

Journal of Image  
and Graphics

# 中国图象图形学报



ISSN1006-8961  
CN11-3758/TB

2013 **3**  
Vol.18 No.

中国科学院遥感应用研究所  
中国图象图形学学会主办  
北京应用物理与计算数学研究所



## 中国图象图形学报

Zhongguo Tuxiang Tuxing Xuebao

2013 年 3 月 第 18 卷 第 3 期(总第 203 期)

## 目次

## 图像处理和编码

沿边局部灰度差分椒盐噪声滤波 ..... 孔勇奇, 潘志庚(249)

## 图像分析和识别

核稀疏保留投影及在步态识别中的应用 ..... 王科俊, 阎涛, 吕卓纹, 唐墨(257)

局部 Gist 特征匹配核的场景分类 ..... 杨昭, 高隼, 谢昭, 吴克伟(264)

## 图像理解和计算机视觉

条件随机场模型的场景描述 ..... 赵龙, 郭立, 谢锦生(271)

基于  $L_1$  与空间点分类的 3 维场景重建 ..... 李静, 杨宜民, 付根平, 陈建平(277)

稀疏表示的 Lucas-Kanade 目标跟踪 ..... 徐如意, 陈靛影(283)

## 医学图像处理

管状特性和主动轮廓的 3 维血管自动提取 ..... 尚岩峰, 汪辉, 汪宁, 陈杰, 姚丽萍, 孙琨(290)

## “第 17 届 CAD&amp;CG 学术会议”专栏

多步攻击的规则树检测及可视化 ..... 胡亮, 赵剑明, 解男男, 努尔布力(299)

显著性边缘引导下的图像抽象化 ..... 王勋, 李红, 刘春晓, 严彩萍(305)

稀疏表达的运动数据压缩 ..... 齐天, 肖俊, 庄越挺(311)

基于语义的用户关注度计算 ..... 陈蕾英, 陈佳舟, 潘斌, 彭群生(318)

九宫格空间框架的图书图像检索 ..... 李宗民, 唐志辉(325)

光学相干层析视网膜体数据的 3 维分割 ..... 樊鲁杰, 孙延奎, 张田, 田小林(330)

混合软影绘制算法 ..... 刘浏, 周炜, 李华(336)

基于场的人群运动仿真 ..... 赵欣欣, 张勇, 孔德慧, 尹宝才(344)

利用大位移视图的自动可信图像修补 ..... 刘春晓, 金剑秋, 彭群生(351)

第八届国际数字地球会议征稿通知 ..... (358)

# Journal of Image and Graphics

(Monthly, Started in 1996)

Vol. 18 No. 3 March 2013

## Contents

### Image Processing and Coding

Salt and pepper noise filtering algorithm based on local border gray-scale differences ..... Kong Yongqi, Pan Zhigeng(249)

### Image Analysis and Recognition

Kernel sparsity preserving projections and its application to gait recognition ..... Wang Kejun, Yan Tao, Lv Zhuowen, Tang Mo(257)

Scene categorization of local Gist feature match kernel ..... Yang Zhao, Gao Jun, Xie Zhao, Wu Kewei(264)

### Image Understanding and Computer Vision

Scene description based on the conditional random fields model ..... Zhao Long, Guo Li, Xie Jinsheng(271)

Three dimensional scene reconstruction based on  $L_1$  and scene point classification .....  
..... Li Jing, Yang Yimin, Fu Genping Chen Jianping(277)

Lucas-Kanade tracking based on sparse representation ..... Xu Ruyi, Chen Jingying(283)

### Medical Image Processing

Three dimensional vessel extraction model based on tubular characters and active contour model  
..... Shang Yanfeng, Wang Hui, Wang Ning, Chen Jie, Yao Liping, Sun Kun(290)

### Issum of the CAD&CG'2012

Multi-step attacks detected by rules tree and visualization ..... Hu Liang, Zhao Jianming, Xie Nannan, Nurbol(299)

Salient edge guided image abstraction optimization ..... Wang Xun, Li Hong, Liu Chunxiao, Yan Caiping(305)

Motion data compression using sparse representation ..... Qi Tian, Xiao Jun, Zhuang Yueting(311)

Semantic-based user interests computation ..... Chen Leiying, Chen Jiazhou, Pan Bin, Peng Qunsheng(318)

Books image retrieval based on grid space frame ..... Li Zongmin, Tang Zhihui(325)

Three dimensional segmentation to detect retinal boundary surfaces from OCT volume data  
..... Fan Lujie, Sun Yankui, Zhang Tian, Tian Xiaolin(330)

Hybrid soft shadow rendering method ..... Liu Liu, Zhou Wei, Li Hua(336)

Field-based crowd simulation ..... Zhao Xinxin, Zhang Yong, Kong Dehui, Yin Baocai(344)

Automatic and accurate image completion from a large displacement view ..... Liu Chunxiao, Jin Jianqiu, Peng Qunsheng(351)

中图法分类号: TP391.41 文献标识码: A 文章编号: 1006-8961(2013)03-0277-06

论文引用格式: 李静, 杨宜民, 付根平, 陈建平. 基于  $L_1$  与空间点分类的3维场景重建[J]. 中国图象图形学报, 2013, 18(3): 277-282.

