
河北工业大学

毕业设计（论文）

作者：王昭然 学号：170762

学院：机械工程学院

系(专业)：车辆工程专业

题目：基于双目视觉的路面信息重建研究与应用

指导者：刘璇 副教授

指导者：许晟杰 工程师

评阅者：

2021 年 5 月 21 日

毕业设计（论文）中文摘要

基于双目视觉的路面信息重建研究与应用

摘要：

随着社会的发展，不同的 3D 应用场景已经深入人们的生活，在生活中周围每天会产生无数的立体图像。将视觉技术应用于物体检测是图像处理和计算机视觉领域的重要研究方向之一。视觉技术也可以应用于车辆、行人、交通标志、视频监控、异常事件分析的目标检测等等。

随着近年来高质量立体图像和影像需求和应用方式逐渐扩大增长，也开始广泛地应用到无人车自动驾驶技术之中。准确高效地搜集和确定路面信息，这样就有助于我们的智能型车载系统更好的了解道路情况，实现主动安全的驾驶，避免了交通中各种可能会遇到的危险，提高交通安全和运输的效率。在机器视觉的领域内，双目立体视觉一种十分重要的分支，具有丰富的内容、较高的精度、成本低及抗干扰性强等突出的特点，成为驾驶汽车感知外界的重要方式。

本文基于双目立体视觉，在 Anaconda 开发环境下，使用 python 语言来完成对路面深度信息的获取。从双目视觉原理入手，进行特征点识别、特征匹配、视角转换来达到与参考坐标系一致的目的。通过对归一化互相关算法及其优化算法的深入理解，来对双目图像进行灰度处理，生成积分图、计算图像的均值和标准差。最后，得到双目图像的视差图，来达到获取深度信息的目的。

关键词： 双目视觉 特征识别 视差图 归一化互相关

毕 业 设 计 （ 论 文 ） 外 文 摘 要

Title Research and application of pavement information
reconstruction based on binocular vision

Abstract

With the development of society, different 3D application scenes have penetrated into people's lives, and countless three-dimensional images will be generated around them every day. Visual-based object detection is one of the important research directions in the field of image processing and computer vision. It can be applied to vehicle, pedestrian, traffic sign and other object detection in the autonomous driving system, abnormal event analysis in video surveillance, and service robot and other fields.

In recent years, with the increasing demand and mode of high-quality stereo image, it has also begun to be applied in the autonomous driving technology of unmanned vehicles. Accurate and efficient determination of road surface information helps the intelligent vehicle system to better understand the road environment, realize active safe driving, avoid possible dangers in traffic, and improve traffic safety and efficiency. Binocular stereo vision as an important form of machine vision, with rich content, high accuracy, low cost and strong anti-interference characteristics, become an important way of driving cars to perceive the outside world.

Based on binocular stereo vision, this paper uses Python language to obtain the pavement depth information under the development environment of Anaconda. Starting from the principle of binocular vision, feature point recognition, feature matching and perspective conversion are carried out to achieve the purpose of consistent with the reference system coordinate system. Based on the in-depth understanding of the

normalized cross-correlation algorithm and its optimization algorithm, the gray processing of binocular image is carried out, integral graph is generated, and the mean value and standard deviation of the image are calculated. Finally, the parallax and depth maps of the binocular images are obtained to obtain the depth information.

Keywords: Binocular vision Feature identification Disparity map
Normalized cross correlation

目 录

1	绪论.....	5
1.1	研究背景.....	5
1.2	研究现状.....	5
1.2.1	研究技术现状.....	5
1.2.2	无源传感器用于三维路面重建的研究现状.....	7
1.3	本文的研究工作.....	9
2	双目立体视觉.....	10
2.1	双目视觉原理.....	10
2.2	双目视觉方法.....	11
2.3	坐标系转换.....	13
2.3.1	世界坐标系与相机坐标系转换.....	14
2.3.2	相机坐标系与图像坐标系转换.....	15
2.3.3	图像坐标系与像素坐标系转换.....	16
3	特征匹配.....	18
3.1	特征点提取.....	18
3.2	特征匹配.....	20
3.2.1	蛮力特征匹配.....	20
3.2.2	间距过滤匹配.....	21
3.3	视角转换.....	21
3.3.1	1PnP 算法.....	22
3.3.2	最小二乘法.....	23
3.3.3	视角转换结果.....	24
4	归一化互相关算法及优化.....	27
4.1	1NCC 算法.....	27
4.1.1	1NCC 算法原理.....	27
4.1.2	2NCC 算法优化.....	28
4.1.3	3NCC 优化算法验证.....	29
4.2	图像处理.....	30
4.2.1	灰度处理.....	30
4.2.2	积分图生成.....	33
4.2.3	均值图和标准差图生成.....	34
5	视差图以及深度信息的获取.....	37
5.1	相关算法介绍.....	37
5.1.1	SR 算法.....	37
5.1.2	绝对差异和算法.....	38
5.1.3	最大相关值验证.....	38
5.2	视差图.....	39
5.3	深度信息获取.....	41

结 论.....	44
参 考 文 献.....	46
致 谢.....	49

1 绪论

1.1 研究背景

随着社会的发展,不同的 3D 应用场景已经深入我们的生活,并且出现在诸多的行业领域,比如电影业、制造业、游戏业等,具有 3D 功能的移动电子设备的出现也进一步促进了它的普及^[1]。在我们周围每天会产生无数的立体图像(3D images/Stereoscopic images)。将视觉技术应用于物体检测是图像处理和计算机视觉领域的重要研究方向之一。视觉技术也可以应用于车辆、行人、交通标志、视频监控、异常事件分析的目标检测等。随着近年来高质量立体图像需求和方式逐渐增长,也开始应用于无人车自动驾驶技术之中。

其中,准确高效地识别和确定路面信息,这将有助于我们利用智能化的车载监控系统来更好地掌握道路情况,实现自动安全的驾驶,避免了在交通中出现各种可能发生的危险,提高了交通安全与运行的效率。

双目立体视觉作为机器视觉的一种重要表现形式,具有丰富的内容、较高的测量精度、成本低及抗干扰性强等突出的特点,已经发展成为自动驾驶和检测汽车各种环境和事件的重要手段。通过对路面信息的感知,通过对路面信息的感知,了解路面特征与障碍物识别,用以机器视觉的路面信息三维重构算法的完成算法的开发。

总的来说,基于双目立体视觉进行路面信息三维重建,获取路面信息,能够让汽车在驾驶过程中,是处于安全的交通环境之中的。反之,在无人车的驾驶过程中,可能会因为对路面信息的不熟悉、对路面的不了解而导致可能出现一些交通事故。因此,路面的微观外观和宏观外观是影响路面性能的主要因素之一。

1.2 研究现状

1.2.1 研究技术现状

路面信息的状态评估能够提供至关重要的安全性,由结构工程师或经过认证的检查员进行的人工视觉检查成本高昂,耗时且麻烦。在过去的十年中,诸如遥感,振动感应和计算机视觉之类的各种技术已越来越多地应用于研究路面信息中,以评估其的

物理和功能状况，并进行三维重建。

3D 重建方法可以分为基于激光扫描仪的三维激光扫描技术，基于微软 Kinect 传感器的和基于无源传感器的三大类方法^[2]。

三维激光扫描技术也可以被理解为实景复制技术。三维激光扫描技术的工作原理是，使用激光扫描仪（如图 1.1），在激光扫描仪的内部通过激光发射器发射激光，光线经过垂直和水平两个反射镜的折射关系，从而射向目标。之后，激光再次反射回来的被扫描仪接收，内部的距离测量模块可以测量反射时间和激光强度，再通过 $S=1/2ct$ （ c 为光速， t 为时间）得到目标点与扫描仪之间的距离，同时各传感器会记录激光发射的角度 θ ，再利用极坐标系与笛卡尔坐标系之间的相互转换关系计算出目标点的 X, Y, Z 坐标，从而得到大量带有三维坐标和反射率信息的点，并经过内部软件转换成数据信息，构建被测物的数字化模型。激光扫描仪收集来自物体的反射激光脉冲，以构建其准确的 3D 模型^[3]。

非接触测量也用于三维激光扫描仪，获取地面形状和复杂对象的几何数据和图像数据，并使用后处理软件对点组数据和图像数据进行处理和分析，将进行转换后的三维空间位置坐标是能够处于绝对坐标系中之中的，并可建立复杂结构及不规则形状的实体三维可视化模型^[4]。尽管它提供了准确的建模结果，但用于路况分析的激光扫描仪设备仍然昂贵，不具有普遍性且成本过高。



图 1.1 激光扫描仪

Fig. 1.1 laser scanner

Microsoft 公司的 Kinect 传感器^[5]，如下图 2 的 Kinect V1，利用人体语音、动作、表情等人机交互方式应用于教学、医疗、休闲娱乐等领域。并在此基础上，应用于各类体感人机交互领域。但基于微软 Kinect 传感器的方法，对于强烈吸收红外光的材料而言，室外环境的深度测量有些无效。而应用与无人驾驶中的三维路面重建而言，

室外环境的深度探测是重要且必要的。通过深度探测路面的宏观和微观信息，能够更好地识别路面进行，以进行无人驾驶工作。



图 1.2 Kinect V1 传感器

Fig. 1.2 Kinect V1 sensor

因此，基于无源传感器的方法（例如，双目立体视觉、单目立体视觉）更能够重建 3D 路面以进行状态评估或损坏检测。成本低、识别准确等特点，使其成为目前的研究焦点。

1.2.2 无源传感器用于三维路面重建的研究现状

（1）基于点云预处理方法的三维重建

点云预处理主要包括去噪、平滑和采样 3 个步骤^[6]。高精度路面三维模型的生成，需要将激光扫描所获点云转化为连续曲面，这一转化过程被称为曲面重构。扫描环境中的灰尘、过往车辆，以及扫描元件的轻微扰动使扫描得到的原始点云中存在较多干扰信息，使用这些原始点云直接进行曲面重构会导致精度较低。对于干扰信息进行剔除和调整的过程称为点云预处理，它是保障路面三维模型精度的必要操作。

魏亚等人^[7]为了将三维重建之后的模型进行优化，提出了路面点云预处理的方法，通过使用基于局部密度的去噪算法处理点云零散噪声，解决了常规去噪方法中出现的复杂度高和适应性差的问题。然后对于去噪后的点云，采用基于移动最小二乘法的平滑和采样连续算法，在区分路面特征区与非特征区的基础上，优化了平滑与采样阶段割裂造成的数据出现的二次劣化问题。最后，进行了实际路面试验验证。改善了传统方法在去噪、平滑和采样三个阶段中存在的问题，并提高了建模精度，达到了数据优化的目的。

（2）基于双目立体视觉的三维重建

这种双目立体视觉系统被认为是在所有计算机视觉系统之中，最为接近一个人的双眼立体视觉系统。双目的学习基本原理就是通过对学习者自身的一些视觉学习原理，

取得诸如深度信息等许多 3 维数据信息[8]。**marr** 的计算机视觉学理论主要是在继承了计算机科学理论的基础上,系统阐述了通过 2 维的图像转化成为 3 维的物体形态的一种基本途径和可能性,这个理论也就是现代计算机视觉学研究领域的必然基础。

立体匹配,是指经过 2 台参数相同的摄像机在同一场景拍摄,然后得到左视图和右视图两个视图,后续运用数学计算模型运算出左视图和右视图中的一一对应的点以及视差的关系,最后获取深度信息的过程^[9,10]。立体匹配在相关领域中的应用,其主要的难点在遮挡的区域、低纹理的区域以及无纹理的区域、深度不连续的区域,还包括噪声等等几个方面,近些年的研究,对立体匹配来说,研究的主要方向是通过引入不同的算法从而得到更优化的匹配效果。**Barnard** 等人^[11]提出了经典的计算机立体视觉系统。经典计算机视觉系统主要包括:图像获取、摄像机参数的标定、特征提取、立体匹配、深度信息获取、内插一共 6 个环节。朱松等人^[12]研究为视觉导航中路面检测方法,总结出一种图像灰度和图像深度相融合的检测路面的方法,通过 DSP 的路面检测技术,选择应用双目立体视觉的方法达到场景深度图像获取的目的,依据图像的灰度值和图像的深度信息来超像素分割,最后通过提取超像素的灰度和深度特征,采取 SVM 分类器来达到对超像素分类的目的,以便于完成路面的检测。此方法通过验证,并可以提高检测精度。

从 20 世纪 80 年代计算机视觉理论框架形成以来^[13],计算机技术的不断迅速发展,使得计算机视觉技术也得到了飞快发展,计算机视觉的应用领域也越来越宽,特别是在普通便携式相机拍摄成像情况下计算机视觉技术的应用,如制造业的缺陷检测、自动化的生产线、机器人视觉、无人驾驶导航、生物特征识别、医学成像、智能视频监控系统、反求工程、军事运用、行星表面巡航车、虚拟现实等领域。双目立体视觉不仅仅是计算机视觉领域的重要部分,而且还是机器视觉中的一种重要形式,是一门有着广阔应用前景的学科^[14]。在无人驾驶技术发展中,机器视觉在车道线检测技术研究中占据主导地位,也是未来的发展方向^[15]。

双目立体视觉的一个主要基本原理为三维视差学的基础原理,同时可以利用各种成像装置和设备通过不同位置分别得到被检测物体的左视图和右侧观察图,通过分析和计算检测到的图像各个节点间的位置和偏差,达到了获得物体的三维几何信息的主要目的。立体化的视觉系统是机器识别世界的一个重要手段。利用左视图和右视图形成的视差能够通过计算,得到像素点的深度信息,从而获得三维的立体感知^[16,17]。

其工作流程主要有 4 步:双目标定、图像校正、立体化匹配和三维模型重建,而

立体匹配是关键步骤^[18]。陈炎等人^[19]围绕基于人工特征和深度学习两个方向对算法的应用进行了综述,对算法在实现中的代价计算、代价集聚、视差计算和视差求精进行了分析和讨论,评估了双目立体算法的准确性和时间复杂度。李宇杰等人^[20]把视觉的三维目标检测方法进行了系统的总结,还做了许多调研的工作,分析出了基于单个目标视觉、双目标视觉以及采用多种深度相机的三维目标检测技术和方法,该技术的主要分类为室内外场景。

