

# 车载前向运动视频的实时全景成像方法

蒋欣兰<sup>1</sup>, 王胜春<sup>2</sup>, 罗四维<sup>1</sup>, 黄雅平<sup>1</sup>, 戴 鹏<sup>2</sup>

<sup>1</sup>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;

<sup>2</sup>中国铁道科学研究院基础设施检测研究所, 北京 100081

**摘要** 传统的视频拼接模型和图像对齐方法中,耗时的图像匹配和复杂的光流计算一直是实时拼接的性能瓶颈。提出了一种车载前向运动视频的实时全景成像方法,该方法充分利用物理场景的几何结构和列车的运动信息来构造视频的拼接区域,并建立几何模型实现了拼接区域的对齐。整个拼接避免了复杂的图像处理过程,实现了铁路全景图的实时获取,同时也提供了一种轻量级的视频全景索引方式,降低了视频数据的存储和访问开销。

**关键词** 机器视觉; 全景成像; 实时拼接; 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/AOS201737.0515003

## Real-Time Panoramic Imaging Method for Train-Borne Forward Motion Video

Jiang Xinlan<sup>1</sup>, Wang Shengchun<sup>2</sup>, Luo Siwei<sup>1</sup>, Huang Yaping<sup>1</sup>, Dai Peng<sup>2</sup>

<sup>1</sup>Beijing Key Laboratory of Traffic Data Analysis and Mining, Beijing Jiaotong University, Beijing 100044, China;

<sup>2</sup>Infrastructure Inspection Research Institute, China Academy of Railway Sciences, Beijing 100081, China

**Abstract** In the traditional video mosaic model and image alignment method, the time-consuming image matching and complicated optical flow computation are the performance bottleneck of real-time mosaic. A real-time panorama imaging method is proposed, which makes use of the geometric structure of the physical scene and the motion information of the train to construct the mosaic region of the video and establishes the geometric model to achieve the alignment of the mosaic region. The entire stitching avoids the complex image processing, realizes the railway panorama real-time acquisition and also provides a lightweight video panoramic index method, which reduces the storage and access cost of the video data.

**Key words** machine vision; panoramic imaging; real-time stitching; train-borne video; forward motion; railway environment

**OCIS codes** 150.0155; 330.7326; 350.4600

## 1 引 言

目前利用前向运动相机获取移动式场景来进行场景监控和目标检测,已广泛应用于智能交通、医疗器械、能源勘探和设备检测等领域。如 An 等<sup>[1]</sup>利用前向运动的车载摄像头实现汽车车道偏离自动预警;王胜春等<sup>[2]</sup>同样使用单个前向运动的车载摄像机采集获取视频,进而生成立体全景图像;Nassu 等<sup>[3]</sup>基于改进的尺度不变特征变换(SIFT),从车载相机获取前向运动视频序列实现自动识别铁路沿线的信号和标志。Ou-Yang 等<sup>[4]</sup>使用径向成像胶囊形内窥镜(RICE)获取前向视频,主要利用皮尔逊相关系数等方法设计图像拼接算法,生成活体动物的十二指肠全景图,有效压缩了医学视频的数据规模,从而缩短了医务人员检查图像

**收稿日期:** 2016-10-27; **收到修改稿日期:** 2017-01-09

**基金项目:** 国家 973 计划(2013CB329400)、国家自然科学基金(61272354,61273364,61300176,61473031,61472029)、北京市自然科学基金(4152042)

**作者简介:** 蒋欣兰(1976—),女,博士研究生,主要从事机器视觉和图像处理方面的研究。E-mail: 09112067@bjtu.edu.cn

**导师简介:** 罗四维(1943—),男,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

结果的时间。Hansen 等<sup>[5]</sup>利用搭载单个摄像设备的机器人运动平台,提出了一种单目视觉里程计,即单目测距法,实现了液化气管道的可视化检测;Niu 等<sup>[6]</sup>提出了一种基于马尔科夫随机场(MRF)的最优化理论的管道视频序列的圆柱形拼接模型,不需要相机校准和跟踪,从前向运动视频中生成了可用于地质检测勘探的全景图像。郑戈毅等<sup>[7]</sup>利用顺着炮管内膛移动的 CCD 相机采集炮管内膛的图像,利用图像无缝拼接技术,实现了内膛表面展开图,从而根据火炮管内膛的瑕疵缺陷检测结果对其使用寿命进行客观评价。

前向运动视频数据存储量大,其算法计算复杂度通常无法满足检测任务实时性要求,为视频数据的存储和检索带来困难。学者们主要从两方面解决问题:1)改进算法从而降低算法时间复杂度;2)利用视频序列间存在信息冗余,通过多视点全景图拼接技术去除相邻图像序列的冗余信息,从而压缩和抽取信息,保留下视频内容的关键信息,降低视频数据存储和检索开销。Peleg 等<sup>[8]</sup>提出了一种根据相机运动自适应映射到管道流形的条带拼接方法,并实现了公路场景下前向运动视频的拼接。Zhang 等<sup>[9]</sup>进一步深化了管状流形拼接理论框架以统一图像拼接方法,利用 Möbius 映射详细阐述了前向和后向运动视频的拼接过程。Zheng 等<sup>[10]</sup>利用前向运动的车载鱼眼摄像机,生成了城市道路沿线的场景隧道全景图,并虚拟绘制了城市场景。Kilicarslan 等<sup>[11]</sup>提出了一种道路轮廓的 U 形曲线采样方法,从前向运动的车载视频中生成了公路驾驶的场景体验全景图,从而实现行车环境安全检测。

以往前向运动视频大都是在相对低速运动条件下获取,其拼接算法多基于光流估计和图像匹配的方法,算法时间复杂度高;又假设检测场景无显著深度变化(封闭的检测区域诸如管道,肠胃,炮膛等场景目标全部位于同一平面),相邻的图像之间不存在运动视差(MP),拼接对应关系可以通过选择图像变换模型;而对于一些开阔的场景诸如铁路等,场景深度不单一,场景目标位于不同深度层,且目标之间的相对位置关系(遮挡关系)随摄像机的移动而变化,此时相邻拼接图像之间不满足单应性变换,无法通过图像变换模型直接求取。针对这些问题,本文以铁路车载前向运动视频为研究对象,提出了基于物理场景几何结构和列车运动信息的实时全景成像方法,无损抽取海量的视频数据的信息,获取轻量级的全景图,降低了视频数据的存储和检索开销,视频转化为一种更适合于人工检索的形式。

