

移动车载视频的立体全景图生成方法

王胜春¹ 罗四维¹ 黄雅平¹ 王亮¹ 戴鹏²

(¹ 北京交通大学交通数据分析与挖掘北京市重点实验室, 北京 100044)
² 中国铁道科学研究院基础设施检测研究所, 北京 100081)

摘要 装载有视频采集设备的巡检列车已被应用于铁路环境监测中。提出了一种从前向运动的列车车载视频中生成铁路环境立体全景图的方法,该方法仅使用一个摄像设备即可获取整个场景的立体全景信息。基于场景的几何先验和标定的相机参数从视频序列中连续地抽取一组梯形窄带作为完全采样拼接区域,窄带的宽度根据车速自动调节,将所有窄带经单应性变换修正后拼接得到全景图像,并据此构建准确的几何模型来生成全景立体像对。实验结果表明,该方法能够快速有效地生成具有立体感的全景图。与原始的视频格式相比,生成的全景图所需存储空间小,具有更广阔的视野和更强的真实感,可实现铁路场景的快速访问和沉浸式浏览。

关键词 机器视觉;立体全景图;条带拼接;车载视频;前向运动;铁路环境

中图分类号 O439; TP391 **文献标识码** A **doi**: 10.3788/AOS201434.1215005

Stereo Panorama Generation from the Train-Borne Video

Wang Shengchun¹ Luo Siwei¹ Huang Yaping¹ Wang Liang¹ Dai Peng²

(¹ Beijing Key Laboratory of Traffic Data Analysis and Mining, Beijing Jiaotong University, Beijing 100044, China)
² Infrastructure Inspection Research Institute, China Academy of Railway Sciences, Beijing 100081, China)

Abstract The patrol train with picture pick-up device has been used for railway environment monitoring. A method to generate the stereo panorama from forward motion train-borne video using only one camera to acquire the panoramic stereo information of the entire scene is presented. On the basis of the geometric priori information of the scene and known camera parameters, a set of trapezoidal stripes is extracted from a video sequence continuously, whose width is automatically adjusted according to the vehicle speed, and the stripes are stitched into a panoramic image after aligning all the stripes by homography transformation. The accurate geometric model is built to generate the panoramic stereo pairs. The experimental results show that the method can effectively and efficiently generate a stereoscopic panorama. Compared with the original video format, the stereo panorama requires less storage space, possesses wider field of view and greater realism, and enables the inspector to perform quick examination and immersive browsing.

Key words machine vision; stereo panorama; stripe mosaic; train-borne video; forward motion; railway environment

OCIS codes 150.0155; 330.1400; 330.7326; 350.4600

1 引 言

近几年来,巡检列车已被广泛使用以确保铁路运行的安全,如意大利的“阿基米德”号综合检测列

车,法国的 IRIS320 综合检测列车,日本的 East-I 新干线综合检测列车以及我国的 0 号高速综合检测列车^[1-2]。通过安装在列车上的视频采集设备获取视

收稿日期: 2014-06-23; **收到修改稿日期**: 2014-07-21

基金项目: 国家自然科学基金(61272354,61273364,61105119,61300176)、教育部博士点基金(20131113)

作者简介: 王胜春(1985—),男,博士研究生,主要从事计算机视觉和全景图生成方面的研究。

E-mail: 09112069@bjtu.edu.cn

导师简介: 罗四维(1943—),博士,教授,博士生导师,主要从事机器学习和视觉感知等方面的研究。

E-mail: swluo@bjtu.edu.cn

本文电子版彩色效果请详见中国光学期刊网 www.opticsjournal.net

频图像,在视频图像上实时叠加线路名、里程和速度等线路信息,并进行存储;之后对视频数据进行人工分析来完成列车运行环境的监测,及时发现影响列车运行安全的各种异常情况。然而,频繁重复的巡检任务将产生海量的视频数据,针对如此巨大的视频数据进行人工分析是一项低效、耗时、劳动力集中的工作,且需要很高的管理花费。除此之外,如何在有限的存储空间中归档每日采集的巡检视频也是一个亟待解决的问题。

三维(3D)立体全景图由于其真实感强、信息量丰富、便于生成、存储量小等优点,提供了一种可行的解决方案。相对于视频数据,立体全景图不仅扩展了空间的成像视野,还包含了空间物体的3D深度信息,而且具有更小的数据格式,对存储空间的要求更小。鉴于上述优势,3D立体全景图技术已广泛应用于基于图像的城市场景绘制以及大场景的环境监测。例如谷歌(google)3D虚拟街景^[3],基于航拍或车载的土地和资源监测管理等^[4-5]。

本文提出了一种从列车车载视频中生成3D立体全景图的方法。在巡检列车前端安装一个正对前方的视频采集设备,以此来获取整个铁路场景的环境信息(包括铁道两侧环境、护栏、电杆、铁道和接触网等)。通过比较和分析现有的方法,建立了立体全景图获取的几何计算模型,并提出了一种适用于前向运动视频的立体全景图生成方法。

2 相关工作

全景图可概括为两类,即单一视点的全景图与多视点的全景图。单一视点的全景图,其成像设备的光学中心是固定不变的,可以通过环绕相机的光学中心旋转拍摄得到。这类单一视点的全景图通常称之为360°全景图,常见的有柱形全景图或球形全

景图^[6-8]。另一类全景图是由平移相机采集的图像序列拼接而成,常用来合成多视点的全景图,如条带全景图^[9-11]和道路全景图^[12]。该方法中,成像设备的光学中心是变化的,相对于前者视点固定的方法,由于视点是变化的,因此全景图的水平视野在空间上可以无限延伸。如图1所示,利用一个平移摄像机获取一系列视频序列,然后在视频序列中的每一帧中抽取一个垂直的窄带,将这些窄带拼合就形成了多视点全景图。

