

城市建筑立面三维“线-面”结构快速重建

王 伟¹, 于 磊¹, 任国恒¹, 陈立勇¹, 董秋雷², 胡占义²

(1. 周口师范学院网络工程学院, 河南周口 466001; 2. 中国科学院自动化研究所, 北京 100190)

摘 要: 本文提出一种有效的基于线段匹配与结构先验的城市建筑立面三维“线-面”结构快速重建算法. 本文算法利用在当前图像中检测的初始线段将图像分割为互不重叠的区域, 然后在点-线约束空间内对每个区域的竖直边进行匹配并通过全局多平面拟合的方式获取建筑初始主平面; 在此基础上, 利用结构先验(如共线、共面等)对图像中潜在线段匹配(或空间线段)进行推断与优化, 进而获取每个区域对应的空间平面. 实验结果表明, 本文算法仅利用两幅图像即可重建以“线-面”形式表达的城市建筑立面完整结构, 整体上具有较高的效率与精度.

关键词: 线段匹配; 城市建模; 平面拟合

中图分类号: TP391.41

文献标识码: A

文章编号: 0372-2112(2021)08-1551-10

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.12263/DZXB.20180752

Rapidly Reconstructing 3D Line-Plane Structures of Urban Building Facades

WANG Wei¹, YU Lei¹, REN Guo-heng¹, CHEN Li-yong¹, DONG Qiu-lei², HU Zhan-yi²

(1. School of Network Engineering, Zhoukou Normal University, Zhoukou, Henan 466001, China;

2. Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China)

Abstract: The paper presents an effective method to model building facades based on line segment matching and structure priors. The proposed method splits the image into several non-overlapping regions using the lines detected in the image, and then performs a rapid matching for the vertical edges of each region within a highly restrained point-line space and extracts dominant building planes using a global multi-plane fitting method. Based on the resulting planes, it further infers and optimizes potential line segment matches (or spatial lines) using structure priors (e.g., collinearity and coplanarity) to obtain the plane associated with each region. Experimental results confirm that the proposed method can reconstruct satisfactory line-plane building facade structures from only two images, and has high accuracy and efficiency on the whole.

Key words: line segment matching; urban scene modeling; plane fitting

1 引言

近年来,基于图像的城市建筑三维重建由于代价低廉、操作灵活等优点而倍受关注;然而,在光照变化、透视畸变等因素的影响下,相关算法的精度与效率往往较低,此问题仍需深入研究与解决.

在城市建筑立面三维重建中,采用不同的重建基元或融入不同层次的结构先验往往对重建精度与效率产生不同程度的影响.具体而言,点基元虽有利于表现建筑结构细节,却不易用于重建弱纹理区域的结构;线基元虽有利于表现建筑结构的边界,却由于在图像中

检测线段的稀疏性以及图像间线段匹配可靠性较低等问题而不易重建完整的结构;面基元虽可解决弱纹理区域的重建问题而获取完整的结构,却不易重建精确的结构边界(如图像区域分割精度引起的直线结构重建偏差)或仅获得过于简单的结构模型.在此情况下,有效地融合不同类型的基元及更丰富的结构先验将有利于提高建筑立面重建的精度与效率.然而,在传统相关算法中,以下问题仍未得以有效解决:(1)不同类型的基元仅被单独应用或仅相互作为初值进行考虑而未得以充分融合,不易获得较高的重建精度或完整的结构;(2)更丰富的先验知识(如直线之间以及平面

之间特定的夹角)未能被充分利用,难以保证较高的重建效率.

为了克服以上问题,本文提出一种基于线段匹配与结构先验的城市建筑立面快速重建算法.本文算法在点、线、面等基元统一的框架下,以线段匹配为基础,在线段共面、平面夹角等结构先验的引导下采用逐步求精的方式对建筑立面进行重建,仅利用两幅图像即可快速获取以“线-面”形式表达的完整结构,整体上具有较高的精度与效率.

2 相关工作

传统线段匹配的目的在于利用几何约束、图像特征与共面结构等信息提高图像之间线段匹配的可靠性.在相关工作中,Fan等人^[1]通过与线段相邻特征点的匹配估计两线段之间仿射变换并依此度量两线段之间的匹配相似度,进而利用特征点之间的旋转量滤除不可靠线段匹配以提高整体线段匹配的性能.Kim等人^[2]根据Manhattan模型假设与平行共面直线不变性对在图像中检测到的线段进行分组并确定每个分组对应的平面,有效提高了室内场景弱纹理区域的重建可靠性.为了提高宽基线图像线段匹配的可靠性,Al-Shahri等人^[3]利用极线几何约束与局部结构共面特征对线段匹配进行约束与验证,Verhagen等人^[4]则提出一种尺度不变的宽基线图像线段描述子,Jia等人^[5]首先检测点-线不变性共面区域并确定相应的单应矩阵,然后在单应矩阵的约束下对线段进行匹配.类似地,基于相邻空间线段共面的假设,Li等人^[6]首先对图像中的尺度与仿射不变性的V型结构区域进行检测并确定相应的单应矩阵,然后求取线段匹配对应的空间线段并对其进行全局优化.

在实际中,尽管以上算法可在特定条件下提高线段匹配的可靠性,但对于场景三维重建问题,由于线段数量较少,特别在诸多干扰因素的影响下,所获取的线段匹配(或空间线段)通常较为稀疏而不足以表达场景的完整结构.

在此情况下,场景分段平面重建算法可在一定程度上解决以上问题.在相关工作中,Bodis-Szomoru等人^[7]在初始稀疏空间点与图像过分割获得的超像素的基础上在能量最小化框架下推断场景的完整结构;该算法虽然速度较快,但其为了通过空间点拟合的方式获取每个超像素对应的初始平面而采用了大尺寸超像素,这在实际中往往会导致较大的错误(大尺寸超像素对应的空间点深度变化通常较大,相应的空间面片并不能简单地近似为平面).Verleysen等人^[8]首先利用DAISY特征描述子^[9]进行图像匹配以生成稠密的空间点,然后在此基础上抽取候选平面集以完成能量最小化框架下的场景结构推断.相对而言,由于稠密空间点

通常蕴含场景更丰富的结构信息,由其生成的候选平面集往往具有较高的完备性,因而可有效保证场景结构推断的可靠性.然而,图像稠密匹配较为耗时,致使该算法整体效率较低.

3 问题描述与算法流程

如图1所示^[10],传统以点、线、面等基元进行场景重建的算法通常存在以下问题:(1)点基元在弱纹理区域重建性能较差,在稠密匹配时效率较低;(2)线基元由于数量较少或匹配可靠性较差而仅获取较为稀疏空间线段而难以表达场景完整的结构(如图1(b)、图1(c)所示);(3)面基元不易重建精确的结构边界或仅获取过于简单的模型(如图1(d)矩形标示区所示不精确的图像过分割往往导致场景分段平面重建算法产生较大的误差).

