



JOURNAL OF IMAGE AND GRAPHICS

主办: 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

中国图象图形学报

2016
01
VOL.21

ISSN1006-8961
CN11-3758/TB



医学图像处理 P69

中国图象图形学报

刊名题字：宋健 | 月刊（1996年创刊）



第21卷第1期（总第237期）
2016年1月16日

中国精品科技期刊
中国国际影响力优秀学术期刊
中国科技核心期刊
中文核心期刊

版权声明

凡向《中国图象图形学报》投稿，均视为同意在本刊网站及CNKI等全文数据库出版，所刊载论文已获得著作权人的授权。本刊所有图片均为非商业目的使用，所有内容，未经许可，不得转载或以其他方式使用。

Copyright

All rights reserved by Journal of Image and Graphics, Institute of Remote Sensing and Digital Earth, CAS. The content (including but not limited text, photo, etc) published in this journal is for non-commercial use.

主管单位 中国科学院
主办单位 中国科学院遥感与数字地球研究所
中国图象图形学学会
北京应用物理与计算数学研究所

主 编 李小文
编辑出版《中国图象图形学报》编辑出版委员会
邮政信箱 北京9718信箱
邮 编 100101
电子信箱 jig@radi.ac.cn
电 话 010-64807995
网 址 www.cjig.cn

广告经营许可证 京朝工商广字第0361号
总 发 行 北京报刊发行局
订 购 全国各地邮局
海外发行 中国国际图书贸易集团有限公司
(邮政信箱: 北京399信箱 邮编: 100048)
印刷装订 北京科信印刷有限公司

Journal of Image and Graphics

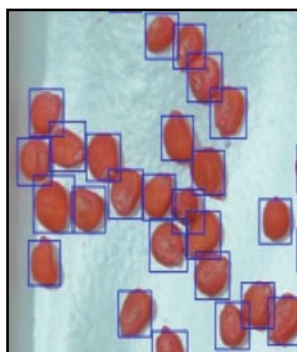
Title inscription: Song Jian | Monthly, Started in 1996

Superintended by Chinese Academy of Sciences
Sponsored by Institute of Remote Sensing and Digital Earth, CAS
China Society of Image and Graphics
Institute of Applied Physics and Computational Mathematics

Editor-in-Chief LI Xiaowen
Editor, Publisher Editorial and Publishing Board of Journal of Image and Graphics
P.O.Box 9718, Beijing, P.R.China
Zip code 100101
E-mail jig@radi.ac.cn
Telephone 010-64807995
Website www.cjig.cn

Distributed by Beijing Bureau for Distribution of Newspapers and Journals
Domestic All Local Post Offices in China
Overseas China International Book Trading Corporation
(P.O.Box 399, Beijing 100048, P.R.China)
Printed by Beijing Kexin Printing Co., Ltd.

CN 11-3758/TB
ISSN 1006-8961
CODEN ZTTXFZ
国外发行代号 M1406
国内邮发代号 82-831
国内定价 60.00元



多尺度特征融合与极限学习机的玉米种子检测
(第0024页)



深度卷积神经网络的显著性检测(第0053页)



双树复小波域的MRI图像去噪(第0104页)

图像处理和编码

空时形状预测与高效编码

朱仲杰, 王玉儿, 蒋刚毅..... 1

第二小方向导数信息熵的兴趣点检测

卢健, 黄杰, 潘峰..... 8

采用面积采样的反走样字符旋转算法

郑凯文, 刘文波..... 17

图像分析和识别

多尺度特征融合与极限学习机的玉米种子检测

柯道, 杜明智..... 24

图像理解和计算机视觉

结合NSCT和压缩感知的红外与可见光图像融合

陈木生..... 39

Plücker直线描述的空间后方交会

盛庆红, 陈姝文, 肖晖, 张斌, 王青, 费利佳..... 45

深度卷积神经网络的显著性检测

李岳云, 许悦雷, 马时平, 史鹤欢..... 53

计算机图形学

以优先点为中心的Delaunay三角网生长算法

尤磊, 唐守正, 宋新宇..... 60

医学图像处理

快速数字影像重建的2维/3维医学图像配准

刘坤, 吕晓琪, 谷宇, 于荷峰, 任国印, 张明..... 69

小波-Lagrange方法进行医学图像层间插值

武士想, 尚鹏, 王立功..... 78

遥感图像处理

多主体框架下结合最大期望值和遗传算法的SAR图像分割

张金静, 李玉, 赵泉华..... 86

第十届图像图形技术与应用学术会议

视差互信息引导下的立体航空影像与LiDAR点云自动配准

吴军, 胡彦君, 饶云, 彭智勇..... 95

双树复小波域的MRI图像去噪

黄学优, 张长江..... 104

快速定位图像尺度和区域的3维跟踪算法

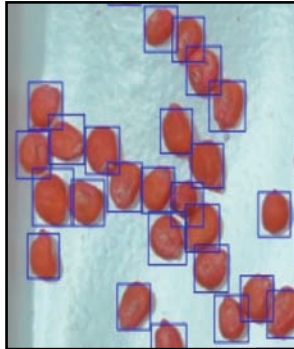
赵奇可, 孙延奎..... 114

局部不变特征点的精度指标

滕日, 周进, 蒋平, 崔雄文..... 122

CONTENTS

JOURNAL OF IMAGE AND GRAPHICS



Detection of maize seeds based on multi-scale feature fusion and extreme learning machine(P0024)