在路面三维重建之外的研究方向,双目立体视觉应用也十分广泛。双目立体视觉在各个领域均有应用,比如应用于自动驾驶系统中的车辆、行人、交通标志等目标检测^[21]、视频监控中出现异常事件进行分析,以及服务机器人等等许多领域。

1.3 本文的研究工作

尽管目前无源传感器已经成为了主流的研究三维重建的方式,但仍然存在很多问题,诸如识别精度以及效率的问题等等。

本文的主要研究工作安排如下:

第一章:绪论。首先介绍了三维重建和机器视觉的研究背景,并展开论述了目前三维重建技术的现状以及无源传感器的应用和工作。

第二章:双目视觉原理。主要介绍了双目视觉的基本原理,以及在本文中采用的双目摄像的方法,最后介绍了在三维世界中转换到二位像素世界坐标点的关系,以及相关理论公式和矩阵的推导。

第三章:特征匹配。本文先对双目图片进行了图像的预处理,通过算法来对特征节点进行了提取,并分析比较了不同算法来进行特征匹配的优劣。最后,生成与参考系一致的视角转换图对应相机参考系,完成视角转换的目的。

第四章:归一化互相关算法及优化。主要介绍了归一化互相关匹配算法,在本章中,使用归一化互相关匹配算法来进行图像的灰度处理。然后传统的归一化互相关匹配算法进行重新排列,并进行验证,进而通过生成积分图、均值图、标准差图来完成对图像的灰度处理。

第五章:视差图以及深度信息的获取。本章介绍了绝对差异和和最大相关值验证算法,以及视差和深度的数学关系。通过推导 SR 算法可以得到一个视差图,从而达到获取深度信息的目的。

2 双目立体视觉

2.1 双目视觉原理

立体视觉往往通过多张图像（一般为两张，少数情况为多张）获取物体几何三维信息的方法。

在生物的视觉系统，往往是用两个眼睛同时观察物体的，这样会有深度或远近的感知。流行的立体电影也是仿照了立体视觉的原理，才能够产生逼真的深度感。在立体电影的拍摄过程中，使用左摄像机和右摄像机来同时进行拍摄，而在放映电影时，将左右摄像机拍摄的画面同时投影到电影屏幕上，在其中应用偏振光的原理，让人们在观看电影时，使得双眼分别看到两个摄像机拍摄的图像，进而产生真实的三维物体的立体感。在计算机视觉系统中，通过使用左右摄像机，从两个以上不同的角度获取周围物体的多幅数字图像，进而通过计算机算法重构物体的三维位置与形状。

假设两个相机的内部参数一致，两个相机的 X 轴方向一致，像平面重叠，光轴严格平行，如图 1 所示，坐标系以左相机为准，右相机相对于左相机是简单的平移，用坐标表示为 $(T, 0, 0)$ ， T 称为基线，根据三角原理，易得出空间中的一点 $P(X, Y, Z)$ 分别在左右像平面上的投影坐标。

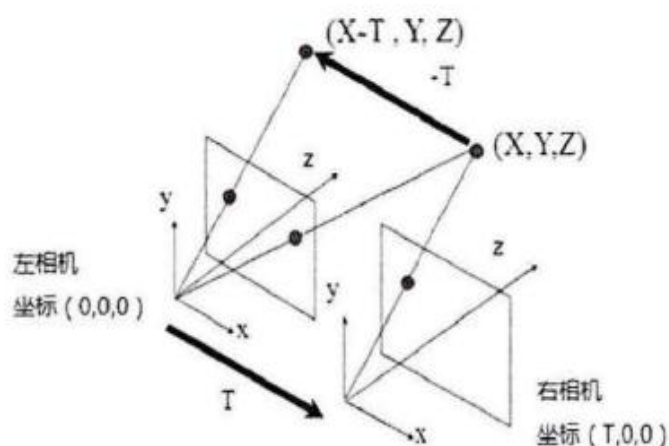


图 2.1 相机坐标系

Fig. 2.1 camera coordinate system

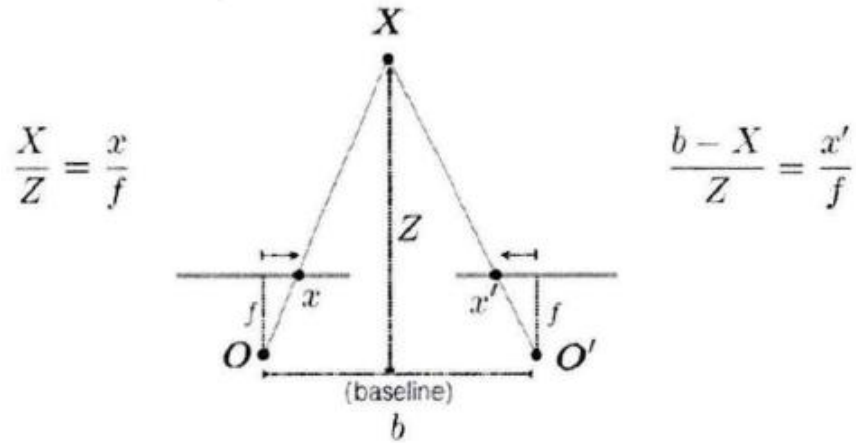


图 2.2 双目相机测距原理

Fig. 2.2 the ranging principle of binocular camera

左相机的像平面坐标点为：

$$x_l = f \frac{X}{Z}, \quad y_l = f \frac{Y}{Z} \quad (2-1)$$

右边相机的像平面坐标点为：

$$x_r = f \frac{T-X}{Z}, \quad y_r = f \frac{Y}{Z} \quad (2-2)$$

双目视差为：

$$d = x_l - x_r = f \frac{X}{Z} - f \frac{T-X}{Z} = f \frac{T}{Z} \quad (2-3)$$

深度值为：

$$Z = \frac{fT}{d} \quad (2-4)$$

双目测距主要是利用了目标点在左右两幅视图上成像的横向坐标直接存在的差异（即视差）与目标点到成像平面的距离 Z 存在着反比例的关系。

2.2 双目视觉方法

本文实现双目视觉方法，采用了双目摄像头。通过双目摄像头能够生成直接的左视图和右视图，并能保证其的水平位置平行。使用的双目相机为 ZED 相机，如下图：



图 2.3 ZED 双目摄像头

Fig. 2.3 ZED binocular camera

ZED 双目摄像头采用了最新的英特尔实感视觉处理器和模块，并还能够提供计算机视觉，ZED 相机的视觉处理器功能十分强大，该处理器使用 28 纳米制程技术并支持达 5 个 MIPI 摄像头串行接口 2 通道，通过计算实时的深度图像，而且能够增加输出速度，加速输出，可以得到高达每秒 90 帧的深度视频流。同时，通过采用高分辨率的图像传感器，能够使图像在深度数据流中得到与原来相比，高于 5 倍的像素数。分辨率足够清晰，能达到 1280×720 ，而首代立体深度摄像头的分辨率为 480×360 。并有专用的颜色图像信号处理器用于图像校正及颜色数据缩放，从而大大提高图像质量。内置的高级立体深度算法和新设计，以实现更准确的深度感知及更远的范围。通过校准，立体深度感知的误差率低至 1%。在适环境中，这些摄像头在室内和室外环境中均可捕获相距达 10 米的数据。最重要的是，它支持新的跨平台开源英特尔实感 SDK 2.0，其中包括许多操作的系统界面、第三方的插件软件、开发环境及编程所采用的计算机语言。可用于本文开发的算法，应用与无人车的环境感知工作中。

使用双目相机，拍摄的路面双目图片如下：



图 2.4 路面左视图

Fig. 2.4 left view of pavement



图 2.5 路面右视图

Fig. 2.5 right view of pavement

本文其余章节，均基于拍摄的路面双目图像进行讨论、分析与处理。通过对双目图片的特征点提取、特征匹配、以及灰度处理到生成视差图，来完成本文的研究目的。

2.3 坐标系转换

在本文中，图像处理、立体视觉等等需要涉及坐标系转换。经常涉及到的坐标系有：世界坐标系、相机坐标系、图像坐标系、像素坐标系。通常来说，构建世界坐标

系有多种多样的方式，但最重要的是为了更好的描述相机的位置。

2.3.1 世界坐标系与相机坐标系转换

通常在双目视觉中一般将世界坐标系原点定在左相机或者右相机或者二者 X 轴方向的中点。将世界坐标系转换为相机坐标系的过程是刚体的转换，在这个过程中，物体不发生形变，因此只有旋转和平移变换。具体转换如下：

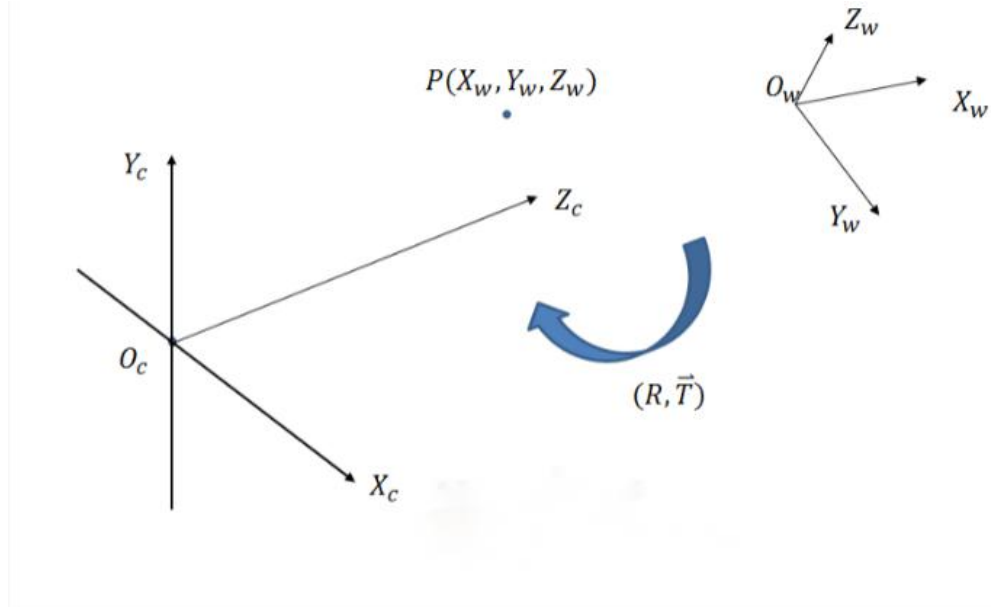


图 2.6 世界坐标系与相机坐标系转换

Fig. 2.6 world coordinate system and camera coordinate system conversion

其中, R 为旋转矩阵, T 为偏移向量。不难发现, 从世界坐标系到相机坐标系, 因为仅仅涉及到旋转和平移。绕着不同的坐标轴旋转不同的角度, 得到相应的旋转矩阵。从而进行转换。不妨认为, 转换后的点为 (x', y', z') 。那么, 围绕 Z 轴旋转转换矩阵如下:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R_1 \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \quad (2-5)$$

围绕 X 轴旋转转换矩阵如下:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & \sin \varphi \\ 0 & -\sin \varphi & \cos \varphi \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R_2 \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \quad (2-6)$$

围绕 Y 轴旋转转换矩阵如下:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos \varphi & 0 & -\sin \varphi \\ 0 & 1 & 0 \\ \sin \varphi & 0 & \cos \varphi \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = R_3 \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} \quad (2-7)$$

可以得到，旋转矩阵为 $R=R_1R_2R_3$ 。

于是，可以得到，由世界坐标系转换为相机坐标系的矩阵为：

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = R \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + T \quad (2-8)$$

2.3.2 相机坐标系与图像坐标系转换

由相机坐标系转换到图像坐标系，是因为在投影过程中，进行了缩小。简而言之，属于透视投影关系。同时，也进行了从 3D 转换到 2D 的维度转变。具体关系如下：

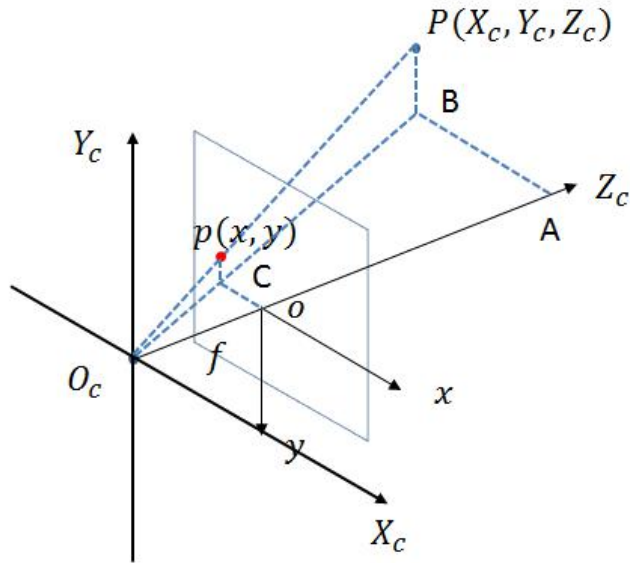


图 2.7 相机坐标系与图像坐标系转换

Fig. 2.7 camera coordinate system and image coordinate system conversion

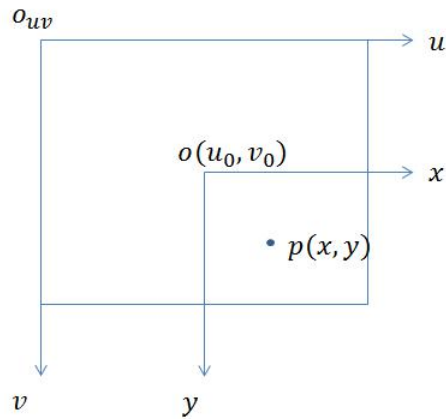
上文已经给出相机内部参数的影响，不在赘述。根据相似三角形原理，得到变换矩阵：

$$Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \\ 1 \end{bmatrix} \quad (2-9)$$

