

虚拟现实基于内容优化的图像拼接方法

颜无瑕^{1*}, 朱荣鑫^{1,2}, 崔燕¹

¹南京特殊教育师范学院 江苏南京

²海南大学 海南海口

【摘要】图像拼接是虚拟现实技术中虚拟场景构建的核心技术之一, 而通过图像拼接得到的全景图像中总是存在伪影、重影、畸变等问题。为了快速鲁棒地拼接图像并得到高质量的全景图, 提出了一种基于内容优化的图像拼接方法。通过局部内容相似性保持和线性相似性保持的优化规则可以保证拼接前后图像中的基本内容以及线段的不变性, 从而最大限度减少全景图像中伪影、重影、畸变的产生。该方法是基于粗匹配+精细匹配的框架, 首先根据匹配特征点获取图像单应变换模型实现图像初步对齐, 然后再根据优化规则进行内容优化以减少伪影和畸变的产生得到高质量的全景图像。通过和相似方法的实验对比结果表明, 提出的方法匹配精度高, 拼接图像产生的伪影畸变少, 能快速获取高质量拼接图像。

【关键词】虚拟现实; 图像拼接; 图像对齐; 特征点检测

【基金项目】2019 年江苏省高校自然科学基金项目: 虚拟现实全景拼接技术研究(19KJB520011); 国家青年自然科学基金项目: 面向手语视觉数据的离散哈希学习(62001236)

【收稿日期】2023 年 11 月 6 日 **【出刊日期】**2023 年 12 月 18 日 **【DOI】**10.12208/j.jer.20230033

Image stitching method based on content optimisation in Virtual Reality

Wuxia Yan^{1*}, Rongxin Zhu^{1,2}, Yan Cui¹

¹Nanjing Normal University of Special Education, Nanjing, Jiangsu

²Hainan University, Haikou, Hainan

【Abstract】 Image stitching is one of the core technologies for virtual scene construction in virtual reality technology, and there are always artifacts, ghosting, distortion and other problems in the panoramas obtained by image stitching. In order to quickly and robustly stitch images and obtain high-quality panoramas, an image stitching method based on content optimisation is proposed. The optimisation rules of local content similarity preservation and linear similarity preservation can ensure the invariance of the basic content as well as the line segments in the image before and after stitching, so as to minimise the generation of artifacts, ghosting, and distortion in the panorama. The method is based on the framework of coarse matching + fine matching, which firstly obtains the image homography transform model according to the matching feature points to achieve the initial alignment of the image, and then performs the content optimisation according to the optimisation rules to reduce the generation of artifacts and distortion to obtain a high-quality panorama. The experimental comparison results with similar methods show that the proposed method can quickly obtain high-quality spliced images with less artefacts and distortions.

【Keywords】 Virtual reality; Image stitching; Image alignment; Feature point detection

引言

工业和信息化部网站发表的《虚拟现实与行业应用融合发展行动计划(2022-2026 年)》中指出虚拟现实技术是新一代信息技术的重要前沿, 是数字

经济的重大前沿领域, 将深刻改变人类的生产生活方式^[1]。虚拟现实是利用计算机硬件和软件及各种交互手段, 让参与者能交互式观察和操作虚拟世界, 获取包括视觉、听觉、触觉等全方位多层次的真切

*通讯作者: 颜无瑕

感受。而图像拼接技术作为基于图像的场景绘制关键技术之一, 可以快速、高效的获取全方位全景图像, 满足虚拟现实中场景建模需求。

图像拼接技术是把带有重叠部分的多两幅或多幅窄视域图像通过相邻图像对应关系及其映射建立图像变换模型得到宽广视域 (Wide Field of View) 全景图像。目前图像拼接的主要技术包括基于特征的方法和基于区域的方法。

基于区域的方法是采用局部或整体的相似性测量进行图像匹配, 其缺点是计算成本较大, 对图像光照、缩放、旋转以及失真的变化较为敏感。基于特征的图像拼接方法是基于对显著特征的检测和描述以及特征匹配, 建立待拼接图像间变换模型, 该方法可以较好地处理几何变形, 具有更高的灵活性和鲁棒性。在基于特征的匹配方法中拼接的关键在于提取特征点, 并通过特征点匹配实现图像对齐的过程。为了提高图像对齐的效果, 目前主流的方法是在建立的图像变换模型的基础上进一步采取相应的优化措施, 即粗匹配+精细匹配。其中, 视差容忍图像拼接方法^[2]结合单应性变换和局部扭曲解决视差和局部失真问题。首先采用随机搜索算法利用迭代生长的特征点对得到局部对齐, 在边缘图上设计接缝生成方法, 通过比较不同对齐假设成本值得到粗对齐结果, 然后采用结构保持的网格优化方法实现精对齐。

