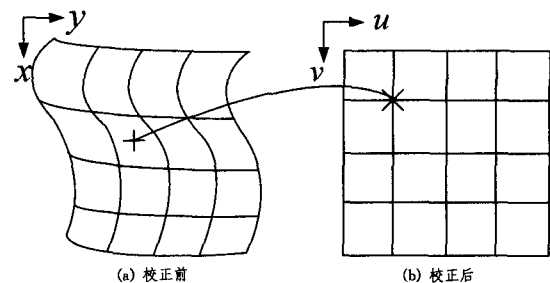


图1 图像校正预处理前后对比图

2 特征点提取与匹配

特征点提取与匹配这一步骤是空间变换模型计算的重点。本文在保证特征点提取与匹配质量的前提下,为了提高

速度,提出了一种基于 Sobel 算子的描述器。

Sobel 算子由于反映的是图像灰度变化的梯度情况,梯度值越大,说明图像灰度变化越剧烈,因此能够敏感地反映出图像中处于边缘位置的像素,保证了从两幅图像中提取出的特征点具有一致性。本文的 Sobel 描述器包含了特征点周围的信息,利用这些邻域信息准确地匹配对应的特征点。

本文提出的基于 Sobel 算子的描述器的描述对象是像素灰度变化,与传统的特征点提取和匹配算法中使用到的 SIFT 算子相比,不存在尺度空间上的检测,因此能够极大提高算法的速度。

2.1 Sobel 卷积

对左右灰度图分别沿着水平方向与垂直方向进行 Sobel 卷积,卷积值记为 $Conv$ 。采用的水平方向的卷积模版为 $[-1, 0, +1; -2, 0, +2; -1, 0, +1]$,垂直方向的卷积模版为 $[+1, +2, +1; 0, 0, 0; -1, -2, -1]$ 。

由于 Sobel 算子利用像素 8 个邻域的灰度值进行加权,将在边缘点处达到极值这一原理作为特征点提取的前提。该方法不但能产生较好的检测效果,而且对噪声具有平滑作用。

2.2 生成 Sobel 描述器

为了有效并区分地对每一个像素进行描述,将 16 个分布位置不同的 Sobel 算子组成 Sobel 描述器,其中沿水平方向的 Sobel 算子为 12 个,沿垂直方向的 Sobel 算子为 4 个。像素本身及其周围像素的沿不同方向的灰度变化情况,能够全面、唯一地反映出像素的特征。

取 5×5 的窗口作为提取 Sobel 描述器的模版, Sobel 描述器的像素分布位置如图 2 所示,其中每个方格表示一个像素的 Sobel 算子,模版中心与被描述的像素重合,方格中的数字表示提取的次数,图 2(a)为沿水平方向的像素提取分布图,图 2(b)为沿垂直方向的像素提取分布图。

0	0	1	0	0
1	0	1	0	1
0	1	2	1	0
1	0	1	0	1
0	0	1	0	0

0	0	0	0	0
0	0	1	0	0
0	1	0	1	0
0	0	1	0	0
0	0	0	0	0

图 2 分别沿水平和垂直方向的像素提取位置分布

将 Sobel 描述器记为 Des_{feat} ,沿水平方向的 Sobel 算子卷积值记为 $uConv$,沿垂直方向的 Sobel 算子卷积值记为 $vConv$,则 Sobel 描述器可表示为:

$$Des_{feature} = \{uConv(0, -2), uConv(-2, -1), uConv(0, -1), uConv(2, -1), uConv(-1, 0), uConv(0, 0), uConv(0, 0), uConv(1, 0), uConv(-2, 1), uConv(0, 1), uConv(2, 1), uConv(0, 2), vConv(0, -1), vConv(-1, 0), vConv(1, 0), vConv(0, 1)\} \quad (1)$$

括号中的坐标是以模版中心为原点的相对坐标。

2.3 提取特征点

Sobel 描述器利用像素本身及其周围像素的沿不同方向的灰度值变化情况进行描述, Sobel 算子的卷积值越大,说明梯度值越大,灰度变化越剧烈,将这样的像素点作为特征点,能够准确地表示出图像的边缘特性。

对描述器中的 16 个 Sobel 算子卷积值进行求和,记为 S_D ,则,

$$S_D = uConv(0, -2) + uConv(-2, -1) + uConv(0, -1) + uConv(2, -1) + uConv(-1, 0) + uConv(0, 0) + uConv(0, 0) + uConv(1, 0) + uConv(-2, 1) + uConv(0, 1) + uConv(2, 1) + uConv(0, 2) + vConv(0, -1) + vConv(-1, 0) + vConv(1, 0) + vConv(0, 1) \quad (2)$$