## 2 环带全景采样模型

### 2.1 完全采样与场景主平面

如图 1 所示,对于一段帧数为  $N$  的前向运动视频, $S_t$  为从第  $t$  帧图像抽取的采样环带, $R_t$  为  $S_t$  在物理空间场景(PS)中的对应采样区域。有如下定义: $S_t$  为环形采样带,由外部拼接环  $OSR_t$  (绿色)和内部采样环  $ISR_t$  (红色)组成; $R_t$  为  $S_t$  在物理空间场景中的采样区域;构造拼接区域:从视频序列中提取一组环带序列  $\{S_1, S_2, S_3, \dots, S_N\}$ , 满足  $R_{t-1} \cap R_t = \emptyset$ , 表示相邻空间采样区域不重合; $\bigcup_{t=1}^N R_t = PS$ , 表示各采样区

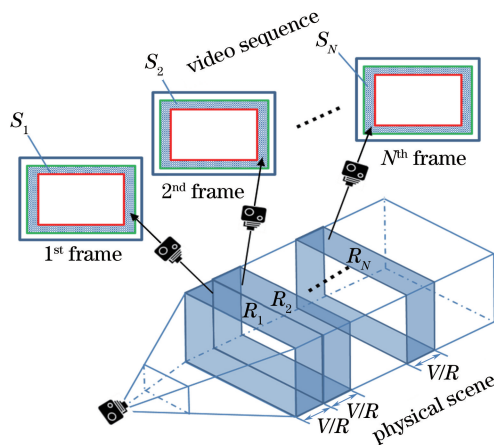


图 1 前向运动视频的环带全景采样模型

Fig. 1 Panoramic annular band sampling model for forward looking video stitching

域之间无间隔,实现对物理场景的全覆盖。全景拼接  $M = \dot{\bigcup}_{t=1}^N T(S_t)$ ,  $T(\cdot)$  为图像几何校正。

基于以上定义,前向运动视频的全景拼接可以描述为从视频序列中提取一组环带序列  $\{S_1, S_2, S_3, \dots, S_N\}$ , 并满足任意相邻的空间采样区域之间既不重叠也无间隔,称之为环带序列对物理场景的完全采样,而场景中执行完全采样的平面就称作场景主平面,将满足条件的环带序列进行几何修正后合并在一起就生成了全景图像。因此,从前向运动的视频中生成全景图的关键问题在于如何在每帧图像中确定外部拼接环  $OSR_t$  (绿色) 和内部拼接环  $ISR_t$  (红色) 的位置。

### 2.2 基于检测区域几何结构确定外部拼接环

之前的研究工作中,为了最大化空间采样信息,要求采样线的构造垂直于像素光流的方向。但这一操作不仅增加了时间开销,而且也影响了生成全景图的分辨率。更重要的是,光流的计算易受环境干扰,对于环境多变的铁路场景,许多情况下由于不能满足光流假设条件即空间一致性、亮度恒定、局部小运动而使得计算结果失真。

如图 2 所示,大多数的实际应用场景都具有固定的几何结构。铁路场景由两侧的护栏电杆,上方的接触网以及地面轨道组成,可以看作一个长方体的结构;隧道场景为拱形的顶部,两侧的墙壁和路面构成的半圆柱几何体;石油运输管道则为弧形管壁包络成的圆柱体。因此,基于场景的检测区域特定的几何形状来构造外部拼接环,避免了复杂的光流计算,提升了拼接效率和拼接稳定性。为了保证全景图具有最佳的分辨率,采样环的位置应位于具有较大空间分辨率的图像边缘处。但列车高速运行条件下,往往导致采集的图像中产生显著的运动模糊,这就要求对分辨率和运动模糊进行折中处理,权衡二者之间的关系。外部拼接环的形状和位置一旦确定,就不再发生变化,即每一帧都具有固定形状和位置的外部拼接环。

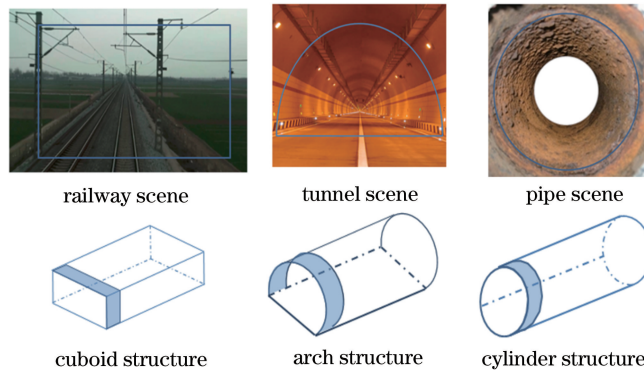


图 2 检测区域的几何结构

Fig. 2 Geometric structure of detected area

### 2.3 基于几何先验的快速光流估计方法确定内部拼接环

如图 3 所示,外部拼接环确定后,就需要进一步求取内部拼接环的位置。不失一般性,以铁路左侧场景为例构造完全采样的拼接区域。这里  $t=1, 2, 3, \dots, N$  为视频帧数,  $A_t B_t$  为第  $t$  帧中外部拼接环的左拼接线,其沿  $x$  轴方向的位置坐标为  $x_1$ ;  $A'_t B'_t$  为内拼接线,其沿  $x$  轴方向的位置坐标为  $x_2 = x_1 - w$ , 这里  $w$  为拼接区域的条带宽度。  $A_t B_t$  和  $A'_t B'_t$  在物理空间的采样线段为  $\bar{A}_t \bar{B}_t$  和  $\bar{A}'_t \bar{B}'_t$ , 构成了空间采样区域  $R_t (\bar{A}_t \bar{B}_t, \bar{A}'_t \bar{B}'_t)$ 。为了使两帧之间的采样满足完全采样,相邻帧之间列车驶过的路程应等于空间采样距离  $W$ , 即

$$W = V/R, \quad (1)$$

式中  $V$  为列车的行驶速度,  $R$  为相机的拍摄帧率。具体来说,就是第  $m-1$  帧中内部采样线  $A'_{m-1} B'_{m-1}$  的位置必须保证在相机运行期间,相邻两帧图像中的环带采样的空间区域  $R_{m-1} (\bar{A}_{m-1} \bar{B}_{m-1}, \bar{A}'_{m-1} \bar{B}'_{m-1})$  和  $R_m (\bar{A}_m \bar{B}_m, \bar{A}'_m \bar{B}'_m)$  之间既不重合又无间隔,即  $\bar{A}'_{m-1} \bar{B}'_{m-1}$  与  $\bar{A}_m \bar{B}_m$  重合,表现在图像中则为第  $m$  帧中的外拼接线  $A_m B_m$  与第  $m-1$  帧中的内拼接线  $A'_{m-1} B'_{m-1}$  相互匹配,是对空间中的同一位置在不同视角下的采样。因此,在确定了外拼接线  $A_m B_m$  后,通过特征匹配求得相邻帧之间的单应性矩阵  $H$ , 可反求外拼接线  $A_m B_m$