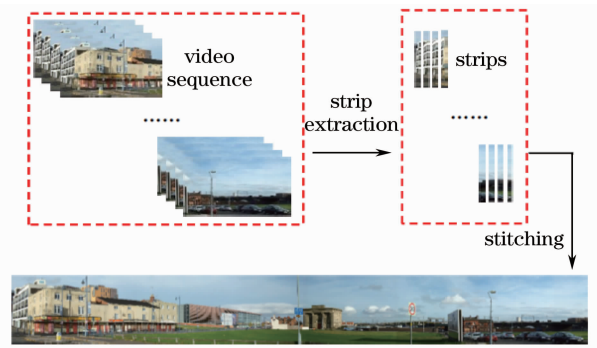


图1 多视点全景图拼接示意图

Fig. 1 Schema of multiple-viewpoint panorama generation

在上述研究的基础上,一些学者提出了立体全景图的生成方法,实现了大场景的3D可视化表示。文献^[13]利用两个绕光轴旋转的相机生成了立体全景图,Peleg等^[14]改进了这一方法,仅利用一个相机就可以生成全局的立体全景图。文献^[15]提出了一种适合于多视点条带全景图的立体信息获取方法,从一个侧向平移的相机中生成了立体全景图像对。文献^[16]提出了一种基于全景环带的立体成像系统,该系统使用两个共轴的环形全景镜头,可以生成两幅具有不同视角的360°全景图。

本研究受到了以上研究的启发,但成像系统的

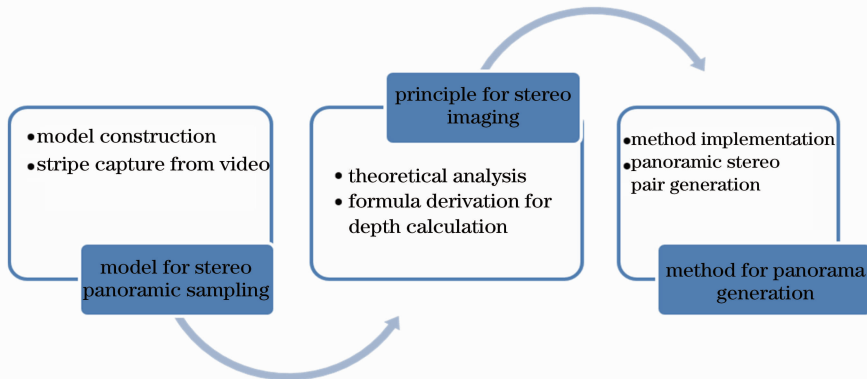


图2 立体全景图生成的流程图

Fig. 2 Flowchart of stereo panorama generation

运动模式不完全相同,生成全景立体像对的几何模型也有很大的区别。当前解有研究工作介绍从前向运动的视频中生成立体全景图像,而这种运动方式是巡检列车执行铁路环境监测任务时的主要图像采集模式。本文提出了一种从列车车载前向运动视频中生成铁路立体全景图的方法,实现铁路场景的3D可视化表示。如图2所示,构建了立体全景采样模型来获取不同宽度的窄带图像,分析了立体全景图成像的原理来推导全景图深度计算公式,提出了全景图拼接方法来生成立体全景图像对。

3 立体全景采样模型

如图3所示,摄像机正对前方移动,即相机的光轴平行于运动方向。对于前向运动的车载成像设备,分别在获取的视频序列帧的边界区域抽取两个拼接窄带,两窄带之间的距离为 d ,则两个窄带可分别表示为 $S_{t,1}(v_1, x_1)$ 和 $S_{t,2}(v_2, x_1 - d)$,其中 v_1 和 v_2 分别为两个窄带的宽度, x_1 和 $x_1 - d$ 分别为所取的两个窄带的位置, t 为当前取到的帧数($t = 1, 2, 3, \dots, N$, N 为视频序列的总帧数)。如图4所示,沿着 t 轴将窄带 $S_{t,1}$ 和 $S_{t,2}$ 按照提出的前向运动视频的全景图生成方法分别进行拼接,可以生成两幅全景图 M_1 和 M_2 (蓝、红虚线所围成的矩形),即全景立体像对。由于这两张全景图具有不同的视角,可以直观地抽象为一个线阵电荷耦合器件(CCD)相机以不同的视角对同一场景的两次扫描。生成的两幅全景图中存在同一空间物理点在两幅全景图中相互对应的两个像素点(这两点称作共轭匹配点),这两点之间的位置差异称之为像素的位置差,简称为像位差(DPL)。像位差和空间深度的关系可以根据之后建立的几何模型推导出来,因此通过图像拼接得