但是，现在转换后的 P 点的单位还是长度单位，并不是像素单位，需要进一步转换到像素坐标系。

2.3.3 图像坐标系与像素坐标系转换

像素坐标系和图像坐标系都在成像平面上，但是原点位置和度量的单位并不相同。图像坐标系的原点为相机光轴与成像平面的交点，通常情况下是成像平面的中点或者叫像素点。图像坐标系的单位是毫米，属于物理单位。而像素坐标系的单位是像素，我们平常描述一个像素点都是行与列的关系。所以这二者之间的转换关系如下：其中 dx 和 dy 表示每一列和每一行分别代表毫米单位的量。



[http:](http://)

图 2.8 图像坐标系和像素坐标系转换

Fig. 2.8 image coordinate system and pixel coordinate system conversion

转换的公式如下：

$$u = \frac{x}{dx} + u_0 \quad (2-10)$$

$$v = \frac{y}{dy} + v_0 \quad (2-11)$$

转换的矩阵如下：

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \quad (2-12)$$

那么通过上面四个坐标系的转换就可以得到一个点从世界坐标系如何转换到像素坐标系的。反之，也是成立的，能够从一个像素点转换为一个三维坐标的点。

下面给出总的转换关系：

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ \vec{0} & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ \vec{0} & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (2-13)$$

其中相机的内参为 $\begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ ，外参为 $\begin{bmatrix} R & T \\ \vec{0} & 1 \end{bmatrix}$ ，这两组参数可以通过张正友

标定获取。

不难发现，在最终的转换关系之中，一个三维中的坐标点，的确可以在图像中找到一个对应的像素点。但是反过来，通过图像中的一个点找到它在三维中对应的点就很成了一个问題，因为我们并不知道等式左边的 Z_c 的值。本文基于双目视觉来探究路面深度信息，就是为了解决深度信息 Z_c 的值，来完成三维重建的目的。

2.4 本章小结

本章主要介绍了双目视觉的基本原理，用成像原理入手，介绍了相机参数的含义，推导了焦距的计算公式和双目测距的反比例的关系。还介绍了本本使用的 ZED 摄像头，其强大的功能和视觉处理能力，能够以较高的精度用于本文的开发环境中。最后，推导了从世界坐标系到相机坐标系，从相机坐标系到图像坐标系，从图像坐标系到像素坐标系的理论关系的推导，其目的为了更好的描述相机的位置。并且发现，在从一个二维像素点到一个三维世界的坐标点转换时，是需要得到深度信息才能够进行转换的。

3 特征匹配

特征匹配通常是指在图片影像中提取特征，将提取的特征信息视为共轭实体，然后将提取得特征属性或描述参数当做是匹配实体进行匹配，实际上是特征的特征，也可以理解为影像的特征。

特征匹配是一种通过计算匹配实体之间的相似性测度以及实现共轭实体配准的影像匹配方法。图像特征提取及特征匹配是计算机视觉、图像处理和计算机图形学的研究热点^[22-23]。

3.1 特征点提取

OpenCV 是一个拥有 BSD 认证许可，已经开源发行的跨平台计算机视觉和机器学习软件库，可以运行在 Linux、Windows、Android 和 Mac OS 操作系统上。它轻量级而且高效，由一系列 C 函数和少量 C++ 类构成，同时提供了 Python、Ruby、MATLAB 等语言的接口，实现了许多通用的图像处理和计算机视觉算法。本文采用 python 语言为计算机语言，通过 OpenCV 来进行图像的识别与匹配，是十分方便与便捷的。

OpenCV 是 Intel 公司开发的开源计算机视觉算法库^[24]，在 OpenCV 中拥有丰富的函数功能以及强大的图像处理和矩阵运算能力，且具有平台无关性和程序运行的实时性，因而成为了机器人实时识别算法库的最佳选择^[25]。

图像中总是有唯一的像素点。这些点可以被认为是图像的特征，成为特征点。在计算机视觉领域中，基于特征点的图像特征匹配是非常重要的，因此如何定义和找出图像中的特征点是非常重要的。

特征点识别的原理为：从图像中选取一些特征点，对图像进行局部分析，只要在视图中有足够的兴趣点可以被探测到，并且这些兴趣点具有不同且稳定的特征，就可以精确的定位并且就可以被识别出来。

特征点识别的算法主要有 sift 和 surf。本文通过 sift 算法进行识别。

尺度不变特征转换 (sift) 是一种电脑视觉的算法用来侦测与描述影像中的局部性特征，它在空间尺度中寻找极值点，并提取出其位置、尺度、旋转不变量。其应用范围包含物体辨识、机器人地图感知与导航、图像缝合、三维模型建立、手势识别、图像追踪和动作比较。

对局部图像特征的描述和检测有助于目标的识别，sift 特征是基于兴趣点的局部外观的对象，并独立于图像的大小和旋转，换言之，就是与影像的大小和旋转无关。对于光线、噪声、一些微视角改变的误差也有一定的余量。

基于以上的一些特性，特征点是非常明显而且相对容易采集和获取的。在数目庞大的特征数据库中，很容易辨识物体而且很少会出现误认的情况。使用 sift 特征描述，局部目标遮挡的检测率也很高，甚至仅仅需要 3 个以上的 sift 物体特征就能够计算出物体的位置与方向。目前，在计算机硬件速度快、特征库小的情况下，识别速度已接近实时运行。sift 特征的信息量大，更加合适在大量数据库中进行快速准确有精度的匹配。

本文识别结果如下：

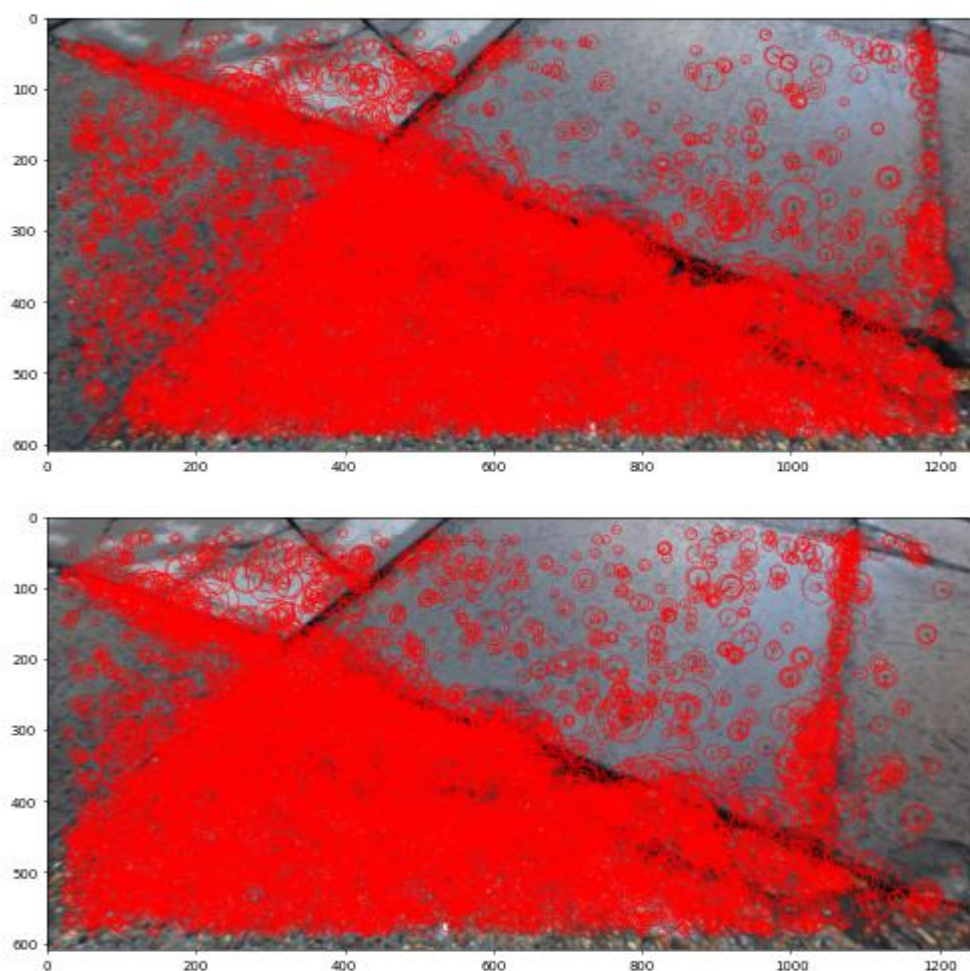


图 3.1 特征点识别

Fig. 3.1 feature point recognition

3.2 特征匹配

3.2.1 蛮力特征匹配

在特征匹配中，OpenCV 提供了两种匹配器，用于特征匹配的 Brute-Force 匹配器和基于 FLANN 的特征匹配的匹配器。蛮力特征匹配是通过 Brute-Force 匹配器完成。

首先使用 cv2 创建对象。它需要两个可选参数，首先指定要使用的距离测量，第二个参数为布尔变量。匹配器返回具有值 (i, j) 的匹配，使得一个集合中的第 i 个描述符，具有另一个集合中的第 j 个描述符作为最佳匹配，反之也成立。也就是说，因为两组中左右图片提供了一致的结果，所以两个特征点应该相互匹配。

本文匹配的结果如下图：

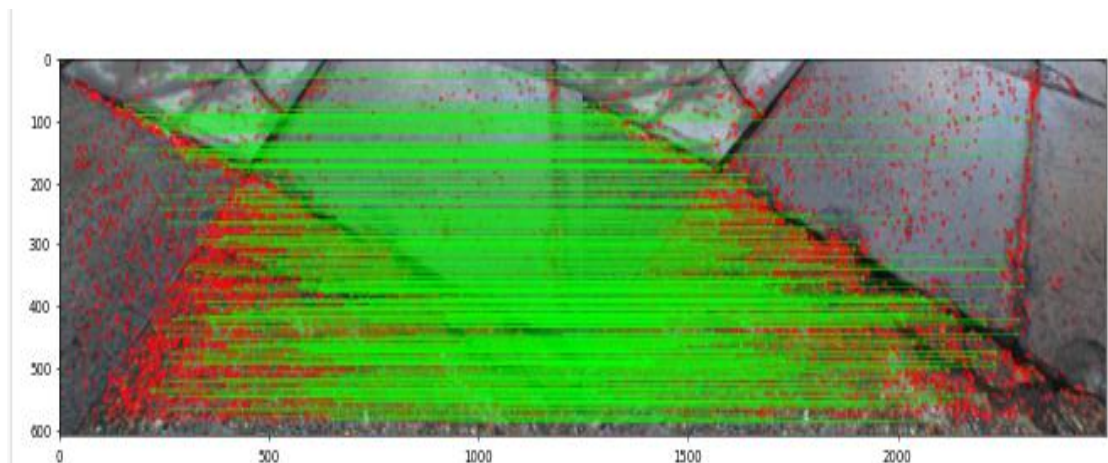


图 3.2 BF 匹配结果

Fig. 3.2 BF matching results

特征点共 10555 个点，通过 BF 匹配器进行蛮力匹配，匹配点为 436 的点。

对于蛮力特征匹配来说，有如下优点：

(1) 蛮力特征匹配所依赖的最基本技术是扫描技术，依次处理元素是蛮力特征匹配的关键。蛮力特征匹配可以作为一种时间性能阈值，更有效地度量同一问题。

(2) 虽然设计很高效的算法很少来自于蛮力特征匹配，但蛮力算法也是一种重要的算法设计技术。许多算法基于此设计。

(3) 理论上，蛮力匹配算法可以解决可计算许多领域的各种问题。对于一些十分基本的问题，比如在一个序列中，求解最大元素；计算多个数之和，蛮力匹配算法是一种十分常用的算法设计技术。

但蛮力特征匹配也是有缺点的：

(1) 通常来说一些较小规模的问题使用蛮力特征匹配来解决比较便捷。如果需要解决的问题规模比较小,用蛮力特征匹配算法的运行速度是没有问题的。但是,处理更复杂、数据更多的问题来说,BF 算法效率就比较低了。

(2) 一些像查找排序的问题和字符串匹配的问题来说,使用蛮力特征匹配可以设计出能够运行的算法,且有实用价值,如果算法过于复杂仍可以运行,但运行时间过长,算法过于复杂。

3.2.2 间距过滤匹配

针对基于 BF 匹配器的蛮力匹配方法搜索效率较低、精度还不够的情况,通过分析特征识别过程,优化方法采用模糊匹配来进行。同时因为图片识别特征匹配距离算法只有较低的性能,提出采用间距过滤规则进行搜索图片,过滤后再采用模糊匹配的方法进行改进。实验结果表明,采用间距过滤匹配的改进方法具有较高的执行效率和精度。通过间距过滤匹配结果如下图:

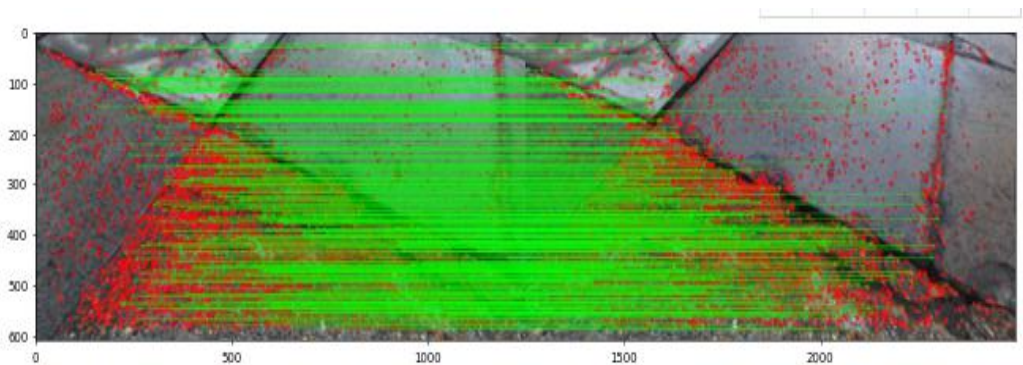


图 3.3 间距过滤匹配结果

Fig. 3.3 spacing filters match results

在特征点共 10555 个点的情况下,通过本文算法进行间距过滤匹配,匹配点为 380 的点。与 BF 匹配器相比,处理速度提高的同时,筛除相关度小的点,并可以保证精度。故采用本算法。