事实上,对于以平面结构为主的城市建筑立面,其中不但包含丰富的直线与平面信息,而且直线之间、平面之间以及直线与平面之间均存在较大的相关性或先验知识:(1)在图像中检测到的直线与建筑结构边界相一致,直线的主方向数量有限且竖直或近似竖直直线(下文简称竖直直线)决定着城市建筑垂直于地面的主要特征;(2)直线与平面之间存在紧密的联系与对应关系(如两平面相交为直线、两相邻平行直线共面等);(3)相邻直线或平面之间存在特定的约束关系(如相邻两平面之间的夹角为特定值).

根据以上分析,提高城市建筑立面重建可靠性的最直接的方法在于:在图像中检测线段并通过竖直线段聚类的方式获得竖直直线,然后利用非竖直线段将相邻竖直直线之间的建筑区域分割为多个两边为竖直线段的区域,进而在以上建筑结构先验的引导下采用线段匹配的方式快速求取每个区域(如图1(g)、图1(h)所示)对应的平面,最终获取建筑立面完整的“线-面”结构.

图1(b)所示为采用文献[11]算法检测的初始线段,通过竖直线段的聚类可获得如图1(e)所示建筑主竖直直线;对于两相邻竖直直线,选择所有与其距离较近的非竖直线段(图1(f)所示)并求取相应的交点,然后根据交点从上到下的顺序将两者位于建筑区域中的部分分割为多个如图1(g)所示的区域与如图1(h)所示的线段.实验中发现,以上方法不但可在一定程度上解决长线段断裂、部分线段检测失败等问题而获得更多线段以用于表达建筑的主体结构,而且相应区域对应的空间面片更可能为平面,因而更有利于通过匹配其两竖直边的方式而重建建筑完整的多平面结构.需要注意的是,在此过程中,本文算法通过语义标注算法^[12]将以上过程约束在建筑区域内完成以提高整体重建过

程的可靠性与效率;此外,虽然以上方法也存在直线将单个平面对应图像区域分割为两部分、天空与建筑以

及地面与建筑的相接边界为非直线情况,此问题并不影响线段的匹配与相应平面的推断.

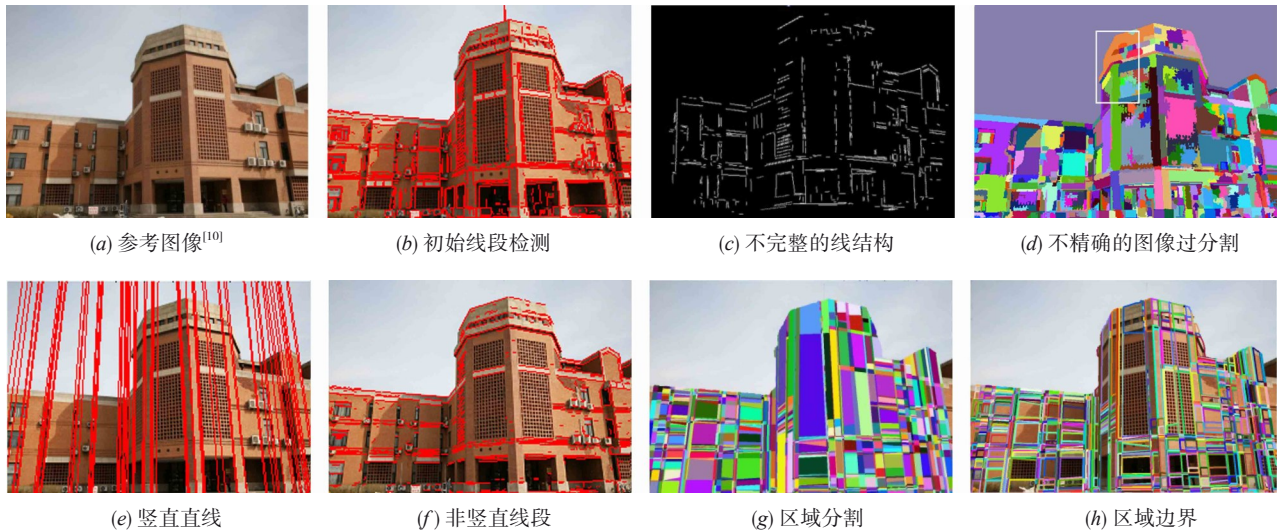


图1 传统算法存在的问题与解决思路

最终,本文采用逐步求精的方式确定图 1(g)、图 1(h)所示区域对应的“线-面”结构,具体流程如图 2 所示.

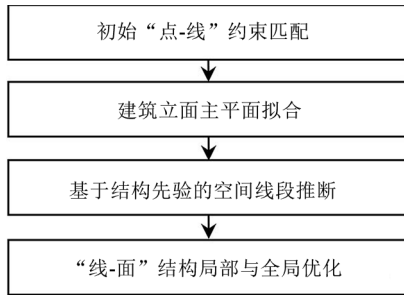


图2 本文算法流程

在本文算法中,线段匹配包括成对图像线段(即在两幅图像中均可检测到实际可被匹配的两线段)的匹配及单个图像线段(即仅在参考图像中检测到的线段)对应空间线段的推断.为方便算法描述,下文将图像线段简称为线段以区别于空间线段.

4 初始“点-线”约束匹配

本文算法以图 1(g)所示区域的两竖直边为对象进行“点-线”约束匹配,进而确定每个区域相应的“线-面”结构;为此,设参考图像 I 与匹配图像 I' 中的竖直直线集分别为 $\mathcal{L}:\{L_i\}(i=1,\dots,k_1)$ 与 $\mathcal{L}':\{L'_i\}(i=1,\dots,k_2)$, 图像 I 中竖直线段集为 $\mathcal{M}:\{l_i\}(i=1,\dots,k_3)$ (即集合 \mathcal{L} 中所有竖直直线 L_i 上的竖直线段的集合).

对于线段 $l_i \in \mathcal{M}$, 本文算法根据 DAISY 特征确定

其在图像 I' 中的匹配直线(相应的匹配线段可根据线段 l_i 中像素的极线与该直线的交点确定). 设线段 l_i 上像素集为 $P:\{p_i\}(i=1,\dots,n)$, 对于像素 $p_i \in P$, 由于其在图像 I' 中的匹配点更可能位于竖直直线上或竖直直线附近,因而,本文算法首先在图像 I' 中求取像素 p_i 的相应极线与集合 \mathcal{L}' 中所有竖直直线的交点,然后分别在其极线上每个交点两侧选择 10 个像素(过多像素可提高可靠性但也增加计算复杂度)的 DAISY 特征与像素 p_i 的 DAISY 特征进行匹配并从中确定最优匹配点 p'_i , 进而确定集合 P 中所有像素对应匹配点的集合 $P':\{p'_i\}(i=1,\dots,n)$. 为了提高整体效率,此过程均在建筑区域内完成.