## 基于 $L_1$ 与空间点分类的3维场景重建

李静<sup>1</sup>, 杨宜民<sup>1</sup>, 付根平<sup>1</sup>, 陈建平<sup>2</sup>

1. 广东工业大学自动化学院, 广州 510090; 2. 肇庆学院计算机学院, 肇庆 526061

**摘要:** 针对多视图的3维重建问题, 提出一种基于  $L_1$  与空间点分类的3维场景重建技术。首先采用效率较高的  $L_1$  方法剔除残留局外点, 以避免重建过程中受到局外点的干扰; 再将空间点分为两类, 第1类为两视图可见空间点, 第2类为多于两视图可见空间点; 针对第1类空间点采用高精度的最优三角形法进行重建, 而第2类空间点采用  $L_\infty$  范数方法进行重建, 并不断调整二分搜索的上下界, 以减少迭代次数。实验结果表明了本文方法的优越性。

**关键词:** 3维场景重建; 空间点分类; 二阶锥规划;  $L_1$  方法;  $L_\infty$  范数; 最优三角形法

## Three dimensional scene reconstruction based on $L_1$ and scene point classification

Li Jing<sup>1</sup>, Yang Yimin<sup>1</sup>, Fu Genping<sup>1</sup>, Chen Jianping<sup>2</sup>

1. School of Automation, Guangdong University of Technology, Guangzhou 510090, China;

2. School of Computer, Zhaoqing University, Zhaoqing 526061, China

**Abstract:** In this paper, we present a new technique for solving geometric structure and motion problems based on  $L_1$  and scene point classification. First, in order to avoid that our reconstruction method is affected by outliers, the outliers are removed using the  $L_1$  approach. Second, the scene points are divided into two classes; the first class contains those scene points which are visible in two views; the second class contains those scene points which are visible in more than two views. The optimal triangulation method is used for the first class. The  $L_\infty$  norm minimization is used for the second class and the gap between the upper and lower bound of the bisection is kept as small as possible, so that the computation time is decreased. The experimental results show the advantages of our method.

**Key words:** 3D scene reconstruction; scene point classification; second order cone programming;  $L_1$  approach;  $L_\infty$  norm; the optimal triangulation method

## 0 引言

3维场景重建技术是计算机视觉关键技术之一。若已知多幅视图上的图像对应点和各个摄像机的投影矩阵, 如何获得精确的3维场景重建结果, 是本文将要解决的问题。若图像对应点精确满足对极几何关系, 则投影线在空间中有唯一的交点, 但是检测到

的图像对应点不可避免地存在误差, 这些投影线在空间中一般不相交, 因此, 有必要为3维空间点求一个最佳解。常见的算法是基于LM (levenberg marquardt) 迭代的黄金标准算法<sup>[1]</sup>, 但在该算法中, 基于  $L_2$  范数的目标函数并非凸函数, 不能保证得到全局最优值。为了得到  $L_2$  范数目标函数的全局最优值, 文献[2-5]采用了分支定界算法, 但是计算量很大。为了避免出现多个局部极小值, 并且简化代价函数公

收稿日期: 2012-09-26; 修回日期: 2012-11-18

基金项目: 广东省自然科学基金自由申请项目 (S2011010004006)

第一作者简介: 李静 (1983—), 女, 现为广东工业大学控制科学与工程专业博士研究生, 主要研究方向为多视图几何和3维重建技术。

E-mail: litangjing61@163.com

式,文献[6-9]采用最小化  $L_\infty$  范数的优化准则,最小化  $L_\infty$  范数产生一个拟凸代价函数,能够用二阶锥规划求解到全局最优值,而且不需要计算初始值。但  $L_\infty$  范数方法精度较低,计算量也很大。考虑到最优三角形法在两视图重建中的优势,具有统计意义,精度高,计算量也较小。将空间点分为两类,两视图可见空间点为第 1 类,多于两视图可见空间点为第 2 类。对于第 1 类空间点采用最优三角形法进行求解,而第 2 类空间点采用  $L_\infty$  范数方法进行求解,并在求解过程中从一个子集出发,并尽量减少迭代次数。由于在第 1 类空间点重建中,最优三角形法无论效率还是精度都高过  $L_\infty$  范数方法,以及在第 2 类空间点重建中,最小化  $L_\infty$  范数从一个子集出发,并调整二分搜索的上下界,减少了运算时间,因此可以提高整体的计算效率和精度,第 1 类空间点越多,精度提高的越多。由于最优三角形法和  $L_\infty$  范数方法对局外点都比较敏感,所以在重建之前先采用效率较高的  $L_1$  方法剔除残留局外点,以避免重建方法受到局外点的干扰。

## 1 二阶锥规划

凸优化一般形式为

$$\begin{aligned} \min_{\mathbf{X}} \quad & f_0(\mathbf{X}) \\ \text{s. t.} \quad & f_i(\mathbf{X}) \leq 0, i = 1, \dots, M \end{aligned} \quad (1)$$