3.3 视角转换

研究路面信息,是基于水平面的,所提出的算法完全集中在可以被视为路面上进行。为了提高立体匹配的精度,首先利用中的地平面约束(GP 约束)的概念^[26-27],

对两幅图像进行透视变换，然后再估计其视差。

GP 约束通常用于广泛的障碍物检测系统，其中一侧的图像设置为参考，另一侧的图像转换为参考视图。由 GP 产生的像素满足相同的仿射变换，而在 GP 之上的对象不会被成功转换，变换后的图像中障碍物的像素被扭曲。然而，转换后的图像中的路面看起来更像它的参考视图。因此，透视变换使得障碍物区域噪声大且不可靠，但大大提高了两幅图像之间的路面相似性。

3.3.1 PnP 算法

PnP 问题在计算机视觉中目标进行空间定位的基本且重要的问题，一般可以描述为在透视投影条件下存在多对 3D 和 2D 匹配点，当预知相机内部参数的前提下，求解这些 3D 点在相机系的空间位置。关于 PnP 问题的求解吸引了很多学者的关注，目前主流算法包括 P3P、直接线性变换（DLT）、EPnP、UPnP、等。EPnP 需要 4 对不共面的 3D 与 2D 匹配点，被公认为求解速度快精度较高的 PnP 求解方法^[28]。

相机投影的关系为：

$$\begin{bmatrix} su \\ sv \\ s \end{bmatrix} = K \begin{bmatrix} R & t \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad \forall f_i \quad (3-1)$$

其中，u 为左视图，v 为右视图，s 指匹配点间的距离。为实现参数的确定，使用 PnP 求解算法来实现。PnP 算法是求解点的匹配问题的，通常是指通过多对 3D 与 2D 匹配点的问题。无论是否知道相机内参的前提下，采取最小化重投影误差来求解相机外参的方法。相机坐标系按照针孔模型建立即可^[29]。SLAM 前端位姿跟踪部分常用 PnP 求解算法来求解问题。PnP 算法通过 OpenCV 来实现。

本文将相机参数进行简化，简化如下：

- (1) K_l 与 K_r 是一样的
- (2) R 为单位矩阵
- (3) 方向与 X 轴方向一致
- (4) 路面为水平面
- (5) 旋转仅围绕 X 轴旋转

3.3.2 最小二乘法

最小二乘是一种被称为最小二乘的数学优化技术。它使误差平方和最小化，以找到数据的最佳函数匹配。最小二乘法使未知数据易于查找，并使所得数据与实际数据之间的平方和最小化。

在曲线拟合中，能够很好的应用最小二乘法。其他最优化问题也可以用最小二乘来表示能量最小化或熵最大化。下列公式为最小二乘法的公式：

$$\frac{\partial \theta}{\partial \alpha_0} = -2 \sum (Y_i - \alpha_0 X_i) = 0 \quad (3-2)$$

$$\frac{\partial \theta}{\partial \alpha_1} = -2 X_i \sum (Y_i - \alpha_0 - \alpha_1 X_i) = 0 \quad (3-3)$$

在本文中，为了估计 α_0 与 α_1 ，来达到视角转换的目的，通过最小二乘法进行插值估计。

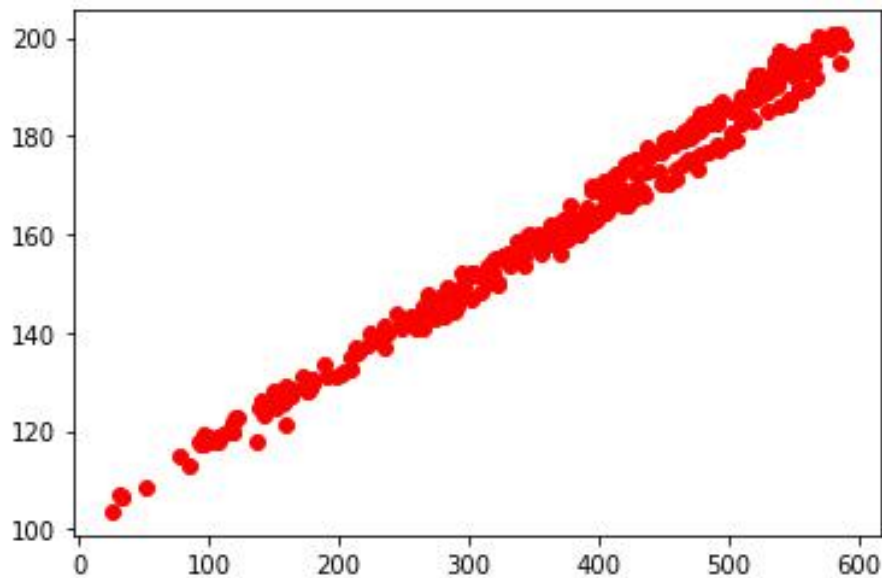


图 3.4 α_0 与 α_1 的线性关系

Fig. 3.4 The linear relationship between α_0 and α_1

通过算法，估计值如下：

$$\alpha_1 = 0.16850709$$

$$\alpha_0 = 99.61006242$$

3.3.3 视角转换结果

通过双目相机拍摄得到的左右视图中，路面总是存在透视失真，这进一步影响了块匹配的准确性。因此，视角转换的目的是使变换后图像中的路面与参照系中的路面相似。

通过对参数 α 的估计来实现视角转换。对 α 定义为：

$$\alpha = [\alpha_0, \alpha_1]^T \quad (3-4)$$

其中， α_0 与 α_1 可以通过求解一组最小二乘的矩阵进行估计。

定义 Q_l 与 Q_r ：

$$Q_l = [p_{l1}, p_{l2}, \dots, p_{lm}]^T \quad (3-5)$$

$$Q_r = [p_{r1}, p_{r2}, \dots, p_{rm}]^T \quad (3-6)$$

在本文中，使用 BRISK 函数进行匹配 Q_l 与 Q_r ，来进行图片的关键点与描述符。BRISK 函数是对 BRIEF 函数的一种改进，相比于 BRIEF 函数的特征，它旋转时具有不变性，且还有尺度不变性以及对待噪声的鲁棒性。

具体的视角转换算法结果如下：



图 3.5 左视图视角转换前

Fig. 3.5 left view before conversion

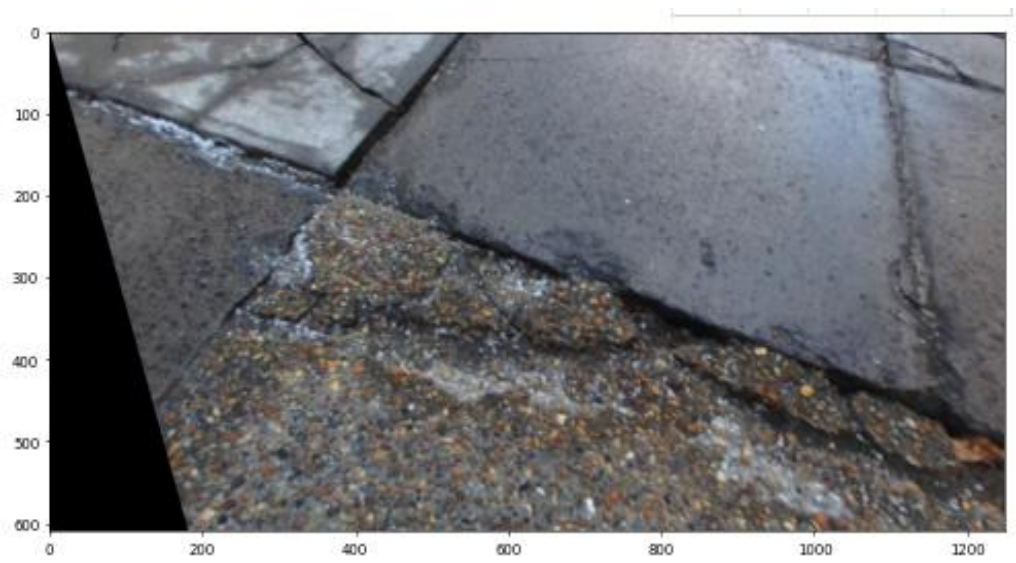


图 3.6 左视图进行视角转换后

Fig. 3.6 left view after conversion

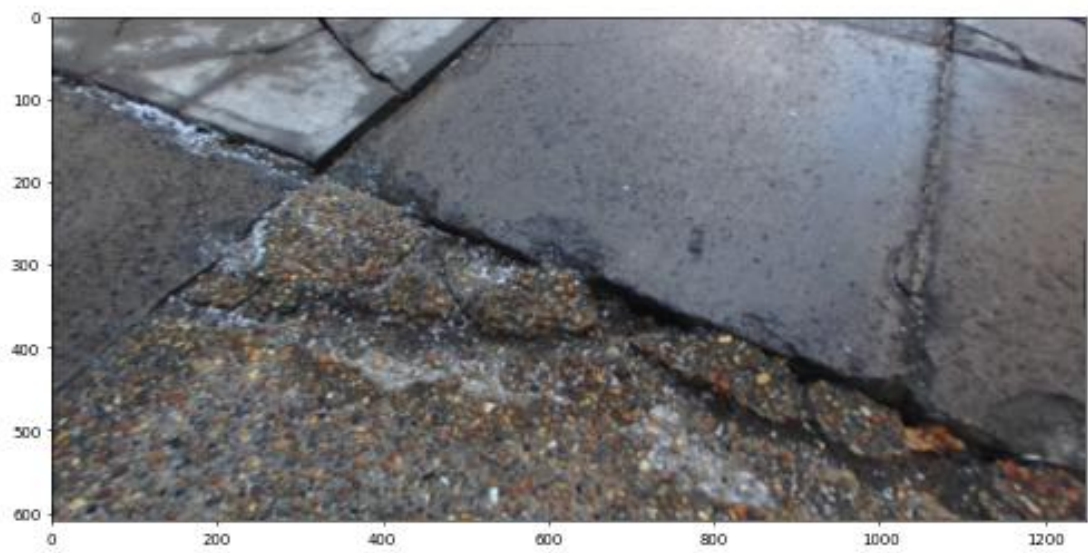


图 3.7 右视图进行视角转换前

Fig. 3.7 right view before conversion



图 3.8 右视图进行视角转换后

Fig. 3.8 right view after conversion

通过视角转换能够达到使变换后图像中的路面与参照系中的路面相似的目的。

3.4 本章小结

本章主要介绍了一些图片预处理工作，完成了对于特征点的提取、特征匹配的算法分析以及视角转换。特征点识别使用尺度不变特征转换算法进行，共识别出 10555 个点。在特征匹配的部分中，先使用了蛮力特征匹配来进行匹配，共对应 436 个点；使用间距过滤匹配来匹配时，对应 386 个点。二者比较来说，间距过滤匹配更快捷，且准确性更好。在视角转换中，引入 PnP 算法和最小二乘法来估计参数，并完成路面图片参考系转换，使之与参考系路面保持一致。解决了拍摄图片透视失真，以及进后续会影响块匹配的准确性等问题。

4 归一化互相关算法及优化

4.1 NCC 算法

4.1.1 NCC 算法原理

归一化互相关（NCC）匹配算法最开始由 Rosenfeld 等^[30]学者提出，属于基于区域内的灰度图像匹配。NCC 也是常用于图像特征匹配中的一种方法。NCC 算法是从统计学原理引入的，后来发现图像的相似性可以通过两个向量的相关性进行类比，假设有向量 $\vec{\alpha}$ 和向量 $\vec{\beta}$ ，根据向量点乘的定义得到公式：

$$\cos \theta = \frac{\vec{\alpha} \cdot \vec{\beta}}{|\vec{\alpha}| \cdot |\vec{\beta}|} \quad (4-1)$$

当 $\cos \theta$ 接近 1 时，可以说明 $\vec{\alpha}$ 和 $\vec{\beta}$ 是相似的^[31]，这个原理同样适用于二维图像的处理中。

上面说过，归一化互相关算法（NCC）是一种基于统计学基础的计算两组样本数据相关性的算法，结果的取值范围为 $[-1, 1]$ 之间。而在图像相似性方面来说，图像中的每个像素点均可以视为 RGB 数值，这样图像整体就可以看成是一个样本数据的集合。如果它有一个子集匹配另一个样本数据，那么它的 NCC 值是 1，这意味着相关性是高的。如果结果为 -1，则表示完全不相关，基于此原理，实现一种基于模板匹配的图像识别算法，该算法的第一步是对数据进行归一化处理，数学公式如下：

$$\tilde{f} = \frac{f - \mu}{\sigma} \quad (4-2)$$

其中， f 表示像素点的灰度值。 μ 表示图像所有像素的平均值， σ 表示图像标准方差的像素值。

完整的 NCC 公式如下：

$$NCC = \frac{1}{n-1} \sum_{x,y} \frac{(i_l(x,y) - \mu_l)(i_r(x,y) - \mu_r)}{\sigma_l \sigma_r} \quad (4-3)$$

其中， n 表示像素总数， $n-1$ 表示自由度。

使用 NCC 算法，根据图像的像素值并计算均值与标准方差、像素与均值。在目标图像上从左到右，从上到下移动窗口，计算每移动一个像素之后窗口内像素与模板像素的 NCC 值，与阈值比较，大于阈值则记录位置并显示结果，通过这样的方式来描述两个像素点的相关性。

4.1.2 NCC 算法优化

在本文中，通过使用归一化互相关（NCC）算法用于计算匹配代价，通过层次迭代得到匹配代价的值，故在 (u, v) 像素的搜索范围是从第 $v+1$ 行，三个估计的相邻差异传播过来的。

NCC 虽然比较常用，但处理速度过慢，不满足实时性的要求。为了增强算法实时性，故对 NCC 进行优化。

为了加速 NCC 的执行，将 NCC 方程重新排列如下：

$$NCC(u, v, d) = \frac{1}{n \sigma_l \sigma_r} \left(\sum_{x=u-\rho}^{x=u+\rho} \sum_{y=v-\rho}^{y=v+\rho} i_l(x, y) i_r(x-d, y) - n \mu_l \mu_r \right) \quad (4-4)$$