其实,通过竖直直线对候选匹配点进行约束,以上过程整体效率通常较高,而对于集合 P' 的可靠性,本文通过以下条件进行度量:

$$T(l_i, L'_m) = \frac{1}{n} \sum_{k=1}^n \delta(d(p'_k, L'_m) < k) \quad (1)$$

其中, $L'_m \in \mathcal{L}'$ 为图像 I' 中当前竖直直线, $d(p'_i, L'_m)$ 表示点 p'_i 到直线 L'_m 的距离, 函数 $\delta(\cdot)$ 当条件为真时取值 1, 否则取 0; k 为距离阈值.

事实上, $T(l_i, L'_m)$ 值较高时,如图 3(d)所示,集合 P' 中像素大部分位于直线 L'_m 附近,因而线段 l_i 对应的匹配直线应为 L'_m , 对应的匹配线段 l'_i 可通过求取集合 P 中像素的极线与直线 L'_m 交点的方式确定;否则,线段 l_i 对应的匹配线段则可能未被检测到或被遮挡,此时应推断其对应的空间线段.

实验中发现,由于噪声、透视畸变等因素的影响,

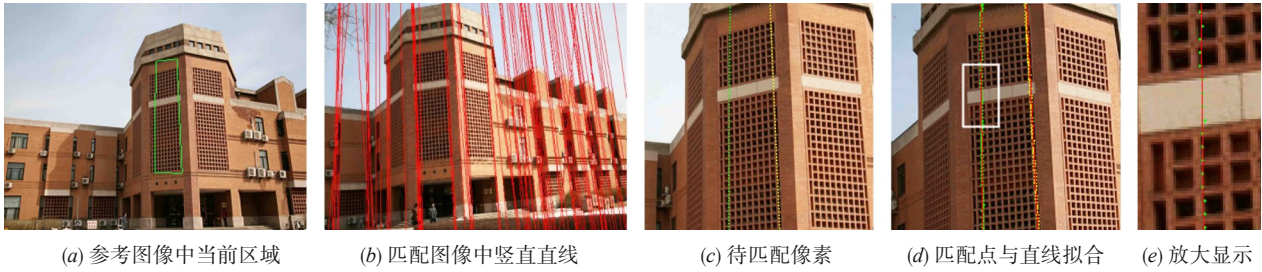


图3 “点-线”约束匹配示例

集合 P' 中不但可能存在少量外点,而且当前直线 L'_m 与真实直线之间可能也存在偏差;为此,如图3(d)所示,本文采用RANSAC(Random Sample Consensus)方法对集合 P' 中的匹配点进行拟合并利用所获取的直线更新直线 L'_m .

在实际中,相对于短线段(如包含少于10个像素),长线段的匹配往往具有更高的可靠性;因而,在“点-线”约束匹配中,本文算法首先匹配长线段,然后在已匹配结果的约束下对短线段进行匹配;其中,对于图像 I' 中直线 L'_m ,参考图像 I 中的短线段 $l_i \in L$ 匹配可靠性度量定义为

$$M(l_i, L'_m) = \max_{L'_m} \left(T(l_i, L'_m) + \mu \sum_{l_j \in N(l_i)} e^{-|s(l_i, l_j) - s(l'_m, l'_j)|} \right) \quad (2)$$

其中, l'_m 为线段 l_i 的位于直线 L'_m 上的候选匹配线段(根

据线段 l_i 上像素的极线与直线 L'_m 的交点确定), $N(l_i)$ 表示与线段 l_i 位于同一直线且已匹配的线段集合,线段 l'_j 为线段 l_j 的匹配线段, $S(x, y)$ 为线段 x 与 y 中点的DAISY特征相似度; μ 为约束项权重.

在以上线段匹配过程完成后,本文算法采用交叉验证方法剔除外点以获取可靠的匹配结果,即对于线段 l_i 在图像 I' 中的匹配线段 l'_i , 如果其通过以上方法在图像 I 中确定的匹配线段 l_i^* 与线段 l_i 之间的平均距离小于指定阈值(实验设置为2),则认为线段 l_i 与 l'_i 为可靠匹配. 在此实验中,如图4(a)所示,集合 \mathcal{M} 中近70%的线段可获得可靠的匹配,其余线段未得以正确匹配的原因在于:(1)由于竖直直线的约束而匹配错误(约5%左右);(2)对应的空间线段在图像 I' 相应视点被遮挡或未在图像 I' 中检测到可被匹配的线段. 为此,本文算法采用建筑结构先验对这些线段对应的空间线段进行推断.



图4 建筑主平面抽取(不同主平面上的线段由不同颜色表示)

5 建筑主平面拟合

在获取可靠线段匹配集 $Z: \{l_i, l'_i\} (i = 1, \dots, k_4)$ 后,本文算法通过三角化方法^[13]生成相应的空间线段集 $\mathcal{F}: \{\ell_i\} (i = 1, \dots, k_4)$ 并采用全局多模型拟合^[14]从中拟合建筑主平面 $\mathcal{H}: \{h_i\} (i = 1, \dots, k_5)$ 以对未匹配线段对应空间线段进行推断,相应的能量函数定义为

$$E(H) = \sum_{\ell_i \in \mathcal{F}} S(h_{\ell_i}) + \lambda \sum_{\ell_k \in N(\ell_i)} S(h_{\ell_i}, h_{\ell_k}) \quad (3)$$

其中, h_{ℓ_i} 表示当前为空间线段 ℓ_i 分配的平面, $N(\ell_i)$ 为

空间线段 ℓ_i 的相邻空间线段集(即线段的左右与上下相邻线段), λ 为正则项权重; $S(h_{\ell_i})$ 表示空间线段 ℓ_i 对应两匹配线段端点的DAISY特征相似性度量均值, $S(h_{\ell_i}, h_{\ell_k})$ 表示两相邻空间线段 ℓ_i 与 ℓ_k 的平滑性约束,即

$$S(h_{\ell_i}, h_{\ell_k}) = \begin{cases} e^{-d(\ell_i, \ell_k)} \cdot e^{-S(\ell_i, \ell_k)}, & h_{\ell_i} \neq h_{\ell_k} \\ 0, & h_{\ell_i} = h_{\ell_k} \end{cases} \quad (4)$$

其中, $d(\ell_i, \ell_k)$ 为相邻空间线段 ℓ_i 与 ℓ_k 中点之间的距离, $S(\ell_i, \ell_k)$ 为空间线段 ℓ_i 与 ℓ_k 在图像 I 中对应线段中

点的 DAISY 特征相似度.