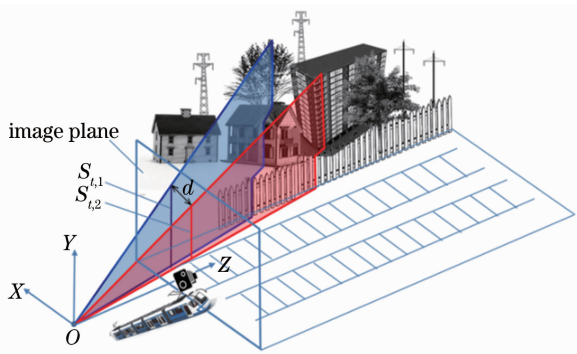


图3 前向运动视频的立体全景采样模型

Fig. 3 Illustration of stereo panoramic sampling from forward motion video

到的两幅全景图存在与深度信息相关的像位差,这两幅全景图构成了全景立体像对,可生成具有立体感的全景图像。

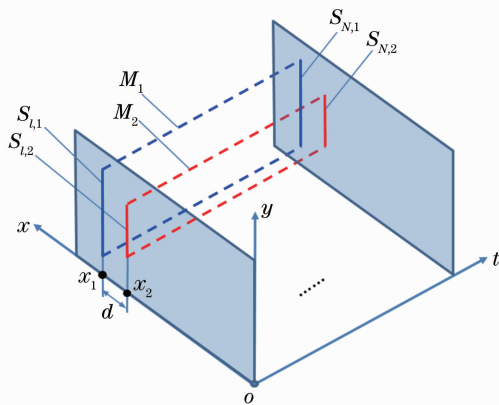


图4 全景立体像对的构造

Fig. 4 Structure diagram of constructing the panoramic stereo pairs

4 立体全景图成像原理

立体成像的几何模型如图5所示,摄像机沿Z轴向前运动,不失一般性,考虑理想情况下,列车运行在平直的轨道上以速度 V 匀速前进,且摄像机没有明显的晃动。令摄像机光心坐标为 $(0, 0, Z_t)$,两条拼接窄带分别为:靠近图像边缘的窄带记作 $S_{t,1}$,宽度为 v_1 ;靠近图像中心的窄带记作 $S_{t,2}$,宽度为 v_2 。当 $t = m$ 时,摄像机位于 $o_m(0, 0, Z_m)$ 处,空间中近处栏杆上的一点 $P(X, Y, Z)$ 被窄带 $S_{t,2}$ 采样到,此处的拼接窄带记作 $S_{m,2}$;而当 $t = n$ 时,摄像机继续前移到 $o_n(0, 0, Z_n)$ 处,该点被窄带 $S_{t,1}$ 采样到,此处的拼接窄带记作 $S_{n,1}$ 。从图5中直观表示为点 $P(X, Y, Z)$ 在第 m 帧中的成像点 $p_m(x, y)$ 随着摄像机的前移而运动到第 n 帧中的成像点 $p_n(x, y)$ 。而空间中远处树木上的一点 $P'(X, Y, Z)$ 在第 m 帧中的成像点 $p'_m(x, y)$ 的移动速度较慢,只有在 $t > n$ 的位置才能到达窄带 $S_{t,1}$ 的采样区域,即 $S_{t,1}$ 在第 n 帧后续的某帧中才能采样到该点。记为 $\Delta Z = Z_n - Z_m$,称为点 $P(X, Y, Z)$ 对应的基线长度。可以看出,深度不同的空间点,对应的基线长度也不同。

令 $p_1(t_n, h_n), p_2(t_m, h_m)$ 分别为 $p_n(x, y), p_m(x, y)$ 在全景图 M_1, M_2 中的位置坐标。则

$$m = \frac{Z_m}{V/R} + 1 = \left(\frac{t_m}{v_2} \right),$$

$$n = \frac{Z_n}{V/R} + 1 = \left(\frac{t_n}{v_1} \right), \quad (1)$$

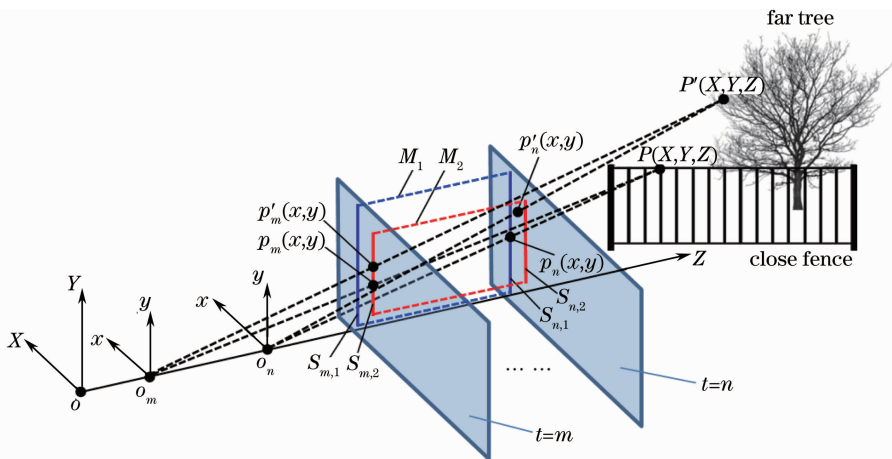


图 5 像位差和深度信息之间的几何模型

Fig. 5 Geometrical model between the DPL and the depth

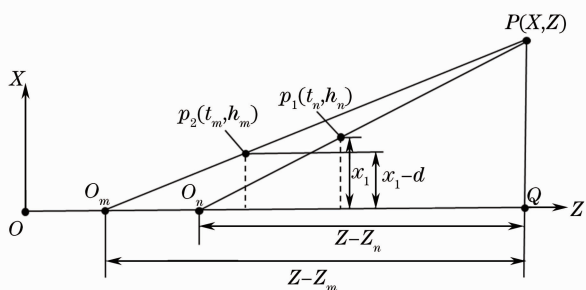


图 6 立体成像几何模型沿 XOZ 平面的切面图

Fig. 6 Sectional drawing of Fig. 5 along the plane XOZ

式中 m, n 分别为摄像机当前采集的帧数, V 为列车(摄像机)行驶的速度, R 为摄像机的采集帧率, V/R 表示两帧之间列车(摄像机)移动的距离。

对于基线长度 $\Delta Z = Z_n - Z_m$, 有

$$\Delta Z = Z_n - Z_m = (n - m)(V/R) = \left[\left(\frac{t_n}{v_1} \right) - \left(\frac{t_m}{v_2} \right) \right] (V/R). \quad (2)$$

如图 6 所示, 从两个视点 O_m 和 O_n 处观察点 $P(X, Y, Z)$ 的视角分别记作 $\alpha = \angle PO_m Q, \beta = \angle PO_n Q$, 则有

$$\tan \alpha = \frac{x_1 - d}{f} = \frac{X}{Z - Z_m},$$

$$\tan \beta = \frac{x_1}{f} = \frac{X}{Z - Z_n}, \quad (3)$$

因此

$$\frac{1}{\tan \alpha} - \frac{1}{\tan \beta} = \Delta Z \frac{1}{X}, \quad (4)$$