式中,  $\mathbf{X} \in \mathbf{R}^n$ , 目标函数  $f_0(\mathbf{X}) : \mathbf{R}^n \rightarrow \mathbf{R}$  和约束函数  $f_i(\mathbf{X}) : \mathbf{R}^n \rightarrow \mathbf{R}$  都是凸函数。如果目标函数  $f_0$  是线性函数,约束函数为

$$\|\mathbf{A}_i \mathbf{X} + \mathbf{b}_i\|_2 - \gamma(\mathbf{c}_i^T \mathbf{X} + d_i) \leq 0 \quad (2)$$

式中,  $\mathbf{A}_i \in \mathbf{R}^{n_i \times n}$ ,  $\mathbf{b}_i \in \mathbf{R}^{n_i}$ ,  $\mathbf{c}_i \in \mathbf{R}^n$ ,  $d_i \in \mathbf{R}$ 。

这种特定的凸优化称为二阶锥规划。很容易证明约束函数是一个凸函数。最小化最大二阶锥向量表示为

$$\begin{aligned} \min_{\mathbf{X}} \max_i \quad & \frac{\|\mathbf{A}_i \mathbf{X} + \mathbf{b}_i\|_2}{\mathbf{c}_i^T \mathbf{X} + d_i} \\ \text{s. t.} \quad & \mathbf{c}_i^T \mathbf{X} + d_i > 0 \quad i = 1, \dots, M \end{aligned} \quad (3)$$

如果考虑单个函数  $\|\mathbf{A}_i \mathbf{X} + \mathbf{b}_i\|_2 / (\mathbf{c}_i^T \mathbf{X} + d_i)$  是一个向量的组成元素,那么式(3)将被认为是最小化  $L_\infty$  范数。如果引入一个附加变量,很容易得到式(3)的等价转化形式为

$$\begin{aligned} \min_{\gamma, \mathbf{X}} \quad & \gamma \\ \text{s. t.} \quad & \frac{\|\mathbf{A}_i \mathbf{X} + \mathbf{b}_i\|_2}{\mathbf{c}_i^T \mathbf{X} + d_i} \leq \gamma, \mathbf{c}_i^T \mathbf{X} + d_i > 0 \end{aligned} \quad (4)$$

若  $\gamma \geq 0$  是已知的,那么式(4)也可以转换为一个二阶锥可行性问题,可写为

$$\begin{aligned} \text{find } \mathbf{X} \\ \text{s. t.} \quad & \|\mathbf{A}_i \mathbf{X} + \mathbf{b}_i\|_2 - \gamma(\mathbf{c}_i^T \mathbf{X} + d_i) \leq 0 \end{aligned} \quad (5)$$

若给定  $\gamma \geq 0$ ,  $\|\mathbf{A}_i \mathbf{X} + \mathbf{b}_i\|_2 - \gamma(\mathbf{c}_i^T \mathbf{X} + d_i) \leq 0$  是一个凸集,那么函数  $e_i(\mathbf{X}) = \frac{\|\mathbf{A}_i \mathbf{X} + \mathbf{b}_i\|_2}{\mathbf{c}_i^T \mathbf{X} + d_i}$  是一个

拟凸函数<sup>[10]</sup>,则  $\max_i \frac{\|\mathbf{A}_i \mathbf{X} + \mathbf{b}_i\|_2}{\mathbf{c}_i^T \mathbf{X} + d_i}$  也是拟凸函数,

拟凸函数在其定义域内只有一个最小值,所以最小化  $L_\infty$  范数能够得到全局最小值。对于固定的  $\gamma$ ,目标函数式(5)中的约束函数是二阶锥规划,因此可利用二分搜索在公开的软件包 SeDuMi 上进行求解来获得最优  $\gamma$  值,以及对应的解  $\mathbf{X}$ 。

## 2 基于 $L_1$ 与空间点分类的 3 维场景重建

### 2.1 局外点的剔除

因为  $L_\infty$  范数方法和最优三角形法对局外点都非常敏感,所以必须先对残留局外点进行剔除。近年来很多文献[11-14]提出通过引入松弛变量来剔除局外点的全局优化方法。若在目标函数中引入一个松弛变量(1-slack)描述为

$$\begin{aligned} \min_{s, \mathbf{X}} \quad & s \\ \text{s. t.} \quad & \|\mathbf{A}_i \mathbf{X} + \mathbf{b}_i\|_2 \leq \gamma(\mathbf{c}_i^T \mathbf{X} + d_i) + s \\ & i = 1, \dots, M \end{aligned} \quad (6)$$

式中,  $s \in \mathbf{R}$ 。由于产生最大误差的数据至少含有一个局外点,可采用一种迭代的策略来剔除局外点。因为每次迭代只能剔除非常有限的局外点,所以需要很多次迭代才能够清除全部局外点,这种方法可靠性好,但是耗时太长,以致于不能够得到广泛应用。若在目标函数中引入  $M$  个松弛变量,可描述为

$$\begin{aligned} \min_{s_i, \mathbf{X}} \quad & s_1 + s_2 + \dots + s_M \\ \text{s. t.} \quad & \|\mathbf{A}_i \mathbf{X} + \mathbf{b}_i\|_2 \leq \gamma(\mathbf{c}_i^T \mathbf{X} + d_i) + s_i \\ & s_i \geq 0, i = 1, \dots, M \end{aligned} \quad (7)$$