在式(4)中,综合采用空间线段距离与图像线段特征旨在增强空间平面标记之间的平滑性约束;相对于在图像中通过三角网格度量空间点距离的方法^[14](实际相距较远的两空间点或线段在图像中的投影点或线段可能相距较近,采用三角网格构造平滑性约束可能会导致较大错误),往往可获得更好的效果.此外,主平面拟合过程中所用候选平面通过相邻空间线段上的空间点拟合生成,可在一定程度上避免随机采样空间点生成候选平面的复杂性.

如图4(b)、图4(c)所示,通过以上方法获取的建筑主平面整体上表达了建筑的主体结构,有利于后续环节对建筑结构细节进行推断.

6 基于结构先验的空间线段推断

在已知建筑主平面的情况下,本文算法根据平面夹角先验对“点-线”约束匹配阶段未匹配线段对应的空间线段进行推断.

具体而言,如图5所示,若线段 l_2 对应的匹配线段未能在图像 I' 中被检测,则无法采用匹配的方式确定其对应的空间线段 ℓ_2 ;然而,当建筑主平面 h_1 与空间线段 ℓ_1 已知时,则可根据相邻直线共面的假设以空间线段 ℓ_1 所在直线为轴将平面 h_1 的法向量旋转 45° 确定平面 h_2 ,进而通过求取线段 l_2 反投影线与平面 h_2 交线的方式确定空间线段 ℓ_2 .在此过程中,虽然平面 h_2 与 h_1 之间的夹角 45° 事先未知,但往往却在指定范围之内;因而,为从根据平面夹角先验生成的空间平面中确定最优者,本文定义如下度量:

$$T(l_i, H_A) = \min_{H_i} \frac{1}{n} \sum_{p_i=1}^n S(p_i, H_A(p_i)) \quad (5)$$

其中, H_A 为集合 \mathcal{H} 中平面根据夹角先验 A_{prior} (设置为 $[-60^\circ, -45^\circ, -30^\circ, 0^\circ, 30^\circ, 45^\circ, 60^\circ, 90^\circ]$)生成的平面, $H_A(p_i)$ 为像素 p_i 的反投影线与平面 H_A 的交点在图像 I' 中的投影, $S(p_i, H_A(p_i))$ 为像素 p_i 与投影 $H_A(p_i)$ 间的 DAISY 特征相似度.

事实上,当线段 l_i 在图像 I' 视点被遮挡时, $T(l_i, H_A)$ 值通常较大;在本文算法中,当 $T(l_i, H_A) > 2\bar{S}$ (\bar{S} 为所有

已匹配线段对应 DAISY 特征相似度的均值)时,不再对当前线段对应的空间线段进行推断;而当区域的两竖直边均被遮挡时,则不再对该区域对应的空间平面进行推断.

7 “线-面”结构优化

在“点-线”约束线段匹配及空间线段推断之后,图1(g)所示区域对应的平面则可随之确定(如拟合区域两竖直边上空间点).然而,由于在图像中相邻线段之间的区域并不一定与真实空间平面相对应或者两相邻线段之间实际可能包含多个平面,因而,最终获取的建筑多平面结构中往往包含一定的错误.如图5所示,空间线段 ℓ_3 与 ℓ_4 之间的平面 h_4 并非真实空间平面,因而需要进行检测与优化.

为了度量空间平面的可靠性,本文算法对分别利用图像 I 与图像 I' 获取的平面的一致性进行了交叉验证.即对于区域 $S \in I$ 与 $V \in I'$ 对应的平面 h_s 与 h_v ,当以下条件满足时,则认为两者是可靠的且为同一平面.

$$T(h_s, h_v) = \left(\frac{\min(R(S, h_v), R(V, h_s))}{\max(|S|, |V|)} > \varepsilon \right) \quad (6)$$

其中, $|S|$ 为区域 S 内所有像素数,阈值 ε 用于判断平面可靠性; $R(S, h_v)$ 为区域 S 中已获取可靠空间点的像素数,其中的可靠空间点定义为:对于像素 $p \in S$ 及其在平面 h_s 诱导下在图像 I' 中投影点 $h_s(p)$,如果投影点 $h_s(p)$ 在平面 h_v 诱导下在图像 I 中的投影点 $h_{sv}(p)$ 与像素 p 之间距离小于指定阈值(实验设置为2),则认为像素 p 已获取可靠空间点(即其反投影线与平面 h_s 的交点).

7.1 局部优化

对于图像 I 中两竖直边分别为 l_i 与 l_k 的区域 Q_{ik} ,若其对应的平面不可靠,本文采用两种方式对其进行局部优化.

空间平面调整 分别以空间线段 ℓ_i 与 ℓ_k 所在竖直直线为轴,利用基于结构先验的空间线段推断方法(参见第6节)推断线段 l_k 与 l_i 在其他平面对应的空间线段 ℓ_k^* 与 ℓ_i^* ,进而根据空间线段 ℓ_k^* 与 ℓ_i 以及 ℓ_i^* 与 ℓ_k 构建满足式(6)所示条件的最优平面作为区域 Q_{ik} 对应的可靠平面.如图5所示,平面 h_4 不可靠,以空间线段 ℓ_4 所在竖直直线为轴确定了位于平面 h_1 中被空间线段 ℓ_3 遮挡的空间线段 ℓ_3^* ;本文在保留空间线段 ℓ_3 并删除平面 h_4 的同时,增加了线段 l_3^* 与相应的空间线段 ℓ_3^* 以使线段 l_3^* 与 l_4 构成的区域与空间平面 h_1 相关联.

图像区域分裂 对于图像 I' 中的线段 l'_m ,若其在图像 I 中未能检测到与之匹配的线段但对其对应的空间

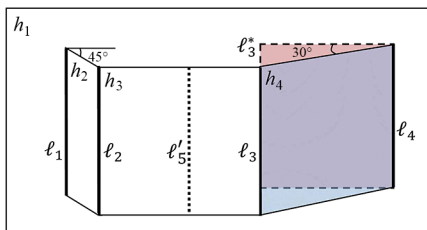


图5 空间线段推断与平面优化

线段 ℓ'_m 进行了推断, 则空间线段 ℓ'_m 在图像 I 中的投影线段 l_m 可能位于区域 Q_{ik} 的内部; 因而, 若 ℓ'_m 与 ℓ_i 与 ℓ_k 可以构成可靠平面, 则将 Q_{ik} 进行分裂并分别确定两分裂区域对应的平面. 如图 5 所示, 空间线段 ℓ'_5 的投影线段 l_5 位于区域 Q_{23} 内部且 ℓ'_5 与 ℓ_2 以及 ℓ'_5 与 ℓ_3 可构成可靠平面, 则将区域 Q_{23} 分裂为区域 Q_{25} 与 Q_{33} 并分别为其分配相应的平面, 同时将线段 l_m 添加至线段集 \mathcal{M} .

