

# 图像拼接方法及其关键技术研究

蔡丽欢, 廖英豪, 郭东辉

(厦门大学 电子工程系, 福建 厦门 361005)

**摘要:** 图像拼接是图像处理技术的一个重要内容, 是一种将多张有衔接重叠的图像拼成一张高分辨率全景图像的技术。该技术广泛应用于显微图像分析、数字视频、运动分析、医学图像处理、虚拟现实技术和遥感图像处理等领域。文中阐述了图像拼接技术的一般问题及其难点, 介绍了图像拼接技术发展过程中几种主要的图像拼接方法, 重点分析了近几年提出的角点检测及尺度不变特征转换的方法, 通过编程实现来讨论这些方法的优缺点, 并给出实验的结果。

**关键词:** 图像拼接; 模板匹配; 特征点匹配; 角点检测; 特征转换

中图分类号: TP391.41

文献标识码: A

文章编号: 1673-629X(2008)03-0001-04

## Study on Image Stitching Methods and Its Key Technologies

CAI Li-huan, LIAO Ying-hao, GUO Dong-hui

(Department of Electrical Engineering, Xiamen University, Xiamen 361005, China)

**Abstract:** Image stitching is one of important technologies for image processing. It is normally used to make up a seamless and high resolution image with a set of partial images in the fields of microscopic inspection, digital video, motion analysis, medical image processing, virtual reality and remote sensing mapping. In this paper, the content and the difficulty of image stitching are reviewed and some typical algorithms are introduced. And according to the results of relative experiments, the advantage and limitations of these algorithms are discussed.

**Key words:** image stitching; template matching; keypoints matching; corner detection; feature transform

## 0 引言

在人们实际生活和工作中往往需要获得宽视角、高分辨率的全景图像, 但是由于摄像设备的机械限制, 一般只能得到局部的成像, 而得到全景图像的硬件设备(全景相机、广角镜头等)一般比较昂贵, 不适合普遍应用, 于是人们提出了利用计算机进行图像拼接来获得全景图的方法。目前图像拼接技术广泛应用于数字视频、运动分析、虚拟现实技术、医学图像分析、遥感图像处理等领域。

图像拼接技术原理是根据图像重叠部分将多张衔接的图像拼合成一张高分辨率全景图。这些有重叠部分的图像一般由两种方法获得: 一种是固定照相机的转轴, 然后绕轴旋转所拍摄的照片; 另一种是固定照相机的光心, 水平摇动镜头所拍摄的照片<sup>[1]</sup>。其中, 前者

主要用于远景或遥感图像的获取, 后者主要用于显微图像的获取, 它们共同的特点就是获得有重叠的二维图像。但是由于场景或机械的不理想的缘故, 所获得的各衔接图像的重叠部分不能完全匹配, 这给图像拼接处理带来严重的问题, 因此, 多年来这种匹配不理想的图像拼接问题受到人们极大关注。近年来, 针对不同场景获得的衔接图像, 人们提出了各种相应的技术方法<sup>[2~15]</sup>。文献[16]也把现有的一些典型的图像拼接方法进行了总结, 并将它们归类为基于自适应流形和基于人工确定流形的两种图像拼接方法, 还就局部对齐、全局调整和图像融合等图像处理技术进行了讨论, 内容比较全面, 但未具体介绍其中拼接原理算法, 且某些新的技术方法也未提及。

为此, 文中从二维图像拼接技术的一般问题出发, 把图像拼接算法归类为基于区域和基于特征两种基本的图像匹配分析方法, 具体介绍了相应的匹配拼接算法, 并分析它们在实际应用中的优缺点。

## 1 图像拼接技术的一般问题

图像拼接的关键是精确找出相邻两张图像中重叠

收稿日期: 2007-06-12

基金项目: 福建省自然科学基金资助项目(A0410007)、福建省科技重点项目(2005F0088); 国家人事部留学人员创业基金项目

作者简介: 蔡丽欢(1982-), 女, 福建漳州人, 硕士研究生, 研究方向为数字图像处理、人工智能等; 郭东辉, 教授, 博士生导师, 主要研究方向为集成电路、人工智能。