式(7)是最小化松弛变量的  $L_1$  范数,因此被称为  $L_1$  方法,其目的是找到使  $s_i$  为 0 的解。如果  $\sum_i s_i = 0$ ,则说明没有局外点;如果有  $s_i \neq 0$ ,表明存在局外点,需要剔除局外点,才能进行后面的求解。由于  $L_1$  方法预先设定了阈值  $\gamma$ ,只需对二阶锥规划进行一次运算,可以一次性无需迭代的剔除残留局外点,



效率较高。存在的缺陷是,当局外点比例较高时, $L_1$ 方法会变得不稳定。

为了寻求可靠性和效率之间的平衡,文献[14]中提出了一种引入 $K$ 个松弛变量( $K$ -slack)的方法,用零和博弈的策略迭代地剔除局外点。 $K$ 代表了局外点数目,一般情况下, $K = [5\%, 20\%]M$ , $K$ -slack方法是 $L_1$ 方法的一种较为普遍的形式,当 $K = M$ 时,即为 $L_1$ 方法。实际上,在局外点比例较低的情况下(小于5%), $K$ -slack方法的优势并不明显,计算精度稍高于 $L_1$ 方法,但是速度慢了将近一倍;而在局外点比例较高的情况下(5%~20%), $K$ -slack的优势较为明显,虽然速度比 $L_1$ 方法慢了很多,但是精度却得到了大大的提高。只是从实际应用的角度来看, $K$ -slack方法的意义并不大, $K$ -slack最大的优势是可以处理局外点比例高达5%~20%的数据。一般情况下,都是先采用随机抽样一致性(RANSAC)算法剔除局外点,只是由于RANSAC算法的随机性,且在全局的匹配跟踪过程容易产生迭代误差,所以会有一些残留局外点存在,数量比较小,一般只有1%左右。因此采用 $L_1$ 方法剔除局外点性价比更高,效率高,精度也不低。

这里需要特别指出的是,无论是1-slack, $L_1$ 方法,还是 $K$ -slack方法都只是一个精制过程,当遇到局外点比例太高的数据时都会变得不稳定。

## 2.2 最优三角形法

最优三角形法只能用于重建两视图可见的空间点,如果高斯噪声模型的假定是正确的,这个方法可以证明是最优的<sup>[1]</sup>。若给定一组检测得到的图像对应点 $\mathbf{x} \leftrightarrow \mathbf{x}'$ ,通过最小化距离平方和,寻找对应点 $\hat{\mathbf{x}} \leftrightarrow \hat{\mathbf{x}}'$ ,表示为

$$\begin{aligned} \min d(\mathbf{x}, \hat{\mathbf{x}})^2 + d(\mathbf{x}', \hat{\mathbf{x}}')^2 \\ \text{s. t. } \hat{\mathbf{x}}' \mathbf{F} \hat{\mathbf{x}} = 0 \end{aligned} \quad (8)$$

式中, $d(*,*)$ 表示图像点之间的欧氏距离,式(8)等价于最小化空间点 $\mathbf{X}$ 的重投影误差, $\mathbf{X}$ 由与 $\mathbf{F}$ 相容的投影矩阵映射到 $\hat{\mathbf{x}}$ 和 $\hat{\mathbf{x}}'$ 。假定图像检测点满足高斯误差分布的情况下, $\hat{\mathbf{x}}$ 和 $\hat{\mathbf{x}}'$ 是关于图像对应点的最大似然估计,一旦获得 $\hat{\mathbf{x}}$ 和 $\hat{\mathbf{x}}'$ ,那么对应的射线将在空间中准确相交,空间点 $\mathbf{X}$ 可以利用线性的三角测量方法获得。最优三角形法是通过极线束的一个适当参数化,即可将式(8)转换为

$$\min d(\mathbf{x}, \mathbf{l}(t))^2 + d(\mathbf{x}', \mathbf{l}'(t))^2 \quad (9)$$

用参数 $t$ 参数化第1幅图像的对极线束,并记第1幅图像上的对极线为 $\mathbf{l}(t)$ ,利用基本矩阵 $\mathbf{F}$ ,计算第

2幅图像上的对极线束 $\mathbf{l}'(t)$ ,因此 $\hat{\mathbf{x}}$ 是极线 $\mathbf{l}(t)$ 上与点 $\mathbf{x}$ 最接近的点, $\hat{\mathbf{x}}'$ 是极线 $\mathbf{l}'(t)$ 上与点 $\mathbf{x}'$ 最接近的点。即将式(9)中的距离函数转换成 $t$ 的一个有理多项式,并求这个函数的最小值。可写为

$$s(t) = \frac{t^2}{1+f^2t^2} + \frac{(ct+d)^2}{(at+b)^2+f'^2(ct+d)^2} \quad (10)$$