7.2 全局优化

局部优化仅对小部分线段、区域及其对应的空间线段与平面进行调整. 为获得整体一致的“线-面”结构, 本文算法在能量最小化框架下进一步对当前结果进行全局优化, 相应的能量函数定义为

$$E(H) = \sum_{s \in R} E_{\text{point}}(s, h_s) + \alpha \sum_{t \in N_1(s)} E_{\text{line}}(h_s, h_t) + \beta \sum_{k \in N_2(s)} E_{\text{plane}}(h_s, h_k) \quad (7)$$

其中, R 与 H 分别表示图像 I 中所有区域及相应的平面, $N_1(s)$ 表示两竖直边位于相同直线且与区域 s 相邻的区域集合, $N_2(s)$ 表示其他与区域 s 相邻的区域集合; α 与 β 为权重参数; $E_{\text{point}}(s, h_s)$ 、 $E_{\text{line}}(h_s, h_t)$ 、 $E_{\text{plane}}(h_s, h_k)$ 分别为点特征度量、共线约束与平面先验约束.

点特征度量 度量为区域 s 分配平面 h_s 的代价, 定义为

$$E_{\text{point}}(s, h_s) = \frac{1}{m} \sum_{i=1}^m D(q_i, h_s(q_i)) \quad (8)$$

其中, 像素 q_i 为区域 s 的顶点 ($m = 4$), $h_s(q_i)$ 为像素 q_i 的反投影线与平面 h_s 的交点在图像 I' 中的投影点, $D(q_i, h_s(q_i))$ 表示 q_i 与 $h_s(q_i)$ 之间 DAISY 特征相似度.

共线约束 鼓励竖直边位于同一竖直直线上的两相邻区域分配相同的平面, 定义为

$$E_{\text{line}}(h_s, h_t) = \begin{cases} e^{-D(s,t)} & h_s \neq h_t \\ 0 & \text{otherwise} \end{cases} \quad (9)$$

其中, $D(s, t)$ 为区域 s 与 t 质心之间的 DAISY 特征相似度.

平面先验约束 为不具备显著共面特征的相邻区域分配属于指定结构先验的平面, 定义为

$$E_{\text{plane}}(h_s, h_k) = \begin{cases} e^{-\frac{D(s,k)}{\sigma}} & A(h_s, h_k) \notin A_{\text{prior}} \\ 0 & \text{otherwise} \end{cases} \quad (10)$$

其中, $A(h_s, h_k)$ 为空间平面 h_s 与 h_k 的夹角, 参数 σ 控制平面先验约束力度.

式(7)所示能量函数的求解属于 NP-hard 问题, 本文算法采用协同优化方法^[15]获取其近似最优解; 最后利用式(6)所定义条件对两幅图像不一致的结构(区域及相应空间平面)进行剔除. 如图 6(a)~图 6(d)所示, 除遮挡线段与区域之外(如实线矩形标示区域), 其他线段与区域皆可得以较好地重建, 位于相同平面上的空间面片在图像中对应的区域与边界(如虚线矩形标示的两平面交线)较为一致, 纹理化后的“线-面”结构较为完整.

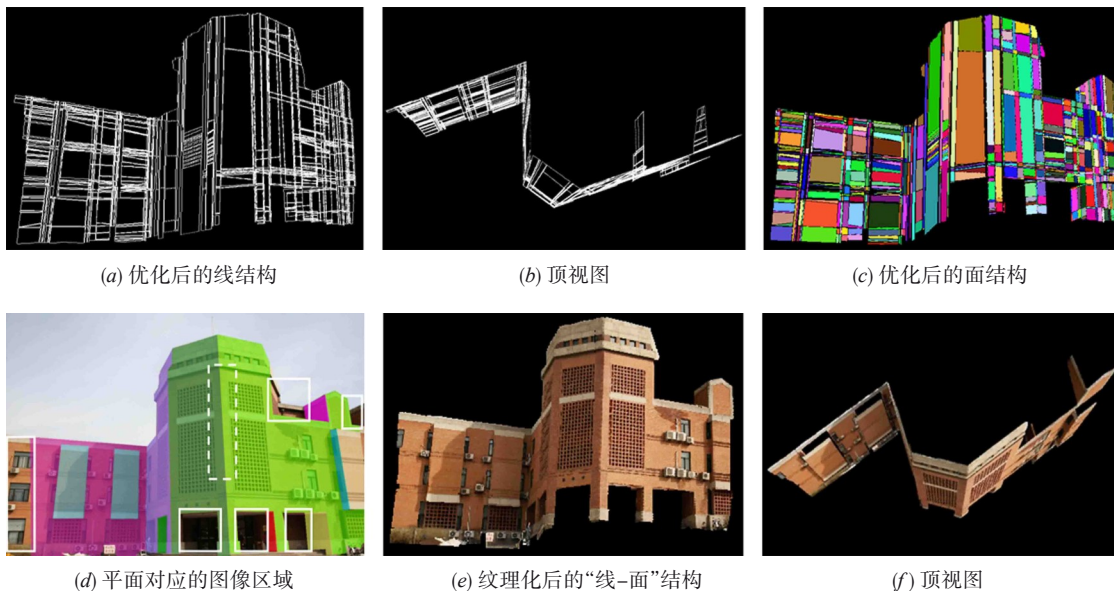


图6 “线-面”结构优化(相同平面对应相同颜色表示的图像区域)

8 实验与分析

为了验证本文算法的性能,除图 1 所示生命科学院 (LSB, 2184×1456) 数据集之外,本文进一步采用如图 7 所示的数据集对其进行测试:

- (1) 牛津大学数据^[16]: Valbonne(512×768);
- (2) 中科院数据^[10]: 清华学堂(TS, 2184×1456);
- (3) 实拍数据: City(1884×1224).



图 7 数据集图像示例

对于当前图像,本文算法采用其一幅相邻图像重建相应的“线-面”结构;此外,本文采用可靠空间点比例 T_1 与可靠平面总数 T_2 衡量算法的可靠性,其中, T_1 为式 (6) 所定义已获取可靠空间点的像素所占建筑区域内所有像素的比例, T_2 为根据式 (6) 确定的建筑区域内可靠平面的总数.

