

文本复制检测报告单(全文标明引文)

ADBD2017R_20170527154302427905889863

检测时间：2017-05-27 15:43:02

检测文献：双目视觉立体匹配算法设计与实现 论文

作者：王灏

检测范围：

中国学术期刊网络出版总库
中国博士学位论文全文数据库/中国优秀硕士学位论文全文数据库
中国重要会议论文全文数据库
中国重要报纸全文数据库
中国专利全文数据库
互联网资源(包含贴吧等论坛资源)
英文数据库(涵盖期刊、博硕、会议的英文数据以及德国Springer、英国Taylor&Francis 期刊数据库等)
港澳台学术文献库
优先出版文献库
互联网文档资源
图书资源
CNKI大成编客-原创作品库
大学生论文联合比对库
个人比对库

时间范围：1900-01-01至2017-05-27

指导教师 宣琦

检测结果

总文字复制比：12.1% 跨语言检测结果：0%

去除引用文献复制比：11.6% 去除本人已发表文献复制比：12.1%

单篇最大文字复制比：3.4%

重复字数：[2839]	总字数：[23506]	单篇最大重复字数：[792]
总段落数：[7]	前部重合字数：[901]	疑似段落最大重合字数：[901]
疑似段落数：[7]	后部重合字数：[1938]	疑似段落最小重合字数：[101]

指标：☐ 疑似剽窃观点 ☒ 疑似剽窃文字表述 ☐ 疑似自我剽窃 ☐ 疑似整体剽窃 ☐ 过度引用

表格：0 脚注与尾注：0

21.5% (901)	中英文摘要等 (总4199字)
10.4% (310)	第1章绪论 (总2988字)
12.9% (451)	第2章双目立体匹配原理 (总3490字)
15.5% (409)	第3章数据集 (总2646字)
10.6% (492)	第4章立体匹配算法设计 (总4647字)
2.2% (101)	第5章立体视觉实现与运行 (总4549字)
17.7% (175)	第6章总结与展望 (总987字)

(注释：无问题部分 文字复制比部分 引用部分)

1. 中英文摘要等

总字数：4199

相似文献列表 文字复制比：21.5%(901) 疑似剽窃观点：(0)

1	基于增强现实游戏中的标识注册技术 杨智 - 《大学生论文联合比对库》 - 2015-06-09	18.9% (792) 是否引证：否
2	振动分析仪的界面设计与实现 陈春帆 - 《大学生论文联合比对库》 - 2015-06-12	17.3% (726) 是否引证：否