部分的位置,然后确定两张图像的变换关系,最后进行拼接和拼缝融合。但是由于照相机受环境和硬件等条件影响,所要拼接的图像往往存在平移、旋转、大小、色差及其组合的形变与扭曲等差别,这就给图像拼接带来一定的难度。

实际上,任何的图像形状变化都是由三种基本坐标变化即:平移、旋转和尺度大小变化组合出来的<sup>[17]</sup>。假设  $A$  为原图像的像素坐标, $A'$  为变换后图像的像素坐标,则三种基本坐标变化可以表示为:

- (1) 平移:  $A' = A + t$ ,  $t$  为平移因子;
- (2) 尺度大小:  $A' = sA$ ,  $s$  为尺度因子;
- (3) 旋转:  $A' = rA$

$$r = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \text{ 为旋转因子。}$$

但是同一图像的不同像素点在变化后图像中的相对像素点,各变化因子的数值不一定都一样,因此如何确定两图像间的像素变化因子,找出两图像间的像素匹配点是图像拼接技术的关键。

## 2 基于区域的图像拼接技术

基于区域的图像拼接技术是指利用两张图像像素间灰度的关系来确定图像间坐标变化的参数。该方法充分利用了图像的信息。

### 2.1 基于空间的像素匹配算法

早先的像素匹配算法是用图像在相邻有重叠部分的图像上平移,然后比较两张图像覆盖区域各个像素的匹配程度,进行所有可能的平移之后选择最匹配的一种情况,该算法运算量极大,而且不能解决图像旋转和尺度缩放的问题。基于使用全局搜索的运算量太大,文献[2]提出选取第一幅图像中的重叠部分的一块矩形区域作为模板,在第二幅图像中搜索相似的对应块,计算其相关度,相关度最大的位置为匹配的位置,该算法运算量有所减少,但仍然很大。

为了进一步减少运算量,文献[3]提出了基于列像素的方法,在第一张图像中选取一定间隔的两列,用该两列像素的比值作为模板,在第二张图像中搜索最佳的匹配,从而确定重叠部分的位置。该方法很大程度上减少了运算量,但是鲁棒性不强,仍然不能解决旋转和缩放等问题。

文献[4]提出了基于金字塔数据结构的多分辨率匹配,按先粗后细的顺序寻找匹配位置,先对图像中领域内的像素点的灰度值取平均(见图1),得到分辨率低一级的图像,照此原理可得到一组金字塔式的图像,对两张待拼接的图像都做了该处理之后,从两图像的某一低的分辨率开始进行搜索,直到搜索到一个最优

的位置。该算法搜索速度比较快,可以在一定程度上解决尺度缩放的问题,但是对于旋转角度较大的情况,匹配效果不是很理想。

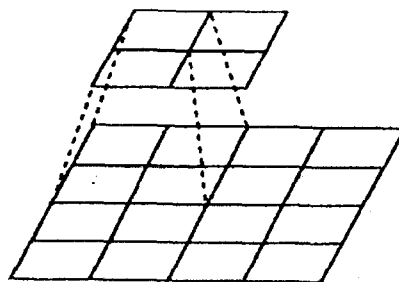


图1 金字塔式数据结构

### 2.2 基于频域的拼接算法

文献[5,6]提出了基于FFT的相关算法,该算法先对两幅待拼接的图像进行二维离散傅里叶变换,假设变换的结果为  $X(u, v)$  和  $Y(u, v)$ ,由相关定理可得相关的离散傅里叶变换:

$$\phi(u, v) = X(u, v) Y^*(u, v) \quad (1)$$

再对  $\phi(u, v)$  进行傅里叶反变换即可得到空间域的相关函数:

$$\phi(j, k) = \sum_{\mu=0}^{M-1} \sum_{v=0}^{M-1} [X(\mu, v) * Y^*(\mu, v)] \omega_M^{j\mu} \omega_M^{vk} \quad (2)$$

由空间域的相关函数即可找出最佳的匹配位置。

文献[7,8]提出以傅里叶变换为基础的相位相关算法,两幅图像的离散傅里叶变换分别为  $X(u, v)$ ,  $Y(u, v)$ ,其互功率谱为  $S(u, v) = X(u, v) Y^*(u, v)$ ,归一化后可得对应的互功率谱的相位谱<sup>[9]</sup>:

$$e^{jQ_{u,v}} \triangleq \frac{X(u, v) Y^*(u, v)}{|X(u, v) Y^*(u, v)|} = e^{j(Q_X - Q_Y)} \quad (3)$$