本文实验环境为 64 位 Windows 7 系统(硬件配置为 Intel 4.0GHz 四核处理器与 32GB 内存),所有算法均采用 Matlab 语言实现.

8.1 参数设置

本文算法对所有数据集的实验均采用相同的参数设置. 具体而言,对于距离阈值 κ ,其值设置过大易引入较多的匹配外点,而设置较小则可能导致获取的可靠匹配数较少;短线段匹配约束项权重 μ 设置较大时趋于为短线段分配与其相邻线段相同的匹配直线,因而有利于获取整体一致的结果. 在获取初始主平面时,正则项权重 λ 设置较大时趋于以较少的平面拟合当前空间线段,而平面可靠性判断阈值 ε 设置过大时则导致较多的平面被判定为不可靠,从而增加“线-面”结构优化的计算复杂度. 在本文实验中,当 $\kappa=2$, $\mu=0.7$, $\lambda=0.6$ 与 $\varepsilon=0.8$ 时,算法整体上表现出较好的性能. 在“线-面”结构优化中,参数 σ 用以控制平面先验约束力度,其值设置较大易趋于强制地为相邻区域分配相同的平面;在实验中,本文首先相对于共线约束力度将其设置为 0.8,然后采用枚举方式获取正则项权重 α 与 β ,即权重 α 在区间 $[0.1, 1]$ 以步长 0.1 依次取值时,求取权重 β 在相同区间以相同方式取值时的最高 T_1 值并将此时的 α 与 β 取值作为候选权重组合,最后从 10 个候选权重组合中选取对应 T_1 值最高者作为权重 α 与 β 的最优取值. 在本文实验中, α 与 β 取 0.6 与 0.4 时,平均 T_1 值基

本达到最大.

8.2 结果分析

本文算法对不同数据的初始化结果(初始线段、区域等)如表 1 所示.

表 1 初始化结果

数据集	初始线段	竖直直线	区域
Valbonne	521	40	723
LSB	1074	76	1655
TB	5296	207	9805
City	4577	219	4531

在第一组实验中,本文采用 Valbonne 数据测试本文算法的可行性. 该场景建筑立面结构相对简单,如表 2 与图 8(a)、图 8(b) 所示,由于在建筑区域检测到的初始线段较少,因而从中仅确定较少的竖直直线,因此对图像进行分割产生的区域也较少. 在线段匹配阶段,由于匹配图像中竖直直线的约束与 DAISY 特征描述子对旋转、尺度等因素具有较好的适应性,建筑立面中的大部分线段皆被可靠地匹配,进而可求取初始可靠空间线段并通过多平面拟合的方式获得可靠的建筑主平面(如图 8(c) 所示). 以此为基础,在建筑结构先验的引导下,未匹配线段对应的空间线段可被有效地推断,同时也保证了后续“线-面”结构优化的可靠性. 最终,通过融合点、线、面之间的几何约束与特征表达,利用协同优化方法可获得全局一致性的“线-面”结构;如图 8(d)、图 8(e) 所示,线结构与面结构相对较为完整,位于相同平面的空间面片对应的图像区域被合并(图 8(f)、图 8(g) 所示),不同平面的面片之间的边界与建筑结构基本一致,整体上具有较高的精度. 相对而言,文献[8]算法虽然通过 DAISY 特征匹配的方式有利于获取建筑相对完整的初始结构. 然而,如图 8(h) 所示,该算法由于在初始阶段难以确定与建筑结构相一致的图像区域,因而在最终获取的建筑分段平面结构中存在边界不精确的问题.

在效率上,如表 3 所示,由于图像稠密匹配非常耗时,文献[8]算法整体运行速度较低. 相对地,本文算法在“点-线”约束匹配阶段,通过竖直直线与建筑语义区域的约束,极大地减少了当前像素对应候选匹配点的数量,有效提高了匹配效率;此外,建筑主平面拟合仅针对数量不多的空间线段,因而未消耗过多时间,而建筑结构先验的引导作用使得空间线段推断与“线-面”全局优化可快速收敛;整体上,本文算法相对具有更高的效率.

在第三组实验中,本文采用 TB 数据集进一步验证算法的鲁棒性. TB 建筑结构虽然相对较为复杂,但其直线、平面等结构特征却更加明显. 因而,利用

表2 不同算法的重建结果

数据集	线段 匹配	主平 面	线段 推断	初始 平面	本文算法		文献[8]算法	
					T_1	T_2	T_1	T_2
Valbonne	414	4	289	649	0.8934	6	0.8737	6
LSB	1447	11	646	1534	0.8690	18	0.8311	12
TB	7988	17	1785	9577	0.9195	27	0.8994	18
City	3899	5	608	4178	0.8005	8	0.6491	8

表3 运行时间/s

数据集	区域 分割	初始 匹配	主平 面	线段 推断	结构 优化	合计	文献[8] 算法
Valbonne	1.1	1.4	0.6	0.7	1.1	4.9	109.6
LSB	3.4	3.1	1.4	2.2	2.7	12.8	354.2
TB	2.3	4.5	2.1	1.9	3.1	13.9	299.1
City	2.1	4.8	1.2	3.1	2.4	13.6	226.4

如图9(a)所示的初始线段对建筑区域进行分割可获得与实际结构相一致的区域(如图9(b)所示),有效保证了后续环节线段匹配、主平面拟合以及“线-面”结构优化的可靠性.整体上,如图9(c)、图9(e)与表2所示,本文算法仍可获得较好的结果(如窗户结构、墙面交线的精确重建).此外,图9(c)矩形标示的植物区域,由于从中检测到的线段不满足竖直性约束条件,因而未对后续环节产生影响;相反地,此区域在“线-面”结构全局优化阶段被分配了建筑区域对应的平面,这在一定程度上表明本文算法具有较好的可靠性.

在此实验中,如图9(h)所示,文献[8]算法仍存在结构边界重建不精确的问题;而在效率上,由于图像分

辨率相对较高,文献[8]算法在稠密像素匹配阶段消耗更多时间,整体效率更低;本文算法由于利用竖直直线对像素的匹配进行了约束,因而线段匹配过程较快,而主平面拟合、空间线段推断与结构优化等过程仅针对数量不多的线段或区域并在结构先验的引导下完成,整体上也未消耗过多时间.