Saliency detection based on deep convolutional neural network(P0053)



MRI denoising based on dual-tree complex wavelet transform(P0104)

Image Processing and Coding

- Spatio-temporal shape prediction and efficient coding
Zhu Zhongjie, Wang Yuer, Jiang Gangyi 1
- Interest point detection by using information entropy of the second small direction derivative
Lu Jian, Huang Jie, Pan Feng 8
- Anti-aliasing algorithm for character rotation based on area sampling
Zheng Kaiwen, Liu Wenbo 17

Image Analysis and Recognition

- Detection of maize seeds based on multi-scale feature fusion and extreme learning machine
Ke Xiao, Du Mingzhi 24

Image Understanding and Computer Vision

- Image fusion of visual and infrared image based on NSCT and compressed sensing
Chen Musheng 39
- Space resection method based on Plücker line
Sheng Qinghong, Chen Shuwen, Xiao Hui, Zhang Bin, Wang Qing, Fei Lijiao 45
- Saliency detection based on deep convolutional neural network
Li Yueyun, Xu Yuelei, Ma Shiping, Shi Hehuan 53

Computer Graphics

- Growth algorithm centered on priority point for constructing the Delaunay triangulation
You Lei, Tang Shouzheng, Song Xinyu 60

Medical Image Processing

- The 2D/3D medical image registration algorithm based on rapid digital image reconstruction
Liu Kun, Lyu Xiaoqi, Gu Yu, Yu Hefeng, Ren Guoyin, Zhang Ming 69
- Inter-slice interpolation for medical images by using the wavelet-lagrange method
Wu Shixiang, Shang Peng, Wang Ligong 78

Remote Sensing Image Processing

- Combining the expectation maximization and genetic algorithms for SAR image segmentation within a MAS framework
Zhang Jinjing, Li Yu, Zhao Quanhua 86

Column of IGIA' 2015

- Automatic registration of aerial stereo imagery to LiDAR data through iterative parallax mutual information computation
Wu Jun, Hu Yanjun, Rao Yun, Peng Zhiyong 95
- MRI denoising based on dual-tree complex wavelet transform
Huang Xueyou, Zhang Changjiang 104
- Three dimensional tracking with fast locating of image scale and area
Zhao Qike, Sun Yankui 114
- Extraction precision of local invariant feature points
Teng Ri, Zhou Jin, Jiang Ping, Cui Xiongwen 122

中图法分类号: P237 文献标识码: A 文章编号: 1006-8961(2016)01-0114-08

论文引用格式: Zhao Q K, Sun Y K. Three dimensional tracking with fast locating of image scale and area[J]. Journal of Image and Graphics, 2016, 21(1):0114-0121. [赵奇可, 孙延奎. 快速定位图像尺度和区域的 3 维跟踪算法[J]. 中国图象图形学报, 2016, 21(1):0114-0121.][DOI: 10.11834/jig.20160114]

快速定位图像尺度和区域的 3 维跟踪算法

赵奇可, 孙延奎

清华大学计算机科学与技术系, 北京 100084

摘要: **目的** 传统的 2 维自然图像的增强现实算法, 对模板图像的各个尺度下的整个图像提取特征点并保存到特征点数组中, 跟踪阶段对模板图像提取出的所有特征点进行匹配, 造成了大量的无效运算, 降低了特征匹配的效率。为了解决这个问题, 将模板各个尺度的图像进行区域划分, 提出了一种快速定位图像尺度和区域的算法, 缩小特征匹配的范围, 加快 3 维跟踪的速度。**方法** 预处理阶段, 通过对图像金字塔每一尺度图像分成小区域, 对模板图像上的特征点进行分层次分区域的管理。在系统实时跟踪阶段, 通过计算几何算法快速确定当前摄像机图像所对应的尺度和区域, 从而减小了特征匹配的范围。**结果** 该方法大幅度缩小了特征匹配的范围, 降低了特征匹配所消耗的时间, 与传统算法相比, 在模板图像分辨率较大的情况下特征匹配阶段时间可以缩短 10 倍左右, 跟踪一帧图像的时间缩短 1.82 倍。系统实时跟踪过程中的帧率总体保持在 15 帧/s 左右。**结论** 提出的快速定位图像尺度和区域算法适用于移动设备上对 2 维自然图像的跟踪, 尤其在模板图像分辨率较大的情况下, 算法能够显著减小特征匹配的范围, 从而提升了实时 3 维跟踪算法的运行效率。