其中，NCC 被定义为计算左右平面的匹配代价。NCC 的计算值与匹配效果成正比关系，NCC 计算的值越大，匹配效果越好。 i_l 与 i_r 表示为左右平面的像素强度。定义图片边缘长度为 $2\rho+1$ ， n 仍表示其中的像素数。

把 (u, v) 与 $(u-d, v)$ 定义为左右匹配块的中心值。从重新排列的方程中，可以看出，均值 μ 和方差 σ 只对从平面中选取的单独的匹配块起作用，换言之，每个匹配块之间是独立的，且根据 d 的值进行匹配。与传统 NCC 算法相比，传统 NCC 的立体

匹配算法总是会重复计算 μ_l 、 μ_r 、 σ_l 和 σ_r 。优化之后，可以先计算 μ 和 σ 的值，并通过程序代码，将它们存储在一个静态程序存储器中，以便直接搜索。这样就可以把 NCC 的计算复杂度简化为点积，使得立体匹配更加有效。实现增加实时性的目的。

4.1.3 NCC 优化算法验证

为了验证经典的 NCC 算法和优化的 NCC 算法的处理图片的速度，采用相同的图片，对二者同时处理一样的图片来进行比较，来分析算法是否达到优化的目的。

以左视图进行视角转换前后进行匹配，选取匹配点如下：

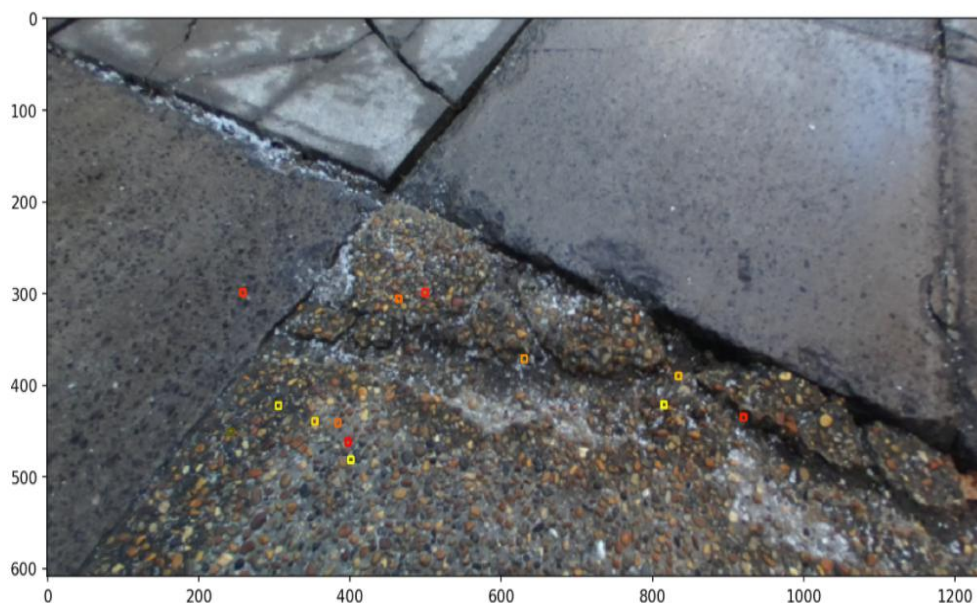


图 4.1 视角转换前匹配点

Fig. 4.1 match point before viewpoint shift

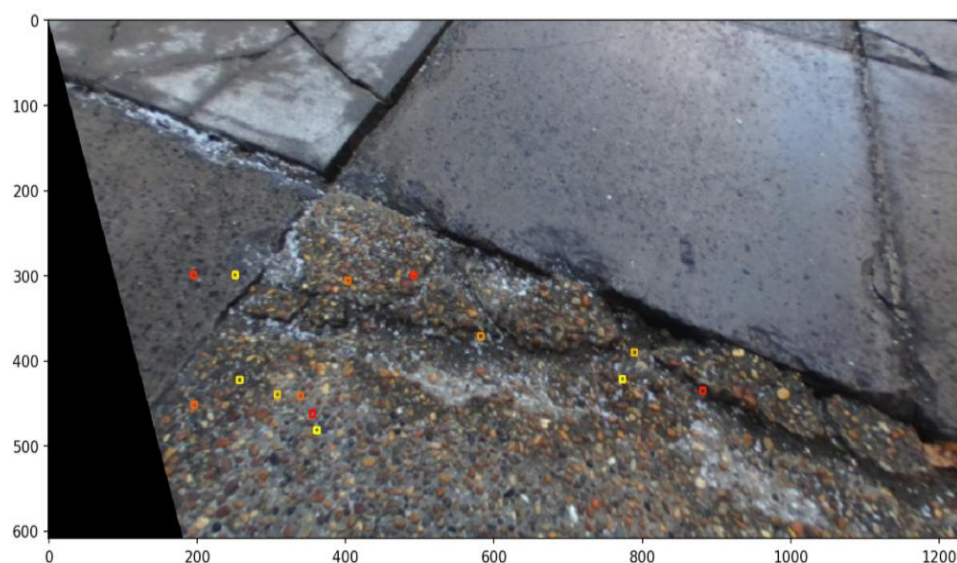


图 4.2 视角转换后匹配点

Fig. 4.2 match points after viewpoint conversion

通过两种算法处理，在识别匹配相同的数量的点的情况下，处理速度有很大的不同。经典的 NCC 算法处理并获得信息，需要 5.8ms，优化的 NCC 算法只需要 0.40ms，处理速度快了很多。

4.2 图像处理

4.2.1 灰度处理

在图像处理过程中，把图片转换成灰度图像是很有必要的。灰度图像上每个像素的颜色值又称为灰度，指黑白图像中点的颜色深度，范围一般从 0 到 255，白色为 255，黑色为 0。所谓的灰度值是指颜色的强度，灰色直方图是指一幅数字图像，对应于每个灰度值统计出的像素数所对应的灰度值。

灰度是没有颜色的，所以 RGB 颜色组件都是相等的。如果是二值灰度图像，其像素值只能为 0 或 1，我们称其灰度为 2。目前大多数彩色图像都是 RGB 颜色模式。在处理图像时，RGB 的三个分量应分别处理。事实上，RGB 并不反映图像的形态学特征，只是从光学原理上对颜色进行匹配。

图像处理的预处理步骤为图像灰度处理，并且可以为图像分割、图像识别和图像分析做准备。灰度处理的方法如下：

(1) 分量法

将彩色图像中三个分量的亮度作为三幅灰度图像的灰度值，可根据应用要求选择一幅灰度图像。公式如下：

$$Gray_1(i, j) = R(i, j) \quad (4-5)$$

$$Gray_2(i, j) = G(i, j) \quad (4-6)$$

$$Gray_3(i, j) = B(i, j) \quad (4-7)$$

(2) 最大值法

最大值法灰度图的灰度值采用的是彩色图像中的三个分量亮度的最大值。公式如下：

$$Gray(i, j) = \max \{R(i, j), G(i, j), B(i, j)\} \quad (4-8)$$

(3) 平均值法

通过对彩色图像中的三分量亮度先求和再求平均值，即为平均值法得到的灰度值。公式如下：

$$Gray(i, j) = \{R(i, j) + G(i, j) + B(i, j)\} / 3 \quad (4-9)$$

(4) 加权平均法

根据重要性、亮度等指标，对三个组成部分按不同的权重进行加权。由于人眼对绿色最敏感，而对蓝色最不敏感，按此公式对 RGB 三个分量进行加权可以得到更合理的灰度图像：

$$Gray(i, j) = 0.299 * R(i, j) + 0.578 * G(i, j) + 0.114 * B(i, j) \quad (4-10)$$

还有很多方法可以对像素信息进行处理，不一一列举，生成的左视图灰度图如下：

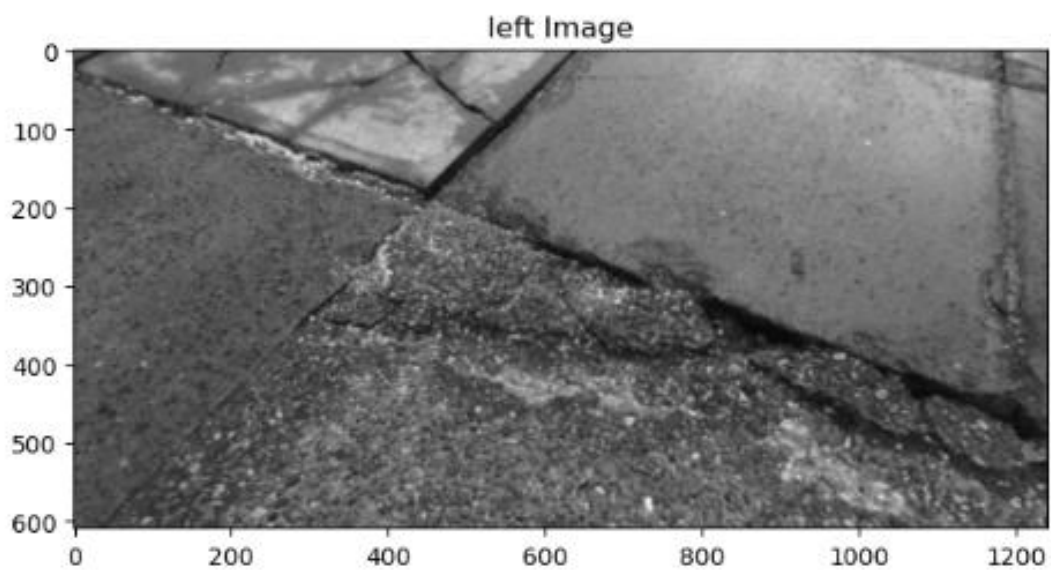


图 4.3 转换前左视灰度图

Fig. 4.3 left view grayscale image before conversion

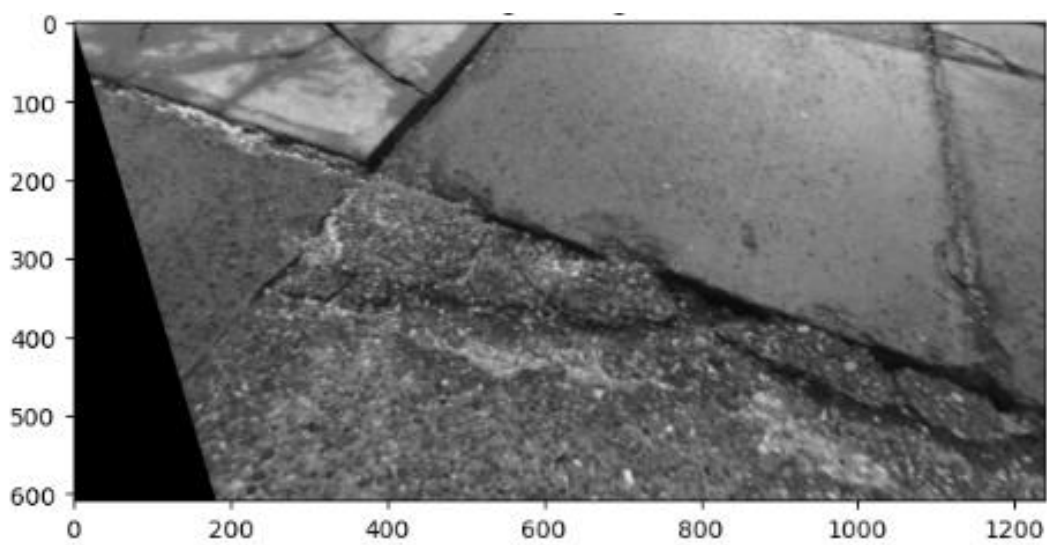


图 4.4 转换后左视灰度图

Fig. 4.4 left view grayscale image after conversion

本文为获得深度信息，在生成视差图中，为了得到准确的信息，才采用生成积分图、均值图与方差图来完成信息的获取。

4.2.2 积分图生成

积分图像是 *Crow* 在 1984 年首次提出，为了提高多尺度透视投影的渲染速度。然后将该技术应用于基于 NCC 的快速匹配、目标检测和 Surf 变换、基于统计的快速滤波等。。

在积分图中，通常用图像左上角到该点坐标所围成的区域的总和来表示某一点的数据。积分图像被认为是一种快速计算矩形区域和的方法，积分图的主要优点如下：

(1) 当计算出积分图像时，我们可以在常量时间内计算图像中任意大小矩形区域的和。

(2) 有了积分图像之后，极大降低了在图像模糊、边缘提取、对象检测的时候的计算量并能够提高计算速度。

在 Viola-Jones 的对象检测框架中首次应用积分图像技术。对于积分图像来说，它的意义是任何位置上的一个像素，表示该点左上角所有像素的总和。积分图有很多生成的算法，不在做举例，本文应用的是优化的 NCC 算法，得出的图像如下：