在第四组实验中,本文采用实拍城市建筑数据集测试本文算法的适应性.如图10(a)所示,由于城市建筑较高、重复纹理较多等特征,在图像中检测到的位于同一竖直直线上的线段通常较短,通过对同一竖直直线上短线段的聚类与合并,更有利于探测建筑主体结构.此外,本文算法首先通过两幅图像建筑区域的检测避免了天空、地面等区域不必要的线段匹配,在一定程度上提高了建筑区域内线段匹配与结构重建的可靠性与效率.如图10(c)所示,根据建筑区域内线段匹配结果,建筑主平面可被有效地拟合;在已知其建筑主平面的情况下,根据特定平面夹角先验,如图10(d)、图10(i)所示,更多平面以及平面之间的交线(如图11所示)被可靠地重建,线结构与面结构 T 均具有较高的可靠性.相对而言,文献[8]算法由于诸多干扰因素的影响而未能获得较好的结果,如图10(j)所示,部分建筑区域由于未包含空间点或空间点与真值偏差较大而未在空间平面推断中获得可靠的结果.在效率上,由于建筑区域在图像中所占比例较小,而本文算法仅针对建筑区域进行重建而且在结构先验的引导下完成线段匹配与结构优化等过程,因而整体效率相对于文献[8]算法更高.

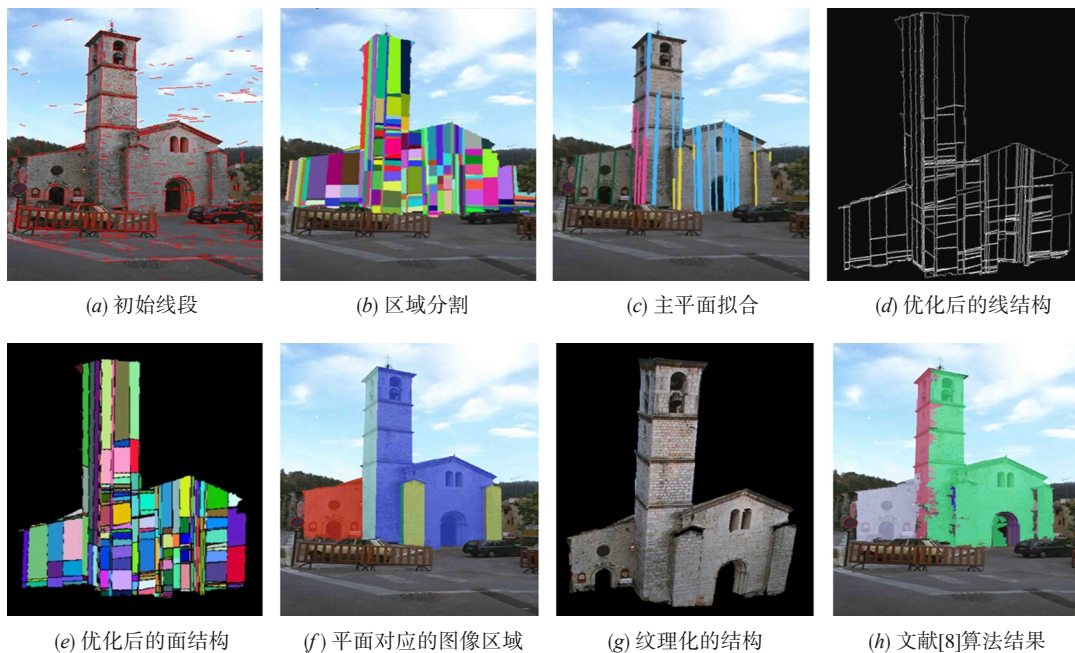


图8 Valbonne 建筑重建结果(相同平面对应相同颜色表示的图像区域)

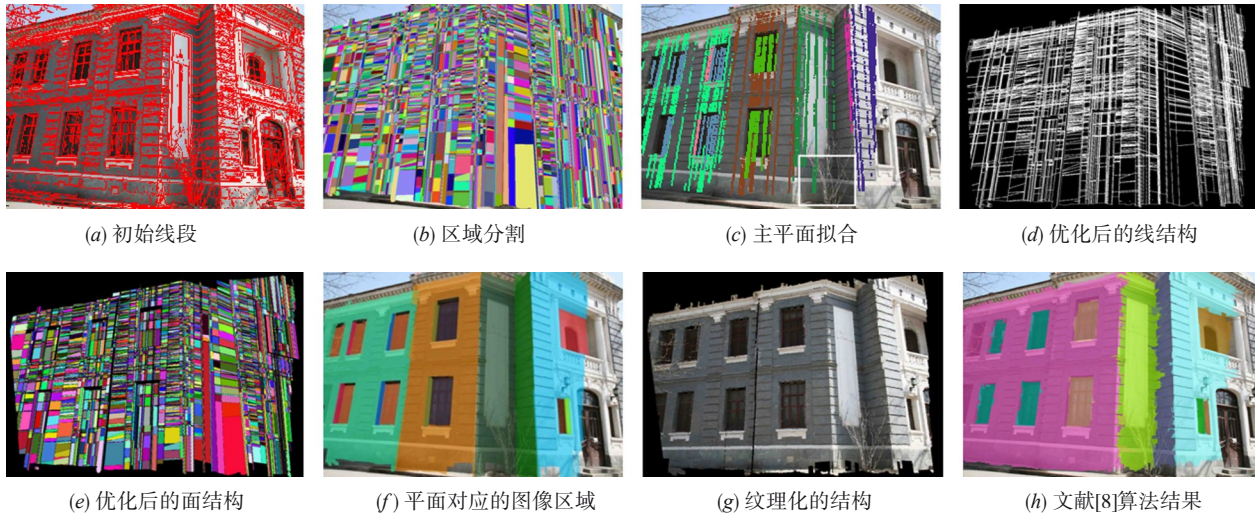


图9 TB建筑重建结果(相同平面对应相同颜色表示的图像区域)