3	振动分析仪的界面设计与实现 陈春帆 - 《大学生论文联合比对库》 - 2015-06-14	17.3% (726) 是否引证：否
4	大并发订票系统的设计与实现 陈洪翔 - 《大学生论文联合比对库》 - 2016-06-01	17.2% (723) 是否引证：否
5	基于单片机的程控信号发生器设计 孙煜烽 - 《大学生论文联合比对库》 - 2015-06-14	17.1% (716) 是否引证：否
6	面向屏幕视频的HEVC帧内编码快速算法研究 尹煜 - 《大学生论文联合比对库》 - 2015-06-12	16.6% (698) 是否引证：否
7	基于嵌入式系统的传感器网络网关设计 陈寒 - 《大学生论文联合比对库》 - 2015-06-10	16.3% (686) 是否引证：否
8	林秀琳_201103120413_自动化 (陈朋, 论文) 林秀琳 - 《大学生论文联合比对库》 - 2015-06-12	16.1% (677) 是否引证：否
9	基于Hadoop的推荐算法分析与应用 王俊杰 - 《大学生论文联合比对库》 - 2015-06-10	15.9% (668) 是否引证：否
10	基于无线网络的火灾早期预警控制台的设计 (硬件) 侯学文 - 《大学生论文联合比对库》 - 2015-06-12	15.8% (663) 是否引证：否
11	谢莹莹_2011031203422_自动化 (陈朋, 论文) 谢莹莹 - 《大学生论文联合比对库》 - 2015-06-12	15.6% (655) 是否引证：否
12	旋转轴系非接触光电通信系统的设计 陆阳 - 《大学生论文联合比对库》 - 2015-06-10	15.4% (648) 是否引证：否
13	基于TMS320VC5509A的简易音频信号分析仪硬件设计 陈昌洪 - 《大学生论文联合比对库》 - 2015-06-14	15.3% (642) 是否引证：否
14	无线血糖管理系统——Android手机客户端程序设计 姚文超 - 《大学生论文联合比对库》 - 2015-06-11	15.1% (636) 是否引证：否
15	民用智能配电箱的硬件设计 戴圣杰 - 《大学生论文联合比对库》 - 2015-06-11	14.9% (626) 是否引证：否
16	民用智能配电箱的硬件设计 戴圣杰 - 《大学生论文联合比对库》 - 2015-06-14	14.9% (626) 是否引证：否
17	基于DSP的永磁同步电机矢量控制器的研究 钟思恒 - 《大学生论文联合比对库》 - 2015-06-15	14.7% (617) 是否引证：否
18	无线静态应变测量系统设计 张晓峰 - 《大学生论文联合比对库》 - 2015-06-09	14.6% (611) 是否引证：否
19	资产管理指令交易软件自动化测试设计 郑亚珍 - 《大学生论文联合比对库》 - 2015-06-10	14.1% (591) 是否引证：否
20	基于光OFDM通信系统的FFT处理器设计及FPGA实现 黄强 - 《大学生论文联合比对库》 - 2015-06-14	13.9% (585) 是否引证：否
21	基于主动式全景视觉的管道内壁貌的检测技术研究 蔡盈梅 - 《大学生论文联合比对库》 - 2016-06-02	13.3% (558) 是否引证：否
22	MANET网络安全路由技术研究 童姜况 - 《大学生论文联合比对库》 - 2015-06-09	13.0% (544) 是否引证：否
23	道路交通状态预测方法研究 陈彦昌 - 《大学生论文联合比对库》 - 2015-06-10	12.4% (522) 是否引证：否
24	IP_RAN综合承载解决方案技术分析 项阳 - 《大学生论文联合比对库》 - 2014-05-23	7.9% (332) 是否引证：否
25	基于模式搜索算法的无线传感器网络资源分配方法设计 柴寅凯 - 《大学生论文联合比对库》 - 2015-06-09	7.4% (311) 是否引证：否
26	基于深度学习的鸟类图像分类系统设计 江汪发 - 《大学生论文联合比对库》 - 2016-03-27	5.1% (216) 是否引证：否

8191504699000

本科毕业设计论文

题目：双目视觉立体匹配算法设计与实现

作者姓名王灏

指导教师宣琦研究员

专业班级通信工程1301

学院信息工程学院

提交日期 2017年6月14日

浙江工业大学本科毕业设计论文

双目视觉立体匹配算法设计与实现

作者姓名：王灏

指导教师：宣琦研究员

浙江工业大学信息工程学院

411670533020002017年6月

Dissertation Submitted to Zhejiang University of Technology
for the Degree of Bachelor

Binocular Stereo Matching Algorithm Design

And Implementation

Student: Hao Wang

Advisor: Associate Professor Qi Xuan

College of Information Engineering

Zhejiang University of Technology

June 2017

浙江工业大学

本科生毕业设计(论文、创作)任务书

专业通信工程班级通信1301 学生姓名/学号王灏/201303090119

一、设计(论文、创作)题目：

双目视觉立体匹配算法设计与实现

二、主要任务与目标：

双目视觉立体匹配系统，实现双目视觉图像数据采集及匹配，实现场景深度信息的计算，具有较快的运行速度。

三、主要内容与基本要求：

1.采集双目视觉图像 2.设计双目视觉立体匹配算法对双目视觉图像进行匹配，完成相关软件程序，达到指定性能指标 3.撰写毕业论文，提交相关的算法程序流程图及程序代码。