式中, $a, b, c, d, f, f'$ 是与基本矩阵 $\mathbf{F}$ 有关的几个参数,均为已知。对式(10)求导得到

$$\begin{aligned} s'(t) = \frac{2t}{(1+f^2t^2)^2} - \\ \frac{2(ad-bc)(at+b)(ct+d)}{((at+b)^2+f'^2(ct+d)^2)^2} \end{aligned} \quad (11)$$

令式(11)等于0,则是一个6次多项式,最多有6个实根。求式(10)取最小值时的 $t_{\min}$ ,利用这个 $t_{\min}$ 求解到 $\hat{\mathbf{x}}$ 和 $\hat{\mathbf{x}}'$ ,并最终获得空间点坐标值 $\mathbf{X}$ 。

## 2.3 最小化 $L_\infty$ 范数方法

针对多于两视图可见的空间点采用最小化 $L_\infty$ 范数方法进行重建。若已知图像平面上的图像点坐标为 $\mathbf{x}_i = [u_i, v_i]^T$ 摄像机投影矩阵为 $\mathbf{P}_i, i = 1, \dots, M$ ,未知空间点坐标为 $\mathbf{X} = [X, Y, Z, 1]^T$ ,则投影点坐标为 $\hat{\mathbf{x}}_i = \left( \frac{\mathbf{P}_{i,1}^T \mathbf{X}}{\mathbf{P}_{i,3}^T \mathbf{X}}, \frac{\mathbf{P}_{i,2}^T \mathbf{X}}{\mathbf{P}_{i,3}^T \mathbf{X}} \right)$ ,其中 $\mathbf{P}_{i,j}^T$ 为第 $i$ 个摄像机投影矩阵 $\mathbf{P}_i$ 的第 $j$ 行。则投影误差函数描述为

$$\begin{aligned} e_i(\mathbf{X}) = \left\| u_i - \frac{\mathbf{P}_{i,1}^T \mathbf{X}}{\mathbf{P}_{i,3}^T \mathbf{X}}, v_i - \frac{\mathbf{P}_{i,2}^T \mathbf{X}}{\mathbf{P}_{i,3}^T \mathbf{X}} \right\|_2 = \\ \frac{\| (u_i \mathbf{P}_{i,3}^T - \mathbf{P}_{i,1}^T) \mathbf{X}, (u_i \mathbf{P}_{i,3}^T - \mathbf{P}_{i,2}^T) \mathbf{X} \|_2}{\mathbf{P}_{i,3}^T \mathbf{X}} = \\ \frac{\| \mathbf{A}_i \mathbf{X} + \mathbf{b}_i \|_2}{\mathbf{c}_i^T \mathbf{X} + d_i} \end{aligned} \quad (12)$$

式中, $\| \cdot \|_2$ 代表欧氏距离, $\mathbf{A}_i$ 是系数矩阵, $\mathbf{b}_i, \mathbf{c}_i$ 是向量, $d_i$ 是个标量。最小化 $L_\infty$ 范数的目的就是在所有图像中最小化最大投影误差值,记为

$$\begin{aligned} \min_{\mathbf{X}} \max_i \frac{\| \mathbf{A}_i \mathbf{X} + \mathbf{b}_i \|_2}{\mathbf{c}_i^T \mathbf{X} + d_i} \\ \text{s. t. } \mathbf{c}_i^T \mathbf{X} + d_i > 0, \quad i = 1, \dots, M \end{aligned} \quad (13)$$

式中,约束条件 $\mathbf{c}_i^T \mathbf{X} + d_i > 0$ 表示所有可视点都在摄像机的前面。可以看出,式(13)同式(3)形式一样,因此可以转换为二阶锥规划问题进行求解。

最小化 $L_\infty$ 只是最小化最大误差值,若能找到一个刚好含有最大误差值的支持集,仅对这个支持集进行最小化求解即可,但支持集并不易得。文献[15]中采用从一个子集依次开始运算,并通过计算后面数据的重投影误差值和当前 $L_\infty$ 误差值的关系,

而确定后面数据是否再添加进来进行计算,有些数据是没有必要再添加进来计算的,以此来减少计算时间。但是,在有些情况下,这种方法是失效的,譬如说一组数据的误差呈阶梯式增长,这样为了找到最大误差数据,要计算后面所有数据的重投影误差值,并都要添加进来进行二分搜索计算,由此不仅不会减少计算时间,反而大大增加了计算量。为了避免这种情况的出现,先利用  $L_1$  方法获得的初始结果代入式(12)中计算保留数据的重投影误差,并从大到小进行排序,相应的投影矩阵也要跟着调整顺序,先对排在前面的误差较大的数据进行最小化  $L_\infty$  优化计算,再依次计算。虽然这样的排序只是一个粗略结果,但是从一定程度上讲可以作为评判标准。先对经过误差排序的前几个对应点集  $\{x_1 \leftrightarrow x_2 \leftrightarrow \dots \leftrightarrow x_k\}$  进行最小化  $L_\infty$  误差,设得到  $L_\infty$  误差值为  $e_k^\infty$ , 对应的空间点坐标为  $X_k^\infty$ , 利用式(12)计算对应  $x_{k+1}$  的误差  $e_{k+1}(X_k^\infty)$ , 比较  $e_{k+1}(X_k^\infty)$  和  $e_k^\infty$  的大小,并分两种情况进行处理。