关键词: 增强现实; 3 维跟踪; 特征匹配; 移动平台; 快速定位; 模板图像

Three dimensional tracking with fast locating of image scale and area

Zhao Qike, Sun Yankui

Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China

Abstract: **Objective** The traditional augmented reality algorithms for natural image detect and extract the natural features of a template image for multi scales. At the tracking stage the algorithms match the screen features with all of the features extracted from the template image, which produces a large amount of useless computation and reduces the efficiency of feature matching. To solve this problem, this article presents an algorithm which can locate the scale and area of the template image fast and accelerate the speed of 3D tracking. **Method** In the preprocess stage, the image in each scale is partitioned into small areas, and the features of the template image are managed according to the scale and area. In the real-time tracking stage, the method locates the scale and area of current camera image quickly using some computation geometry algorithms so that the scope of feature matching is decreased. **Result** The search field of feature matching, cuts down the time spent by feature matching, compared with the traditional method, the time spent by feature matching is cut short by about 10 times and the time spent by tracking one frame is cut short by 1.82 times when the resolution of template image is high. The

收稿日期: 2015-05-06; 修回日期: 2015-05-29

基金项目: 国家高技术研究发展计划(863) 基金项目(2013AA013702)

第一作者简介: 赵奇可(1990—), 男, 清华大学计算机科学与技术系硕士研究生, 主要研究方向为计算机视觉、增强现实。

E-mail: zhaoque@163.com

通信作者: 孙延奎, 副教授, E-mail: syk@mail.tsinghua.edu.cn

Supported by: National High Technology Research and Development Program of China(2013AA013702)

frame rate of the 3d tracking algorithm keeps at 15 frames per second on average. **Conclusion** Our method can be used to the tracking of natural image on mobile platform, especially when the resolution of image is high, the method can narrow the search field of image scale and area and promote the efficiency of real-time 3d tracking algorithm.

Key words: augmented reality; three dimensional tracking; feature matching; mobile platform; fast locating; template-image

0 引言

增强现实是当前非常热门的一个发展领域,其中,对2维平面自然图像进行增强现实是非常重要的一个方向。随着近几年移动设备的性能快速发展,移动增强现实技术也取得了快速的发展。一方面,移动设备在计算能力,存储能力和续航能力等方面相比PC端存在着比较大的差距,另一方面,移动设备具有更好的便携性,并且配备了更多的传感设备,例如GPS、加速计、陀螺仪等。这些传感设备使得移动设备可以从外界获得更多的方位和移动信息。由于移动设备具有这样的优缺点,人们在移动设备上实现了增强现实算法,并对算法进行了计算上的优化,使得增强现实系统可以在移动设备上流畅的运行^[1]。

传统的2维图像增强现实算法大体可以分为两个阶段:预处理阶段和实时跟踪阶段。预处理阶段从模板图像中提取特征点,并存放在一个特征点数组中。传统的3维跟踪算法为了实现特征的尺度不变性,对模板图像构建金字塔结构,对不同的尺度图像分别提取特征点。提取特征点之后,将各个不同尺度图像上提取到的特征点保存在一个特征点数组中。在实时跟踪阶段,对获取到的每一帧摄像机图像提取特征点,并与预处理阶段得到的模板特征点数组进行特征匹配,进而计算摄像机位姿^[2]。

由于移动设备在计算能力方面相对不高,人们提出了很多方法提升3维跟踪算法的效率。这方面的改进主要集中在特征提取和特征匹配方面。Wagner等人^[3]将传统的SIFT^[4]和Ferns^[5]算法进行了改进,提出了适用于移动端的PhonySIFT和PhonyFerns算法,加快了特征提取的速度,并在提取到特征点之后,使用PatchTracker的跟踪策略。这种策略利用相邻图像帧之间的相似性,在上一帧图像特征点的邻域内寻找相同特征点,加快了特征检测和匹配的速度。Yang和Cheng^[6]从SSD(self simi-

larity descriptor)出发提出了一种性能更高的LDB二进特征点算法,与BRIEF算法^[7]相比,该LDB算法在特征匹配的速度和准确率方面都有较大的提升。Ufkes等人^[8]使用了ORB^[9]作为特征提取算法,使用map数据结构保存不同位置拍摄到的模板图像。在实时跟踪阶段,使用词袋模型确定匹配程度最高的map,然后使用层次聚类树算法^[10]快速匹配特征点。使用NEON指令集重写了特征提取和匹配算法,并在获取了前一帧信息的前提下使用光流跟踪算法^[11]进行跟踪,有效提升了3维跟踪算法的效率。

目前算法通常将模板图像上提取到的特征点存放在一个数组中,没有按照特征点在模板上出现的位置和尺度分别存放管理。所以在实时跟踪过程中,摄像机图像特征点需要与模板图像上单尺度或多尺度下的整个图像的特征点进行匹配。而在实际中,摄像机图像可能只会与模板图像的某个尺度或某个区域匹配成功。其余的大部分特征匹配都是无效的。这无疑浪费了大量的计算资源。本文提出了一种可快速定位图像尺度和区域的特征匹配算法(FLISA),大大提升了特征匹配的效率,而且,该算法具有尺度不变性。

1 快速定位关键帧算法

整个算法大体上可以分为预处理和实时跟踪两个部分。预处理部分的主要工作是对将要跟踪的模板图像进行处理,生成相应的数据结构供后续的跟踪阶段使用。这个数据结构里面包含了模板的不同的尺度和区域信息。实时跟踪阶段主要是使用快速算法确定当前摄像机图像所处的尺度和区域,从而实现特征的快速匹配。

