

# 基于 SURF 的序列图像快速拼接方法

杨云涛,冯莹,曹毓,陈运锦

(国防科学技术大学 光电科学与工程学院,湖南 长沙 410073)

**摘要:**针对车载摄像平台的序列图像拼接问题,提出一种基于 SURF 的序列图像快速拼接方法。首先,对经预处理后的序列图像提取 SURF 特征点,采用最近邻法匹配特征点;其次,利用 RANSAC 算法估计相邻帧的变换关系;再采用改进的全局配准策略,完成图像配准;最后,采用一种最大值融合法,去除了拼接缝和光照明暗变化的影响,实现序列图像的无缝拼接。实验结果表明该方法适用于车载摄像平台序列图像的快速、无缝拼接,具有较强的鲁棒性,有一定的实用价值。

**关键词:**序列图像拼接;SURF;RANSAC;最大值融合法

**中图分类号:**TP391.4

**文献标识码:**A

**文章编号:**1673-629X(2011)03-0006-04

## Fast Method for Image Sequences Mosaic Based on SURF

YANG Yun-tao, FENG Ying, CAO Yu, CHEN Yun-jin

(College of Optoelectronic Science and Engineering, National University of  
Defense Technology, Changsha 410073, China)

**Abstract:** A fast method based on SURF is proposed for image sequences mosaic of vehicle camera platform. First, SURF features are extracted from preprocessing images and matched by using the nearest neighbor method. Second, RANSAC is used to solve for the homography between adjacent frames. Then, a global model is found with improved image registration method. Finally, seamless image sequences mosaic is obtained with image fusion of maximum, which is against seams and illumination variations. The experimental results show that the method is fast for seamless image sequences mosaic of vehicle camera platform, is robust and has certain practical value.

**Key words:** image sequences mosaic; speeded up robust features; random sample consensus; image fusion of maximum

## 0 引言

图像拼接技术是将具有重叠区域的一系列图像拼接为一幅具有较宽视角图像的技术,这种技术克服了一般成像设备视场受限和广角镜头失真严重且价格较贵的弊端,在不改变硬件条件的前提下极大地拓展了人们的视野。目前已广泛应用于虚拟现实、遥感图像处理、机器人视觉、医学图像处理等领域。

图像配准是图像拼接的关键。图像配准方法可分为基于灰度、基于变换域和基于特征三类。基于灰度的配准方法直接利用图像像素的灰度信息实现配准,但计算量大,对光照变化和几何畸变敏感。基于变换域的配准方法利用傅里叶变换将图像变换到频域,依据傅里叶变换的特性找到图像间对应关系再配准,但当重叠比例过小时,会造成平移量估计错误。基于特

征的配准方法对图像提取出几何特征,根据特征的匹配关系对图像进行配准,按特征类型不同又可分为基于特征点(如角点、极值点等)、特征线(如直线、边缘、轮廓等)、特征面(如湖泊或建筑物等有区域特征的结构)三种类型。基于特征线、特征面的配准方法对特征线、特征面的提取本身有难度,且会出现特征线断裂或特征面被遮盖导致提取失败的情况。相比之下,基于特征点的配准方法更为简便、稳健,是目前国内外研究的热点,Lowe<sup>[1]</sup>提出了具有对图像缩放、旋转、仿射变换、光照不变性且对噪声不敏感的 SIFT (Scale Invariant Feature Transform) 算法。蔡丽欢等人<sup>[2]</sup>对图像拼接算法及其关键技术进行了研究,并实现了基于 SIFT 的图像拼接,国内其他学者将 SIFT 算法成功应用于不同场合的图像拼接,如李云霞等人<sup>[3]</sup>提出了一种基于 SIFT 特征匹配的图像拼接算法,高超等人<sup>[4]</sup>实现了一种基于 SIFT 特征的航拍图像序列自动拼接方法,张朝伟等人<sup>[5]</sup>将 SIFT 特征匹配用于监控图像自动拼接,这些都充分表明了 SIFT 算法的鲁棒性。但 SIFT 算法的计算量大,速度慢,直接影响了拼接时间,Bay 等人<sup>[6]</sup>提出基于积分图像的 SURF (Speeded Up

收稿日期:2010-08-01;修回日期:2010-11-05

基金项目:中国人民解放军总后勤部资助项目(编号略)