#### 2.4 基于 $L_1$ 与空间点分类的 3 维场景重建

基于  $L_1$  与空间点分类的 3 维场景重建实现步骤如下:

1) 输入  $N$  个空间点对应的图像点数据,摄像机投影矩阵  $P_i, i=1, \dots, M$ , 设第  $j$  个空间点对应的图像点数据为  $I_j(j=1, \dots, N)$ , 令当前为第 1 个空间点对应的数据  $I_j(j=1)$ 。

2) 将图像点数据  $I_j$  和相应的摄像机投影矩阵代入式(7)中进行计算,即利用 2.1 节的  $L_1$  方法剔除局外点,删除  $s_i > 0$  的数据点。

3) 判断空间点是两视图上能看到的,还是在多于两视图上能看到的。

4) 若为两视图上能看到的空间点,则利用 2.2 节的最优三角形法进行重建,并输出空间坐标值,转至步骤 6)。

5) 若为多于两视图上能够看到的空间点,则利用  $L_\infty$  范数方法进行重建,实现如下:

(1) 按照 2.3 节提到的方式将对应点集排序为  $U_p = \{x_1 \leftrightarrow \dots \leftrightarrow x_p\}$ , 相应的投影矩阵也跟着调整顺序。

(2) 选择前面的  $k$  个对应点集  $U_k = \{x_1 \leftrightarrow \dots \leftrightarrow x_k\}$  (一般情况下取  $k=2$ ) 进行最小化  $L_\infty$  误差,设得到  $L_\infty$  误差值为  $e_k^\infty$ , 对应的空间点坐标为  $X_k^\infty$ 。

(3) 利用式(12)计算对应  $x_{k+1}$  的误差

$e_{k+1}(X_k^\infty)$ , 并比较  $e_{k+1}(X_k^\infty)$  和  $e_k^\infty$  的大小,此时可分为两种情况进行处理:

① 若  $e_{k+1}(X_k^\infty) \leq e_k^\infty$  则令  $X_{k+1}^\infty = X_k^\infty, e_{k+1}^\infty = e_k^\infty$ , 无需再将  $x_{k+1}$  并入子集中进行计算;

② 若  $e_{k+1}(X_k^\infty) > e_k^\infty$ , 将  $x_{k+1}$  并入子集中,对  $\{x_1 \leftrightarrow x_2 \leftrightarrow \dots \leftrightarrow x_{k+1}\}$  再利用二分搜索进行迭代计算,并令二分搜索的上界  $\gamma_u = e_{k+1}$ , 下界  $\gamma_l = e_k^\infty$ 。

(4) 直到将  $U_p/U_k(U_p$  去掉  $U_k$  的剩余数据)中的所有数据计算完毕,输出空间点的全局最优值,并转至步骤 6)。

6)  $j=j+1$ , 如果  $j \leq N$ , 返回步骤 2), 否则循环结束。

7) 将得到的空间点坐标值作为初始值,再进行一次集束调整,得到最终重建结果。

### 3 本文方法性能分析

本文方法是在  $L_1 - L_\infty$  方法的基础上又做了改进,因此主要分析一下本文方法和  $L_1 - L_\infty$  方法之间的不同。先分析一下耗时,假设  $L_1$  方法总耗时(单位 s)为  $t$ , 最优三角形法重建一个两视图可见空间点耗时为  $t_1$  s, 本文中加速的  $L_\infty$  方法重建多于两视图的所有空间点总耗时为  $t_2$ ,  $L_\infty$  方法重建一个两视图可见空间点耗时为  $t_3$ , 而  $L_\infty$  方法重建多于两视图可见所有空间点总耗时为  $t_4$ 。假设两幅图像上可见空间点为  $m$  个。本文方法较  $L_1 - L_\infty$  方法多了一个判断的步骤,然而这个步骤的耗时相对于整个重建时间来说非常小,几乎可以忽略不计。因此,本文方法总耗时为  $t + mt_1 + t_2$ , 而  $L_1 - L_\infty$  方法总耗时为  $t + mt_3 + t_4$ 。对于两视图看见点的重建,最优三角形法的耗时少于  $L_\infty$  方法,而且本文中加速的  $L_\infty$  方法比普通  $L_\infty$  方法耗时也少,那么本文方法总耗时肯定比  $L_1 - L_\infty$  方法少,共少了  $m(t_3 - t_1) + t_4 - t_2$ 。

再来分析一下精度,由于局外点一般产生较大的误差,本文方法和  $L_1 - L_\infty$  都利用  $L_1$  方法剔除了残留局外点,所以精度会比直接进行重建高出很多。另外,重建精度与最小化的误差有关,虽然  $L_\infty$  范数方法产生的拟凸函数能够获得全局最优值,但是  $L_\infty$  范数方法是最小化最大误差值,当检测误差服从高斯分布时,  $L_\infty$  范数方法的优化准则不是最大似然估计。而最优三角形法是最小化误差的平方和,当检测误差服从高斯分布时,优化准则为最大似然估计。而且通过建立一个 6 次多项式能找到一个全局最优解,避免了