1.1 模板训练方法

预处理部分的主要工作是对将要跟踪的模板图像进行处理,得到图像多尺度金字塔和关键帧信息,并将这些信息有序的组织到关键帧数据结构中,提

供给后续的实时跟踪阶段使用。这部分可以分为两个阶段,分别为构建金字塔阶段和生成关键帧阶段。

1.1.1 构建金字塔方法

构建金字塔流程如下:

1) 确定降采样频率。通过对模板图像不断降采样得到不同尺度上的图像。其中降采样最常用的是2的幂次。在本文算法中,发现采用2的幂次时两个尺度之间的间隔太大,在尺度间过渡时容易跟踪失败。故采用了文献[12]中的采样方法,对原图像先进行一次1.5倍采样,然后对原图像和1.5倍采样图像分别不断进行2采样并将这两个序列交叉放置。其中 l 尺度下相对模板原图像的缩放比例 s_l 为

$$s_l = \begin{cases} 2^{\frac{l}{2}} & l \text{ 是偶数} \\ 1.5 \times 2^{\frac{l}{2}} & l \text{ 是奇数} \end{cases} \quad (1)$$

式中, l 表示当前尺度。

2) 确定分层数。确定了模板图像的降采样频率后,可以确定模板图像的尺度数。在本文算法中,希望金字塔最大尺度下的图像分辨率刚好小于摄像头获取的图像分辨率,需要计算模板图像尺度相对于摄像机图像分辨率的倍数 M ,即

$$M = \frac{\frac{w_p}{w_c} + \frac{h_p}{h_c}}{2} \quad (2)$$

式中, w_p, h_p 分别为模板图像的长和宽方向的分辨率, w_c, h_c 分别为摄像机图像的长和宽方向的分辨率。所以对原始模板图像分割的最大层数为

$$L = \lceil 2 \times \log_2 M + 1 \rceil = \lceil 2 \times \log_2 \left(\frac{w_p}{w_c} + \frac{h_p}{h_c} \right) - 1 \rceil \quad (3)$$

3) 构建金字塔。对模板图像按照之前确定的降采样频率不断进行降采样,获得每一个尺度下的图像。直达到最大尺度 L 后停止。

4) 提取特征点并还原位置。对金字塔中每一尺度下的图像,检测并提取特征点,在本文中,特征检测算法使用了高效快速的FAST^[13]算法,特征提取算法使用了BRIEF算法。但是每一层金字塔图像经过缩放之后,这些特征点的位置已经不是原始模板图像中的位置,需要将这些特征点的位置乘以当前尺度图像帧相对于原始模板图像缩放的倍数,才能得到特征点在原始模板图像中的位置。

1.1.2 生成关键帧算法

在对模板图像构建金字塔之后,得到了模板图像的 L 个尺度的金字塔结构。然后对金字塔的每一层进行分块构建关键帧。具体的构建关键帧方法如下:

1) 确定关键帧位置。对于每一尺度下的图像,使用相同大小的矩形区域用相同的间隔排在尺度图像上,使得这些矩形区域能够尽量均匀地覆盖在尺度图像上。将每个矩形区域的位置和其中的子图像以及区域内的特征点综合起来成为一个关键帧结构。关键帧的大小采用最大尺度第 L 层的图像大小,这样最大尺度 L 层下的图像只有一个关键帧。对于 l 层,横向和纵向上的关键帧数目

$$n = \lceil \frac{w_l}{w_L} \rceil \quad (4)$$

式中, w_l 和 w_L 分别为模板金字塔第 l 层和第 L 层的图像宽度。如果 w_l 和 w_L 不是整数倍关系,则这一层上的关键帧之间会有一些重叠。第 l 层中横向上第 i 列,纵向上第 j 行关键帧的起始位置的横坐标和纵坐标分别为

$$x_{l(i,j)} = i \times \frac{w_l - w_L}{n - 1} \quad (5)$$

$$y_{l(i,j)} = j \times \frac{h_l - h_L}{n - 1} \quad (6)$$

式中, h_l 和 h_L 分别为模板金字塔第 l 层和第 L 层的图像高度。这样就得到了尺度图像上每一行和每一列关键帧的起始点位置,也就得到了每一个关键帧的位置。本文使用北京市地图作为模板图像。模板图像如图1所示。



(a) 关键帧有重合的划分 (b) 关键帧无重合的划分

图1 模板图像关键帧划分

Fig. 1 Keyframe partitioning of template image ((a) keyframe overlapping partitioning; (b) keyframe no overlapping partitioning)

图1分别显示了模板图像的两个尺度下的关键帧区域。图1(a)中,每个关键帧的长和宽分别为当前尺度图像的 $2/3$,所以4幅关键帧之间有一定的重合区域。图1(b)中,每个关键帧的长和宽分别为

当前尺度图像的1/2,所以4幅关键帧之间没有重合区域,恰好将尺度图像分成4部分。

2) 分配特征点到关键帧中。知道了每一个关键帧的位置,可以将每一尺度下所有的特征点按照是否该在关键帧范围内分配到当前尺度下相应的关键帧中去,由关键帧保存和管理这些特征点。