作者简介:杨云涛(1986-),男,河北唐山人,硕士生,研究方向为图像处理;冯莹,教授,博士生导师,研究方向为光纤光源、光学生物传感及光电检测系统等。

Robust Features)算法,通过基于积分图像的方形滤波器来近似拉普拉斯高斯算子,构造了一种 Fast-Hessian 矩阵,提高了检测速度。张锐娟等人<sup>[7]</sup>研究表明在特征点检测速度上 SURF 算法要快于 SIFT 算法,在性能上 SURF 算法与 SIFT 算法相当。

将图像拼接技术应用于车载摄像平台不同于相机绕固定三角架垂直旋转的架设方式,也区别于航拍时相机平行于地面的架设方式。在车载摄像平台系统中,相机以一定的高度和姿态角固定在车辆上,相机不垂直于平台也不平行于地面。由于车辆行驶过程中颠簸和转弯的影响会导致拍摄视角不一致,另外拍摄到的地面序列图像数目多、存在畸变而且特征不明显,这些因素都加大了图像拼接的难度。

文中提出一种基于 SURF 的序列图像快速拼接方法,将其应用于车载摄像平台的序列图像拼接中,实现了序列图像的快速、无缝拼接,并通过实验对方法进行了验证。

## 1 思路框架

图像拼接技术主要分为三个步骤:图像预处理、图像配准和图像融合。图像预处理是指对图像进行几何畸变校正、抑制噪声或图像增强等操作,为图像拼接做准备;图像配准就是根据相邻帧图像重叠区域的对应关系,确定两帧图像间的变换关系,将待配准图像变换到基准图像空间,实现空间上的对齐;图像融合是对配准后的图像进行平滑处理,实现重叠区域的自然过渡。

文中提出基于 SURF 的序列图像快速拼接方法整体思路如下:

- (1) 图像预处理,消除图像镜头畸变和透视畸变的影响;
- (2) 对全部序列图像提取 SURF 特征点并匹配相邻帧图像的 SURF 特征点;
- (3) 用 RANSAC<sup>[8]</sup> (RANDOM SAMPLE CONSENSUS) 算法去除匹配 SURF 特征点对中的外点,利用最大内点集估计出变换矩阵;
- (4) 求解全局配准模型,将每帧图像变换到基准图像坐标系下;
- (5) 图像融合,实现序列图像无缝拼接。

## 2 图像预处理

相机以一定的高度和角度固定在车载平台上,除相机本身的镜头畸变外,由于车载平台的姿态在行驶中存在变化,因而拍摄的地面图像存在严重的透视畸变。采集到的原始图像不能直接用于拼接,需要对图像进行预处理。首先,对原始序列图像进行镜头畸变校正;其次,在实验中,车载平台上架设了姿态角传感

器,可实时感知平台的姿态,参照文献[9]中提出的基于序列图像自测速的精确模型,根据车辆姿态参数和相机架设参数,计算相机在地面坐标系中的旋转矩阵,将透视图变换为俯视图。经预处理后的序列图像作为图像拼接的输入图像,如图 1 所示。



(a)原始图像 (b)俯视图

图 1 图像预处理

## 3 SURF 算法

SURF 算子与 SIFT 算子一样采用 Hessian 矩阵来提取特征点,基于尺度空间的 Hessian 矩阵定义为:

$$H(x) = \begin{bmatrix} L_{xx}(x) & L_{xy}(x) \\ L_{xy}(x) & L_{yy}(x) \end{bmatrix} \quad (1)$$

其中  $L_{xx}$ 、 $L_{yy}$  和  $L_{xy}$  称为拉普拉斯高斯算子,

$L_{xx}(x)$  是高斯函数二阶导数  $\frac{\partial^2 g(\sigma)}{\partial x^2}$  与图像在  $(x, y)$  点处的卷积,  $L_{yy}$  和  $L_{xy}$  的含义类似,  $x = (x, y, \sigma)$ ,  $\sigma$  为尺度因子。

Bay<sup>[10]</sup>采用基于积分图像的方形滤波器来近似拉普拉斯高斯算子,构造了一种 Fast-Hessian 矩阵:

$$H_{\text{approx}} = \begin{bmatrix} D_{xx}(x) & 0.9D_{xy}(x) \\ 0.9D_{xy}(x) & D_{yy}(x) \end{bmatrix} \\ = D_{xx}(x) D_{yy}(x) - (0.9D_{xy}(x))^2 \quad (2)$$

Fast-Hessian 矩阵与 SIFT 算法采用的 Hessian 矩阵相比,以损失部分精度的代价来提升速度。通过 Fast-Hessian 矩阵的值和本征值来判别是否为极值点,若 Fast-Hessian 矩阵的值为正,且两个本征值不同时为正或为负,则认定为极值点。