在上一帧中的位置  $A'_{m-1}B'_{m-1}$ , 记作

$$A'_{m-1}B'_{m-1} = \mathbf{H}^{-1} \cdot A_m B_m. \quad (2)$$

但图像匹配过程较为耗时, 且场景中存在变化的深度, 相邻帧之间图像匹配的准确性受运动视差影响<sup>[12-14]</sup>。列车运行相对平稳, 一定时间内相机的姿态较为固定, 相邻帧之间的图像变换可近似为相似变换。因此, 通过计算图像的光流<sup>[15]</sup>, 可以得到外拼接线上像素移动的速度和方向, 按照光流的大小和方向可以求得上一帧中内部拼接环的位置, 记作

$$A'_{m-1} = A_m - (u_{A_m}, v_{A_m}), \quad B'_{m-1} = B_m - (u_{B_m}, v_{B_m}), \quad (3)$$

式中  $(u_{A_m}, v_{A_m})$  和  $(u_{B_m}, v_{B_m})$  分别代表点  $A_m$  和  $B_m$  处的光流, 即图像中像素的移动方向和大小。

之前的光流计算模型由于易受环境影响且计算复杂度高, 造成拼接效率低且结果不稳定。本文针对具体的铁路应用场景, 提出了基于场景几何结构的光流估计方法, 避免了各种光流计算中涉及到的频域分析、梯度微分计算或者复杂的图像匹配, 仅仅通过建立场景的几何模型实现光流的快速计算。

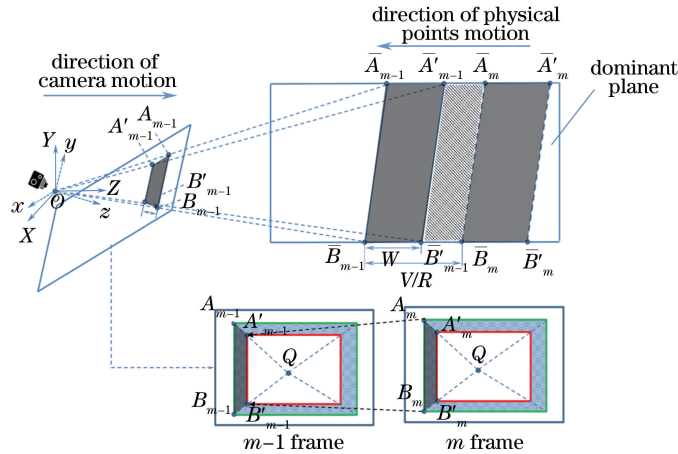


图 3 构造完全采样拼接区域

Fig. 3 Construction of just-sampling stitching area

光流的反向扩展交点(FOE)可以看作是空间中的物体沿着运动方向投影到图像中会聚而成的交点, 而图像消失点是空间中的平行线投影到图像中会聚而成的焦点。若空间中有一组线段平行于运动方向, 则 FOE 与消失点  $Q$  在图像中重合, 而这类平行线段在许多场景里广泛存在, 如图 4 所示。基于这一基本设定, 提出了基于场景消失点和列车速度的光流计算方法。令  $P$  位于第  $m$  帧的外拼接线  $A_m B_m$  上, 则记  $P$  点的光流为

$$P_{\text{flow}} : (v, \mathbf{QP}), \quad (4)$$

式中  $v$  为与列车车速有关的图像中的像素移动距离, 表示光流的大小;  $\mathbf{QP}$  为从消失点  $Q$  指向外拼接线  $A_m B_m$  上的  $P$  点的方向矢量, 表示光流方向。



图 4 场景中平行于运动方向的结构线

Fig. 4 Structure line parallel with direction of motion

如铁路场景中, 前置车载相机随铁轨向前移动, 相机运动方向与铁轨延伸方向同向, 空间中所有平行于铁轨延伸方向的直线, 投影到图像中都会聚于消失点  $Q$ 。如图 3 所示, 第  $m$  帧图像中的外拼接线即为在空间中的采样点  $\bar{A}_m$  和  $\bar{B}_m$  形成的外采样线, 其运动方向如平行于铁轨, 则它们在图像平面中的投影点  $A_m$



和  $B_m$  所形成的外拼接线的光流方向一定聚集经过消失点  $Q$ 。因此,确定消失点  $Q$  的位置,就可以求得场景主平面中每一像素点的光流方向,而光流的大小可以由外拼接线  $A_{m-1}B_{m-1}$  和内拼接线  $A'_{m-1}B'_{m-1}$  之间的距离  $v$  得到。列车在平直轨道运行时间长,运行在弯道、坡道和隧道场景时间短,所以消失点的位置确定之后相对稳定。为使拼接结果更准确,对每帧图像的消失点位置进行估算,并投票选出集中聚集消失点的位置,称该点为不变的全局消失点。 $v$  表示相邻图像帧之间像素平移的大小,与列车行驶车速  $V$  密切相关,将其称为像素的图像速度(IV)。

### 3 前向运动视频的拼接方法

为了解决传统视频拼接方法中的时效性问题,基于场景的消失点以及列车的运行速度提出了前向运动视频的拼接方法。如图 5 所示,相邻帧之间的像素运动方向  $n$  可由场景的消失点  $Q$  来确定,而像素之间的位移大小  $v$  与列车的运动速度  $V$  密切相关。因此,相邻两帧图像中对应于空间中同一位置的投影,可以通过  $(n, v)$  实现帧之间的像素对齐。