记纹理阈值为 $T_{texture}$,当 $S_D > T_{texture}$ 时,则提取该描述器所对应的像素点作为图像的特征点。 $T_{texture}$ 越低,提取的特征点越多,计算出的空间变换模型越精确,但是消耗的时间越长,因此需要采用适当的阈值来平衡精确性与时长消耗。

图 3 为使用 Sobel 描述器对图像进行特征点提取的效果图。可以看出, Sobel 描述器对于图像边缘的敏感特征使得两幅图像中的特征点具有一致性。

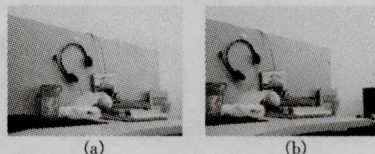


图 3 特征点提取

2.4 特征点匹配

以图 3(a)为参考图像,在图 3(b)中查找与左图特征点相匹配的特征点。为了确保特征点的准确匹配,本文在 Sobel 描述器的基础上,再次提取特征点 4 邻域的 Sobel 描述器作为相似性描述器,记为 $Desc_{similar}$,则

$$Desc_{similar} = \{Desc(-2, -2), Desc(2, -2), Desc(-2, 2), Desc(2, 2)\} \quad (3)$$

其中括号中的坐标是以被描述特征点为原点的相对坐标。

以图 3(a)的一个特征点为参考点,分别对图 3(b)中每个特征点的相似性描述器进行相减运算,并取绝对值之和,即 SAE。SAE 值越小,说明两个特征点的可能匹配度越高,若第一匹配与第二匹配的比值小于比例阈值 T_{th} ,则确定第一匹配所对应的特征点为参考特征点的匹配点。若第一匹配与第二匹配的比值大于比例阈值 T_{th} ,则说明与参考特征点类似的特征点有两个或多个,为避免匹配错误,放弃对该参考点的匹配。按照此法对参考图像的每个特征点进行匹配。

图 4 为使用相似性描述器对提取到的特征点进行匹配的结果。可以看出,使用 Sobel 描述器提取到的特征点能够得到较好的匹配,能够全面、唯一地反映出像素的特征。

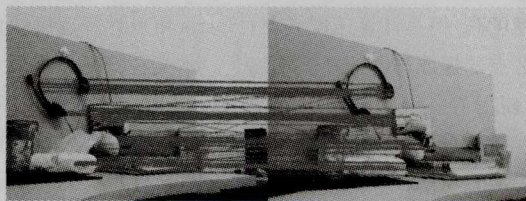


图 4 特征点匹配

3 变换模型计算

根据两幅图像的匹配特征点,可以求得两幅图像的空间变换模型,以矩阵形式表示,即变换矩阵 H 。变换矩阵中有 8 个独立参数,则至少需要 4 对匹配点才可以求解出变换矩阵的 8 个参数^[9]。但是在得到两幅图像的匹配特征点之后,由

于存在一定的误匹配,若从中直接抽取 4 对匹配点来计算变换矩阵,必然会导致变换矩阵的不稳定性和较大的计算误差。因此本文采用 PROSAC 算法^[10,11],计算出具有较大有效集的变换模型。

与传统的模型估计算法 RANSAC^[12]不同,PROSAC 算法的采样集不是整个数据集,而是匹配度最高的子集,子集随着抽样次数的增加而扩大,并最终达到整个数据集。这种模型估计的方法与 RANSAC 算法本质相同,但是其可以快速对拟合过程进行收敛,提高了变换矩阵的计算速度。

假设整个匹配点对个数为 N ,集合为 U_N ,包含 n 个点对的子集为 U_n ,则 $U_n \subset U_N$ 。 T_n 为假设使用 RANSAC 算法在 U_n 中随机抽取 4 个匹配点对的平均采样次数。PROSAC 算法步骤如下:

1. 将数据集按照匹配度进行降序排列。
2. 确定采样集。

n 的初始值为 4,即估计变换矩阵所需的最小匹配点对数。采样集长度 $n=n+1$ 。

3. 在采样集中抽取样本。

当采样次数 $t > T_n$ 时,在子集 U_{n-1} 中随机抽取 3 个点对,另外 1 个点对为第 n 个点对,由这 4 个点对共同构成估计变换矩阵的样本。

当 $t \leq T_n$ 时,在集合 U_n 中随机抽取 4 个匹配点对作为估计变换矩阵的样本。

4. 计算变换矩阵。

根据 4 对样本集进行变换矩阵的估计。

5. 采样终止判断。

采样终止的条件有两个:有效点对与总点对的比值大于规定阈值;有效点对的增长率小于规定阈值。如果满足终止条件,则退出,否则返回步骤 2。

采用 PROSAC 算法不仅能够较为准确地对匹配点中的误匹配进行筛选,得到包含最多有效匹配点或达到有效点数量的变换模型,而且计算速度快。

4 实验结果

为了说明基于 Sobel 描述器进行空间变换模型计算的有效性和快速性,本文采用图像拼接的方式对其进行验证。图 5 是本实验采用的 3 组具有一定空间变换关系的场景图,第一组场景图采用的图像大小为 463×370 ,其余场景图像大小为 640×480 。