1.2 选取关键帧

实时跟踪阶段需要快速确定能够和当前摄像机图像匹配成功的模板关键帧。选取关键帧的流程如图2所示。系统首先使用上一帧图像位姿作为当前帧位姿的估计,使用计算几何方法验证这个估计位姿的正确性。如果不正确,则使用传统的3维跟踪方法进行跟踪。如果正确,则将所有的关键帧矩形投影到当前帧上并计算与当前帧的重合程度。最后选取与摄像机图像重合度最高的若干关键帧作为候选关键帧序列。

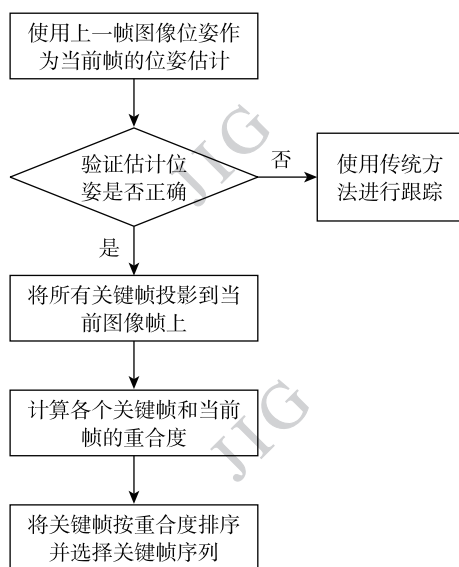


图2 选取关键帧算法

Fig. 2 Keyframe selection method

选取关键帧算法的具体流程如下:

1) 估计当前帧位姿。根据摄像机移动的连续性,连续两幅图像的摄像机位姿很接近,故将上一帧摄像机图像的摄像机位姿作为当前位姿的估计值。

2) 投影正确性判断。每个关键帧在模板图像上的边框都为矩形,根据仿射变换的性质,经过仿射变换投影到摄像机图像上的边框仍然为一个简单的凸四边形。且因为这些关键帧都在一个平面上,所以投影后的形状也是相似的,所以只需要对一个关键帧进行判断。在本文中,对第 L 层的唯一一个关

键帧进行判断,如果投影后的四边形为复杂多边形或凹多边形,则证明估计的位姿错误。需要使用传统方法进行跟踪进行错误修正。简单多边形的判断算法是:对多边形的任意两条不相邻边求交点,如果都没有交点则证明是简单多边形。凸多边形的判断算法是:求解多边形顶点集合的凸包,判断凸包的面积和多边形的面积是否相等,相等则证明是凸多边形,如果凸包面积大于多边形面积,则证明是凹多边形。

3) 关键帧投影。若位姿估计正确,则利用之前估计好的位姿将所有的关键帧的范围投影到当前图像帧上去。

4) 求解相交多边形。将关键帧投影到摄像机图像上之后对关键帧投影四边形和摄像机图像4个顶点构成的四边形求相交部分,根据凸集的封闭性看,凸多边形的交仍然是一个凸多边形。求相交多边形的方法是:首先摄像机图像形状是个矩形,使用简单的包围盒技术排除一些肯定不可能相交的情况。如果不能排除,则计算两个四边形的所有边相交的点和一个多边形在另一个多边形当中的顶点。这些顶点就是相交多边形的顶点。求这些多边形的凸包,则为相交多边形。

5) 计算重合因子。用 S_{kf} 、 S_c 、 S_{inter} 分别来表示关键帧多边形、摄像机图像矩形和相交多边形这3个多边形的面积。面积的计算方式是: N 边形,分为 $N-2$ 个三角形,求每个三角形求面积,最后相加得出多边形面积。定义摄像机图像和关键帧的重合因子为

$$\sigma = \frac{\frac{S_{inter}}{S_{kf}} + \frac{S_{inter}}{S_c}}{2} \quad (7)$$

重合因子 σ 度量了关键帧和当前摄像机图像的重合程度。

6) 选取关键帧序列。当计算出所有的关键帧与摄像机图像的重合因子后,对重合因子进行排序,挑选出重合度最高的若干关键帧作为候选关键帧,进行后续的特征匹配。本文选取了3帧重合度最高的候选关键帧,这样可以使摄像机的空间位置在两个相邻尺度中间或者同一尺度下两帧相邻关键帧之间时也能得到较好的匹配。

图3显示了算法在运行过程中选取关键帧的结果。图中的白线是所有的关键帧在当前帧上的投

影,深色矩形是筛选出的与当前摄像机图像重合度最高的关键帧。从图3可以看到,由于当前摄像机图像位于模板图像的一角,故系统选取出了这个角落3个与摄像机图像最接近的尺度的关键帧,这3个关键帧与摄像机图像的重合度最高。