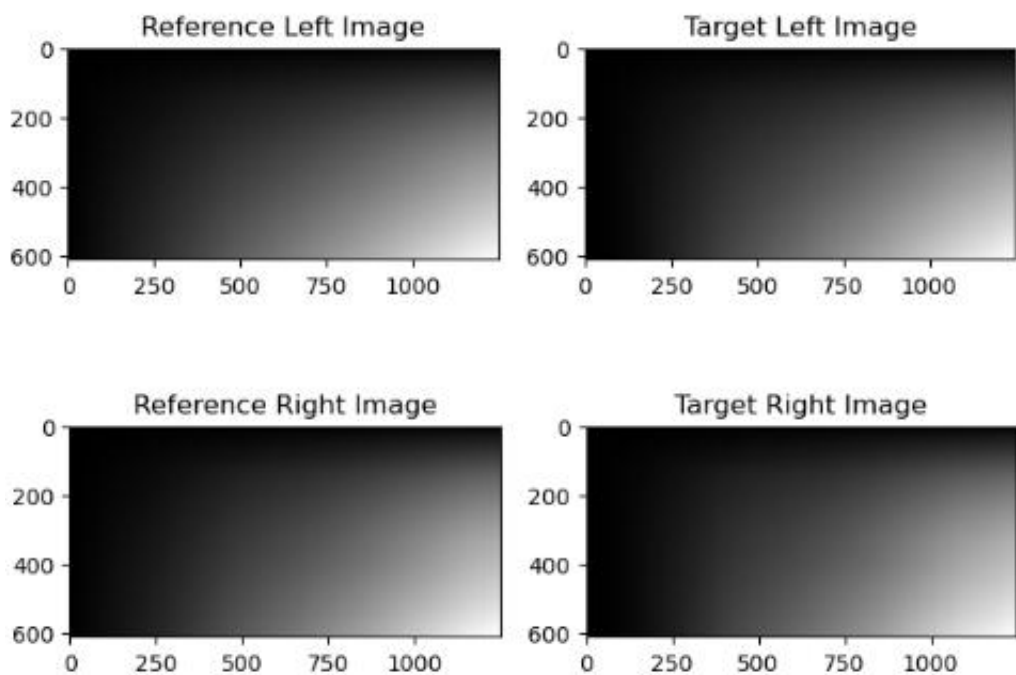


图 4.5 积分图

Fig. 4.5 integrogram

4.2.3 均值图和标准差图生成

为了获得更多的信息，来增强视差图和深度图的准确性和精度，图像像素的均值和标准差是图像处理中常用的方法。图像对比度也可以通过处理均值和标准差图像来提高，达到显著的效果。图像的亮度反应依据是均值，均值越大往往意味着图像亮度越大，反之越小；图像像素值与均值的离散程度依靠标准差进行反映，标准差越大往往意味着图像的质量越好。

对左视图的像素进行计算均值，并生成如下：

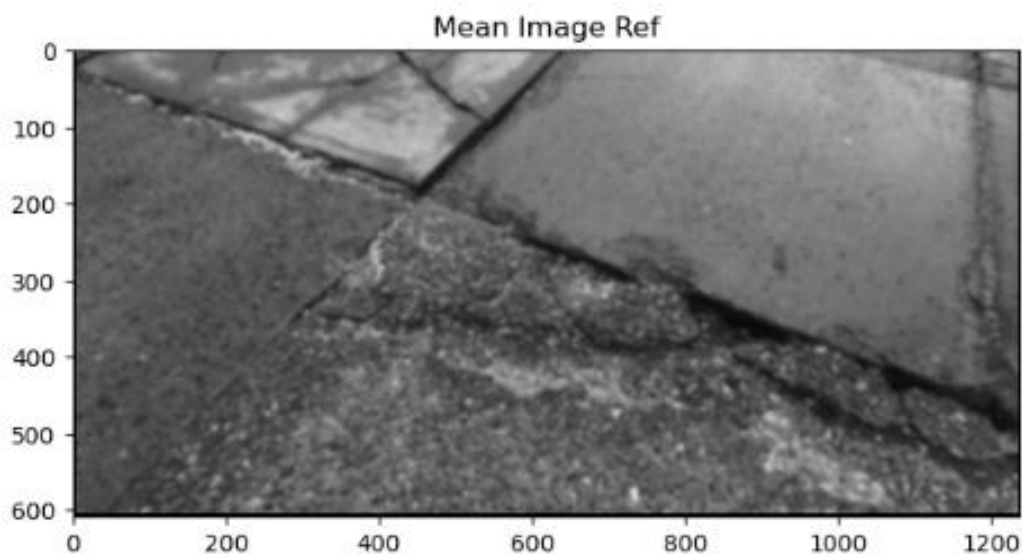


图 4.6 转换前均值图

Fig. 4.6 the preconversion mean graph

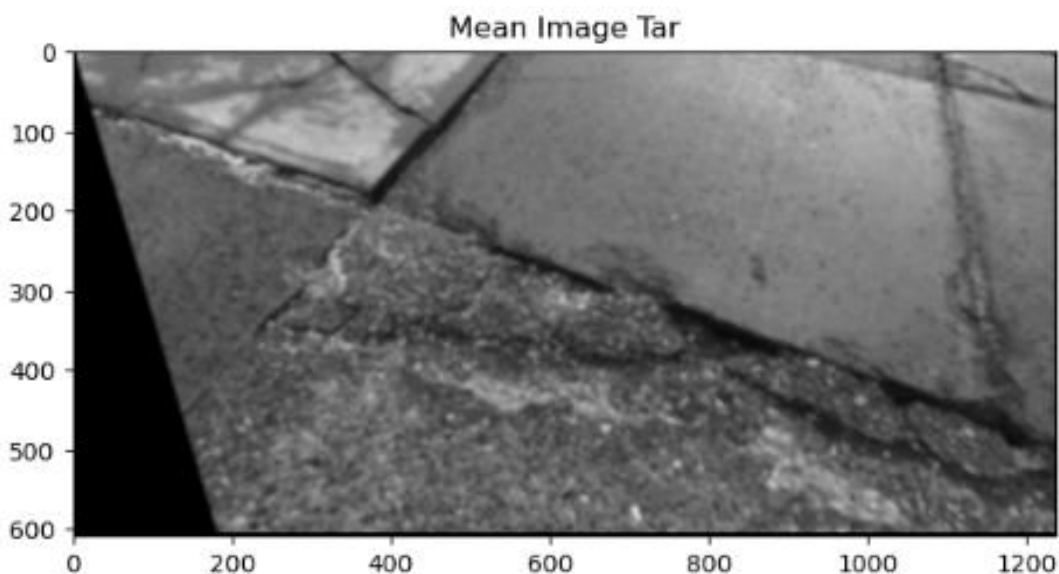


图 4.7 转换后均值图

Fig. 4.7 the conversion mean graph

对左视图的像素进行计算标准差，并生成如下：

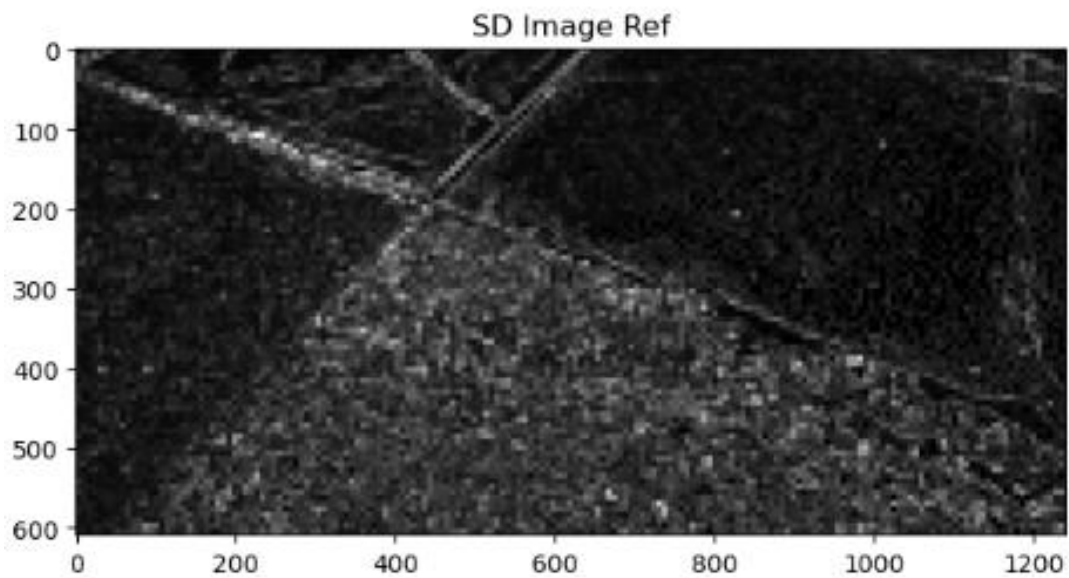


图 4.8 转换前标准差图

Fig. 4.8 standard deviation graph before conversion

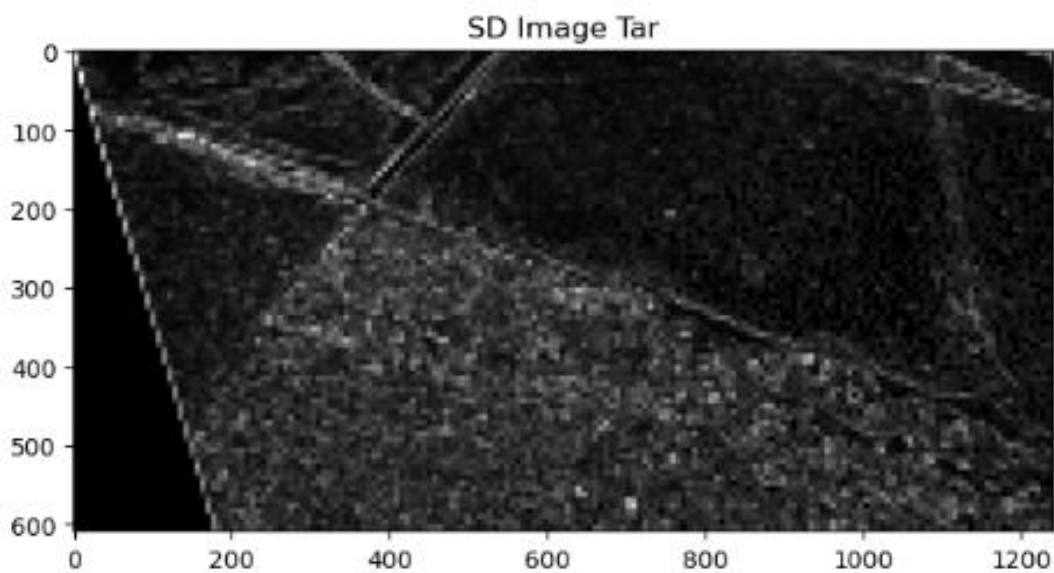


图 4.9 转换后标准差图

Fig. 4.9 standard deviation chart after conversion

4.4 本章小结

在本章中，对 NCC 算法展开了介绍，从理论入手，到程序算法中的实现。首先介绍了 NCC 的由来，根据向量相关引入 NCC 算法，进而推导了传统 NCC 算法，这是一种基于统计学基础的计算两组样本数据相关性的算法，本文基于此原理，实现一种基于模板匹配的图像识别算法。并且对 NCC 进行优化，解决处理速度过慢，不满足实时性等问题，来增强算法的实时性。将 NCC 方程重新排列，以加速 NCC 的执行速度。并设计了程序来分析优化结果，发现重新排列的方程能够增强实时性。并在图像之中，把图片转换成灰度图像，同时生成积分图、均值图、标准差图。积分图能够极大降低图像模糊、边缘提取等的过程的计算量，均值图、标准差图的生产能够提高计算速度来完成方程中的参数的获取，以便于下一步的视差图和深度信息的获取。

5 视差图以及深度信息的获取

经立体校正得到左右图像后，匹配点在同一直线上。利用 OpenCV 中的 BM 算法或 SR 算法计算视差映射，由于 SR 算法的性能明显优于 BM 算法，因此利用 SR 算法计算视差图。

5.1 相关算法介绍

5.1.1 SR 算法

在立体匹配算法中，在可靠特征点集合的指导下，可以有效地估计半稠密或准稠密的视差图，称之为搜索范围传播 (SRP) 理论。

由于视差误差主要发生在图像的均匀区域，传统的视差估计算法如 GCS 等在左右两幅图像上均使用 Harris 角点检测器来进行特征点的获取，并来确定鲁棒的视差点来对应这些特征点，得到它们的差异，从而得到视差。

本文将路面视为一个从图像底部到顶部差异逐渐变化的 GP。在估计一个准确的稠密视差图方面，设计算法将搜索范围 SR 从图像的底部逐行迭代传播到图像的顶部。在第一次迭代中，视差估计执行一个完整的搜索范围。

定义 SR 算法为：

$$SR = \bigcup_{k=u-1}^{u+1} \{sr \mid sr \in [l(k, v+1) - \tau, l(k, v+1) + \tau]\} \quad (5-1)$$

然后，任意像素点 (u, v) 处的 SR 通过上式，从三个估计的相邻差传播，其中 τ 是 SR 的边界，在本文中，设定为 1。

SR 算法将初始点以及到其邻近区域进行匹配，并将搜索范围扩展到该点的左右 2 个点的邻近区域。每个邻近区域将获得一个视差值，使数值相似的函数在传播的搜索范围内最小。然后，将其添加到列表中，除非最小代价低于阈值，因为这意味着不可靠的匹配点，或者得到的视差值违反唯一性约束，即无法进行左右一致性检查，这是因为有遮挡区域。按照图像相似度的顺序绘制对应的点，并重复此过程，直到该集合为空集。通过这种方式，匹配关系从每个点扩散到整个图像的邻近区域，进而完成匹

配。

5.1.2 绝对差异和算法

在数字图像处理中，绝对差异和(SAD)是图像块之间相似性的度量。这是通过取原始块中每个像素与该块中相应像素之间的绝对差值进行比较来计算的。添加这些差异来创建一个简单的块相似性度量，差异图像的 L1 范数或两个图像块之间的曼哈顿距离。很多过程可以应用绝对差异和，例如物体识别，生成立体图像的视差图以及运动估计中的视频压缩。

绝对差异和（SAD）是匹配块算法的一种，具体方法为：将右视图所选模板块与左视图所选模板块对应的像素进行相减，将所有差值取绝对值，然后求和。求解的和表示两个块之间的匹配度，匹配度越小。下图为一个 SAD 算法的例子：