图 5 3 组具有空间变换关系的场景

表 1 是分别使用 3 种不同的组合算法进行图像拼接的时间对比,组合 3 即为本文所提出的算法,括号里表示在左右图中提取到的特征点数量。

表 1 不同组合算法下的图像拼接时间

	组合 1:SIFT+ RANSAC	组合 2:Sobel+ RANSAC	组合 3:Sobel+ PROSAC
图像 1	764ms (178/188)	77ms (265/117)	22ms (265/117)
图像 2	862ms (78/54)	67ms (95/85)	19ms (95/85)
图像 3	965ms (234/203)	106ms (246/217)	24ms (246/217)

可以看出,在提取特征点数量相近的情况下,采用本文所提出的算法能够有效地提高计算速度。

在拼接质量上,如图 6 所示,基于 Sobel 描述器的拼接结果基本可以达到拼接要求,甚至比基于 SIFT 算子的拼接质量好。



图 6 不同场景基于 SIFT 算子和 Sobel 描述器的拼接结果

经过实验的对比与分析可知,本文提出的基于 Sobel 描述器的空间变换模型的计算方法能够在保证计算精度的前提下,对变换模型进行快速计算,计算速度可以达到实时性的要求,为增强现实和图像拼接等领域提供了前提^[13]。

结束语 本文研究了一种基于 Sobel 描述器的快速空间变换模型计算的新方法,利用 Sobel 算子对边缘的敏感性,提出了 Sobel 描述器,用来描述图像像素及其周围像素的梯度总体变化情况,该方法能够有效并区分地标识出每个像素点的特征,从而对两幅图像中具有相似特征的特征点进行匹配,并计算出空间变换模型。

实验表明,采用本文提出的空间模型计算方法能够快速并准确地确定图像之间的空间变换模型。Sobel 描述器直接反映了像素的梯度变化,操作简单,容易实现,计算速度快,对于图像的平移变换能够准确提取特征点并匹配,具有较高的稳定性和鲁棒性。因此,本文提出的基于 Sobel 算子的快速空间变换模型计算的新方法为后续研究如增强现实和图像拼接提供了理论和实验基础。

参考文献

- [1] Höllerer T, Wither J, DiVerdi S. Anywhere augmentation: towards mobile augmented reality in unprepared environments [C] // Lecture Notes in Geoinformation and Cartography. Springer Verlag, 2007: 93-416
- [2] Zitova B, Flusser J. Image registration methods: a survey [J]. Image and Vision Computing, 2003, 21: 977-1000
- [3] Wagner D, Mulloni A, Schmalstieg D. Real-Time detection and tracking for augmented reality on mobile phones [J]. IEEE Transactions on Visualization and Computer Graphics, 2010, 16: 355-368
- [4] 孙源,陈靖. 智能手机的移动增强现实技术研究[J]. 计算机科学, 2012, 39(6A): 439-498

(下转第 292 页)