使用迭代搜索算法(譬如LM算法)陷入局部最优值的危险。对于两幅图像的重建,最优三角形法的精度要高于 $L_\infty$ 方法,且 $m$ 数量越多,精度提高的越多。

## 4 实验与结果分析

为了测试本文方法的性能,对真实图像做了实验,实验需要用到软件包SeDuMi并在matlab7.0环境下进行。实验采用Oxford Dinosaur图像数据和Lund Cathedral图像数据,Dinosaur图像共有36幅,有4983个空间点的图像坐标,检测到的图像点共有16432个,图1给出了其中3幅图像。Cathedral图像共17幅,有16961个空间点的图像坐标,检测到的图像点共有46045个,图2给出了其中3幅图像。 $L_1$ 方法的阈值设为5个像素值,剔除局外点的数目如表1所示。

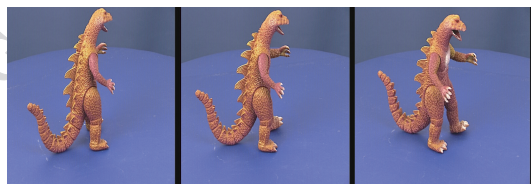


图1 Oxford dinosaur 的3幅图像

Fig. 1 Three views of Oxford dinosaur



图2 Lund cathedral 的3幅图像

Fig. 2 Three views of Lund cathedral

表1  $L_1$  方法剔除局外点数目

Table 1 The number of outliers is removed by  $L_1$  method

| 图像数据      | 剔除局外点数目 |
|-----------|---------|
| Dinosaur  | 348     |
| Cathedral | 679     |

为了证明本文方法的优势,还对相同的数据,分别采用其他两种方法进行了实验,第1个是 $L_2$ 方法,即文献[1]中的重建方法,第2个是 $L_1 - L_\infty$ 方法,即先利用 $L_1$ 方法剔除局外点,再利用 $L_\infty$ 进行重建,即文献[11]中的重建方法。3种方法的总体性能比较如表2所示,即3种方法各自的重建时间以

及所有图像上可见点的RMS误差值。可以看出,本文方法和 $L_1 - L_\infty$ 方法的精度相比 $L_2$ 方法有了较大程度的提高,这是由于本文方法和 $L_1 - L_\infty$ 方法多了一个剔除局外点的步骤。虽然只是剔除了少量数据,但是精度却有了很大程度的提高。由于事先剔除了残留局外点,可以基本上认定此时只存在噪声误差。由于最优三角形法的精度和效率都高过 $L_\infty$ 范数方法,所以,对于两视图可见空间点采用最优三角形法进行重建,比利用 $L_\infty$ 范数方法进行重建,精度和效率都会有所提高,虽然在对多视图可见空间点采用加速的 $L_\infty$ 范数只是提高了效率,但是整体的精度还是有所提高,而且两视图可见空间点越多,精度提高的越多。在Dinosaur图像中,两视图可见空间点为2328个,约占总数的46.7%。在Cathedral图像中,两视图可见点为10406个,约占总数的61.3%。由于在Cathedral图像中,两视图可见空间点所占比例较大,且数量也较多,所以精度提高的较多。本文方法中,无论是对于两视图可见空间点的重建,还是对多于两视图可见空间点的重建,效率都有所提高。所以本文方法的效率高过 $L_1 - L_\infty$ 方法。虽然本文方法的总体耗时比 $L_2$ 方法多一些,但是精度却提高了很多。综合比较之下,本文方法的性能最好。两组图像的3维重建结果如图3和图4所示。

表2 3种重建方法的性能

Table 2 The performance of three reconstruction methods

| 图像数据      | 方法               | 重建时间/s | RMS 误差/像素 |
|-----------|------------------|--------|-----------|
| Dinosaur  | $L_2$            | 265    | 1.236 4   |
|           | $L_1 - L_\infty$ | 1 329  | 0.338 5   |
|           | 本文方法             | 681    | 0.295 6   |
| Cathedral | $L_2$            | 654    | 3.180 1   |
|           | $L_1 - L_\infty$ | 4 315  | 0.893 1   |
|           | 本文方法             | 1 935  | 0.726 5   |

由于 $L_\infty$ 方法、最优三角形法和 $L_2$ 方法对局外点都非常敏感,所以需要先将一些残留局外点剔除掉,再进行3维重建。另外,本文方法中将空间点分为两类,第1类为两视图可见空间点,第2类为多视图可见空间点,对于第1类空间点采用最优三角形法可获得最优解,对于第2类空间点,采用 $L_\infty$ 范数方法进行重建, $L_\infty$ 范数方法简化了代价函数,且产生的拟凸函数能够得到全局最优值,虽然 $L_\infty$ 优化准则不是最大似然估计,但可以避免陷入局部最优的危险。