四、计划进度：

2016.12.20-2017.2.15 收集相关资料文献，学习数据库以及编程知识；完成外文翻译、文献综述；熟悉课题，做好开题准备，有初步设计方案；2017.2.16-3.5 完成开题报告，参加开题交流；2017.3.6-4.23 完成数据采集，算法实现，并开发软件系统，接收中期检查；2017.4.24-5.25 算法测试与改进，做出最终设计成品，撰写毕业论文初稿；2017.5.26-6.12 论文修改，毕业答辩，提交相关文档资料。

五、主要参考文献：

[1]N.Mayer,E.Ilg,P.H?usser,P.Fischer,D.Cremers,A.Dosovitskiy,T.Brox.A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation[C].2015:arXiv preprint arXiv:1512.02134

[2]J.Zbontar,Y.LeCun.Stereo matching by training a convolutional neural network to compare image patches[J].Journal of Machine Learning Research.2016.17:1-32

[3]J.Zbontar,Y.LeCun.Computing the stereo matching cost with a convolutional neural network[C].IEEE Conference on Computer Vision and Pattern Recognition.2015:1592-1599

任务书下发日期 2016 年 12 月 20 日

设计(论文、创作)工作自 2016 年 12 月 20 日至 2017 年 6 月 20 日

设计(论文、创作)指导教师宣琦

系主任(专业负责人)

主管院长

双目视觉立体匹配算法设计与实现

摘要

立体视觉技术一直是计算机视觉领域研究的一个重要分支，特别是近年来，随着无人驾驶、工业自动化、机器人技术的发展，如何从二维图像中提取出三维的深度信息成为一个重要的命题。一个完整的双目视觉系统包括离线标定，双目矫正，立体匹配和三维重建。其中立体匹配是双目视觉技术的核心问题，也是本文研究的重点。立体匹配算法的精度和速度对于双目立体视觉系统的性能起着至关重要的作用。

近年来深度学习技术开始被广泛应用到各个领域，特别是在图像处理领域，深度学习的特征提取特性表现尤为优秀。随着深度学习技术的深入发展，研究人员开始尝试在一些早已被充分研究的问题中应用神经网络，试图提高图像处理的效率和准确率。

本文主要通过卷积神经网络来计算立体匹配代价，并通过左右一致性检查，半全局匹配，中值滤波和双边滤波等后处理步骤进一步优化视差图。本文使用立体视觉数据集KITTI 2012和KITTI 2015构造混合数据集来进行卷积神经网络模型的训练、验证和测试。此外，本文结合OpenCV和Matlab工具箱进行离线标定，基于TensorFlow框架构建卷积神经网络模型，最后利用OpenCV计算机视觉库进行三维重建。同时，我们使用Flask框架和Ngrok服务搭建Web服务器，并基于Qt框架构建客户端程序实现双目数据上传和结果显示。

关键词：双目立体视觉，立体匹配，卷积神经网络，TensorFlow，OpenCV

BINOCULAR STEREO MATCHING ALGORITHM DESIGN AND IMPLEMENTATION

ABSTRACT

Stereo vision technology has always been an important branch of computer vision research. Especially in recent years, with the development of self-driving, industrial automation and robot technology, how to extract three-dimensional depth information from two-dimensional images has become an important topic. A complete binocular vision system includes off-line calibration, binocular rectification, stereo matching and three-dimensional reconstruction. Stereo matching is the core of binocular vision technology and the focus of this paper. The accuracy and speed of the stereo matching algorithm play a crucial role in the performance of the binocular stereo vision system.

In recent years, deep learning technology has been widely applied in various fields. Especially in the field of image processing, deep learning performed quite well in extracting feature. With the development of deep learning, researchers began to try using neural networks in some issues that have already been fully studied in an attempt to improve the efficiency and accuracy of image processing.

In this paper, we compute the stereo matching cost with a convolution neural network, and optimize our disparity map with a left-right consistency check, semiglobal matching, a median filter and a bilateral filter. We construct a hybrid data set with KITTI 2012 and KITTI 2015 stereo data sets to training, evaluating and testing our network model. Besides, we calibrate camera with Matlab and OpenCV, build our network with TensorFlow, and finally use OpenCV computer vision program library for three-dimensional reconstruction. In the meanwhile, we build the web server with Flask and Ngrok service, and build the client based on Qt for binocular data upload and result display.