为了保证特征点具有尺度不变性,需在尺度空间检测 SURF 特征点。尺度空间通常是指构造一系列图像金字塔。传统的方法需将输入图像与高斯核卷积,并降采样,再与高斯核卷积并降采样,反复迭代来实现,降采样就是对图像进行缩放,计算量很大。SURF<sup>[6]</sup>算法采用基于积分图像的方形滤波器构造尺度空间,不同于 SIFT<sup>[11]</sup>算法将输入图像及其降采样图像与高斯核卷积,不对图像进行缩放,而是改变方形滤波器的尺寸,节省了时间。

为了得到稳定且准确的特征点作为候选点,需要对提取的特征点进行精炼:

- (1) 设定阈值: 去除阈值以下的特征点;
- (2) 非极大值抑制: 将每个像素在尺度空间与相邻 26 个像素比较, 得到空间极值点;
- (3) 通过采用拟和三维二次函数<sup>[12]</sup>, 在邻域内插值得到亚像素精度的特征点位置。

SURF 描述子用来描述每个经 Fast-Hessian 检测的特征点在尺度空间邻域内的像素值分布。方法与 SIFT 类似, 但利用基于积分图像的 Haar 小波滤波器在  $x, y$  方向的响应。如图 2 所示: 左为  $x$  方向, 右为  $y$  方向, 黑色区域的权重为 -1, 白色区域的权重为 1。

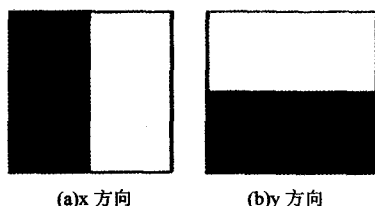


图 2 Haar 小波滤波器

为使描述子具有旋转不变性, 构造以特征点为圆心, 圆心角为  $\pi/3$ 、半径为  $6\sigma$  ( $\sigma$  为特征点所在的尺度因子) 的扇形区。求出区域内各点的 Haar 小波在  $x, y$  方向的响应。赋予响应值以不同的高斯权重系数, 靠近特征点的响应权重重大, 远离特征点的响应权重小, 对区域内的响应进行矢量求和形成新的矢量, 滑动扇形区遍历整个圆域, 找出最大矢量的方向为 SURF 描述子的方向。

构造以特征点为中心、边长为  $20\sigma$  的正方形窗口, 窗口方向为描述子主方向。将窗口划分为  $4 \times 4$  的子区域, 计算子区域内 25 个均匀采样点的 Haar 小波变换响应, 如图 3 所示,  $x$  方向和  $y$  方向的响应值分别记为  $dx, dy$ , 每个子区域求出 4 维分量:

$$v_{\text{subregion}} = [\sum dx, \sum dy, \sum |dx|, \sum |dy|] \quad (3)$$

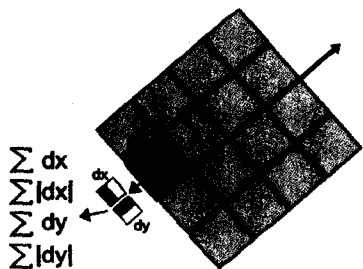


图 3 SURF 描述子

由此, 最终的 SURF 描述子维数为  $4 \times 4 \times 4 = 64$ , 具有旋转、尺度不变性, 再进行归一化后, 具有光照不变性。使得 SURF 算子的性能与 SIFT 算子相当, 能够对图像提取出稳定而丰富的特征点。

SURF 特征匹配采用基于欧氏距离的最近邻法<sup>[11]</sup>, 利用 k-d 树在待配准图像中搜索, 找到与基准

图像中的特征点欧式距离最近的前两个特征点, 如果最近的距离除以次近的距离少于某个比例阈值, 则接受这一对匹配点。降低这个比例阈值, SURF 匹配数会减少, 但更加稳定。

#### 4 去除误匹配和变换矩阵估计

图像经预处理, 消除了镜头畸变和透视畸变的影响, 生成了俯视图, 认为相邻帧图像间仅存在旋转、缩放、平移关系, 选取相似变换模型:

$$H = \begin{bmatrix} s \cos \theta & -s \sin \theta & x \\ s \sin \theta & s \cos \theta & y \\ 0 & 0 & 1 \end{bmatrix} \quad (4)$$

其中  $s$  为缩放因子,  $\theta$  为旋转角度,  $x, y$  为平移量。相似变换矩阵具有 4 个未知量, 理论上选取 2 对不共线特征点可解出。