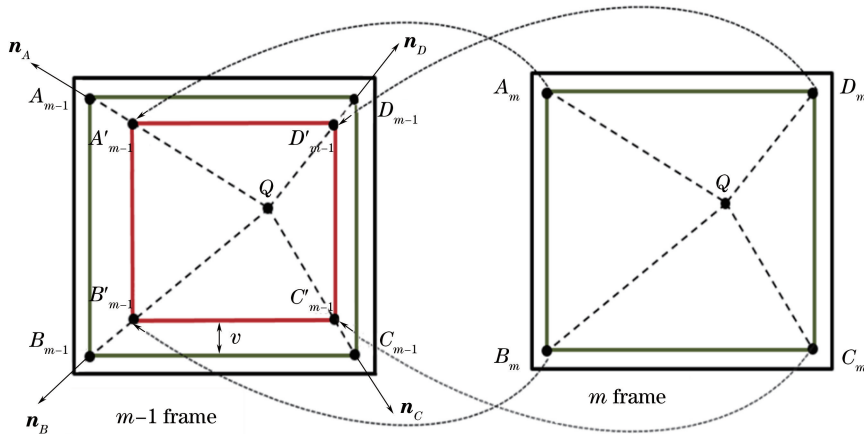


图 5 相邻两帧之间的图像对齐

Fig. 5 Image alignment between the adjacent two frames

铁路环境监测关注的目标集中在封闭的铁路场景内,包括接触网、护栏、声屏障、钢轨及电务设备等。因此,只需在包含这些铁路基础设施的平面上执行完全采样即可生成理想的铁路场景全景图。具体拼接流程如下:

1) 给定一段视频序列  $\{I_1, I_2, \dots, I_{N-1}, I_N\}$  ( $N$  为视频帧数), 通过确定外部矩形  $OR_{m-1} : A_{m-1}B_{m-1}C_{m-1}D_{m-1}$  以及内部矩形  $IR_{m-1} : A'_{m-1}B'_{m-1}C'_{m-1}D'_{m-1}$  在图像帧  $I_{m-1}$  中构造完全采样的拼接区域

2) 在图像中经消失点  $Q$  分别引出四条穿过两侧电杆顶部和底部的射线,并将铁路场景划分为左侧、右侧、天空以及轨道 4 部分。根据场景的几何结构确定外部矩形  $OR_{m-1} : A_{m-1}B_{m-1}C_{m-1}D_{m-1}$  的形状和位置, 4 个顶点  $A_{m-1}, B_{m-1}, C_{m-1}, D_{m-1}$  分别落在 4 条射线上,且在每一帧中都保持固定不变。摄像机的运动方向平行于钢轨的延伸方向,图像中各像素点光流的 FOE 将会聚于消失点  $Q$  处。因此,4 个顶点  $A_{m-1}, B_{m-1}, C_{m-1}, D_{m-1}$  的像素移动方向即  $OR_{m-1}$  在相邻帧之间的缩放方向可以由消失点  $Q$  来确定,记作

$$\begin{cases} \mathbf{n}_{m-1}^A : A_{m-1}(x_a, y_a) - Q_{m-1}(x_q, y_q) \\ \mathbf{n}_{m-1}^B : A_{m-1}(x_b, y_b) - Q_{m-1}(x_q, y_q) \\ \mathbf{n}_{m-1}^C : A_{m-1}(x_c, y_c) - Q_{m-1}(x_q, y_q) \\ \mathbf{n}_{m-1}^D : A_{m-1}(x_d, y_d) - Q_{m-1}(x_q, y_q) \end{cases}; \quad (5)$$

3) 列车的运行速度  $V$  和相机的采集帧速  $R$  为已知,则相机在相邻两帧的间隔内采样的空间距离可记为  $V/R$ 。若相机的内部参数已知,则通过建立相应的几何计算模型即可求得图像中像素的移动速度即图像速度  $v$ , 详见 3.2 节;

4) 与传统的基于图像匹配和光流计算的图像对齐方法不同,第  $m$  帧中外部矩形  $OR_m : A_mB_mC_mD_m$  在

第  $m - 1$  帧中对应的坐标位置  $IR_{m-1}:A'_{m-1}B'_{m-1}C'_{m-1}D'_{m-1}$  可由像素移动方向  $\mathbf{n}_{m-1} = (\mathbf{n}_{m-1}^A, \mathbf{n}_{m-1}^B, \mathbf{n}_{m-1}^C, \mathbf{n}_{m-1}^D)$  以及像素图像速度  $v$  直接求得;

5) 内部矩形  $IR_{m-1}$  和外部矩形  $OR_{m-1}$  构成了第  $m - 1$  帧的环状拼接区域  $S_{m-1}$ , 沿着像素缩放方向  $\mathbf{n}_{m-1}^A, \mathbf{n}_{m-1}^B, \mathbf{n}_{m-1}^C, \mathbf{n}_{m-1}^D$  可以将  $S_{m-1}$  划分为 4 个梯形条带区域, 记作  $S_t, S_b, S_l, S_r$ ;

6) 利用图像扭曲变换(IW)<sup>[16]</sup>对梯形条带  $S_t, S_b, S_l, S_r$  进行几何校正, 将其校正为规则的矩形条带  $S'_t, S'_b, S'_l, S'_r$ ;

7) 每帧图像都重复执行上面的步骤, 可以生成 4 组条带序列, 将其分别依次拼合在一起就生成了铁路场景 4 个面的全景图。

因此, 场景的消失点估计以及像素的图像速度计算是实现本文提出的前向运动视频的拼接方法的关键。

### 3.1 基于消失点检测的场景划分

成像设备装载在列车上, 且运动方向平行于轨道延伸方向, 因此图像中像素运动方向的 FOE 将会聚于两条钢轨在图像上的交点即消失点处。如图 6 所示, 利用直线检测算法提取铁路场景中诸如钢轨、电杆、边缘等线段成分来确定场景的消失点<sup>[17]</sup>, 并利用投票规则筛选消失点的常见位置, 将该位置记作不变的全局消失点。



图 6 基于直线检测确定场景消失点。(a) 视频中随机选取的一帧; (b) 直线检测结果; (c) 消失点检测结果

Fig. 6 Locating vanishing point based on the straight line detection. (a) One frame randomly chosen from the video; (b) line detection result; (c) vanishing point detection result

如图 7 (a) 所示, 经消失点  $Q$  引出 4 条穿过左右两侧电杆顶部和底部的射线, 将图像划分为垂直方向的左右两侧以及水平方向的天空和地面 4 个区域, 拼接区域由内部矩形和外部矩形围成。选取外部矩形的位置应综合考虑图像分辨率和运动模糊, 平衡二者对全景图拼接效果的影响, 如图 7 (b) 所示, 按照虚线框确定的外部矩形将导致生成的全景图分辨率过低。计算内部矩形的位置必须确保其与外部矩形构成的拼接区域实现对空间场景的完全采样, 即采样的内容既无重合又无间隔。