整理得

$$X = \frac{1}{\tan(\beta - \alpha)} \Delta Z - \frac{f}{d}. \quad (5)$$

传统的景深定义为物点到像平面的垂直距离, 但对于本文提出的全景成像模型, 如图 5 所示, 全景

图像平面与实际的成像平面相互垂直, 因此不同于传统的景深定义, 这里将物点到全景图像平面的垂直距离定义为全景图深度, 大小为 $|X|$; 两窄带位置 x_1 和 $x_1 - d$ 给定, 相机焦距 f 已知; $\beta - \alpha = \angle O_m P O_n$, 表示视差; ΔZ 表示“基线”长度。和传统的双目视觉的极面几何不同, 在立体成像几何模型中, 由于 $\tan(\beta - \alpha)$ 只与 x_1, d, f 有关, 因此一旦这三个量确定, 视差是固定的且和深度 $|X|$ 无关。而基线长度并不是固定不变的, 如(2)式所示, 它与两幅全景图中的共轭匹配点 $p_1(t_n, h_n)$ 和 $p_2(t_m, h_m)$ 的像位差有关。对于这两点的水平坐标 t_m, t_n , 给定全景图中的一点 $p_2(t_m, h_m)$, 经图像匹配可以在另一全景图中找到与之匹配的点 $p_1(t_n, h_n)$ 。 v_1, v_2 为两拼接窄带的宽度, 在执行全景图拼接时可以求得它们的值。因此, 全景图的深度只与基线长度成比例关系, 一旦求得一对共轭匹配点的像位差, 则由(2)式可求得基线的长度, 再由(5)式可以得到空间点 $P(X, Y, Z)$ 的全景图深度。

5 全景图拼接方法

建立了从前向运动的视频中生成多视点的铁路全景图的几何模型, 通过构造完全采样的拼接区域 S_t , 实现对铁路场景的完整采样。基于完全采样的拼接模型, 可以从视频图像序列中构造两个“完全采样”的拼接区域 $S_{t,1}(v_1, x_1)$ 和 $S_{t,2}(v_2, x_1 - d)$, 用来生成两幅具有显著像位差的全景立体像对, 从而合成具有立体感的铁路全景图。

5.1 “完全采样”拼接模型

图 7 为前向运动的全景拼接模型, 有如下定义:

1) S_t 为从第 t 帧图像抽取的窄带, 图像中的像

素点 $p_{x,y} \in S_t$, 则全景图可记为 $M = \bigcup_{t=1}^N T(S_t)$, 这里 N 为视频序列的总帧数, $T(\cdot)$ 为单应性变换^[17], 在图像拼接中常用来执行图像对齐操作, 以便实现窄带之间的无缝拼接;

2) R_t 为与 S_t 相对应的空间采样区域, $P_{x,y,z} \in R_t$, 为与像素点 $p_{x,y}$ 相对应的空间物点。 FL_t, ML_t, CL_t 分别为空间中不同的侧向深度层, 即 $P_{x_f,y,z} \in FL_t, P_{x_m,y,z} \in ML_t, P_{x_c,y,z} \in CL_t$;

3) 假设 $R_t = ML_t$, 即空间中的采样点全部位于中间层时, 相邻的采样区域之间既没有重叠部分又没有信息丢失, 即完全采样; 若 $R_t = FL_t$, 空间中的采样点全部位于远层, 相邻的采样区域相互重叠, 即

过采样; 若 $R_t = CL_t$, 空间中的采样点全部位于近层, 相邻的采样区域之间有信息丢失, 即欠采样。

对于从实际场景中采样得到的图像窄带 S_t , 由于存在不同的深度层, 因此不可能在每一深度层上都执行完全采样。指定某一固定的深度层作为标准的拼接平面, 仅在该层执行完全采样, 则生成的全景图中, 远处的景物将由于过采样而产生拉伸的效果, 近处的景物将由于欠采样而产生挤压的效果, 这一扭曲现象被称为静态模糊^[18]。对于本文的研究工作, 基于以下几个原则选取护栏所处的平面为固定的深度层:

1) 铁路两旁的护栏和电杆近似位于一个深度层, 且深度值已知;

2) 由于护栏平面为距离相机最近的深度层, 因此可以保证生成的全景图中不存在因欠采样而引起的信息缺失;

3) 护栏和电杆为铁路场景的重要组成部分, 也是铁路环境监测的重点关注区域, 在该层上执行完全采样有利于后期的检测和分析。

如图 8 所示, 构造了保证摄像机在固定的深度层执行完全采样的拼接区域 $S_t(l_1, l_2)$ 。 l_1 为指定的固定位置的外拼接线, 每一帧中的位置坐标都固定为 x_1 , l_1 的选取应该综合考虑场景的视野和分辨率; l_2 为位置不定的内拼接线, 其位置坐标为 $x_2 = x_1 - v$, v 为待求的拼接窄带的宽度。