特征匹配后难免存在误匹配, 应用 RANSAC 算法去除误匹配特征点对, 通过迭代找到满足变换模型的最大内点集并估计出变换矩阵  $H$ 。

以相邻两帧图像为例, 设匹配特征点对数目为  $N$ , 匹配特征点集合分别记为  $P_{1,N}, P_{2,N}$ , 其中  $P_{1,N}$  为基准图像的特征点集合,  $P_{2,N}$  为待匹配图像的特征点集合, 具体过程如下:

(1) 从初始  $N$  对匹配特征点中随机选取 2 对匹配特征点;

(2) 由选取的 2 对匹配特征点计算出基准图像和待匹配图像间的相似变换矩阵  $H_{12}$ , 利用变换矩阵  $H_{12}$  对待匹配图像的特征点集合  $P_{2,N}$  中剩余  $N-2$  个特征点  $P_{2,N-2}$  变换到基准图像坐标系下, 记为  $P'_{2,N-2}$ ;

(3) 计算变换后的特征点  $P'_{2,N-2}$  与特征点  $P_{1,N-2}$  之间的坐标误差;

(4) 从  $N$  对匹配特征点中找出坐标误差在一定误差阈值内的特征点对个数, 记为  $i$ ;

(5) 反复迭代 (1) ~ (4) 步  $n$  次, 找到  $i$  最大的集合为最大内点集即为内点, 其余  $N-i$  为误匹配点即为外点, 利用最大内点集估计出相似变换模型的最小二乘解, 作为当前相邻两帧图像的变换矩阵  $H$ 。

#### 5 全局配准

传统全局配准方法利用相邻帧的变换矩阵, 通常以第一帧为基准图像, 通过变换矩阵连乘, 得到每帧图像与基准图像的全局变换关系。这种方法实现简单, 但矩阵连乘会引入配准误差的累积, 且累积误差会随拼接图像数目的增多而增大, 导致图像拼接整体效果变差, 甚至失败。文中采用一种改进的全局配准策略<sup>[13]</sup>, 可以避免变换矩阵连乘导致的误差累积。利用

改进后的全局变换模型,将每帧图像进行双线性插值,变换到基准图像坐标系下,完成全局配准。

## 6 图像融合

对配准后的图像直接拼接,由于拍摄时的光照条件不同,拼接处会存在明显的拼接缝。为消除拼接缝,需采用一定的图像融合手段对图像进行平滑处理,常见的图像融合方法有直接平均法、加权平均法、多分辨率样条法等。文中采用一种最大值融合法对图像进行融合,认为图像中灰度值高的部分为有效信息,对相邻图像重叠部分,选取最大值作为融合后图像的像素值:

$$I(x, y) = \max(I_1(x, y), I_2(x, y))$$

其中  $I(x, y)$  为融合后图像,  $I_1(x, y)$ 、 $I_2(x, y)$  为配准后图像,且  $(x, y) \in I_1 \cap I_2$ 。

## 7 实验结果及分析

实验程序在 Matlab7.6(2008a)上开发,在配置为 CPU 主频 2.94GHz,内存 512MB 的计算机上运行,对经预处理后两组分别为 62 帧和 84 帧的序列图像进行拼接,图像大小均为  $217 \times 348$ 。其中第一组 62 帧对应车辆直线行驶,第二组 84 帧对应车辆转弯行驶。不考虑图像预处理部分的耗时,通过实验来验证算法的可行性。

图 4 为对 62 帧序列图像的拼接结果,耗时为 57.61 秒;图 5 为对 84 帧序列图像的拼接结果,耗时为 72.94 秒。

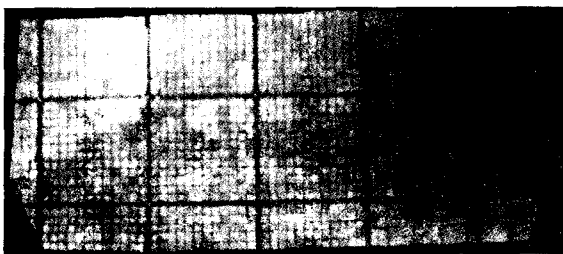


图 4 62 帧序列图像的拼接结果

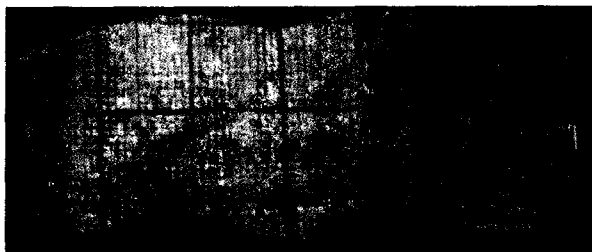


图 5 84 帧序列图像的拼接结果

拼接效果:从由图 4 和图 5 的整体效果来看,采用最大值融合法对图像进行融合,去除了拼接缝和光照明暗变化的影响,融合后的拼接图像没有明显拼缝,地面砖缝清晰可见,但图 4 左侧出现部分模糊;拼接速