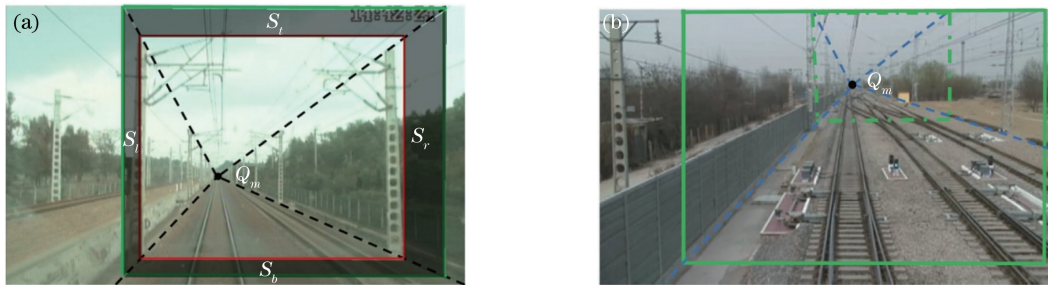


图 7 铁路场景中的矩形拼接区域。(a) 封闭的铁路环境; (b) 宽阔的道岔环境

Fig. 7 Rectangle stitching area in the railway scene. (a) Closed railway environment; (b) wide railroad switch

### 3.2 图像速度计算

通过消失点估计确定了像素的移动方向后, 为了确保无缝拼接, 需要根据列车的运行速度  $V$  来推导像素的图像速度  $v$ , 以此确定内部矩形的位置。如图 8 所示, 建立世界坐标系  $O-XYZ$ , 其中  $Z$  轴正方向为列车的运行方向,  $Y$  轴垂直于轨道平面。  $o-xyz$  为摄像机坐标系且与  $O-XYZ$  的坐标轴之间存在三个夹角, 包括

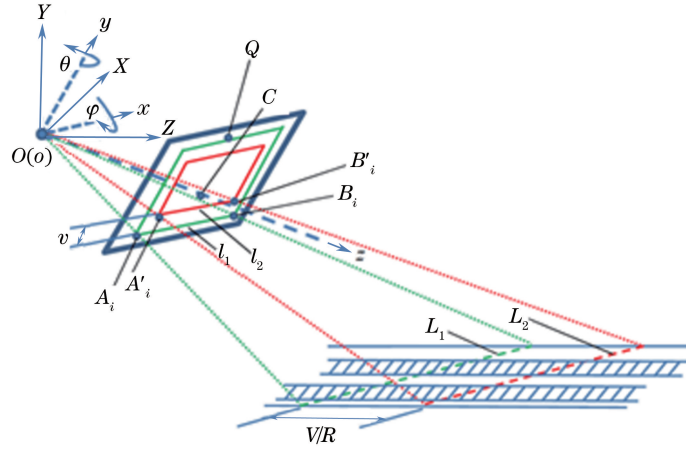


图 8 图像速度  $v$  和列车速度  $V$  之间的关系

Fig. 8 Variation of image velocity  $v$  with train speed  $V$

相机绕  $x$  轴上下俯仰产生的角  $\varphi$ , 绕  $y$  轴左右平转产生的角  $\theta$  和绕  $z$  轴旋转产生的角  $\gamma$ 。由于列车大部分时间都运行在相对平直的轨道上, 摄像机的旋转角  $\gamma$  可近似为 0。此外, 为了使相机的视野尽可能地覆盖整个铁路场景, 在设备安装到车体时就预设了两个较小的俯仰角  $\varphi$  和平转角  $\theta$ 。

根据消失点的位置坐标  $Q(x_0, y_0, f)$  和图像的中心点坐标  $C(0, 0, f)$ , 可以估算出摄像机的俯仰角  $\varphi$  和平转角  $\theta$  分别为

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

式中  $f$  为已标定的摄像机焦距。

第  $i$  帧图像中外拼接线  $l_1$  和内拼接线  $l_2$  在坐标系  $o\text{-}xyz$  下的方程式为

$$\begin{cases} l_1 : A_i(x_{a1}, y_1, f) B_i(x_{b1}, y_1, f) \\ l_2 : A'_i(x_{a2}, y_2, f) B'_i(x_{b2}, y_2, f) \end{cases}, \quad (7)$$

式中  $y_1$  是  $l_1$  的位置坐标且已知,  $y_2$  是内拼接线  $l_2$  的位置坐标但未知 ( $y_2 = y_1 + v$ )。如图 8 所示, 一旦求得直线  $l_2 = y_2$  的解析式, 就可求得其与射线  $QA_i$ 、 $QB_i$  的交点  $A'_i$ 、 $B'_i$ , 四边形  $A_i A'_i B'_i B_i$  就构成了轨道平面的完全采样拼接区域。

外拼接线  $l_1$  和内拼接线  $l_2$  在坐标系  $O\text{-}XYZ$  下的方程式为

$$\begin{cases} l_1 : A_i(X_{a1}, Y_1, Z_{a1}) B_i(X_{b1}, Y_1, Z_{b1}) \\ l_2 : A'_i(X_{a2}, Y_2, Z_{a2}) B'_i(X_{b2}, Y_2, Z_{b2}) \end{cases} \quad (8)$$

建立从  $o\text{-}xyz$  到  $O\text{-}XYZ$  的坐标系变换关系, 有

$$\begin{cases} A_i : (X_{a1}, Y_1, Z_{a1}) = (x_{a1}, y_1, f) \mathbf{M}(\varphi) \mathbf{M}(\theta) \\ B_i : (X_{b1}, Y_1, Z_{b1}) = (x_{b1}, y_1, f) \mathbf{M}(\varphi) \mathbf{M}(\theta) \\ A'_i : (X_{a2}, Y_2, Z_{a2}) = (x_{a2}, y_2, f) \mathbf{M}(\varphi) \mathbf{M}(\theta) \\ B'_i : (X_{b2}, Y_2, Z_{b2}) = (x_{b2}, y_2, f) \mathbf{M}(\varphi) \mathbf{M}(\theta) \end{cases}, \quad (9)$$

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

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

经过  $l_1$  的视平面  $O A_1 B_1$  的法向量为  $\mathbf{N}_1 = OA_1 \times OB_1 = (X_{a1}, Y_1, Z_{a1}) \times (X_{b1}, Y_1, Z_{b1})$ , 则  $O A_1 B_1$  的解析式可记为

$$[X \ Y \ Z] \begin{bmatrix} Y_1 Z_{b1} - Z_{a1} Y_1 \\ Z_{a1} X_{b1} - X_{a1} Z_{b1} \\ X_{a1} Y_1 - Y_1 X_{b1} \end{bmatrix} = 0. \quad (11)$$