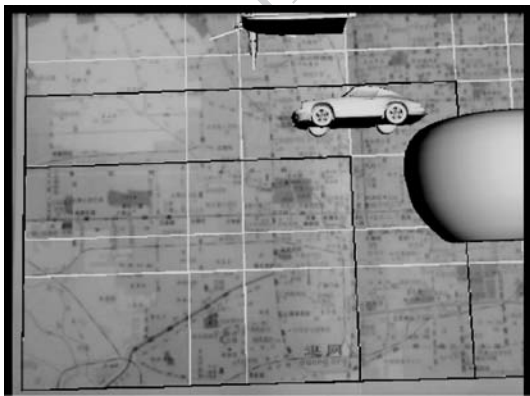


图3 跟踪过程中选取关键帧

Fig. 3 Keyframe selection in tracking stage

2 系统集成

为了验证算法性能,对上述算法进行了集成,实现了2维图像基于自然特征点的增强现实。

系统总体分为预处理阶段和实时跟踪阶段。预处理阶段与之前的模板训练方法相同。预处理结束后,模板图像包含若干尺度图像,每一尺度图像内包含若干关键帧,每一个关键帧内包含若干特征点,形成了将特征点按尺度和区域管理的数据结构,将之保存下来供后续的实时跟踪阶段使用。实时跟踪阶段实现了一个两分支的实时跟踪算法。两个分支分别为特征匹配分支和光流分支,如图4所示。系统会自动根据当前的状态选择合适的分支进行跟踪。

当系统启动或者光流失败的情况下,进行特征匹配分支。本文中当使用光流算法跟踪的帧数达到一个阈值或者当前帧中可以跟踪的点数过少的时候,判定光流跟踪失败。特征匹配分支使用了本文提出的加速算法。在获得一帧摄像机图像后,检测并提取特征点,使用1.2节中的算法选取候选关键帧序列进行特征匹配。匹配成功后求解单应性矩阵进而求解出当前图像帧的摄像机位姿。最后进行虚实融合。

当已有前一帧的特征点集的情况下,执行光流跟踪分支。光流算法并不对当前图像进行特征检

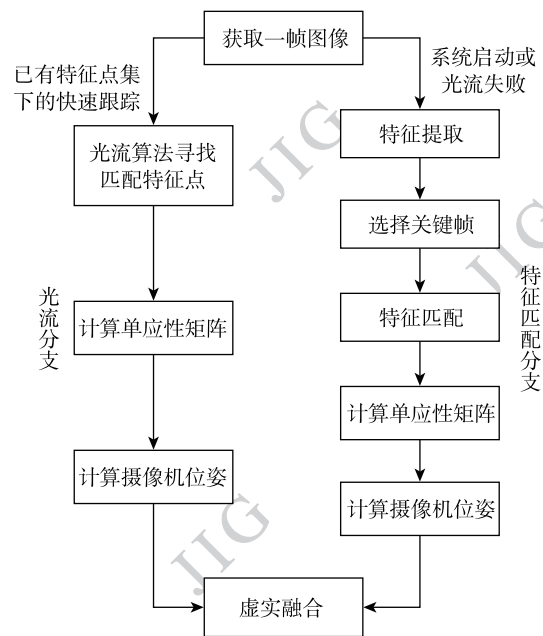


图4 系统流程

Fig. 4 System flow

测,而是利用上一帧图像中的特征点位置,在当前帧中的邻近区域内寻找相同的特征点。从而加快了特征检测的速度,并且找到特征点后直接就得到了与上一帧特征点的匹配。从而可以直接进行摄像机位姿的计算。通过使用光流算法,提升系统的总体效率,并消除了虚拟模型的抖动。

通过这样两分支的设计,实现了系统的鲁棒性和效率的均衡。

3 实验结果与分析

在Android平台上实现了算法和系统并在Android设备上运行了系统。运行过程中为了能够减小每一帧图像的计算量,取得较好的跟踪效率,在初始化阶段要将摄像机分辨率设置为一个较低的值。设定摄像机分辨率后,预处理阶段可以使用这个分辨率求解模板图像的分层数。而在3维跟踪成功后,需要将摄像机图像显示在设备屏幕上。由于目前大部分设备的屏幕分辨率都较高,因此在绘制阶段,需要自动检测设备的屏幕分辨率,根据屏幕分辨率,在保持图像长宽比不变的前提下,尽可能放大摄像机图像至尽可能充满整个屏幕。放大倍数为

$$f = \min\left(\frac{w_w}{w_c}, \frac{h_w}{h_c}\right) \quad (8)$$

式中, w_w 、 h_w 分别为屏幕的长和宽方向的分辨率, w_c 、 h_c 分别为摄像机图像的长和宽方向的分辨率。

对系统进行测试,测试平台是最新的一加 (OnePlus) 手机。一加手机采用高通骁龙 801 主频 2.5 GHz 的 4 核 CPU, 3 G 的 LPDDR3 内存, 16 G 的 eMMC5.0 高速闪存, Adreno 330 的 GPU, 1300 W 高清摄像头和分辨率为 $1\,920 \times 1\,080$ 像素的屏幕。

实验中将设备的摄像机分辨率设置为 640×480 像素。模板图像的分辨率为 $2\,077 \times 1\,379$ 。预处理阶段,模板图像被分为 5 个尺度,共生成了 34 帧关键帧。为了在系统鲁棒性和效率之间取得一个平衡,对获得的每一帧屏幕图像和每一帧关键帧保留 200 个特征点。选取 200 个特征点是基于效率和稳定性的考虑。特征点数目越少,匹配效率越高,但是特征点数目小于 200 个以后,系统匹配失败的概率会增大。主要进行以下几个方面的实验:

1) 消耗时间对比。对本文中使用了关键帧的 FLISA 算法和不使用关键帧的传统 3 维跟踪算法分别进行实验,统计实验过程中特征检测、特征提取、选取关键帧、特征点匹配以及消除误匹配这几个主要环节消耗的时间,以及整个流程消耗的总时间,并进行对比。对比结果如表 1 所示。

表 1 3 维跟踪算法使用关键帧和不使用关键帧的时间对比
Table 1 Time comparison of 3D tracking with keyframe and without keyframe

	特征 检测	特征 提取	关键帧 匹配	特征点 匹配	消除误 匹配	总时间 /ms
传统算法	17.59	6.49	N/A	51.29	45.56	120.93
FLISA	17.81	7.17	0.55	4.84	35.94	66.31
提升倍数	N/A	N/A	N/A	10.59	1.27	1.82

从表 1 可以看出,系统的计算效率在特征点匹配环节有显著提高,因为通过选取关键帧,算法需要匹配的特征点少了 10 倍左右,所以算法在这个环节的效率也提升了 10 倍左右。另一方面是因为将特征点限制在一个较小的范围内匹配,减少了误匹配,所以消除误匹配算法的效率也有了一定的提升。另一方面,系统提升主要集中在特征匹配阶段,所以模板图像越大,图像金字塔的层数越多,关键帧选取算法带来的性能提升也就越明显。

2) 帧率对比。对本文中使用了关键帧的 FLISA

算法和不使用关键帧的传统 3 维跟踪算法分别进行实验,测试两种情况下的算法运行帧率并进行比较。实验过程中的帧率如图 5 所示。

从图 5 可以看到,不使用关键帧的情况下,算法的运行帧率保持在 5 ~ 6 帧/s 左右,加入了关键帧选取功能后,算法的运行帧率保持在 10 帧左右。因为系统从摄像机获取图像需要一定的时间,经测算大概需要 30 ms 左右,所以这里测出的每一帧图像耗时要比表 1 中的耗时多一些。通过测试两种情况下的帧率,验证了关键帧选取算法对系统效率的提升作用是明显的。

3) 增强现实系统运行的帧率统计。对系统进行集成,加入了光流算法后,测试了集成系统的运行帧率。在本实验的光流算法分支中,当跟踪特征点少于 10 个或者光流跟踪的帧数已经超过 50 帧,则判定光流跟踪失败。系统运行过程中的帧率如图 6 所示。

从图 6 可以看到,在加入光流算法之后的系统运行过程中,系统大部分时间保持在 15 帧/s 左右。

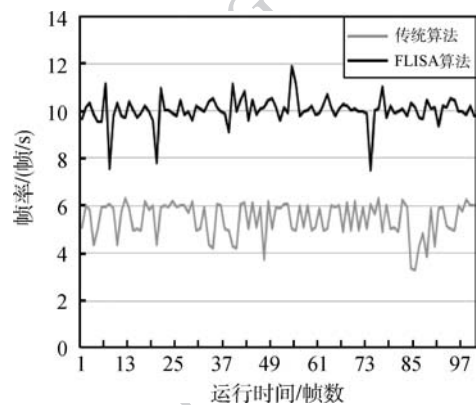


图 5 使用关键帧和不使用关键帧的帧率比较

Fig. 5 Frame rate of method with keyframe and without keyframe

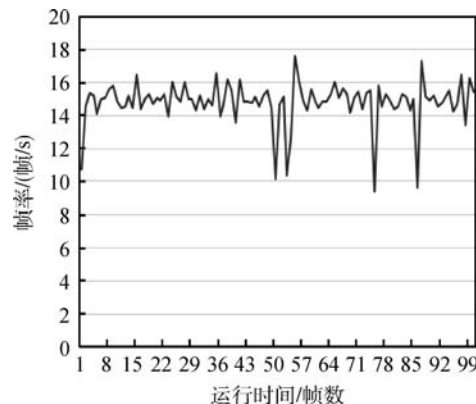


图 6 集成系统的运行帧率

Fig. 6 Frame rate of integrated system

有时光流跟踪失败需要使用特征匹配分支进行恢复。这时运行帧率会短暂下降到 10 帧/s 左右。

4) 系统运行过程中的误差统计。为了测试系统的稳定性,对系统运行过程中的误差进行统计。为了统计系统误差,在每一帧图像计算出对应的摄像机位姿之后,使用这个位姿,将模板图像中匹配成功的那些特征点投影到当前屏幕上,计算这些投影点与当前图像上检测出的对应特征点之间的距离。对这些距离求平均,得到这一帧图像的误差值。系统运行过程中的误差如图 7 所示。

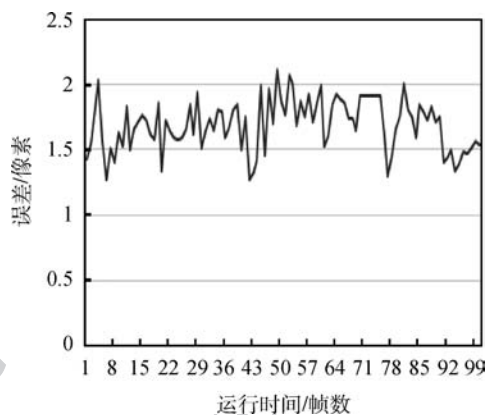


图 7 系统运行误差

Fig. 7 Error of integrated system

从图 7 可以看到,在系统平稳运行过程中,误差基本保持在 1.5 ~ 2 像素之间,证明系统拥有比较好的稳定性。

需要说明的是,由于光流算法在运行过程中会出现特征点漂移和丢失现象,在摄像机快速移动的情况下这种情况更为明显,所以快速移动下,光流算法容易失效,集成系统的效率会因此受到影响;另一方面,摄像机快速移动时,特征匹配分支也有可能失败,这时系统会使用传统的 2 维图像跟踪算法进行错误修正,此时系统的帧率会进一步降低。