度:实现了对两组分别为 62 帧和 84 帧序列图像的快速拼接,每帧图像平均耗时小于 1 秒,拼接速度可通过提高硬件进一步提升;结果分析:图 4 拼接结果部分模糊是由于序列图像间重叠比例过大,导致拼接图像的一个像素对应多个源图像的同位像素,可称为“一对多”的情况。因为姿态角传感器测量误差的存在,会影响变换俯视图的精度,导致配准误差的增大,及图像配准算法本身亚像素级误差的客观存在,再采用最大值融合法会出现该像素值被错位的最大值填充的情况,地面砖缝就会模糊。如果能够从序列图像中按相邻帧的重叠比例提取关键帧,再对关键帧进行拼接,可避免出现“一对多”的情况,提高拼接图像的质量并提升拼接的速度。

## 8 结束语

文中提出一种基于 SURF 的序列图像快速拼接方法,应用于车载摄像平台的序列图像拼接中,通过实验验证了算法的有效性,对车辆直线行驶或转弯行驶均可实现快速拼接,采用最大值融合法的融合效果较好,拼接后图像自然,没有拼缝,实验表明该方法快速、有效,具有一定的应用价值。由于姿态角传感器的测量误差和图像配准算法本身亚像素级误差的存在均会导致图像拼接质量的下降,在硬件上,改用更高精度的姿态角传感器代价大;在算法上,可通过提取关键帧来解决“一对多”的问题。下一步需要实现对序列图像提取关键帧再拼接,来提升算法的性能。

### 参考文献:

- [1] Lowe D G. Object recognition from local scale-invariant features [C]//Proc. of the International Conference on Computer Vision. [s. l.]:[s. n.], 1999:1150-1157.
- [2] 蔡丽欢,廖英豪,郭东辉. 图像拼接方法及其关键技术研究[J]. 计算机技术与发展,2008,18(3):1-4.
- [3] 李云霞,曾毅,钟瑞艳,等. 基于 SIFT 特征匹配的图像拼接算法[J]. 计算机技术与发展,2009,19(1):43-49.
- [4] 高超,张鑫,王云丽,等. 一种基于 SIFT 特征的航拍图像序列自动拼接方法[J]. 计算机应用,2007,27(11):2789-2792.
- [5] 张朝伟,周焰,吴思励,等. 基于 SIFT 特征匹配的监控图像自动拼接[J]. 计算机应用,2008,28(1):191-194.
- [6] Bay H, Tuytelaars T, Van Gool L. SURF: speeded up robust features [C]//Proceedings of the European Conference on Computer Vision. [s. l.]:[s. n.], 2006:404-417.
- [7] 张锐娟,张建奇,杨翠. 基于 SURF 的图像配准方法研究[J]. 红外与激光工程,2009,38(1):160-165.
- [8] Fischler M, Bolles R. Random sample consensus: a paradigm for model fitting with application to image analysis and auto-