$L_1$  为轨道平面  $Y = -H$  ( $H$  为相机的安装高度) 与视平面  $O A_1 B_1$  的交线, 则  $L_1$  的解析式可记为

$$\begin{cases} [X \ Y \ Z] \begin{bmatrix} Y_1 Z_{b1} - Z_{a1} Y_1 \\ Z_{a1} X_{b1} - X_{a1} Z_{b1} \\ X_{a1} Y_1 - Y_1 X_{b1} \end{bmatrix} = 0 \\ Y = -H \end{cases} \quad (12)$$

进一步对(9)式进行推导化简, 结合(12)式有

$$\begin{cases} [X \ Y \ Z] \begin{bmatrix} (y_1 \cos \varphi - f \sin \varphi) \sin \theta \\ -(y_1 \sin \varphi + f \cos \varphi) \\ (y_1 \cos \varphi - f \sin \varphi) \cos \theta \end{bmatrix} = 0 \\ Y = -H \end{cases}, \quad (13)$$

即

$$L_1: Z = -\tan \theta \cdot X + \frac{f \cos \varphi + y_1 \sin \varphi}{(f \sin \varphi - y_1 \cos \varphi) \cos \theta} \cdot H, \quad (14)$$

$L_1$  在  $Z$  轴( $Z$  轴平行于列车运行方向)上的截距  $Z_1$  为

$$\begin{cases} Z = -\tan \theta \cdot X + \frac{f \cos \varphi + y_1 \sin \varphi}{(f \sin \varphi - y_1 \cos \varphi) \cos \theta} \cdot H \\ X = 0 \end{cases}, \quad (15)$$

即

$$Z_1 = \frac{f \cos \varphi + y_1 \sin \varphi}{(f \sin \varphi - y_1 \cos \varphi) \cos \theta} \cdot H. \quad (16)$$

同理  $L_2$  在  $Z$  轴的截距为

$$Z_2 = \frac{f \cos \varphi + (y_1 + v) \sin \varphi}{[f \sin \varphi - (y_1 + v) \cos \varphi] \cos \theta} \cdot H, \quad (17)$$

$L_1$  和  $L_2$  之间的直线距离为

$$D_{L_1 L_2} = Z_2 - Z_1. \quad (18)$$

为了实现相邻帧之间的无缝拼接, 则采样距离  $D_{L_1 L_2}$  应满足完全采样, 即  $D_{L_1 L_2}$  大小应等于列车在相邻两帧的时间间隔移动的距离, 记作

$$D_{L_1 L_2} = V/R, \quad (19)$$

式中  $V$  为列车速度,  $R$  为相机帧率。

综合(18)、(19)式, 有

$$Z_2 - Z_1 = \left\{ \frac{f \cos \varphi + (y_1 + v) \sin \varphi}{[f \sin \varphi - (y_1 + v) \cos \varphi] \cos \theta} - \frac{f \cos \varphi + y_1 \sin \varphi}{[f \sin \varphi - y_1 \cos \varphi] \cos \theta} \right\} H = V/R, \quad (20)$$

由此可得

$$v = \frac{(f \sin \varphi - y_1 \cos \varphi)^2 HR \cos \theta \cdot V}{fHR + (f \sin \varphi - y_1 \cos \varphi) \cos \varphi \cos \theta \cdot V}. \quad (21)$$

由(21)式的计算结果可得到内拼接线  $l_2$  的位置, 如此  $l_1$  和  $l_2$  就构成了梯形状的完全采样拼接区域, 且条带的宽度  $v$  随列车的运行速度变化。图像速度  $v$  和列车的速度  $V$  之间的关系如图 9 所示, 随着车速的提高, 图像速度收敛到一个固定的值。

图像速度和消失点分别确定了光流的大小和方向, 并可进一步得到内部矩形的位置。如图 8 所示, 根据图像速度  $v$  确定了内部矩形的底边  $l_2$  后,  $A_i$  和  $B_i$  沿着从消失点引出的射线方向即光流方向与  $l_2$  分别相交于点  $A'_i$  和  $B'_i$ , 四边形  $A_i A'_i B'_i B_i$  构成了底部拼接区域  $S_b$ , 同理可以构建左右两侧的拼接条带  $S_l$ 、 $S_r$  以及顶部条带  $S_t$ 。

为了提升拼接的实时性,  $v$  和  $V$  之间的变化关系预先通过几段不同车速条件采集的视频样本进行标定, 并建立图像速度查询表来存储标定结果。之后的拼接过程中, 只需根据列车的速度执行查表操作就可得到



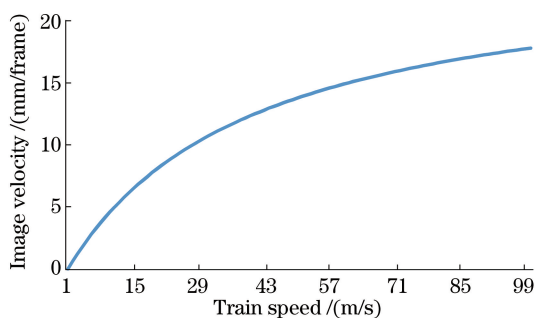


图 9 列车速度与轨道上的图像速度之间的关系( $\theta, \varphi = 0$ )

Fig. 9 Relationship between the train speed and image velocity at the rail track for a forward camera ( $\theta, \varphi=0$ )

拼接位置,避免了对每一帧执行重复的图像速度计算。

## 4 实验和讨论

### 4.1 实验数据

图 10 所示为车载相机采集的三段实验视频数据,列车在采集过程中保持相对匀速,相机帧率为 25 frame/s。三段视频分别记录了铁路直道环境、道岔场景以及列车过弯道的场景,详细数据信息如表 1 所示。



图 10 前向运动视频中的一帧。(a)封闭环境;(b)宽阔的铁路道岔;(c)弯道

Fig. 10 Video frames captured in the forward direction. (a) Closed environment; (b) wide railway switch; (c) curve road