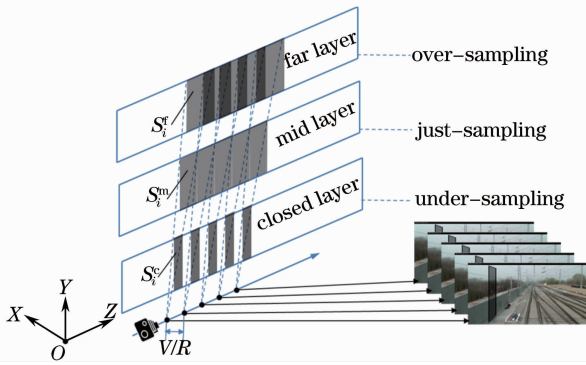


图 7 前向运动的全景拼接模型

Fig. 7 Diagram of the panorama acquisition from forward motion video

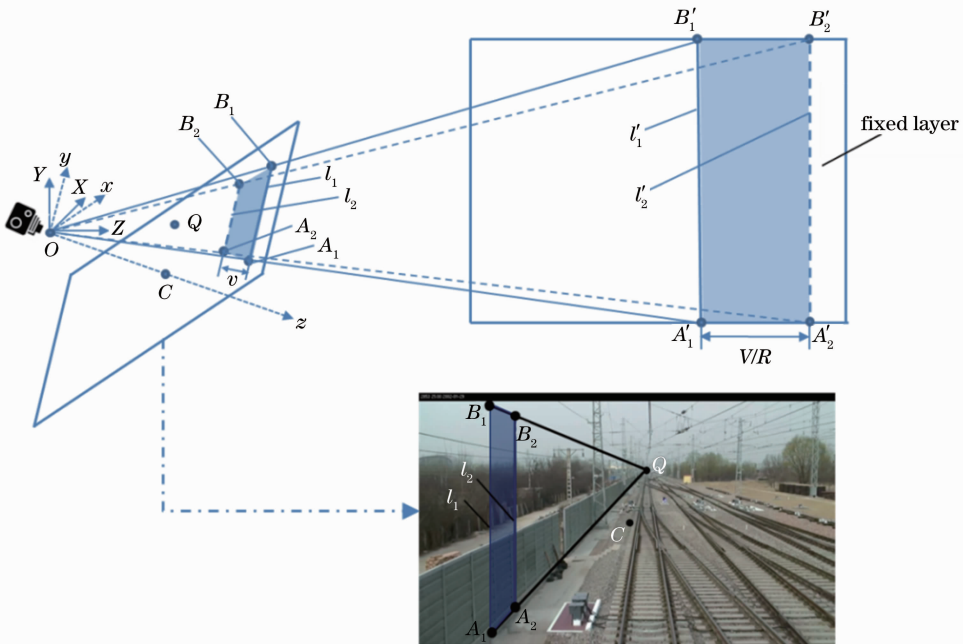


图 8 构造完全采样拼接区域

Fig. 8 Diagram of constructing the just-sampling stitching region

l_1 和 l_2 在固定的深度层的采样线分别为 l'_1 和 l'_2 , 它们构成了空间采样区域 $R_i(l'_1, l'_2)$ 。为了保证相邻两帧之间的采样为完全采样, l'_1, l'_2 之间的距离应等于列车在相邻两帧之间移动的距离, 即 $D_{l'_1 l'_2} = V/R$, V 为列车速度, R 为相机的采集帧率。

列车沿着 Z 轴向前移动, 因此对于空间中平行于 Z 轴的直线, 其在图像平面上的投影必定经过消失点 Q 。如图 8 所示, 线段 A_1A_2 和 B_1B_2 分别为空间中平行于 Z 轴的两线段 $A'_1A'_2$ 和 $B'_1B'_2$ 在图像中的投影, 即 A_1A_2 和 B_1B_2 的延长线汇聚于消失点 Q 。因此, 只要知道消失点 Q 的位置并且求得两条拼接线 l_1 和 l_2 之间的距离 v , 即可得到内拼接线 l_2 在图像中的准确位置。由于列车运行在平直的轨道上, 因此消失点的位置像对稳定。为了使拼接结果更加准确, 采用文献[19]中的算法为每一帧图像计算消失点的位置, 并根据投票规则统计消失点最常出现的位置, 将该点作为全局不变的消失点。 v 直观反应了当列车以速度 V 行驶时, 相邻两帧之间像素平移的值, 称之为图像速度。

5.2 图像速度计算

如图 8 所示, $O-XYZ$ 为世界坐标系, Z 轴为列车以速度 V 运行的方向, 并且 Y 轴垂直于地面。摄像机坐标系 $\sigma-xyz$ 与世界坐标系 $O-XYZ$ 之间存在三个角度, 分别为摄像机绕 x 轴产生的俯仰角 φ , 绕 y 轴产生的平转角 θ 和绕 z 轴产生的旋转角 γ 。列车运行在平直的轨道上(无弯道), 因此获取的视频片段中, 摄像机的旋转角 γ 近似为 0。根据消失点的位置 $Q(x_0, y_0, f)$, 可以估算出摄像机的俯仰角 φ 和平转角 θ ,

$$\varphi = \arctan \frac{y_0}{f}, \quad \theta = \arctan \frac{x_0}{\sqrt{f^2 + y_0^2}}, \quad (6)$$

式中 f 为已标定的摄像机焦距, $C(0, 0, f)$ 为图像的中心点坐标。

给定 l_1 和 l_2 在图像坐标系 $\sigma-xyz$ 下的方程式, 记作

$$\begin{cases} l_1 : A_1(x_1, y_{a1}, f)B_1(x_1, y_{b1}, f) \\ l_2 : A_2(x_2, y_{a2}, f)B_2(x_2, y_{b2}, f) \end{cases}, \quad (7)$$

式中 x_1 已知, 为外拼接线 l_1 的位置坐标, x_2 未知 ($x_2 = x_1 - v$), 为内拼接线 l_2 的位置坐标。如图 8 所示, 一旦直线求得 $l_2 = x_2$, 则它与从消失点 Q 引出的两条射线 QA_1, QB_1 的交点 A_2, B_2 就可求得, 这样四边形 $A_1A_2B_2B_1$ 就构成了完全采样拼接区域。

l_1 和 l_2 在列车坐标系 $O-XYZ$ 下的方程式记作

$$\begin{cases} l_1 : A_1(X_1, Y_{a1}, Z_{a1})B_1(X_1, Y_{b1}, Z_{b1}) \\ l_2 : A_2(X_2, Y_{a2}, Z_{a2})B_2(X_2, Y_{b2}, Z_{b2}) \end{cases}, \quad (8)$$