- Sampling-Rate High Efficiency Video Coding for Full-Chroma Screen Content [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2013, 23(1): 173-185
- [6] Meuel H, Schmidt J, Munderloh M, et al. Analysis of coding tools and improvement of text readability for screen content[C]// Picture Coding Symposium. Krakow, 2012: 469-472
- [7] Lin T, Hao Peng-wei. Compound image compression for real-time computer screen image transmission[J]. IEEE Transactions on Image Processing, 2005, 14(8): 993-1005
- [8] Lin T, Hao Peng-wei, Lee S U. Efficient coding of computer generated compound images[C]// IEEE International Conference on Image Processing. Genoa, 2005: 561-64
- [9] Shen Yan-fei, Li Jin-tao, Zhu Zhen-min, et al. Classification-Based Adaptive Compression Method for Computer Screen Image [C]// IEEE International Conference on Multimedia and Expo Workshops. 2012: 7-12
- [10] Lan Cui-ling, Wu Feng, Shi Guang-ming. Compress Compound Images in H. 264/MPGE-4 AVC by Exploiting Spatial Correlation[J]. IEEE Transactions on Image Processing, 2010, 19(4): 946-957
- [11] Mrak M, Xu Ji-zheng. Improving screen content coding in HEVC by transform skipping[C]// Signal Processing Conference. Bucharest, 2012: 1209-1213
- [12] Nam J, Sim D, Bajic I V. HEVC-based adaptive quantization for screen content videos[C]// IEEE International Symposium on Broadband Multimedia Systems and Broadcasting. Seoul, 2012: 1-4
- [13] Wang Shu-hui, Lin Tao. A Unified LZ and Hybrid Coding for Compound Image Partial-Lossless Compression[C]// 2nd International Congress on Image and Signal Processing. Tianjin, 2009: 1-5
- [14] Wang Shu-hui, Lin Tao. United Coding for Compound Image Compression[C]// 3rd International Congress on Image and Signal Processing. Yantai, 2010: 566-570
- [15] Wang Shu-hui, Lin Tao. United Coding Method for Compound Image Compression [J/OL]. Multimedia Tools and Applications, published online; <http://link.springer.com/article/10.1007/s11042-012-1274-y/fulltext.html>, Springer, 2012
- [16] Lin Tao, Zhang Pei-jun, Wang Shu-hui, et al. 4:4:4 Screen Content Coding using Dual-coder Mixed Chroma-sampling-rate (DMC) Techniques[C]// JCT-VC-9th Meeting, JCTVC-I0272. Geneva, 2012
- [17] Zhang Pei-jun, Lin Tao, Wang Shu-hui, et al. Subjective Quality Comparison between DMC Coding and Chroma- subsampled 420 Coding Using YUV444 Screen Content Test Sequences[C]// JCT-VC-9th Meeting, JCTVC-I0336. Geneva, 2012
- [18] Zhang Pei-jun, Lin Tao, Chen Xian-yi, et al. HM software implementation and source code for JCTVC-K0133 [C]// JCT-VC-11th Meeting, JCTVC-K0134. Shanghai, 2012
- [19] Gary S, Jens-Rainer O. Meeting report of the fifth meeting of the Joint Collaborative Team on Video Coding [C]// JCT-VC-5th Meeting, JCTVC-E600. Geneva, 2011
- [20] Silcock P, Sharman K, Saunders N, et al. Extension of HM7 to Support Additional Chroma Formats[C]// JCT-VC-10th Meeting, JCTVC-J0191. 2012
- [21] Lin Tao, Wang Shu-hui, Zhang Pei-jun, et al. Full-chroma (YUV444) dictionary + hybrid dual-coder extension of HEVC [C]// JCT-VC-11th Meeting, JCTVC-K0133. Shanghai, 2012
- [22] Lin Tao, Zhang Pei-jun, Wang Shu-hui. Syntax and semantics of Dual-coder Mixed Chroma-sampling-rate (DMC) coding for 4: 4:4 screen content[C]// JCT-VC-10th Meeting, JCTVC-J0233. Stockholm, 2012
- [23] Wang Shu-hui, Lin Tao, Zhou Kai-lun. YUV444 and RGB screen content test sequences [C]// JCT-VC-12th Meeting, JCTVC-L0317. Geneva, 2013
- [24] Yu Hao-ping, Cohen R, Gao Wen. New 4:4:4 screen-content sequences for HEVC extension development[C]// JCT-VC-12th Meeting, JCTVC-L0301. Geneva, 2013
- [25] Budagavi M. Cross-check of JCTVC-K0133 (dictionary + hybrid dual-coder extension of HEVC) [C]// JCT-VC-12th Meeting, JCTVC-K0329. 2012
- [26] Bossen F. Common test conditions and software reference configurations, JCT-VC-10th Meeting [C]// JCTVC-J1100. Stockholm, 2012

(上接第 281 页)

- [5] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60 (2): 91-110
- [6] Harris C, Stephens M. A combined corner and edge detector [C]// Proceedings of the Fourth Alvey Vision Conference. 1988: 147-151
- [7] 马颂德, 张正友. 计算机视觉 [M]. 北京: 清华大学出版社, 2000
- [8] Zhang Z Y. Flexible camera calibration by viewing a plane from unknown orientations [J]. Proceedings of the Seventh IEEE International Conference on Computer Vision, 1999, 1(1): 666-673
- [9] 阮鹏. 多图像拼接算法研究[D]. 长沙: 中南大学, 2009
- [10] Chum, Metas J. Matching with PRO SAC-progressive sample consensus [J]. Conference on Computer Vision and Pattern Recognition (CVPR), 2005, 1: 220-226
- [11] 肖娟. 一种渐进式的抽样策略及其性能分析[J]. 计算机工程与应用, 2008, 44(27): 41-42
- [12] Fischler M, Bolls R. Random sample consensus: a paradigm for modal fitting with application to image analysis and automated cartography [J]. Communications of the ACM, 1981, 24(6): 381-395
- [13] Wagner D, Reitmayr G, Mulloni A, et al. Real-time detection and tracking for augmented reality on mobile phones [J]. IEEE Transactions on Visualization and Computer Graphics, 2010, 16 (3): 255-368