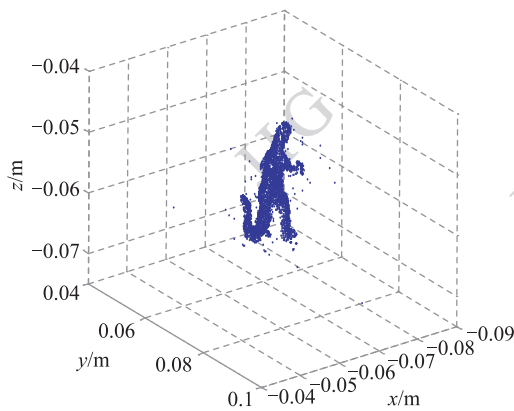


图3 Dinosaur的3维场景重建结果

Fig. 3 3D scene reconstruction of Dinosaur

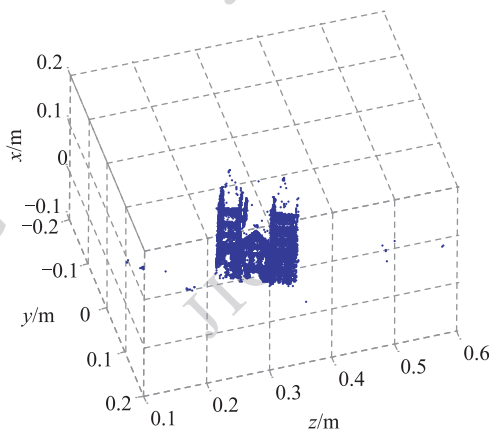


图4 Cathedral的3维场景重建结果

Fig. 4 3D scene reconstruction of Cathedral

## 5 结论

采用一种效率较高的  $L_1$  方法将残留局外点剔除,并考虑到最优三角形法在两幅图像重建中的优势,将场景中的空间点进行分类,对两视图可见空间点采用最优三角形法进行重建,对于多视图可见空间点采用  $L_\infty$  范数方法重建,并从一个子集出发,不断调整二分搜索的上下界,以此减少迭代次数。实验结果表明本文方法精度较高,效率也较高。

## 参考文献 (References)

- [1] Hartley R, Zisserman A. Multiple View Geometry in Computer Vision[M]. Cambridge: Cambridge University Press, 2003.
- [2] Olsson C, Kahl F, Oskarsson M. Branch and bound methods for euclidean registration problems[J]. IEEE Trans. Pattern Analysis and Machine Intelligence, 2009, 31(5): 783-794.
- [3] Kahl F, Agarwal S, Chandraker M K, et al. Practical global optimization for multiview geometry[J]. International Journal of Computer Vision, 2008, 79(3): 271-284.
- [4] Lu F F, Hartley R. A fast optimal algorithm for L2 triangulation [C]//Proceedings of Asian Conference on Computer Vision. Berlin Heidelberg: Springer-Verlag, 2007: 279-288.
- [5] Olsson C, Kahl F, Hartley R. Projective least-squares: global solutions with local optimization[C]//Proceedings of Conference on Computer Vision and Pattern Recognition. Miami: IEEE Computer Society, 2009: 1216-1223.
- [6] Hartley R, Schaffalitzky F. L-infinity minimization in geometric reconstruction problems [C]//Proceedings of Conference on Computer Vision and Pattern Recognition. Washington DC: IEEE Computer Society, 2004: 504-509.
- [7] Kahl F. Multiple view geometry and the L-infinity norm [C]//Proceedings of International Conference on Computer Vision. Beijing: Microsoft Asia Research, 2005: 1002-1009.
- [8] Kahl F, Hartley R. Multiple-view geometry under the L-infinity norm[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(9): 1603-1617.
- [9] Olsson C, Eriksson A. Triangulating a plane [C]//Proceedings of Scandinavian Conference on Image Analysis. Berlin Heidelberg: Springer-Verlag, 2011: 13-23.
- [10] Ke Q, Kanade T. Quasiconvex optimization for robust geometric reconstruction[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(10): 1834-1847.
- [11] Lee H, Seo Y, Lee S W. Removing outliers by minimizing the sum of infeasibilities[J]. Image and Vision Computing, 2010, 28: 881-889.
- [12] Seo Y, Lee H, Lee S W. Outlier removal by convex optimization for L-infinity approaches [C]//The 3rd Pacific Rim Symposium on Advances in Image and Video Technology. Berlin Heidelberg: Springer-Verlag, 2009: 203-214.
- [13] Olsson C, Eriksson A, Hartley R. Outlier removal using duality [C]//Proceedings of Conference on Computer Vision and Pattern Recognition. San Francisco: IEEE Computer Society, 2010: 1450-1457.
- [14] Yu J, Eriksson A, Chin T J, et al. An adversarial optimization approach to efficient outlier removal [C]//Proceedings of International Conference on Computer Vision. Barcelona: IEEE, 2011: 399-406.
- [15] Seo Y D, Hartley R. A fast method to minimize L-infinity error norm for geometric vision problems [C]//Proceedings of International Conference on Computer Vision. Rio de Janeiro: IEEE, 2007: 1-8.