通过从 $\sigma-xyz$ 到 $O-XYZ$ 的坐标变换, 有

$$\begin{cases} A_1 : (X_1, Y_{a1}, Z_{a1}) = (x_1, y_{a1}, f)\mathbf{M}(\varphi)\mathbf{M}(\theta) \\ B_1 : (X_1, Y_{b1}, Z_{b1}) = (x_1, y_{b1}, f)\mathbf{M}(\varphi)\mathbf{M}(\theta) \\ A_2 : (X_2, Y_{a2}, Z_{a2}) = (x_2, y_{a2}, f)\mathbf{M}(\varphi)\mathbf{M}(\theta) \\ B_2 : (X_2, Y_{b2}, Z_{b2}) = (x_2, y_{b2}, f)\mathbf{M}(\varphi)\mathbf{M}(\theta) \end{cases}, \quad (9)$$

式中 $\mathbf{M}(\varphi)$ 和 $\mathbf{M}(\theta)$ 为旋转矩阵,

$$\mathbf{M}(\varphi) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & \sin \varphi \\ 0 & -\sin \varphi & \cos \varphi \end{bmatrix},$$

$$\mathbf{M}(\theta) = \begin{bmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{bmatrix}. \quad (10)$$

通过 l_1 的视平面 OA_1B_1 的法向量 $\mathbf{N}_1 = OA_1 \times OB_1 = (X_1, Y_{a1}, Z_{a1}) \times (X_1, Y_{b1}, Z_{b1})$, 即 OA_1B_1 的解析式可记为

$$[X \ Y \ Z] \begin{bmatrix} Y_{a1}Z_{b1} - Z_{a1}Y_{b1} \\ Z_{a1}X_1 - X_1Z_{b1} \\ X_1Y_{b1} - Y_{a1}X_1 \end{bmatrix} = 0, \quad (11)$$

这个平面与护栏所在的平面 $X = H$ (H 为摄像中心距护栏平面的距离, 为已知) 相交于 l'_1 , 因此 l'_1 的解析式可记为

$$\begin{cases} [X \ Y \ Z] \begin{bmatrix} Y_{a1}Z_{b1} - Z_{a1}Y_{b1} \\ Z_{a1}X_1 - X_1Z_{b1} \\ X_1Y_{b1} - Y_{a1}X_1 \end{bmatrix} = 0 \\ X = H \end{cases}. \quad (12)$$

通过进一步计算化简(9)式, (12)式可写作

$$\begin{cases} [X \ Y \ Z] \begin{bmatrix} x_1 \sin \varphi - f \cos \varphi \\ (f \sin \varphi + x_1 \cos \varphi) \sin \theta \\ (f \sin \varphi + x_1 \cos \varphi) \cos \theta \end{bmatrix} = 0 \\ X = H \end{cases}, \quad (13)$$

即

$$l'_1 : Z = -\tan \theta \cdot Y + \frac{f \cos \varphi - x_1 \sin \varphi}{(f \sin \varphi + x_1 \cos \varphi) \cos \theta} H, \quad (14)$$

l'_1 在 Z 轴 (Z 轴平行于列车行驶方向) 的截距 Z_1 记为

$$\begin{cases} Z = -\tan \theta \cdot Y + \frac{f \cos \varphi - x_1 \sin \varphi}{(f \sin \varphi + x_1 \cos \varphi) \cos \theta} H \\ Y = 0 \end{cases}, \quad (15)$$

即

$$Z_1 = \frac{f \cos \varphi - x_1 \sin \varphi}{(f \sin \varphi + x_1 \cos \varphi) \cos \theta} H. \quad (16)$$

同理 l'_2 在 Z 轴的截距 Z_2

$$Z_2 = \frac{f \cos \varphi - (x_1 - v) \sin \varphi}{[f \sin \varphi + (x_1 - v) \cos \varphi] \cos \theta} H. \quad (17)$$

l'_1 和 l'_2 之间的距离可表示为

$$D_{l'_1 l'_2} = Z_2 - Z_1. \quad (18)$$

$$Z_2 - Z_1 = \left\{ \frac{f \cos \varphi - (x_1 - v) \sin \varphi}{[f \sin \varphi + (x_1 - v) \cos \varphi] \cos \theta} - \frac{f \cos \varphi - x_1 \sin \varphi}{(f \sin \varphi + x_1 \cos \varphi) \cos \theta} \right\} H = V/R, \quad (20)$$

由此可得

$$v = \frac{V(f \sin \varphi + x_1 \cos \varphi)^2 \cos \theta}{fHR + V(f \sin \varphi + x_1 \cos \varphi) \cos \varphi \cos \theta}. \quad (21)$$

因此,内拼接线的位置可以由(21)式给出,这样由两条拼接线 l_1, l_2 围成的梯形窄带就构成了完全采样拼接区域,且窄带的宽度根据车速自动调节。

对于前向运动的视频序列,由于构造的完全采样拼接区域为一组不规则的梯形窄带图像,相邻两个窄带在图像拼合处的大小尺度并不一致,如果将这些梯形窄带直接拼接在一起,则生成的全景图在竖直方向上存在严重的像素错位。因此,如图9所示,通过单应性变换^[17]将不规则的梯形窄带转换为尺度大小统一的矩形窄带,实现了窄带之间的无缝拼接。