表 1 视频数据详细信息

Table 1 Detailed information of video data

Route section	Speed / (km/h)	Resolution / (pixel×pixel)	Code stream / (kb/s)	Video size / (avi, MB)
Changsha South— Xinyu North(50 km)	150	720×576	1536	137
Huaihua—Loudi (50 km)	50	1280×720	1536	572
Loudi—Zhuzhou (50 km)	50	1280×720	3072	1145

### 4.2 全景图生成的结果

图 11 为长沙南—新余北路段的全景图生成结果。可以观察到对于高速条件下拍摄的低分辨率视频,实验上获得了令人满意的拼接效果。虽然远处的电杆因为采样率过高影响光流预测而发生了横向的扭曲拉升,但是铁路环境的监测目标如近处的护栏和电杆处,由于执行完全采样并没有丢失信息且失真较小。此外,倘若关注的监测目标位于远处的场景,则可以通过调节拼接条带的宽度来克服这种拉升变化。

道岔场景较为宽阔,包括丰富的铁路设施如轨道、电杆、声屏障、电务设备和标志牌。图 12 所示为怀化—娄底路段道岔场景的全景图生成结果,可以观察到图中各设施目标清晰可见,效果逼真。因此,利用图像处理算法对视频全景图进行自动的目标检测和识别,可以将复杂耗时的视频分析转化为简单快速的图像处理过程,实现铁路环境的自动化监测。

图 13 为娄底—株洲段列车过弯道时的全景图生成结果。由于弯道场景图像在拼接区域处仍近似于直道,因此按照直道场景的图像对齐方法进行拼接仍然可以得到正确的拼接结果,如图 13(a)和 13(b)所示。而对于轨道平面而言,拼接结果也正确反应了弯道的弯曲形态,如图 13(c)所示。

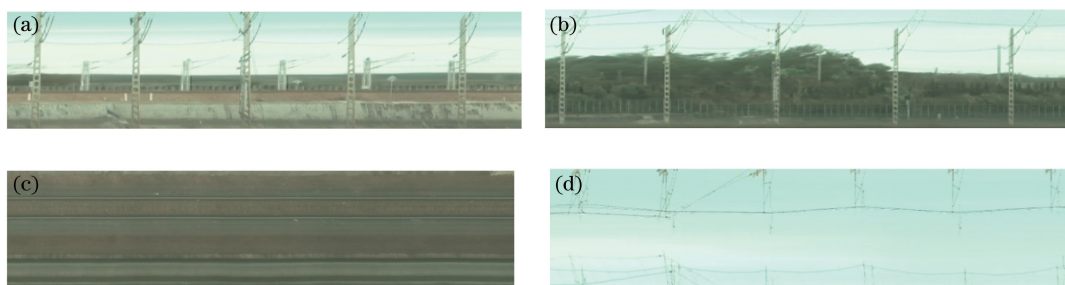


图 11 从封闭铁路环境中产生的全景图。(a)左侧全景图;(b)右侧全景图;(c)底部全景图;(d)上部全景图

Fig. 11 Panoramas generated from closed railway scene. (a) Left panorama; (b) right panorama; (c) bottom panorama; (d) top panorama

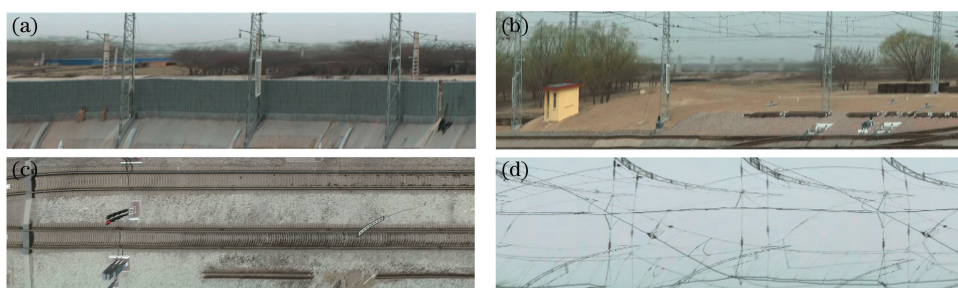


图 12 从道岔场景中生成的全景图。(a)左侧全景图;(b)右侧全景图;(c)底部全景图;(d)上部全景图

Fig. 12 Panoramas generated from road switch scene. (a) Left panorama; (b) right panorama; (c) bottom panorama; (d) top panorama

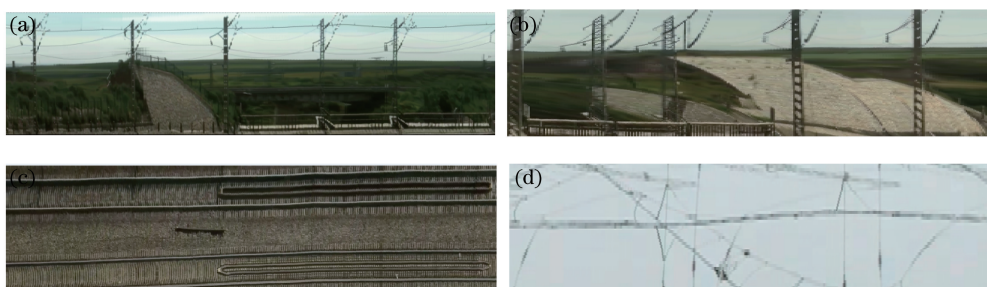


图 13 从弯道场景中生成的全景图。(a)左侧全景图;(b)右侧全景图;(c)底部全景图;(d)上部全景图

Fig. 13 Panoramas generated from curve road scene. (a) Left panorama; (b) right panorama; (c) bottom panorama; (d) top panorama

表 2 所示为全景图与原始视频的数据量对比。视频的数据量大小与采集时间和视频码流率成正比,而全景图的拼接过程类似于线阵推扫,线扫描相机采集的图像大小只与线阵 CCD 的分辨率和扫描长度有关,而与采集时间和视频码流率无关。因此,全景图的大小理论上应该只与拼接条带的长度(单位 pixel)和行驶里程有关。

表 2 全景图与原始视频的数据量对比

Table 2 Data size comparison between video and panorama

Route section	Video size /MB	Panorama size /MB	Compression ratio /MB
Changsha South—Xinyu North	137	13.6	10:1
Huaihua—Loudi	572	17.7	32:1
Loudi—Zhuzhou	1145	17.3	64:1

为了验证方法的有效性和实时性,针对不同的图像对齐方法进行了拼接速度的对比。实验计算平台配置如下: Intel i5 3.30 GHz 处理器, Windows 7 操作系统, 内存 8 GB。实验视频数据共 87632 frames, 摄像