其中,  $Q_X$  和  $Q_Y$  分别是两幅图像的傅里叶变换的相位。由公式(2)可看出,相位谱是一个位于两图偏移  $(u, v)$  处的  $\delta$  脉冲函数,因此可度量两图之间的相似程度。若用极坐标的方式表示两幅图像,则可用相同的方法计算出图像间的旋转角度。

基于频域的算法可克服相关性噪声和频率噪声,可以大大减小几何失真对匹配性能的影响,计算速度快,对小平移量、旋转及变尺度图像的拼接较适合,但是在两张图像重叠部分不大的情况下结果较差。

## 3 基于特征的图像拼接技术

与基于区域的图像拼接技术不同,基于特征的图像拼接是利用图像的明显的特征来估计图像之间的变换,而不是利用图像全部的信息。

### 3.1 基于轮廓特征的算法

基于轮廓特征的算法是指对图像进行轮廓提取,然后再对提取的轮廓进行匹配,从而确定重叠位置。

文献[9]中轮廓提取分两步进行,首先,图像  $I(x, y)$  与 Laplacian-of-Gaussian(LoG)算子进行卷积,检测出图像像素符号变化的点(过零点)作为边界点,其中 LoG 算子可分解为两个独立滤波算子(拉普拉斯算子和高斯算子),加快运算的速度;第二步,利用卷积后图像在  $x$  和  $y$  两个方向上的梯度( $k_x, k_y$ ),计算每个过零点的边缘强度  $S(x, y)$ :

$$S(x, y) = \begin{cases} \sqrt{k_x^2 + k_y^2} & (x, y) \text{ 为过零点} \\ 0 & \text{其它} \end{cases} \quad (4)$$

然后通过双阈值的方法对过零点进行处理,形成图像的轮廓曲线。提取轮廓之后,利用链码对轮廓进行描述,在匹配阶段,用闭合轮廓的五个形状特征进行匹配:周长,离质心最长和最短的边界点, Hu 变换<sup>[10]</sup>中的一阶和二阶不变矩。

文献[11]在轮廓提取阶段,先对卷积图进行增强,用区域增长的方法进行辅助校正,改善提取效果,用形状签名代替链码,提高了计算的速度,在一定程度上降低了噪声和镜头形变的影响。文献[12]把基于轮廓特征和相位相关方法结合在一起进行拼接。

基于轮廓特征的算法可以改善光照的影响,解决旋转的问题,但是要求提取出比较精确的明显的轮廓特征,而轮廓特征的提取易受噪声的干扰,不适用于轮廓特征不明显或噪声干扰较大的图像。

### 3.2 角点检测算法

该算法是通过在图像中沿圆弧曲线扫描得到的角点的信息(夹角、边缘方向)来进行拼接的。该算法最常见的是 Harris 角点检测算法<sup>[13]</sup>。

具体的检测算法为:

(1) 利用角点检测算子检测图像中的角点,角点检测算子为:

$$E(u, v) = \sum_{x, y} w(x, y) [I(x+u, y+v) - I(x, y)]^2 \quad (5)$$

其中,  $w(x, y)$  为窗函数,  $[I(x+u, y+v) - I(x, y)]^2$  为图像灰度的梯度值。 $w(x, y)$  可为矩形窗或高斯窗。对于每个小的位移量( $u, v$ ),检测算子可用双线性近似:

$$E(u, v) \cong [u, v] M \begin{bmatrix} u \\ v \end{bmatrix} \quad (6)$$

其中

$$M = \sum_{x, y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \quad (7)$$

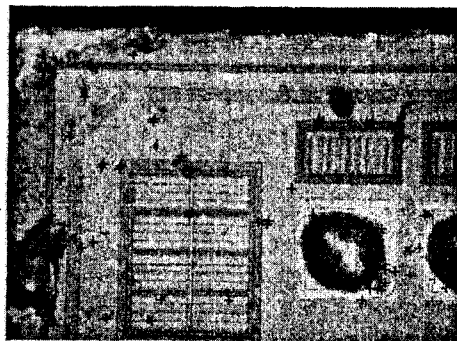
角点检测的响应值  $R$  是算法中角点判定的一个度量值,定义如下:

$$R = \det M - k(\text{trace } M)^2 \quad (8)$$

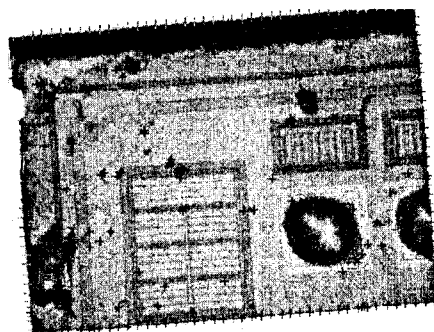
(2) 根据检测出的角点的周围灰度信息进行匹配。

检测出角点后,可利用上文介绍的各种方法对角点进行匹配。

对一张显微镜下拍的芯片图及其旋转缩放后的图像进行角点检测的结果如图2所示。由检测结果可见,该算法对图像间的旋转角度没有限制,配准精度较高,计算量较小。但是对于边缘信息少的图像,效果不理想。



(a) 原始图像检测结果



(b) 旋转缩放后图像的检测结果

图2 角点检测结果

### 3.3 SIFT 尺度不变特征变换算法

SIFT (Scale Invariant and Feature Transform) 算法<sup>[14,15]</sup>是 David G. Lowe 在 1999 年首先提出的用于目标识别的一个方法,首先对两幅图像在尺度和灰度两个空间进行特征检测,利用拉普拉斯金字塔<sup>[16]</sup>的性质消除尺度变化的影响,确定关键点的位置和所处的尺度,然后用关键点邻域梯度的主方向作为该点的方向特征,以实现算子对尺度和方向的无关性,生成关键点的 SIFT 特征向量,利用该向量进行匹配。

具体步骤如下:

(1) 检测灰度和尺度两个空间的极值,初步选定关键点:

图像  $I(x, y)$  在不同尺度下的尺度空间可表示为图像与高斯核的卷积:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (9)$$

其中  $\sigma$  为尺度空间因子,  $G(x, y, \sigma)$  为二维高斯函数。利用 DoG (Difference-of-Gaussian) 算子对两幅图像

在灰度和尺度两个空间进行局部极值检测,DOG 算子即两个不同尺度的高斯核的差分,如下式所示:

$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (10)$$

对于每一个初步选定的关键点,通过具体的模型精确确定关键点的位置和尺度,去除低对比度和边缘响应不稳定的关键点。

(2) 指定关键点领域像素的梯度方向为关键点方向,去除旋转的影响。对于每个  $L(x, y)$ , 关键点的梯度幅值和方向为:

$$m(x, y) = [(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2]^{\frac{1}{2}} \quad (11)$$

$$\theta(x, y) = \arctan((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y))) \quad (12)$$

(3) 生成关键点的特征向量。首先将坐标轴旋转为关键点方向,然后以关键点为中心取  $8 \times 8$  的窗口,计算每个  $4 \times 4$  的小块上 8 个方向的梯度方向直方图,每个梯度方向的累加值形成一个种子点。实际计算中,每个关键点可用 16 个种子点来描述,即关键点的特征向量 128 维。

(4) 对关键点的特征向量进行匹配。一般采用特征向量的欧式距离作为相似性的判定度量。

对两张显微镜下拍的芯片图进行 SIFT 特征点提取并拼接的结果如图 3 所示。该算法有效地解决了图像旋转和缩放的问题,拼接精度高,对视角变化、仿射变换、噪声也保持了一定程度的稳定性,但是误匹配仍然存在。目前该方法广泛应用于图像拼接、图像配准及目标识别等领域。

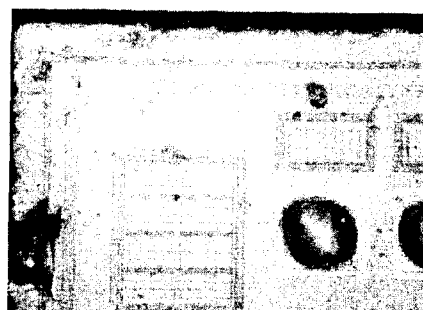
## 4 结束语

文中讨论了主要的几种匹配拼接技术,分析了各自的优缺点。除了对图像进行匹配拼接,全景图的生成还需对图像拼缝进行对齐和融合,一般的融合方法有两种:一种是基于灰度的方法<sup>[18]</sup>,另外一种方法是采用 Laplacian 金字塔对重叠部分进行分级融合<sup>[19]</sup>。