4 结 论

提出了一种对模板图像构造关键帧,引入计算几何方法快速定位关键帧从而实现快速匹配特征点的 3 维跟踪算法。在预处理阶段,对模板图像进行分层分区处理,将模板图像中的特征点按照尺度和区域有序管理。在实时跟踪阶段,将计算几何方面的算法融入到实时跟踪算法中,计算关键帧与当

前屏幕的重合程度,实现了关键帧序列的快速选取,快速缩小了特征匹配范围。极大地提升了后续特征匹配环节的效率,从而显著提升了整个 3 维跟踪流程的效率。通过实验结果验证,本文算法可以在移动设备上实现对 2 维自然图像的实时跟踪。即使在模板图像分辨率比较高的情况下,仍然可以在基本不降低跟踪效率的前提下实现具有尺度不变性的实时跟踪。由于目前在特征匹配环节使用的仍然是简单暴力匹配,所以图像特征点的数量会显著影响特征匹配环节的效率。如何用更好的数据结构(如树结构)组织特征点来进一步加快特征匹配的效率,将会是下一步的研究重点。

志 谢

感谢北京理工大学光电学院陈靖专家给予的有益讨论和帮助!

参考文献 (References)

- [1] Lin Q, Yang K, Wang Y T, et al. Key issues study for mobile augmented reality system[J]. Journal of Image and Graphics, 2009, 14(3): 560-564. [林惊, 杨珂, 王涌天, 等: 移动增强现实系统的关键技术研究[J]. 中国图象图形学报, 2009, 14(3): 560-564.] [DOI: 10.11834/jig.20090328]
- [2] Song C Y, He H W, Chen H E, et al. Implementation of markerless augmented reality registration algorithm based on android[J]. Computer Simulation, 2014, 31(8): 432-437 [宋春雨, 何汉武, 陈和恩, 等. Android 的无标识增强现实注册算法实现[J]. 计算机仿真学报, 2014, 31(8): 432-437.] [DOI: 10.3969/j.issn.1006-9348.2014.08.094]
- [3] 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): 355-368. [DOI: 10.1109/TVCG.2009.99]
- [4] Lowe D G. Distinctive image features from scale-invariant keypoints[J]. International Journal of Computer Vision, 2004, 60(2): 91-110. [DOI: 10.1023/B:VISI.0000029664.99615.94]
- [5] Özuysal M, Fua P, Lepetit V. Fast keypoint recognition in ten lines of code[C]// Proceedigs of IEEE Conference on Computer Vision and Pattern Recognition. Minneapolis: IEEE Xplore, 2007: 1-8 [DOI:10.1109/CVPR.2007.383123]
- [6] Yang X, Cheng K T. LDB: an ultra-fast feature for scalable Augmented Reality on mobile devices[C]//IEEE International Symposium on Mixed and Augmented Reality. Atlanta: IEEE

- Xplore, 2012;49-57. [DOI: 10.1109/ISMAR.2012.6402537]
- [7] Calonder M, Lepetit V, Strecha C, et al. BRIEF: binary robust independent elementary features[C]//Proceedings of the European Conference on Computer Vision. Berlin Heidelberg: Springer 2010;778-792. [DOI: 10.1007/978-3-642-15561-1_56]
- [8] Ufkes A, Fiala M. A markerless augmented reality system for mobile Devices [C]//Proceedings of International Conference on Computer and Robot Vision. Regina: IEEE Xplore, 2013;226-233. [DOI: 10.1109/CRV.2013.51]
- [9] Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alternative to SIFT or SURF[C]//Proceedings of IEEE International Conference on Computer Vision. Barcelona: IEEE Xplore, 2011; 2564-2571 [DOI:10.1109/ICCV.2011.6126544]
- [10] Muja M, Lowe D G. Fast matching of binary features[C]//Proceedings of the 9th Conference on Computer and Robot Vision. Toronto: IEEE Xplore, 2012; 404-410. [DOI: 10.1109/CRV.2012.60]
- [11] Lucas B D, Kanade T. An iterative image registration technique with an application to stereo vision[C] //Proceedings of Imaging Understanding Workshop. Vancouver: British Columbia, 1981; 674-679.
- [12] Leutenegger S, Chli M, Siegwart R. BRISK: binary robust invariant scalable keypoints[C] //Proceedings of IEEE International Conference on Computer Vision. Barcelona: IEEE Xplore, 2011;2548-2555. [DOI: 10.1109/ICCV.2011.6126542]
- [13] Rosten E, Porter R, Drummond T, Faster and better: a machine learning approach to corner detection[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(1): 105-119. [DOI: 10.1109/TPAMI.2008.275]