机采集帧率为 25 frame/s,是列车行驶在 150 km/h 速度下获取的低质视频数据(720 pixel×576 pixel)。

图 14 所示为基于 4 种不同的图像对齐方法进行全景图拼接所得的拼接速度对比图,其中,蓝色折线为拼接耗时,红色折线为拼接速度。本文方法从拼接速度看可达到 42 frame/s,远远超过视频采集速率(每秒 25 帧),达到了实时拼接的效果。本文方法基于物理场景的几何结构和相机的运动信息进行图像对齐,避免了耗时的特征估计和光流计算,并且构造完全采样拼接区域的全局不变消失点,以及图像速度等关键参数都已离线存储在查询表中,使得图像对齐类似于模拟快速查表的过程。

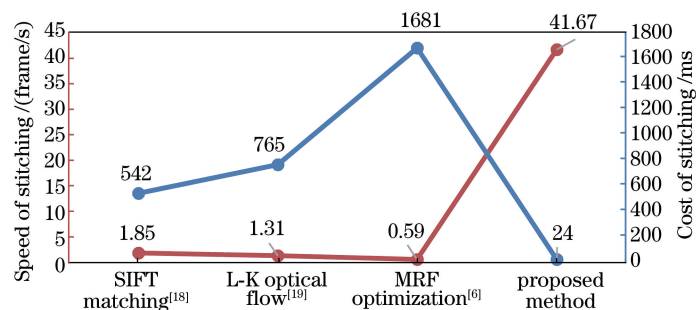


图 14 不同拼接方法的结果对比

Fig. 14 Comparison of time cost using different alignment methods

## 5 结 论

面向铁路环境视频存储与监测的要求,针对传统的视频拼接方法中的拼接时效性问题,提出了一种车载前向运动视频的实时全景成像方法。该方法没有进行复杂耗时的图像分析,仅利用铁路场景的几何结构先验以及列车的运动信息实现了拼接图像的对齐,为铁路环境视频提供了一种轻量级的视频全景索引方式。铁路全景图不仅降低了视频数据的存储和访问开销,而且将视频转化为一种更适合于人工视查或计算机自动分析的形式。

由于光照条件、气候因素及背景噪声等因素造成的图像质量退化,实际铁路场景的拼接效果还有待提升。后续的工作中,将预先对视频数据进行图像增强和复原处理,克服图像退化对拼接的影响,有利于获取高质量的全景图,便于后续的视觉检测分析。除此之外,对生成的多幅历史全景图进行图像变化检测,将有助于实现铁路场景异常变化的自动化监测。

## 参 考 文 献

- [1] An X, Wu M, He H. A novel approach to provide lane departure warning using only one forward-looking camera[C]. IEEE Collaborative Technologies and Systems, 2006: 356-362.
- [2] Wang Shengchun, Luo Siwei, Huang Yaping, *et al.* Stereo panorama generation from the train-borne video[J]. Acta Optica Sinica, 2014, 34(12): 1215005.  
王胜春, 罗四维, 黄雅平, 等. 移动车载视频的立体全景图生成方法[J]. 光学学报, 2014, 34(12): 1215005.
- [3] Nassu B T, Ukai M. Automatic recognition of railway signs using sift features [C]. IEEE Intelligent Vehicles Symposium (IV), 2010: 348-354.
- [4] Ou-Yang M, Jeng W D, Wu Y Y, *et al.* Image stitching and image reconstruction of intestines captured using radial imaging capsule endoscope[J]. Optical Engineering, 2012, 51(5): 057004.
- [5] Hansen P, Alismail H, Rander P, *et al.* Towards a visual perception system for pipe inspection: monocular visual odometry[R]. Pittsburgh, 2010: CMU-RI-TR-10-22/CMU-CS-QTR-104.
- [6] Niu C, Zhong F, Xu S, *et al.* Cylindrical panoramic mosaicing from a pipeline video through MRF based optimization [J]. The Visual Computer, 2013, 29(4): 253-263.
- [7] Zheng Geyi, Wang Guohui, Gao Jiaying, *et al.* Research on image jointing algorithm for the inner surface of the artillery bore[J]. Journal of Armored Force Engineering Institute, 2007, 21(4): 28-31.  
郑戈毅, 王国辉, 高嘉英, 等. 炮管内膛图像拼接检测算法研究[J]. 装甲兵工程学院学报, 2007, 21(4): 28-31.
- [8] Peleg S, Rousso B, Rav-Acha A, *et al.* Mosaicing on adaptive manifolds[J]. IEEE Transactions on Pattern Analysis

- and Machine Intelligence, 2000, 22(10): 1144-1154.
- [9] Zhang H, Li T, Geng J. Manifold modeling and its application to tubular scene manifold mosaicing algorithm [J]. Journal of Mathematical Imaging and Vision, 2012, 44(1): 80-98.
- [10] Zheng J Y, Zhou Y, Mili P. Scanning scene tunnel for city traversing [J]. IEEE Transactions on Visualization and Computer Graphics, 2006, 12(2): 155-167.
- [11] Kilicarslan M, Zheng J Y. Visualizing driving video in temporal profile [C]. IEEE Intelligent Vehicles Symposium Proceedings, 2014: 1263-1269.
- [12] Brown M, Lowe D G. Automatic panoramic image stitching using invariant features [J]. International Journal of Computer Vision, 2007, 74(1): 59-73.
- [13] 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.
- [14] 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.
- [15] Weinzaepfel P, Revaud J, Harchaoui Z, *et al.* Deepflow: large displacement optical flow with deep matching [C]. IEEE International Conference on Computer Vision, 2013: 1385-1392.
- [16] Li L, Yao J, Xie R, *et al.* A unified framework for street-view panorama stitching [J]. Sensors, 2016, 17(1): 1.
- [17] Yu Xinyan, Luo Siwei, Xu Tingfa, *et al.* Attitude estimation of the train-borne camera from the railway-environment [J]. Optical Technique, 2015, 41(3): 197-203.  
于欣妍, 罗四维, 许廷发, 等. 铁路运行环境下的车载相机姿态估计 [J]. 光学技术, 2015, 41(3): 197-203.
- [18] Li L, Geng N. Algorithm for sequence image automatic mosaic based on SIFT feature [C]. IEEE Information Engineering (ICIE), 2010, 1: 203-206.
- [19] Baker S, Matthews I. Lucas-kanade 20 years on: a unifying framework [J]. International Journal of Computer Vision, 2004, 56(3): 221-255.