Key Words: binocular stereo vision, stereo matching, convolutional neural network, TensorFlow, OpenCV

目录

摘要	I
ABSTRACT	II
目录	III
第1章绪论	1
1.1 研究背景及意义	1
1.2 国内外研究现状	2
1.3 主要研究内容与章节安排	3
第2章双目立体匹配原理	5
2.1 双目立体视觉基本原理	5
2.1.1 摄像机成像模型	5
2.1.2 对极几何	7
2.1.3 视差理论	8
2.2 立体匹配的经典步骤	9
2.3 立体匹配的约束条件	10
2.4 立体匹配的难点	10
2.5 本章小结	13
第3章数据集	14
3.1 立体视觉数据集	14

3.1.1 KITTI数据集	14
3.1.2 Middlebury数据集	15
3.2 构造数据集	15
3.3 本章小结	16
第4章立体匹配算法设计	17
4.1 卷积神经网络	17
4.2 网络结构	18
4.3 后处理步骤	19
4.3.1 左右一致性检查	20
4.3.2 半全局匹配	20
4.3.3 滤波	22
4.4 本章小结	22
第5章立体视觉实现与运行	23
5.1 图像矫正	23
5.1.1 双目摄像机选取	23
5.1.2 摄像机离线标定	24
5.1.3 双目矫正与三维重建	26
5.2 立体匹配	27
5.2.1 实验环境	27
5.2.2 模型训练	28
5.2.3 立体匹配细节	28
5.3 服务器搭建与客户端设计	31
5.3.1 服务端搭建	31
5.3.2 客户端设计	32
5.4 本章小结	36
第6章总结与展望	37
6.1 论文总结	37
6.2 论文研究的不足与展望	37
参考文献	38
致谢	41

指 标

疑似剽窃文字表述

1. 完成相关软件程序，达到指定性能指标 3.撰写毕业论文，提交相关的算法程序流程图及程序代码。
四、计划进度：
2016.12.20-2017.2.15 收集相关资料文献，学习数据库以及编程知识；完成外文翻译、文献综述；熟悉课题，做好开题准备，有初步设计方案；2017.2.16-3.5 完成开题报告，参加开题交流；2017.3.6-4.23 完成数据采集，算法实现，并开发软件系统，接收中期检查；2017.4.24-5.25 算法测试与改进，做出最终设计成品，撰写毕业论文初稿；2017.5.26-6.12 论文修改，毕业答辩，提交相关文档资料。

2. 第1章绪论

总字数：2988

相似文献列表 文字复制比：10.4%(310) 疑似剽窃观点：(0)

1	基于双目立体视觉的旋转体三维重建技术研究 - 豆丁网 - 《互联网文档资源 (http://www.docin.com) 》 - 2016	5.9% (177) 是否引证：否
2	基于双目立体视觉的旋转体三维重建技术研究 胡龙(导师：邢英杰) - 《大连理工大学硕士论文》 - 2007	4.8% (143) 是否引证：否
3	基于计算机图像定位的伺服平台研究 聂华;- 《电子设计工程》 - 2013	4.2% (125) 是否引证：否

4	立体视觉技术的应用与发展 周星;高志军; - 《工程图学学报》 - 2010	3.9% (118) 是否引证：是
5	月球车的自主视觉导航 丁同才;高志军;胡震宇; - 《机器人技术与应用》 - 2008	3.9% (118) 是否引证：否
6	计算机立体视觉研究的进展 钱曾波,邱振戈,张永强 - 《测绘学院学报》 - 2001	3.5% (104) 是否引证：否
7	猫狗图像分类算法设计及软件实现 陈官富 - 《大学生论文联合比对库》 - 2016	3.2% (97) 是否引证：否
8	立体视觉与体视匹配 李作新,胡广宁 - 《小型微型计算机系统》 - 1990	2.9% (87) 是否引证：否
9	双目视觉技术在散货堆场中的应用研究 吴振华; - 《自动化博览》 - 2015	1.8% (55) 是否引证：否
10	南极科考模拟支撑平台上位机系统设计与实现 高凡(导师：张侃健) - 《东南大学硕士论文》 - 2016	1.4% (41) 是否引证：否
11	双目立体视觉测距算法研究 明祖衡(导师：李银林) - 《北京理工大学硕士论文》 - 2008	1.1% (32) 是否引证：否