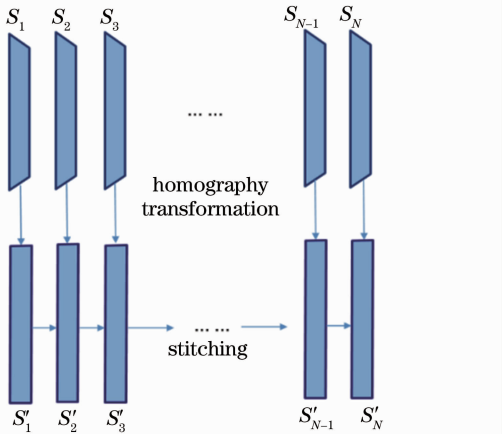


图9 通过单应性变换的图像对齐

Fig. 9 Image alignment using homography transformation

6 实验和讨论

本文的目标是沿着铁路沿线获取完整的铁道环

如5.1节所述,为了保证相邻两帧之间的采样为完全采样,则 l'_1, l'_2 之间的距离应等于列车在相邻两帧之间移动的距离,即

$$D_{l'_1 l'_2} = V/R, \quad (19)$$

式中 V 为列车速度, R 为相机的采集帧率。

综合(18)式和(19)式,有

境立体全景图。整个视频从前向运动的巡检车上采集,为了保证曝光充足,总是保持列车的前灯处于开启状态。图10为列车在平滑轨道上获取的前向运动视频中的一帧,摄像机的采集帧率为25 frame/s。图像的分辨率为1024 pixel×768 pixel,在图像的左右两侧分别构造两个拼接窄带,仅以左侧为例进行说明,蓝色标记的区域称为左侧的外拼接窄带,在图像中的位置为 $x_1 = 480\Delta\gamma$;红色标记的区域称为左侧的内拼接窄带,在图像中的位置为 $x_1 - d = (480 - 60)\Delta\gamma = 420\Delta\gamma$, $\Delta\gamma$ 为使用的成像设备的像素尺寸,大小为 $7.4 \mu\text{m}/\text{pixel}$ 。内外拼接窄带之间的间距 d 设为60 pixel的长度,列车的速度 V 由速度传感器实时记录,因此由(21)式可以得到随车速变化的拼接窄带宽度。

基于上述已知条件,按照第4节中构建的全景图生成模型,可以得到完全采样的拼接窄带,由内外两条窄带拼接产生的全景图分别称为内全景图和外全景图。因此,对于前向运动的铁路场景,可以得到两对全景立体像对,分别生成铁道左右两侧场景的立体全景显示效果。需要注意的是,内外拼接窄带之间的距离 d 的选取应权衡像位差和分辨率的关系,

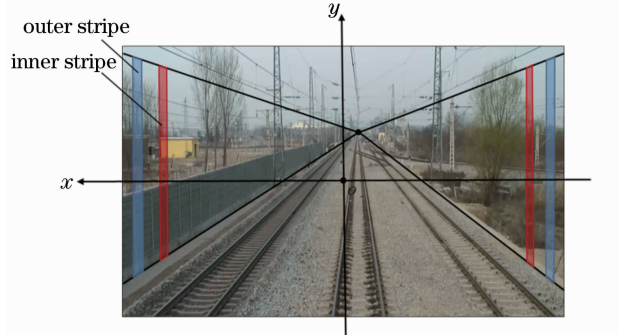


图10 前向运动的视频帧中内外拼接窄带的构造

Fig. 10 Construction of the inner stripe and outer stripe in one frame of forward motion video

若 d 过小,则生成的全景立体像对中像位差变化较小,立体感不明显;若 d 选择过大,则生成的内全景图将会由于分辨率过低而影响立体全景图的合成。

生成的道路全景图的部分结果如图 11 所示。该方法可以生成令人满意的全景图像。以左侧全景图[图 11(a)、(b)]为例,可以清楚地观察到护栏和电杆,没有丢失信息且图像失真较小,远处的景色由于过采样发生了明显的拉伸扭曲。因为远处的场景具有低的图像速度,因此在未来的工作中,可以通过

降低采样率来克服这种失真,以便实现铁路沿线风景的虚拟浏览。

使用的实验视频数据时间间隔接近 1 h,大小为 3024 MB。与原始的 avi 压缩格式相比,产生的全景图具有更小的数据量,jpg 格式的全景图像大小约为原始压缩视频的 1/32。频繁重复的铁路巡检任务将会产生巨大的视频数据量,压缩的数据格式节省了存储空间,且更易于传输到网络中实现数据共享和协同工作。



图 11 铁路场景生成的全景立体像对。(a)铁道左侧环境的外全景图;(b)铁道左侧环境的内全景图;
(c)铁道右侧环境的外全景图;(d)铁道右侧环境的内全景图

Fig. 11 Stereo panorama pairs from railway environment. (a) Outer panorama from the left side of railway environment; (b) inner panorama from the left side of railway environment; (c) outer panorama from the right side of railway environment; (d) inner panorama from the right side of railway environment

图 12(a)为图 11(a)所示的外全景图和图 11 (b)所示的内全景图合成的铁道左侧的红蓝立体图

像,图 12(b)则为图 11(c)和(d)合成的铁道右侧的红蓝立体图像,其中,蓝色区域为外全景图,红色区

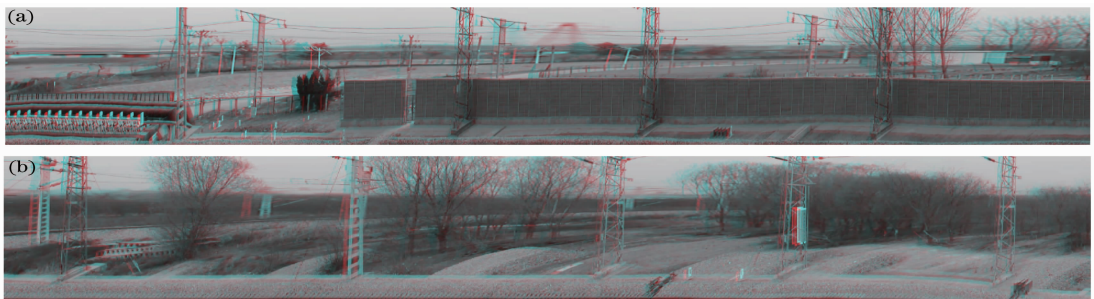


图 12 立体全景图显示(需佩戴红蓝眼镜观看)。(a)铁道左侧;(b)铁道右侧

Fig. 12 Stereo panoramic display (red-blue anaglyph spectacles required). (a) Left side of the railway; (b) right side of the Railway