SEAGULL 方法^[3]为了解决全局对其可能会引入的伪影问题, 使用估计的拼接缝隙引导优化局部对齐的过程, 在每次迭代中改善拼接缝隙的问题。首先采用基于超像素的特征分组方法实现初步对齐, 然后采用结构保持方法对目标图像变形并基于彩色边缘图像估计和优化拼接缝, 实现精对齐。Xi Shen 等^[4]通过特征提取和特征匹配的参数化方法实现粗匹配, 然后在网格化的基础上实现像素级的优化对齐过程。陈殷齐等^[5]为了解决缺少匹配特征点区域的对齐问题, 提出了二步网格优化的图像对齐算法。首先, 通过匹配特征点较多区域得到图像的单应变换模型, 然后将单应变换系数扩散到相邻网格进行优化, 获得更好的对齐效果和拼接结果。

上述方法为解决拼接图像中出现伪影、拼接缝、畸变等问题提出了一种在粗对齐的基础上进一步采取相应措施提高图像对齐精度的方法, 本文综合利

用上述方法的优点, 首先在全局上对所有检测到的特征点做粗匹配建立图像变换模型; 然后采用优化方法对重叠区域进行网格化, 通过优化网格顶点函数提高图像对齐精度减少畸变产生。

1 提出的方法

1.1 建立全局单应性模型

假设 I_1 表示待拼接目标图像, I_2 表示拼接参考图像, 目标是根据 I_1 和 I_2 的重叠区域中匹配的特征点 $p_1 = [x_1, y_1]^T$, $p_2 = [x_2, y_2]^T$ 之间的关系建立图形变换模型:

$$p_2 = h(p_1) \quad (1)$$

$$\begin{cases} h_x(p_1) = \frac{h_1x + h_2y + h_3}{h_7x + h_8y + h_9} \\ h_y(p_1) = \frac{h_4x + h_5y + h_6}{h_7x + h_8y + h_9} \end{cases} \quad (2)$$

在单应坐标系中 $\hat{p}_2 \sim H\hat{p}_1$, 其中 $\hat{p}_1 = [x_1, y_1, 1]^T$, $\hat{p}_2 = [x_2, y_2, 1]^T$, $H \in R^{3 \times 3}$ 是单应变换矩阵。 H 按照列展开表示为 $h_1 = [h_1, h_4, h_7]^T$, $h_2 = [h_2, h_5, h_8]^T$, $h_3 = [h_3, h_6, h_9]^T$ 。

通过对公式两边做向量叉积并展开:

$$O_{3 \times 1} = \hat{p}_2 \times H\hat{p}_1 = \begin{bmatrix} O_{3 \times 1} & -\hat{p}_1^T & y' \hat{p}_1^T \\ \hat{p}_1^T & O_{3 \times 1} & -x \hat{p}_1^T \\ -y \hat{p}_1^T & x' \hat{p}_1^T & O_{3 \times 1} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} \quad (3)$$

令 $h = \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}$, 公式 (3) 中的 3×9 矩阵中只有两行

是线性独立的, 分别用 $a_{i,1}$ 和 $a_{i,2}$ 来表示, 则对于 N 对匹配特征点 $\{\hat{p}_{1,i}\}_{i=1}^N$ 和 $\{\hat{p}_{2,i}\}_{i=1}^N$ 可以通过下列公式估计 h 的值:

$$h = \arg \min_h \sum_{i=1}^N \left\| \begin{bmatrix} a_{i,1} \\ a_{i,2} \end{bmatrix} h \right\|^2 = \arg \min_h \|Ah\|^2 \quad (4)$$

由于单应矩阵有 8 个自由度, 上述公式同时满足约束 $\|h\|^2 = 1$ 。采用移动 DLT 方法^[6], 通过多目标物体增加权重值来估计单应矩阵:

$$h_j = \arg \min_h \sum_{i=1}^N w_{i,j} \left\| \begin{bmatrix} a_{i,1} \\ a_{i,2} \end{bmatrix} h \right\|^2 = \arg \min_h \|W_j Ah\|^2 \quad (5)$$

其中 $W_j = \text{diag}([\omega_{1,j} \quad \omega_{1,j} \dots \omega_{N,j} \quad \omega_{N,j}])$,

$$\omega_{i,j} = \max(\exp(-\frac{\|p_i - p_j\|^2}{\sigma^2}), \gamma), \quad \gamma \in [0,1].$$