原文内容

第1章绪论

1.1 研究背景及意义

众所周知，自然界中的物体都是三维立体的，人类可以通过双眼产生的二维图像来获取三维世界信息，并对各种物体和信息进行分类和识别。但是一般的摄影系统只能把三维的物体以二维的形式保存和记录下来，丢失了大量的深度信息。为了让计算机也能像人类一样“看到”并理解这个世界，人类一直在深入的研究自身的视觉系统，并衍生出来计算机视觉这一学科。立体视觉是计算机视觉的一个重要分支，是一种从二维平面图像中恢复出三维深度信息的技术。

立体视觉技术在军事、工业领域中有着广泛的应用背景。在军事方面，现代化的战争中，通过侦察卫星，单兵侦察等手段获取二维图像数据，利用立体视觉技术进行三维重构，模拟出三维的战场信息，可以为作战指挥提供强有力的信息支持。在工业领域中，立体视觉技术为机器人、无人驾驶汽车提供了导航、避障和物体识别等应用，特别在工业控制领域，大量的应用如工业机器人、机械手臂等都离不开立体视觉技术的支持。在智能安防中，立体视觉技术提取的深度信息还能帮助传统监控摄像头实现环境信息的结构化，从而更快预警，提高智能识别的准确性。

图1-1 激光雷达供应商Velodyne部分产品列表

虽然自二十世纪六十年代以来，不断有人提出不同的方法来提取二维图像中的三维深度信息。但是目前准确率较高的传统方法如激光雷达，毫米波雷达等，提取设备的繁琐复杂，成本高昂，且视角存在局限性，无法得到真正的普及，如图1-1所示。而设备简单的传统立体视觉方法，如单目，双目的立体视觉方法，受限于立体匹配算法的不成熟，又无法提供准确的视差数据。在无人驾驶，机器人导航等技术迅猛发展的今天，我们亟需设计一个能从成本低廉的设备中提取的准确率较高的立体视觉算法。

1.2 国内外研究现状

计算机立体视觉的开创性工作是从20世纪60年代中期开始的，美国麻省理工学院的Robert把2维图像分析推广到3维景物分析，标志着计算机立体视觉技术的诞生，并在随后的20年中迅速发展成一门新的学科[1]。特别是20世纪70年代末，Marr等人通过对生理学、计算机学、物理学和心理学等领域的研究与概括，提出了计算机视觉理论框架，将二维图像信息与三维场景信息联系在一起，对立体视觉的发展产生了巨大影响[1-2]。后来的研究者在此基础上建立了从图像获取到最终的景物可视表面重建的完整体系[2]。

立体匹配是双目视觉技术的核心问题，也是立体视觉研究中的难点。立体匹配算法的精度和速度对于双目立体视觉系统的性能起着决定性的作用。国内外的许多研究者对立体匹配算法进行了深入的研究。

国内的立体匹配研究起步较晚，研究成果主要集中在算法应用方面，如北理工开发的双目视觉空间测距平台，可应用于穿戴计算机人机交互系统中,以手指尖作为特征匹配点，根据三角测距方法确定空间手指坐标；哈工大开发的移动视觉系统，实现了足球机器人的自动导航；浙大利用双目视觉对机器臂进行改进，实现了动态定位与检测。