域为内全景图。可以观察到图像中存在明显的随着全景图深度变化的水平像位差。未来的工作中,可以通过像素匹配的方法求出全景立体像对的水平像位差,并由(5)式求得全景图深度值。

7 结 论

提出了一种从单个车载成像设备中生成铁路环境立体全景图的方法。分析了从前向运动的视频场景中获取立体全景图的成像原理,给出了一种快速准确的从前向运动的车载视频中生成道路全景图的算法来生成全景立体像对。产生的铁路环境全景图是一个完整且无冗余信息的连续图像,与原始视频相比,全景图具有更小的数据量和更加全面直观的显示效果,可以作为视频的索引帮助快速地定位铁路目标,且合成的立体全景图像模拟了真实的3D场景,沉浸感强。

未来的工作主要集中在以下两个方面:1)通过图像匹配计算两个全景立体像对之间的对应关系,求得像位差,从而得到全景图的深度值;2)在得到全景图深度值的基础上,针对远处因过采样而产生的静态模糊的各深度层,可以根据全景图的深度分别对其进行图像修正,从而消除全景图中的失真。

参 考 文 献

- 1 侯卫星. 0号高速综合检测列车[M]. 北京:中国铁道出版社, 2010.
- 2 中国铁道科学研究院. 高速综合检测列车技术交流总结[R]. 2006.
- 3 B Micusik, J Kosecka. Piecewise planar city 3D modeling from street view panoramic sequences [C]. IEEE Conference on CVPR, 2009. 2906-2912.
- 4 Liao Chao, Wang Guijin, Shen Yongling, *et al.*. Aerial video stitchign via multi-direction strips [J]. Journal of Astronautics, 2012, 33(11): 2065-2073.
廖超, 王贵锦, 沈永玲, 等. 航拍视频的多向条带拼接算法[J]. 宇航学报, 2012, 33(11): 2065-2073.
- 5 Z Zhu, A R Hanson, H Schultz, *et al.*. Stereo mosaics from a moving video camera for environmental monitoring [C]. International Workshop on Digital and Computational Video, 1999. 45-54.
- 6 S E Chen. Quick time VR: An image-based approach to virtual

- environment navigation [C]. 22nd annual conference on Computer Graphics and Interactive techniques, 1995. 29-38.
- 7 Wu Xianxiang, Guo Baolong, Wang Juan. Cylindrical panoramic image automatic mosaic algorithm based on phase correlation [J]. Acta Optica Sinica, 2009, 29(7): 1824-1829.
吴宪祥, 郭宝龙, 王娟. 基于相位相关的柱面全景图像自动拼接算法[J]. 光学学报, 2009, 29(7): 1824-1829.
- 8 Wang Huifeng, Liu Shangqian, Wang Dabao, *et al.*. Panoramic image mosaic method for rotary scanning serial image [J]. Acta Optica Sinica, 2009, 29(5): 1211-1215.
王会峰, 刘上乾, 汪大宝, 等. 旋转扫描序列图像的全景图拼接方法[J]. 光学学报, 2009, 29(5): 1211-1215.
- 9 S Peleg, J Herman. Panorama mosaics by manifold projection [C]. IEEE Conference on CVPR, 1997. 338-343.
- 10 S Peleg, B Rousso, A Rav-Acha, *et al.*. Mosaicing on adaptive manifolds [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2000, 22(10): 1144-1154.
- 11 Jiang Chun, Shi Jiaoying, Pan Zhigeng. Rendering using multiple panoramic mosaics of slit images with depth [J]. Journal of Zhejiang University (Engineering Sciece), 2002, 36(2): 162-165.
蒋纯, 石教英, 潘志庚. 基于多深度狭缝图像全景拼图的场景绘制[J]. 浙江大学学报(工学版), 2002, 36(2): 162-165.
- 12 J Y Zheng. Digital route panoramas [J]. IEEE MultiMedia, 2003, 10(3): 57-67.
- 13 H C Huang, Y P Hung. Panoramic stereo imaging system with automatic disparity warping and seaming [J]. Graphical Models and Image Processing, 1998, 60(3): 196-208.
- 14 S Peleg, B E Moshe. Stereo panorama with a single camera [C]. IEEE Conference on CVPR, 1999. 1-8.
- 15 Z Zhu, A R Hanson, E M Riseman. Generalized parallel-perspective stereo mosaics from airborne video [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2004, 26(2): 226-237.
- 16 Tian Yanbing, Bai Jian, Huang Zhi. Depth estimation with a panoramic stereo imaging system [J]. Acta Optica Sinica, 2013, 33(6): 0611002.
田延冰, 白剑, 黄治. 基于全景环带立体成像系统的深度信息估计[J]. 光学学报, 2013, 33(6): 0611002.
- 17 Zhang Chunyu, Wang Wen, Qiu Yate, *et al.*. Algorithm for optimal homography matrix in video mosaic [J]. Journal of Jilin University (Engineering and Technology Edition), 2013, 43(4): 1116-1120.
张春雨, 王文, 邱亚特, 等. 视频拼接中最优自适应单应性矩阵求解算法[J]. 吉林大学学报(工学版), 2013, 43(4): 1116-1120.
- 18 J Y Zheng, M Shi. Scanning depth of route panorama based on stationary blur [J]. International Journal on Computer Vision, 2008, 78(2-3): 169-186.
- 19 P Moghadam, J A Starzyk, W S Wijesoma. Fast vanishing-point detection in unstructured environments [J]. IEEE Transaction on Image Processing, 2012, 21(1): 425-430.

栏目编辑: 张浩佳