1.2 优化方法

通过全局匹配建立的图像转换模型得到的图像拼接结果存在较多问题, 比如重影、变形较大等, 为了改进全景图像拼接质量, 进一步采用内容优化策略对得到的全景图像进行优化。为了保持没有匹配点区域的变换引入相似变换约束规则^[7], 同时减少计算量, 本文对图像重叠区域进行网格化后, 采用该约束测量每个网格从初始形状到变换后位置的偏差, 测量原理如下所示。每个网格可以划分为两个三角形, 局部坐标系下每个三角形中顶点 $V_a(u, v)$ 和其他两个顶点 V_b 和 V_c 的关系如下:

$$V_a = V_b + u(V_c - V_b) + vR_{90}(V_c - V_b) \quad (6)$$

其中 $R_{90} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ 。

为了能够保证图像中每一个单元格都能参与到相似变换中, 在相同的局部坐标系表示下, 需要保证点 V_a 能够由 V_b 和 V_c 依据同样的方法从初始局部坐标 (u, v) 计算得到。因此, 局部内容相似性保持的像素变换约束规则^[4]定义如下:

$$E_1(\hat{V}) = \sum_{i=1}^{N_i} \left\| \hat{V}_1^i - (\hat{V}_2^i + u(\hat{V}_3^i - \hat{V}_2^i) + vR_{90}(\hat{V}_3^i - \hat{V}_2^i)) \right\|^2 \quad (7)$$

其中 N_i 是所有网格划分的三角形的总数。

由于只有局部内容相似性保持约束规则不能够保持图像变换后整体的内容, 因此还需要加入线性保持约束规则, 这里只采用直线段的约束规则^[8], 采用线段检测技术对目标图像 I_2 进行线段检测, 在得

到的线段上均匀采样得到特征点。每个特征点在 1D 坐标系下由线段两个顶点坐标表示出, 从而计算得出 u 值。为了保持变换后目标图像中线段的直线度, 需要建立图像变化模型的特征点在变换后也能够相同的坐标系下采用 u 值来表示。所以线性保持约束规则定义如下:

$$E_2(\hat{V}) = \sum_{i=1}^{N_l} \sum_{j=1}^{N_k} \left\| L_{key}^{i,j} - (L_b^i + u(L_c^i - L_b^i)) \right\|^2 \quad (8)$$

其中 N_l 表示线段的总数, N_k 表示在每个线段 l 上的关键特征点数目, 同时 $L_{key}^{i,j}$ 和线段端点 L_b^i 、 L_c^i 进一步采用线性插值得到优化。

对拼接后全景图像采用方程 (9) 进行优化:

$$E(\hat{V}) = \lambda_1 E_1(\hat{V}) + \lambda_2 E_2(\hat{V}) \quad (9)$$

其中 \hat{V} 表示匹配后的点坐标, E_1 为相似性度量约束, 权重 λ_1 取值 $0.2 \sim 0.5$, E_2 为线性保持约束, 权重值 $\lambda_2 = 1.0$ 。

2 实验与分析

2.1 实验数据来源及实验环境

为了验证本文提出方法的有效性和鲁棒性, 将本文提出的方法和 APAP^[6]以及 SEAGULL^[3]图像拼接方法进行对比实验。为了比较实验效果真实性, 本实验中采用的数据图片均来自原文献。如图 1 所示, 其中 01-08 来自文献[6], 09-16 来自文献[3]。

图 2 是采用图像 09 的实验对比结果, 其中右边一列是对应拼接结果的放大细节图。所有的实验都是在 Thinkpad 笔记本上进行, 配置参数为 Intel core i7 CPU、16G 内存, 64 位操作系统上进行。



图 1 实验中图像数据示意图



图 2 图像 09 的对比实验结果

2.2 实验结果及分析

由于图像拼接方法缺乏原始全景图像真实值 (ground truth) 做对比, 本文主要通过计算位于重叠区域的匹配特征点对的局部相似性来评价拼接后的图像质量。本文需要检测重叠区域匹配像素点对的正确性、准确率, 正负样本分布不均匀, 采用 $F1$ 分数 ($F1\text{-score}$) 值来评估所提方法和相似方法的性能。 $F1$ 分数定义如下:

$$F1 = \frac{2 \times P \times R}{P + R} \quad (10)$$

其中, P (Precision) 是精确率, R (Recall) 是召回率, 分别定义如下:

$$\begin{cases} P = \frac{TP}{TP + FP} \\ R = \frac{TP}{TP + FN} \end{cases} \quad (11)$$

通过对重叠区域匹配特征点对的准确性, 计算

其 TP (True Positives)、 TN (True Negatives)、 FP (False Positives) 和 FN (False Negatives) 值, 代入公式 (11) 求出 P 和 R , 并最终求出每种方法在实验中的 $F1$ 分数值。