国外的研究起步较早，相对领先国内研究。2002年，Scharstein等人首次提出一个比较系统的立体匹配理论框架，将一个双目视觉立体匹配算法分为四个步骤：匹配代价计算，代价聚合，优化，视差精化，并提出了一个立体视觉评测数据集，最终衍变成Middlebury数据集[3]。2004年Kong等人提出可以通过平方和（SAD）来初始化匹配代价，并训练了一个模型来预测三个类别的概率分布：初始视差正确的，由于前景目标过大导致初始视差不正确的，并且由于其他原因导致初始视差不正确的。预测概率被用来调整初始匹配代价[4]。2007年，Zhang等人提出了一种替代优化算法来估算马尔科夫随机场超参数的最优值

[5]。2007年，Scharstein等人构建了一个新的30个立体对的数据集，并使用它来得到条件随机场的参数[6]。2008年，Li等人提出了一个带非参数代价函数的条件随机场模型，并使用结构化支持向量机来得到模型参数[8]。2012年，Peris等人通过AD-Census方法初始化匹配代价，并使用多类线性判别分析来得到计算所得的匹配代价到最终视差的映射[7]。2012年，德国卡尔斯鲁厄理工学院和丰田美国技术研究院共同提出立体视觉评测数据集KITTI，该数据集是目前国际上最大的自动驾驶场景下的计算机视觉算法评测数据集，KITTI包含市区、乡村和高速公路等场景采集的真实图像数据，可用于评测立体图像，光流，视觉测距，3D物体检测和3D跟踪等计算机视觉技术在车载环境下的性能，并在2015年进一步完善[12,13]。2013年，Haeusler等人使用了一种随机森林分类器来组合若干置信度度量方式[9]。2014年，Spyropoulos等人训练了一个随机森林分类器来预测匹配代价的置信度，并且使用预测结果作为马尔科夫随机场中的软约束来减少后处理的误差[10]。2015年，Zbontar等人通过训练Siamese网络来计算图像块的匹配代价，并通过基于交叉的代价聚合和半全局匹配来优化视差，取得了良好的测试结果[11]。

近年来，随着深度学习技术开始被广泛应用到各个领域，其在图像处理领域的特征提取表现尤为优秀。目前在立体匹配算法的主流评测库KITTI[12,13]中，排名靠前的算法几乎都开始使用了卷积神经网络计算初始匹配代价。

1.3 主要研究内容与章节安排

本篇文章以双目立体视觉中的核心步骤立体匹配算法作为主要研究内容。本文通过训练卷积神经网络进行小图像块的相似性度量，初始化匹配代价，并通过左右一致性检查、半全局匹配、中值滤波和双边滤波等一系列后处理步骤来精化视差图。本文使用的数据集是来自KITTI 2012和KITTI 2015的混合数据集。此外，本文结合OpenCV和Matlab工具箱进行离线标定，基于TensorFlow框架构建卷积神经网络模型，最后利用OpenCV计算机视觉库进行三维重建。同时，我们使用Flask框架和Ngrok服务搭建Web服务器，并基于Qt框架构建客户端程序实现双目数据上传和结果显示。根据以上研究内容和步骤，本文的主要架构如下：

第一章：绪论

该章节详细介绍了双目立体视觉立体匹配算法的研究背景及意义、国内外研究现状，并简单介绍了本文的主要内容和章节安排。

第二章：双目立体视觉原理

该章节主要介绍了双目立体视觉的相关原理，重点介绍了双目立体视觉的基本原理，简要介绍了立体匹配的经典步骤，约束条件以及实现难点。

第三章：数据集

该章节主要介绍了目前主流的KITTI和Middlebury立体视觉数据集，并介绍了本文基于KITTI 2012和KITTI 2015数据集构造的训练、验证和测试数据集。

第四章：立体匹配算法设计

该章节简单介绍了卷积神经网络的发展历程，随后重点介绍了本文使用的网络结构，并详细介绍了立体匹配的后处理步骤。

第五章：双目立体视觉实现与运行

该章节介绍了整个立体视觉实现的过程，简单说明了双目立体视觉的离线标定、双目矫正和三维重建过程，详细描述了立体匹配算法的实现细节，最后简单介绍了服务端的搭建过程和客户端的UI设计。