些相关技术本身具有某些不适合 Web 服务特性的地方,需要进行相应改进。

#### 参考文献:

- [1] 蔡月茹, 柳西玲. Web Service 基础教程 [M]. 北京:清华大学出版社, 2005:2-5.
- [2] Web Service Description Language (WSDL) 1.1 [EB/OL]. 2001-03. <http://www.w3.org/TR/wsdl/>.
- [3] Simple Object Access Model (SOAP) 1.1 [EB/OL]. 2000-05. <http://www.w3.org/TR/soap/>.
- [4] UDDI Version 3.0.1, [EB/OL]. 2003-10. <http://uddi.org/pubs/uddi-v3.0.1-20031014.htm>.
- [5] Sivashanmugam K, Verma K, Sheth A. Discovery of Web Services in a Federated Registry Environment [C]// Proceedings of IEEE International Conference on Web Services (ICWS'04). California, USA: IEEE Computer Society, 2004:270-278.
- [6] Verma K, Sivashanmugam K, Sheth A, et al. METEOR-S WSDI: A Scalable Infrastructure of Registries for Semantic Publication and Discovery of Web Services [J]. Journal of Information Technology and Management, 2005, 6(1):17-39.
- [7] Kashani F B, Chen C C, Shahabi C. WSPDS: Web Services Peer-to-Peer Discovery Service [C]// Proceedings of the International Conference on Internet Computing. Las Vegas, USA: CSREA Press, 2004:733-743.
- [8] Schmidt C, Parashar M. A Peer-to-Peer Approach to Web Service Discovery [J]. World Wide Web, 2004, 7(2):211-229.
- [9] Web Service Semantics - WSDL-S [EB/OL]. 2005-11. <http://www.w3.org/Submission/WSDL-S/>.
- [10] OWL-S: Semantic Markup for Web Services [EB/OL]. 2004-11. <http://www.w3.org/Submission/OWL-S/>.
- [11] Web Service Modeling Ontology (WSMO) [EB/OL]. 2005-06. <http://www.w3.org/Submission/WSMO/>.
- [12] Wang Y, Stroulia E. Flexible Interface Matching for Web-Service Discovery [C]// Proceedings of Web Information Systems Engineering (WISE'03). Rome, Italy: IEEE Computer Society, 2003:147-156.
- [13] Platzer C, Dustdar S. A vector space search engine for web services [C]// Proceedings of European Conference on Web services (ECOWS'05). [s.l.]:[s.n.], 2005.
- [14] Ma J, Zhang Y, He J. Web Services Discovery based on Latent Semantic Approach [C]// Proceedings of IEEE International Conference on Web Services (ICWS'08). Beijing, China: IEEE Computer Society, 2008:740-747.
- [15] Paolucci M, Kawamura T, Payne T, et al. Semantic Matching of Web Services Capabilities [C]// Proceedings of International Symposium on Wearable Computers (ISWC'02). Sardinia, Italy: Springer, 2002:333-347.
- [16] Paolucci M, Kawamura T, Payne T R, et al. Importing the Semantic Web in UDDI [C]// Proceedings of International Workshop on Web Services, E-Business, and the Semantic Web. Toronto, Canada: CAISE, 2002:225-236.
- [17] Xu B, Chen D. Semantic Web Services Discovery in P2P Environment [C]// Proceedings of International Conference on Parallel Processing Workshops (ICPPW'07). Xi'an, China: IEEE Computer Society, 2007:60-65.
- [18] Bansal S, Vidal J M. Matchmaking of Web Services Based on the DAML-S Service Model [C]// Proceedings of International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS'03). Melbourne, Victoria, Australia: Proceedings. ACM, 2003:926-927.
- [19] Grigori D, Corrales J C, Bouzeghoub M. Behavioral match-making for service retrieval [C]// Proceedings of IEEE International Conference on Web Services (ICWS'06). Chicago, Illinois, USA: IEEE Computer Society, 2006:145-152.
- [20] Mahleko B, Wombacher A, Fankhauser P. A Grammar-Based Index for Matching Business Processes [C]// Proceedings of IEEE International Conference on Web Services (ICWS'05). Orlando, FL, USA: IEEE Computer Society, 2005:21-30.
- [21] Dong X, Halevy A, Madhavan J, et al. Similarity Search for Web Services [C]// Proceedings of the 30th International Conference on Very Large Data Bases. Toronto, Canada: Springer, 2004:372-383.
- [22] Rajagopal S, Selvi S, Rajagopalan M R, et al. Semantic Grid Service Discovery Approach using Clustering of Service Ontologies [C]// Proceedings of IEEE TENCON 2006. Hong Kong: IEEE Computer Society, 2006:1-4.

(上接第 9 页)

- ated cartography [J]. Communications of the ACM, 1981, 24(6):381-395.
- [9] 于起峰, 尚 洋. 摄像测量学原理与应用研究 [M]. 北京: 科学出版社, 2009:197-198.
- [10] Bay H. From wide-baseline point and line correspondences to 3D [D]. Zurich: ETH Zurich, 2006.
- [11] Lowe D G. Distinctive image features from scale-invariant keypoints [J]. International Journal of Computer Vision, 2004, 60(2):91-110.
- [12] Brown M, Lowe D G. Invariant features from interest point groups [C]// British Machine Vision Conference. [s.l.]:[s.n.], 2002:656-665.
- [13] 李庆忠, 耿晓玲, 王 冰. 大视场视频全景图快速生成方法 [J]. 计算机工程, 2009, 35(22):170-172.