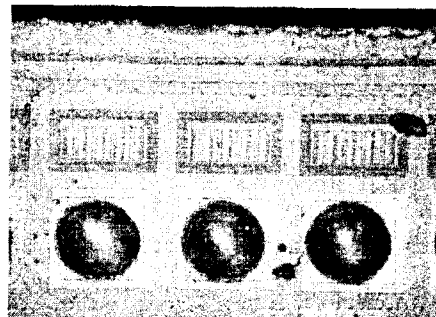
除了对已有的一些经典算法进行改进,该方面的研究人员不断提出一些新的算法,如近年来就有些研究人员从图像的纹理方面<sup>[20]</sup>及研究相机的本身参数<sup>[21]</sup>来实现图像拼接,但由于获取的图像存在的各种几何形变和亮度差异给图像拼接带来的难度,目前尚无一种方法可以实现任意图像序列完全精确的图像自动拼接。

### 参考文献:

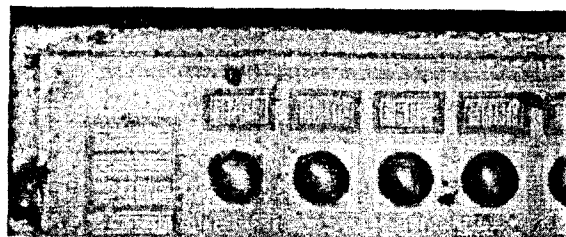
[1] 付金红,宋立新,王武军.图像序列拼接算法的研究[J].哈



(a) 第一张图像提取的特征点



(b) 第二张图像提取的特征点



(c) 拼接结果

图 3 SIFT 算法试验结果

尔滨理工大学学报,2004,9(6):37-40.

- [2] Lucas B D, Kanade T. An iterative image registration technique with an application in stereo vision[C]//In: Seventh International Joint Conference on Artificial Intelligence (IJCAI-81). Vancouver: [s. n.], 1981: 674-679.
- [3] 钟力,胡晓峰.重叠图像拼接算法[J].中国图像图形学报,1998,3(5):367-370.
- [4] 陈果,左洪福.显微图像的一种快速拼接算法[J].仪器仪表学报,2003,24(4):343-347.
- [5] Castro E D, Morandi C. Registration of Translated and Rotated Images Using Finite Fourier Transforms[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI, 1987, 9(5): 700-703.
- [6] Anuta P E. Spatial Registration of Multispectral and Multitemporal Digital Imagery Using fast Fourier Transform Techniques[J]. IEEE Trans Geo. Elec., 1970(8): 353-368.
- [7] Kuglin C D, Hines D C. The Phase Correlation Image Alignment Method[C]//In Proc of IEEE Int Conf. on Cybernetics and Society. New York: [s. n.], 1975: 163-165.
- [8] 许雷,张恒义,虞亚军,等.一种基于相位相关法及数学

(下转第 20 页)

## 4 系统评价

该系统通过把高校中的各个独立的业务子系统改造成 WebService, 从而实现了应用集成。该系统针对高校中存在的应用集成难问题, 充分利用了 WebService 的特点(平台、语言无关性, 低耦合性, 灵活性等), 为高校应用集成提出了一个可行的解决方案。

该系统实现了基本的应用集成, 还有许多深层次的问题需要解决。WebService 的安全性问题, 该系统的安全性是由传输层保证的, 可以改进为采用 WS-Security 体系结构来保证消息层的安全性; WebService 的效率问题, 采用 XML 语言, 可以实现平台、语言无关性, 但是采用 XML 传输的效率问题始终是一个问题。

## 5 结束语

WebService 提出的时候, 人们关注的只是服务本地化资源的使用。随着时间的推移, 人们开始把注意力集中在: 分布式环境中, 如何按需获得服务。这样, 网络计算和 WebService 技术走到了一起, 下一代网络服务的开放标准体系结构 OGSA (Open Grid Service Architecture) 的提出, 就体现了这种发展趋势。这也将成为研究应用集成(无论是企业, 还是高校)的下一个研究方向。