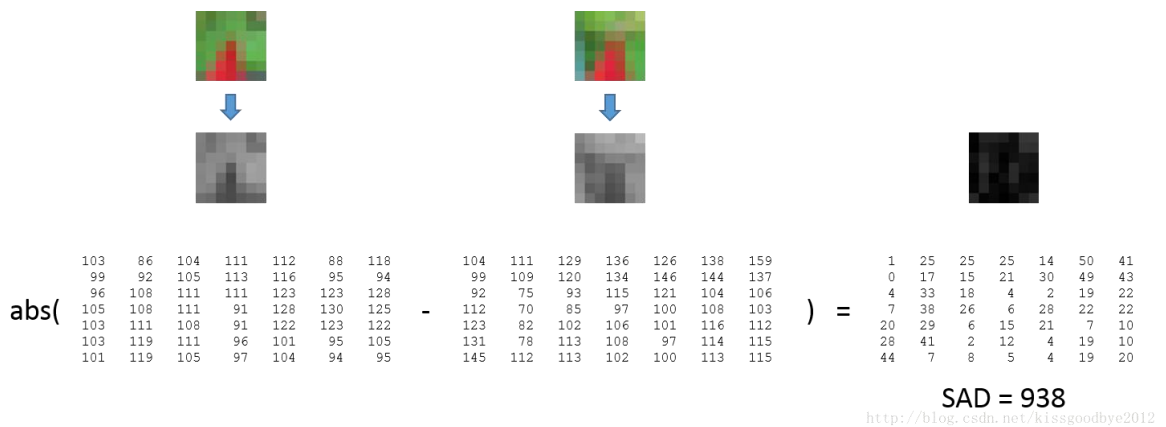


图 5.1 绝对差异和

Fig. 5.1 sum of absolute difference

5.1.3 最大相关值验证

本文未采用 SAD 算法，而是引入最大相关值验证(CMV)算法，来进一步进行验证算法的合理性。由于搜索范围使用 SR 算法进行迭代，当 $c(u,v, d-1)$ 或 $c(u,v, d+1)$ 未计算并与 $c(u,v, d)$ 进行比较时，可能会在亚像素增强中出现误差。为了增强精度并减小误差，通过将三个值进行计算并比较的方法，也就是最大相关值验证的办法（CMV），来进行验证。CMV 算法将运行到视差的相关代价为局部极大值为止。

5.2 视差图

在双目图像中，两个匹配块中心像素之间的水平距离的差异意味着视差。视差使用相同的颜色来表示物体离摄像机的位置相同，视差越大，表示物体离摄像机越近。通过位置的远近，用不同的颜色表示，以便得到位置信息。

生成视差图步骤如下：

- (1) 双目标定进行图像矫正；
- (2) 将双目图像灰度化；
- (3) 搜索范围和方向；

在第二、三章中，通过坐标系转换与视角转换，已经完成了双目标定进行图像矫正，并在第四章中，完成了图像灰度化处理，并在本章前面部分，推导了搜索与迭代的 SR 算法。

根据通过经典的 NCC 算法，用 SR 算法进行迭代，得出的视差图如下：

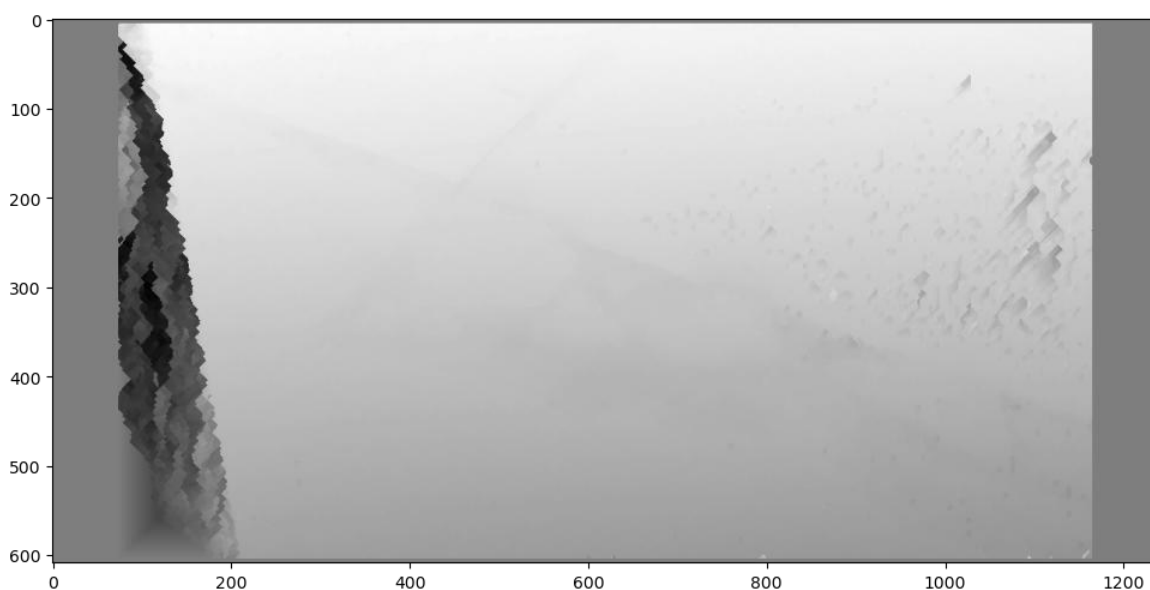


图 5.3 左视图视差图

Fig. 5.3 left view disparity map

Python 中调用的绘图库为 Matplotlib。Matplotlib 库可与 NumPy 一起使用，也可以一起使用其他的图形工具包。灰度视差图虽然能识别视差信息，不能够明显识别视差信息，将视差图在 cmap 下进行颜色转化：

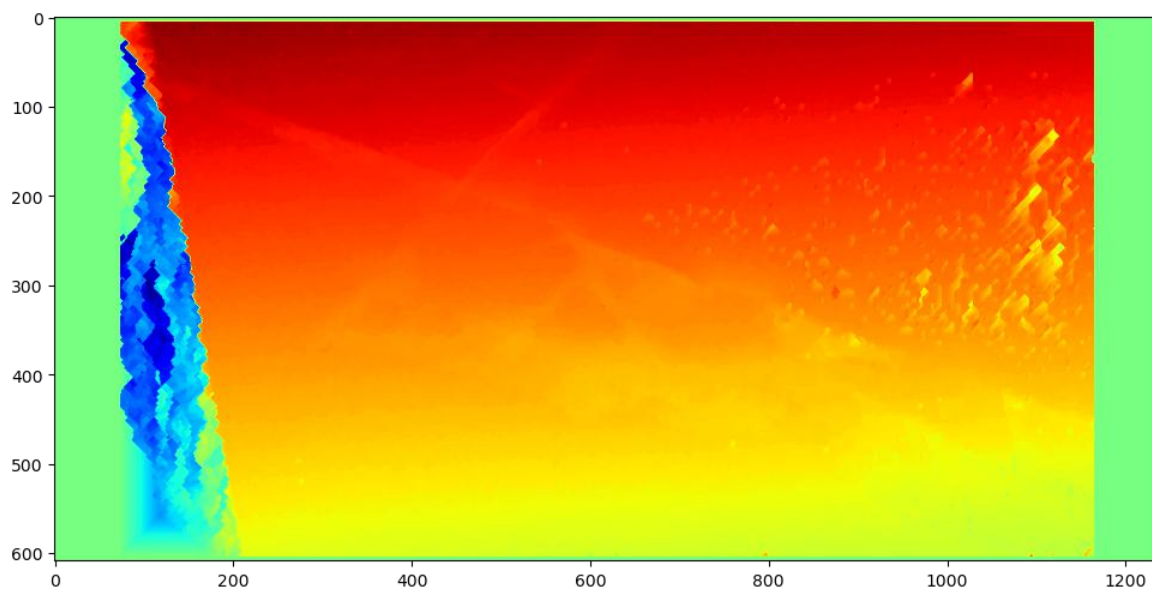


图 5.4 左视图转换后视差图

Fig. 5.4 left view disparity map after conversion

通过重新排列的 NCC 算法运算，性能会增强，计算出来的图像精度更高，用 SR 算法进行迭代，得出的视差图如下：

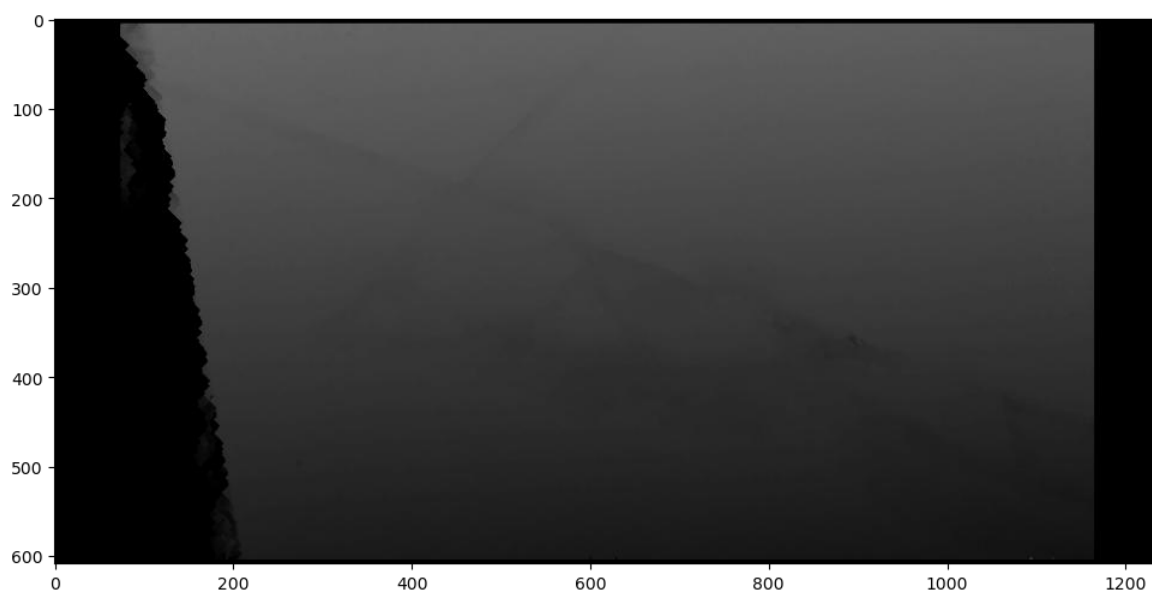


图 5.5 更新的左视图视差图

Fig. 5.5 new left view disparity map after conversion

通过 Matplotlib 库，将视差图在 cmap 下进行颜色转化：

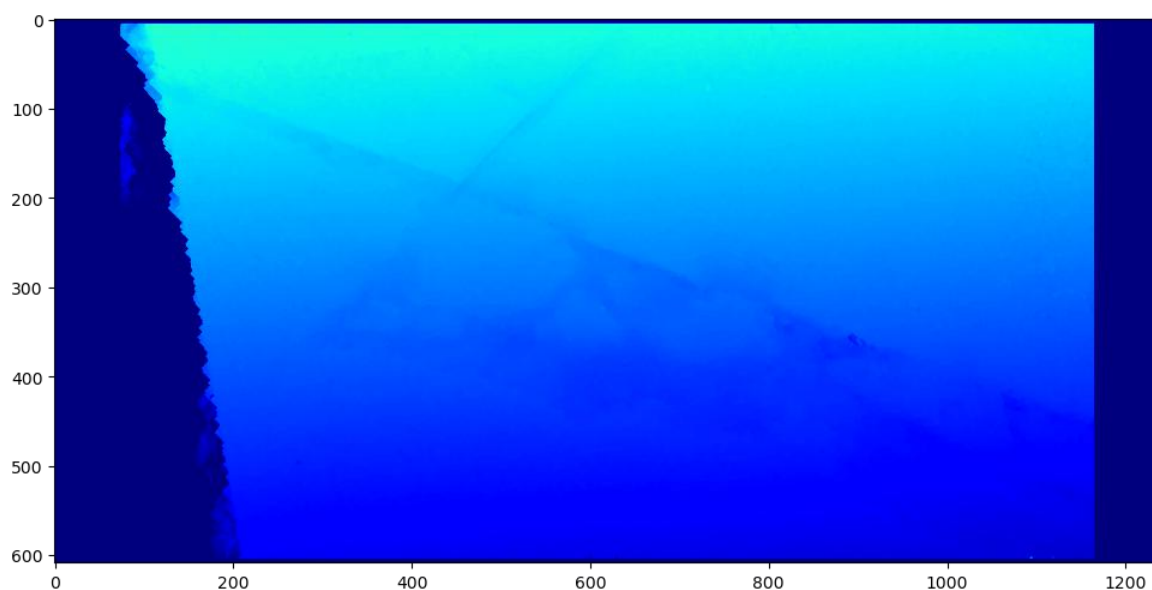


图 5.6 更新的转换后左视图视差图

Fig. 5.6 new left view disparity map after conversion

5.3 深度信息获取

深度映射又称深度图像、距离图像，是将图像从图像收集器到场景的距离(深度)值作为像素值。

深度图的获取方法有：结构光法、激光雷达深度成像法、坐标测量机法、莫尔条纹法、计算机立体视觉成像等等。深度图相比于点云图来说，还是有区别的。点云图是指，当一束激光击中物体表面时，反射的激光会携带方向和距离等信息。

如果激光束沿一定路径扫描，由于扫描过程非常精细，可以获得大量的激光点，从而形成激光点云图案，因此会记录反射的激光光斑信息。对于深度图像来说，深度图像与点云数据之间的转换时可以通过坐标转换来计算的。还有，具有规则和必要信息的点云数据可以被反演成深度图像。在一定条件下，利用 PCL 库从点云中提取深度图，将两者相互转化，利用相机参数将深度图转化为点云图。

视差与深度的关系如下：

像素(pixel)是视差的单位，毫米(mm)是深度的单位。视差与深度的单位不同，为了达到获取深度信息的目的，我们要将视差单位转换为深度单位。像素尺寸与深度范围成反比，随着像素尺寸的增大，深度范围减小。深度与视差的关系如下：