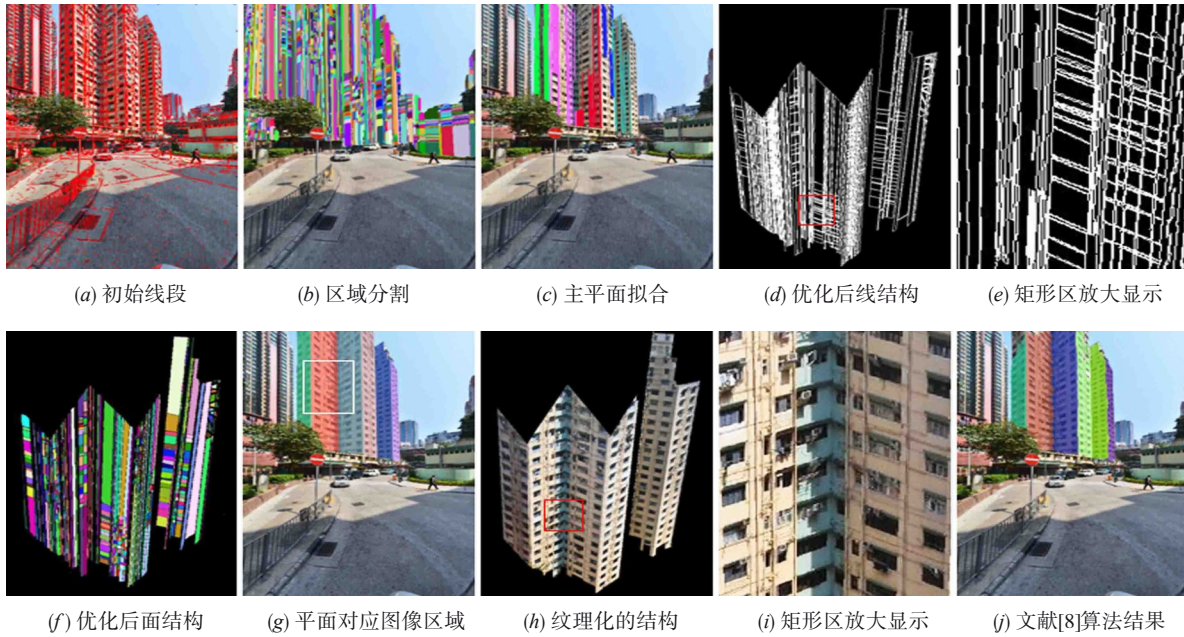


图10 实拍建筑重建结果(相同平面对应相同颜色表示的图像区域)

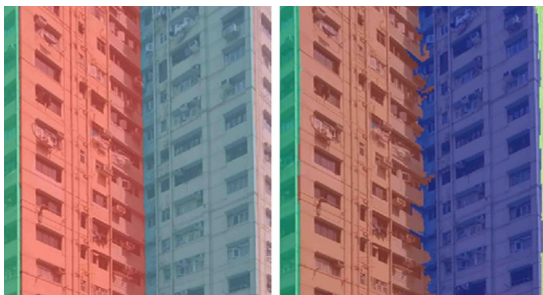


图11 图10(g)矩形区结构(左:本文算法,右:文献[8]算法)

以上实验结果表明,在城市建筑立面的重建中,以线段匹配为基础将点、线、面等基元进行融合更有利于快速获取其完整的结构,而建筑结构先验的融合则可

有效提高整体重建的效率与可靠性.

9 结论

为快速重建城市建筑立面完整的结构,本文算法利用在图像中检测的线段将建筑区域进行了分割,然后在点-线约束下对分割所得区域的竖直边进行了匹配并全局地拟合了建筑主平面,进而利用建筑结构先验对潜在的空间线段进行了推断并在点、线、面统一的框架下对建筑“点-面”结构进行了优化.实验表明,本文算法可有效克服传统算法效率低、边界重建精度差等缺点,利用两幅图像即可获得较好的结果,整体上具有较高的性能.

参考文献

- [1] Fan B, Wu F, Hu Z Y. Robust line matching through line-point invariants[J]. Pattern Recognition, 2012, 45(2): 794 – 805.
- [2] Kim C, Manduchi R. Planar structures from line correspondences in a manhattan world[A]. Proc of Asian Conference on Computer Vision[C]. Cham,GER: Springer, 2015. 509 – 524.
- [3] Al-Shahri M, Yilmaz A. Line matching in wide-baseline stereo: a top-down approach[J]. IEEE Transactions on Image Processing, 2014, 23(9):4199 – 4210.
- [4] Verhagen B, Timofte R, Gool L V. Scale-invariant line descriptors for wide baseline matching[A]. Proc of IEEE Winter Conference on Applications of Computer Vision[C]. Washington, USA: IEEE Computer Society, 2014. 493 – 500.
- [5] Jia Q, Gao X, Fan X, et al. Novel coplanar line-points invariants for robust line matching across views[A]. Proc of European Conference on Computer Vision[C]. Cham, GER: Springer, 2016. 599 – 611.
- [6] Li K, Yao J. Line segment matching and reconstruction via exploiting coplanar cues[J]. ISPRS Journal of Photogrammetry & Remote Sensing, 2017, 125: 33 – 49.
- [7] Bodis-Szomoru A, Riemenschneider H, Gool L V. Fast, approximate piecewise-planar modeling based on sparse structure-from-motion and superpixels[A]. Proc of Computer Vision and Pattern Recognition[C]. Washington, USA: IEEE, 2014. 23 – 28.
- [8] Verleysen C, Vleeschouwer C D. Piecewise-planar 3D approximation from wide-baseline stereo[A]. Proc of Computer Vision and Pattern Recognition[C]. Washington, USA: IEEE, 2016. 3327 – 3336.
- [9] Tola E, Lepetit V, Fua P. Daisy: An efficient dense descriptor applied to wide-baseline stereo[J]. Pattern Analysis and Machine Intelligence, 2010, 32(5): 815 – 830.
- [10] Gao W. CASIA datasets[EB/OL]. <http://vision.ia.ac.cn/data/index.html>, 2014-12-26.
- [11] Almazan E J, Tal R, Qian Y, et al. MCMLSD: A dynamic programming approach to line segment detection[A]. Proc of Computer Vision and Pattern Recognition[C]. Washington, USA: IEEE, 2017. 5854 – 5862.
- [12] Zhao H, Shi J, Qi X, et al. Pyramid scene parsing network [A]. Proc of Computer Vision and Pattern Recognition[C]. Washington, USA: IEEE, 2016. 6230 – 6239.
- [13] Hartley R, Zisserman A. Multiple View Geometry in Computer Vision[M]. Cambridge, UK: Cambridge University Press, 2004.
- [14] Delong A, Osokin A, Isack H N, et al. Fast approximate energy minimization with label costs[J]. International Journal of Computer Vision, 2012, 96(1):1 – 27.
- [15] Huang X F. Cooperative optimization for energy minimization: a case study of stereo matching [EB/OL]. <http://arxiv.org/pdf/cs.CV/0701057>, 2007-01-09.
- [16] Zisserman A. VGG datasets[EB/OL]. <http://www.robots.ox.ac.uk/~vgg/data/data-mview.html>, 2018-03-16.

作者简介



王 伟 男,1976年10月生于河南省周口市. 2011年获西南交通大学硕士学位,2015年获中国科学院自动化研究所博士学位. 现为周口师范学院副教授. 研究方向为计算机视觉、机器学习.
E-mail:wangwei@zknu.cn



于 磊 男,1984年6月生于河南省周口市. 2017年获河南大学硕士学位. 现为周口师范学院讲师. 研究方向为计算机视觉、大数据分析.