在企业应用集成方面, 面向服务的架构(SOA)已

经被深入研究, IBM 更是提出了 ESB(企业服务总线)的概念。Mule 等 ESB 中间件也开始被人们应用在企业中。在高校应用集成方面还远远落后于企业应用集成, 所以 ESB 等这一应用集成方面的优秀技术也是高校应用集成的下一个研究方向。

## 参考文献:

- [1] Moitra D, Ganeshb J. Web services and flexible business processes: towards the adaptive enterprise[J]. Information & Management, 2005, 42: 921 - 933.
  - [2] Grefen P, Ludwig H, Dan A, et al. An Analysis of Web Services Support For Dynamic Business Process Outsourcing[J]. Information and Software Technology, 2006, 48: 1115 - 1134.
  - [3] Wahli U, Tomlinson M, Zimmermann O, et al. IBM red books: Web Services Wizardry with WebSphere Studio Application Developer [EB/OL]. 2002 - 04. [http://www. redbooks. ibm. com/redbooks. nsf/redbooks/](http://www.redbooks.ibm.com/redbooks.nsf/redbooks/).
  - [4] 翁晓翔, 罗 键. 基于 Web Service 的企业应用集成方案[J]. 自动化技术与应用, 2005, 24(10): 19 - 21.
  - [5] 段文敏, 向长喜, 孙洪淋. 基于 WebService 的高校毕业设计管理信息系统的设计[J]. 企业技术开发, 2006(9): 19 - 21.
  - [6] 苗雪兰. 一种基于角色的授权管理安全模型的研究与实现[J]. 计算机工程, 2002(9): 113 - 115.
- 
- (上接第 4 页)
- 形态学方法的眼底血管图像自动拼接方法[J]. 生物医学工程杂志, 1998, 15(3): 286 - 290.
- [9] Li H, Manjunath B S, Mitra S K. A contour based approach to multisensor image registration[J]. IEEE Trans Image Processing, 1995(4): 320 - 334.
  - [10] Hu M K. Visual Pattern Recognition by Moment Invariants [J]. IEEE Transaction on Information Theory, 1962, 8(1): 179 - 187.
  - [11] 俞 辉, 侯在克, 何旭莉, 等. 一种基于轮廓特征的图像拼接算法设计与实现[J]. 石油大学学报: 自然科学版, 2003, 27(2): 114 - 118.
  - [12] 张 静, 胡志萍, 刘志泰, 等. 基于轮廓相位相关的图像自动拼接[J]. 大连理工大学学报, 2005, 45(1): 68 - 74.
  - [13] Harris C, Stephens M. A combined corner and edge detector [C]//In Proc. Alvey Conf Vision. England: Manchester University Press, 1988: 189 - 192.
  - [14] Brown M, Lowe D G. Invariant Features From Interest Point Groups [C]// In British Machine Vision Conference. Cardiff, Wales: [s. n. ], 2002: 656 - 665.
  - [15] Lowe D G. Distinctive Image Features from Scale - Invariant Keypoints [J]. International Journal of Computer Vision, 2004, 60(2): 91 - 110.
  - [16] Burt P J, Adelson E H. The laplacian pyramid as a compact image code [J]. IEEE Transaction on Communications, 1983, com - 31(4): 532 - 540.
  - [17] 解 凯, 郭恒业, 张田文. 图像 Mosaics 技术综述[J]. 电子学报, 2004, 32(4): 630 - 634.
  - [18] Horn B K P. Determining lightness from an image [J]. Computer Graphics and Image Processing, 1974, 3(1): 277 - 299.
  - [19] Burt P J, Adelson E H. A multiresolution spline with applications to imagemozaics [J]. ACM Transactions on Graphics, 1983, 2(4): 217 - 236.
  - [20] Wang L, Kang S B, Szeliski R, et al. Optimal texture map reconstruction from multiple views [C]//In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR' 2001). Kauai, Hawaii: [s. n. ], 2001: 347 - 354.
  - [21] Duffin K L, Barrett W A. Fast Focal Length Solution in Partial Panoramic Image Stitching [C]//In: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Kauai, Hawaii: [s. n. ], 2001: 690 - 695.