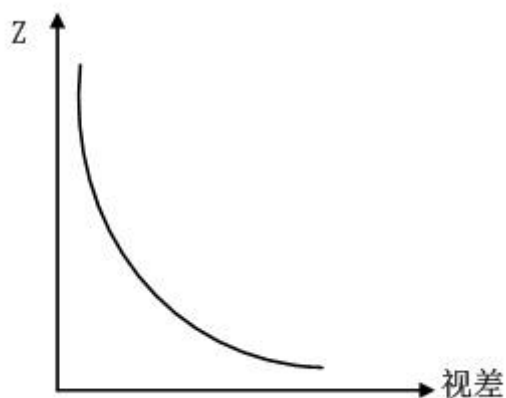


图 5.5 深度和视差的关系

Fig. 5.5 depth and disparity

可以通过绝对差值法来实现视差和深度单位的转换。绝对差值图的计算方法如下：

$$D = |L - R| \quad (5-2)$$

式中，L、R 和 D 分别代表左视差视图、右视差视图和对应的深度值的亮度值。

通过绝对差值法来计算图像深度，得到的深度图不是严格意义上的深度图。但是它的计算方法最为简单，速度快，它给出的结果可以作为参考。如下图，给出一个深度图的参考：

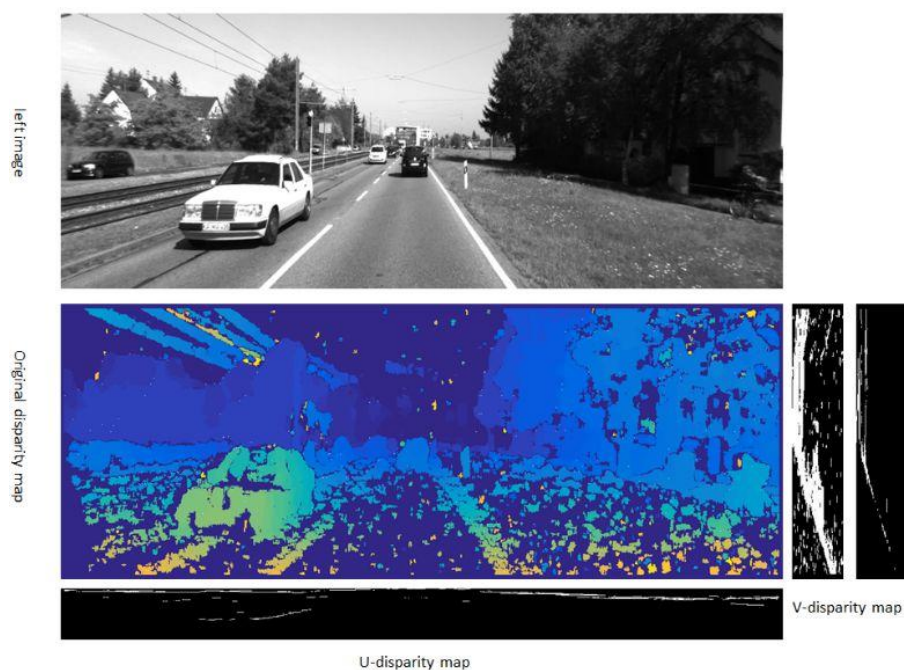


图 5.6 视差图转换为深度信息

Fig. 5.6 The disparity map is converted to a depth information

5.4 本章小结

本章主要介绍了 SR 算法来进行视差图的迭代运算，以及通过 NCC 算法和 SR 算法来得到视差图的过程、并介绍了结对差异和算法和最大相关值验证算法，以便后续对处理结果的验证，并给出了视差和深度的关系，以便于深度信息的获取。在 SR 算法中，本文将路面视为一个从图像底部到顶部差异逐渐变化的，在估计一个准确的稠密视差图的过程中，设计算法将搜索范围 SR 从图像的底部逐行迭代传播到图像的顶部。同时，通过传统 NCC 算法和优化 NCC 算法两种不同的算法来运行，得到视差图，不难发现，两者大体相同，但仍存在差异。优化 NCC 算法生成的视差图明显精度好于传统算法。

结 论

本文讨论了基于双目视觉的路面信息的研究目的和研究意义，通过分析目前关于三维重建的研究方法的研究现状，提出用无源传感器的双目立体视觉为主要研究方法来展开研究，通过 Anaconda 开发环境下的 python 语言来实现算法程序，分析对比不同算法，以及相关优化，来得到双目图片的深度信息。根据本文的研究，得出结论如下：

1. 双目立体视觉系统在机器视觉领域应用许多，这是计算机视觉系统中最接近的立体视觉系统。通过模仿人类视觉原理，获得了包括深度信息在内的多种三维数据。立体匹配通过引入不同的算法可以解决遮挡区域、低纹理或无纹理区域、深度不连续区域和噪声等问题。本文通过双目视觉立体匹配完成了对路面信息的获取处理并得到了深度信息。

2. 本文通过对双目图片对特征点的提取和对特征点的匹配，完成了特征匹配。通过对 BF 算法与间距过滤匹配算法的分析比较，得到较高精度的匹配算法。对左右视图的特征点匹配时，完成了目标要求。

2. 本文将 PnP 算法与最小二乘法进行比较，采用最小二乘法来估计参数，并对相机模型进行了简化，来达到视角转换的目的，使变换后图像中的路面与参照系中的路面相似。解决了双目相机拍摄得到的左右视图中，路面总是存在透视失真，从而进一步影响了块匹配的准确性的问题。

3. 对归一化互相关算法进行优化，通过使用优化的归一化互相关算法被用来计算匹配代价，准确的匹配代价值能够通过分层迭代的方式得到，结果较为精确，并提高了处理图像的速度。并通过对 SR 算法的理论推导和计算。论述了 SAD 与 CMV 算法可以为后续来验证图片的相关度。最终得到了视差图，进而能够推导深度信息，达到研究目的。

除以上研究结论外，本文仍存在一些不足需要进一步研究和讨论：

1. 算法的实时性还不够，通过对 NCC 算法的重新排列以及优化，算法的实时性增强，但程度仍然不够，需要进行后续的优化处理。

2. 本文在研究过程中仅得到深度信息，但未对深度信息的准确性进行算法验证，以及对出现的误差问题的校正处理。

3. 本文的下一步研究计划为：通过使用马尔科夫模型来对视差图进行全局细化处理，并为了增强图片的精度，通过设计算法来达到实现视差图图像亚像素精度的目的。

参 考 文 献

- [1]王宽, 杨环, 潘振宽, 等. 基于单目和双目视觉信息的全参考立体图像质量评价模型[J]. 计算机工程. 2021. 11-24. 10.19678/j.issn.1000-3428.0060191.
- [2]Fan Rui, Xiao Ai, and Naim Dahnoun. Road surface 3D reconstruction based on dense subpixel disparity map estimation[J]. IEEE Transactions on Image Processing, 2018, 27(6): 3025-3035.
- [3]张洪玮, 李鑫. 三维激光扫描技术在建筑数字化模型中的应用[J]. 安徽建筑, 2021, 28(01):15-16.
- [4]金菁, 李运璋, 曹毅. 三维激光扫描仪在建筑领域的应用[J]. 建筑与文化, 2014(04):168- 171.
- [5]王瑶, 项鹏, 孟春阳, 马利兵, 杨秀秀, 赵川. 新一代 Kinect 传感器关键技术综述[J]. 黑龙江科技信息, 2015(24):84-86.
- [6]Wen Yin-fang. Data point cloud preprocessing and feature corner detection algorithm[D]. Harbin:College of Automation, Harbin Engineering University, 2007.
- [7]魏亚, 肖庸, 闫闯等. 基于点云预处理的路面三维重构数据优化[J]. 吉林大学学报(工学版), 2020, 50(03):987-997.
- [8]陈 华, 王立军, 刘刚. 立体匹配算法研究综述[J]. 高技术通讯 2020, 30(2): 157-165.
- [9]刘志花. 面向立体视觉的匹配算法研究 [D] . 北京: 中国科学院研究生院, 2011: 19-28
- [10]Lim KB, Kee WL. Geometrical-analysis-based algorithm for stereo matching of single-lens binocular and multi-ocular stereo vision system [J] . Journal of Electronic Science and Technology, 2012, 10(2) : 107-112
- [11]Barnard S T, Fishler M A. Computational stereo[J]. ACM Computing, Surveys, 1982, 14(4) : 553-572

- [12]朱松,曹丹华,吴裕斌.视觉导航中路面检测方法与 DSP 实现[J].仪表技术与传感器,2015(12):110-112.
- [13]施敏.基于双目立体视觉的路面三维形貌重构算法研究[D].东南大学,2017.
- [14]贺鹏艳.基于双目视觉的倒车障碍物检测及预警方法研究[D].西安理工大学,2014.
- [15]吴一全,刘莉.基于视觉的车道线检测方法研究进展[J].仪器仪表学报,2019,40(12):92-109.
- [16]LEMAIRE T, BERGER C, JUNG I, et al. Vision-based SLAM: stereo and monocular approaches[J]. International Journal of Computer Vision, 2007, 74(3): 343-364.
- [17]亢宇欣, 谌贵辉, 邓宇, 等. 多测度融合的立体匹配. 算法研究[J]. 图学学报, 2019, 40(4): 711-717.
- [18]TIPPETTS B, LEE D, LILLYWHITE K, et al. Review of stereo vision algorithms and their suitability for resource-limited systems[J]. Journal of Real-Time Image Processing, 2016, 11(1): 5-25.
- [19]陈炎,杨丽丽,王振鹏.双目视觉的匹配算法综述[J].图学学报,2020,41(05):702-708.
- [20]李宇杰,李焯鹏,张为公.基于视觉的三维目标检测算法研究综述[J].计算机工程与应用,2020,56(01):11-24.
- [21]路雪,刘坤,程永翔.一种深度学习的非机动车目标检测算法[J].计算机工程与应用,2019,55(8):182-188.
- [22]张震,张照崎,苗志滨,朱留存,李修明,麦冬,张坤伦,周瑞凯.基于 Harris-Hist 的特征匹配及目标定位算法[J].吉林大学学报(理学版),2021,59(02):333-341.
- [23]赵文彬,张艳宁.角点检测技术综述[J].计算机应用研究,2006(10):17-19+38.
- [24]GANDE R W, GOLUB G H, STREBEL R. Least-squares fitting of circles and ellipses [J]. BIT, 1994, 34(4): 558-578.
- [25]吴晓阳.基于 OpenCV 的运动目标检测与跟踪 [D]. 杭州:浙江大学,2008.
- [26]Hiroshi Hattori and Atsuto Maki, "Stereo without depth search and metric calibration," in Computer Vision and Pattern Recognition, 2000.Proceedings. IEEE Conference on. IEEE, 2000, vol. 1, pp. 177 - 184.

- [27] Hiroaki Nakai, Nobuyuki Takeda, Hiroshi Hattori, Yasukazu Okamoto, and Kazunori Onoguchi, “A practical stereo scheme for obstacle detection in automotive use,” in Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on. IEEE, 2004, vol. 3, pp. 346 – 350.
- [28] 李占旗, 高继东, 彭凯, 刘全周, 许小伟. 基于 PnP 算法的前方车辆测距方法[J]. 国外电子测量技术, 2020, 39(12):104-108.
- [29] 曹诗雨, 刘跃虎, 李辛昭. 基于 FastR-CNN 的车辆目标检测[J]. 中国图象图形学报, 2017, 22(5): 671—677.
- [30] Roche A , Malandain G , Pennec X , et al. The Correlation Ratio as a New Similarity Measure for Multimodal Image Registration[C]// International Conference on Medical Image Computing & Computer-assisted Intervention. Springer Berlin Heidelberg, 1998.
- [31] 张宇驰. 一种改进的基于 Harris 角点的图像拼接方法[J]. 机械工程师, 2017(10).

致 谢

时间真的很快，一转眼我即将毕业了。

还记得刚来河北工业大学的时候，红桥的生活很快乐，加上刚来大学的无畏和冲劲，让我对未来的生活充满希冀。转眼一年过去了，在北辰校区度过了我剩下的三年本科时光，图书馆、ABCD 教学楼、宿舍、食堂，这些充满我回忆的地方，真的因为即将毕业变得让人不舍、让人难忘。

感谢刘璇老师与中汽研的许晟杰工程师，有两位老师的悉心指导，才能让我完成我的毕业设计。老师们耐心地回答我的疑问，并且在我遇到瓶颈前进较难时帮我修正思路，使我少走了许多弯路。周末或下班时间我有问题询问时，亦第一时间对我提供帮助。自毕设开始起，保证每周 1 至 2 次会议汇报进度，督促我按时完成任务。

还要感谢与我同组的毕设同学们，我们一起努力、互相帮助、互相打气、互相学习，使我们互相监督、不虚度时光、按时完成毕设任务。尤其感谢汽研的邢王源同学和王彭超同学，在汽研中我们经常交流，解决了彼此的疑惑，顺利完成毕业设计任务。

还要感谢机械工程学院车辆工程系，感谢学院和车辆系给我们本次校企联合毕设的机会，在企业的时光是难忘的，收获很多，学习知识的同时，磨炼了我的意志，使我更加坚韧。

同时，感谢机械工程学院辩论队和河北工业大学校辩论队，让辩论贯穿了我四年的大学生涯，让我在辩论之中收获了许多友谊和快乐。尤其是，在最后一个学期，让我还有机会参与到校赛之中，虽然结果不进人如意，输掉了比赛，但比赛的过程尤其珍贵。在这个赛场，要感谢的人太多太多，思思和旺旺，我院队最好的队友；铁铁和花花，我校队最好的搭档，还有各个学长姐和学弟学妹们对我的影响，真的让我四年的课余生活，温暖又充实。

一直总觉得自己好小，没想到转眼到了应该读研的年纪。很过事情都应该自己学会独自面对。这四年，真的，逐渐学会了独立，很多事情能够自己一个人去做，开始成长，成为更好的自己。

最后的最后，再次感谢河北工业大学。当我真的写下这段文字的时候，我意识到，我终将毕业。这四年的美好时光对我的人生来说，有非常重要的意义。“勤慎公忠”是母校校训，大抵是希望每个河工学子能够勤勤恳恳、踏实做人，成为对社会、对国家有用的人才。在这四年中，在学习中，在娱乐中，在校园的每个角落，都或多或少的充满着我的回忆。我即将面对新的研究生生活，对未来充满的不确定因素感到迷茫和彷徨，但在母校四年的培养下，我有底气、也有实力面对未来的新生活！