在实验中 APAP 和 SEAGULL 方法通过作者给出的意见调节参数达到最好的拼接结果。得到的 P 、 R 和 $F1$ 的结果如表格 1 所示。

从表 1 可以看出, 本文提出的方法在大多数情况下结果都是好的, 其中 APAP 方法在图像对 03、06 和 07 中表现较好, 是因为这几幅图像基本结构相似, 整幅图像可以分为建筑物和地面大两部分, 在 APAP 方法中根据不同区域部分的匹配特征点对进行分组并计算得到最优全局相似变换的结果。SEAGULL 方法对图像对 09、12、14 和 16 中 $F1$ 分值最好, 这是因为该方法是采用超像素分割并采用曲线和线约束来保持结构相似性。本文的方法是在全局变换的基础上进一步对拼接结果进行局部内容相似性保持和线性相似保持的优化结果。

表 1 实验结果对比

	APAP			SEAGULL			Ours		
	P	R	$F1$	P	R	$F1$	P	R	$F1$
01	93.35	92.71	93.03	92.58	93.54	93.06	93.73	94.19	93.96
02	94.38	95.43	94.90	95.43	96.73	96.08	97.47	96.24	96.85
03	97.56	96.83	97.19	95.71	96.65	96.18	97.13	95.37	96.24
04	95.26	93.96	94.61	96.76	97.14	96.95	96.54	98.82	97.67
05	94.18	95.26	94.72	96.43	90.53	93.39	98.17	95.37	96.75
06	96.71	95.78	96.24	96.54	95.78	96.16	92.25	93.63	92.93
07	97.36	97.27	97.31	97.93	96.57	97.25	96.39	97.14	96.76
08	96.25	95.23	95.74	94.35	95.54	94.94	95.73	96.42	96.07
09	96.58	95.78	96.18	96.84	97.16	97.00	97.41	96.46	96.93
10	94.15	94.62	94.38	95.43	96.18	95.80	95.74	97.43	96.58
11	92.55	91.65	92.10	94.76	92.42	93.58	92.54	96.65	94.55
12	95.34	94.24	94.79	96.46	97.31	96.88	97.34	96.24	96.79
13	97.27	85.32	90.90	95.43	95.98	95.70	95.40	96.73	96.06
14	95.23	96.17	95.70	97.32	97.50	97.41	95.87	97.41	96.63
15	87.46	89.43	88.43	90.43	89.54	89.98	90.37	92.53	91.44
16	92.45	94.23	93.33	96.54	97.14	96.84	93.15	92.67	92.91

3 结论

本文提出了一种基于特征点匹配的图像拼接方

法, 在建立图像间变换模型的基础上基于图像内容相似性保持和线性变换保持的规则对拼接结果进行

优化, 得到了较高质量的全景图像, 可以满足绝大多数的图像拼接质量要求。同时, 通过和其他相似拼接方法的对比, 表现出较强的有效性和鲁棒性。

参考文献

- [1] 工业和信息化部、教育部、文化和旅游部国家广播电视总局、国家体育总局关于印发《虚拟现实与行业应用融合发展行动计划（2022-2026 年）》的通知[EB/OL][2023-11-23]https://www.gov.cn/zhengce/zhengceku/2022-11/01/content_5723273.htm.
- [2] Fan Zhang, Feng Liu. Parallax-Tolerant Image Stitching[C].IEEE Conference on Computer Vision and Pattern Recognition. 2014: 3262-3269.
- [3] Kaimo Lin, Nianjuan Jiang, Loong Fah Cheong, et al. SEAGULL : Seam-guided local alignment for parallax-tolerant image stitching[C].European Conference Computer Vision. 2016:70-385
- [4] Xi Shen, Francois Darmon, Alexei A. Efros, et al. RANSAC-Flow: Generic Two-Stage Image Alignment[C]. European Conference Computer Vision. 2020:618-637.
- [5] 陈殷齐,郑慧诚,严志伟等.基于单应性扩散约束的三步网格优化视差图像对齐[J].自动化学报.2022, 48(x): 1-14.
- [6] Julio Zaragoza, Tat Jun Chin, Michael S. Brown, et al. As projective-as-possible image stitching with moving DLT[C]. IEEE Conference on Computer Vision and Pattern Recognition. 2013:2339-2346.
- [7] Feng Liu, Micheal Gleicher, Hailin Jin, et al. Content-preserving warps for 3D video stabilization[J]. ACM Trans. Graph. 2009,44:1-9 .
- [8] Lin K, Jiang N, Liu S. Direct Photometric Alignment by Mesh Deformation[C].IEEE Conference on Computer Vision and Pattern Recognition. 2017:2701-2709.

版权声明: ©2023 作者与开放获取期刊研究中心(OAJRC)所有。本文章按照知识共享署名许可条款发表。

<http://creativecommons.org/licenses/by/4.0/>



OPEN ACCESS