第六章：总结与展望

该章节总结了本论文所做的工作，并提出了不足之处以及在后续设计中可以改进的内容。

指 标		
疑似剽窃文字表述		
1. 根据以上研究内容和步骤，本文的主要架构如下： 第一章：绪论 该章节详细介绍了双目立体视觉立体匹配算法的研究背景及意义、国内外研究现状，并简单介绍了本文的主要内容和章节安排。 第二章：双目立体视觉原理 该章节主要介绍		
3. 第2章双目立体匹配原理		总字数：3490
相似文献列表 文字复制比：12.9%(451) 疑似剽窃观点：(0)		
1	基于双目立体视觉的立体匹配算法的研究	5.8% (203)

	宁静静(导师：孔令德) - 《中北大学硕士论文》 - 2012-06-05	是否引证：否
2	双目立体视觉深度感知与三维重建若干问题研究 罗桂娥(导师：杨欣荣) - 《中南大学博士论文》 - 2012-05-01	2.8% (96) 是否引证：否
3	基于MTK平台手机待机系统 叶忠文(导师：王树梅) - 《南京理工大学硕士论文》 - 2013-01-01	2.6% (89) 是否引证：否
4	接触轨检测中摄像机标定的分析与研究 唐崇伟(导师：陈唐龙) - 《西南交通大学硕士论文》 - 2012-05-01	2.1% (73) 是否引证：否
5	一种摄像机标定系统的实现 江祥奎;纪旭; - 《西安邮电大学学报》 - 2015-09-10	2.0% (69) 是否引证：否
6	摄像测量的温度补偿方法和位姿传递像机网络研究 晁志超(导师：于起峰) - 《国防科学技术大学博士论文》 - 2011-12-01	1.7% (60) 是否引证：否
7	视觉坐标测量机仿真模型的研究 许凌羽(导师：方漪) - 《青岛大学硕士论文》 - 2006-05-01	1.4% (50) 是否引证：否
8	毫米波无源成像系统建模及预处理算法研究 叶弘毅(导师：杨建宇) - 《电子科技大学硕士论文》 - 2011-04-01	1.1% (39) 是否引证：否
9	3DGIS环境中多监控摄像机空间布局的设计 张玲(导师：谈国新;钟正) - 《华中师范大学硕士论文》 - 2012-05-01	1.0% (36) 是否引证：否
10	桌面式视线跟踪系统的研究与实现 史小坤(导师：王军宁) - 《西安电子科技大学硕士论文》 - 2015-11-01	1.0% (35) 是否引证：否
11	目标的视觉检测、分析与描述 赵杰(导师：王林) - 《上海交通大学硕士论文》 - 2013-02-01	1.0% (34) 是否引证：否
12	H型钢截面尺寸线结构光测量 刘燕(导师：王刚) - 《天津大学硕士论文》 - 2012-12-01	0.9% (30) 是否引证：否

原文内容

第2章双目立体匹配原理

2.1 双目立体视觉基本原理

双目立体视觉系统通过模拟人的双眼成像原理，一般由两摄像机从不同角度同时获取目标场景的两幅图像，并基于视差原理恢复出场景的三维几何信息，重建场景中物体的三维轮廓及位置。相对于单目和多目的立体视觉方法，双目立体视觉设备较简单，成本低廉，提取深度信息较方便，故被作为立体视觉技术实现的首选。本节主要介绍了双目立体视觉的实现原理。

2.1.1 摄像机成像模型

摄像机的成像模型是一个将三维空间物体投影成二维数字平面图像的过程，光线从空间物体开始，通过透镜到达摄像机的成像平面，从而形成摄像机捕捉到的数字图像。摄像机的成像模型可以通过图像坐标系、世界坐标系和摄像机坐标系之间的空间几何关系，由二维数字平面图像反推出三维空间信息。下面我们将详细介绍这三种坐标系及其相互关系。

(1) 图像坐标系

图像坐标系可以分为两种：图像物理坐标系和图像像素坐标系，他们的区别主要在于原点的位置不同。

图2-1 图像坐标系

如图2-1所示，图像像素坐标系原点O0位于图像左上方的，它的u，v轴分别表示图像矩阵的列和行。而图像物理坐标系的原点O1位于摄像机光轴和图像像素平面的交点，它的X，Y轴平行于u，v轴。设图像像素坐标系下的某点(u0, v0)，图像物理坐标系下每个像素在横轴X和纵轴Y上的物理尺寸为dX，dY，则两坐标系之间的转换关系如公式2-1所示：

$$u=XdX+u_0v=YdY+v_0 \quad (2-1)$$

为方便后续计算，式2-1转换关系转换为矩阵表达形式如式2-2所示：

$$uv1=1dX0u001dYv0101XY1 \quad (2-2)$$

其逆变换形式如下：

$$XY1=dX0-u0dX0dY-v0dY001uv1 \quad (2-3)$$

(2) 世界坐标系

为描述摄像机在真实场景中的相对位置信息，我们引入世界坐标系，如图2-2所示，Ow为世界坐标系的原点，坐标轴为Xw、Yw和Zw。

图2-2 世界坐标系

(3) 摄像机坐标系

摄像机坐标系是联系图像坐标系和世界坐标系的桥梁。摄像机坐标系以摄像机镜头焦点为原点，摄像机光轴为z轴，其

x, y轴平行于图像坐标系的X, Y轴。具体如图2-3所示, OO1表示摄像机焦距。

图2-3 摄像机坐标系

摄像机坐标系和世界坐标系之间的转换需要经过矩阵平移和旋转操作, 设某一像素点的摄像机坐标为 $x, y, z, 1^T$, 其对应的世界坐标系为 $X_w, Y_w, Z_w, 1^T$, 则两坐标系的关系如下:

$$xyz1 = RTOT1XwYwZw1 \quad (2-4)$$

其中, R为 3×3 的正交单位旋转矩阵, T为三维平移向量, OT为0,0,0

2.1.2 对极几何

对极几何描述了两幅图像之间的视觉几何关系, 与外部场景无关, 只依赖于摄像机的内参数和左右摄像机的相对空间位置。对极几何是视图几何理论的基础, 在双目视觉的双目矫正和三维恢复中有着广泛的应用。具体原理图如图2-4所示。

图2-4 对极几何原理图

图中, O1和O2为左右两摄像机的光心, 平面I1和I2是两摄像机的成像平面, P点为现实空间中的某个待观测点, P1和P2为P点在两成像平面上的成像点。光心O1O2所在的直线和观测点P共同确定的平面为极平面, O1和O2在成像平面上的投影称为极点, 极平面与两成像平面的交线为极线, 极线I1和I2分别是P2和P1的外极线。根据对极几何原理, 待匹配点的对应点一定在其对应的外极线上, 即P1点的匹配点, 一定在其对应的外极线I2上, P2点的匹配点, 一定在其对应的外极线I1上。极线几何约束大大的减少了匹配点的搜索范围, 降低了算法复杂度。

2.1.3 视差理论

假设存在一个理想的双目视觉模型, 左右两摄像机平行摆放, 摄像机内参一致, 左右摄像机成像平面位于同一平面中。

图2-5 理想双目视觉模型

如图2-5所示, P点为现实世界中的某观测点, b为两光心O1和O2的间距, z为点P的深度距离, 两摄像机焦距为f。通过三角形相似原理, 可以得出如下公式:

$$bz = b + x_2 - x_1 z - f \quad (2-5)$$

由公式2-5可以推导出深度距离为:

$$z = bfx_1 - x_2 = bfd \quad (2-6)$$

其中 $x_1 - x_2$ 为视差值d, f和b通过标定得到。这就是双目立体视觉的基本公式。我们常用视差图来表示立体匹配的结果, 视差图中灰度值越小, 目标点越远。

2.2 立体匹配的经典步骤

立体匹配算法是双目立体视觉技术的核心。立体匹配就是从左右摄像机获取的两幅图中找出每个像素的对应关系, 从而提取出视差图。立体匹配算法的精度和速度常被用来评价立体匹配算法的性能。

立体匹配算法一般可以分为以下四个步骤:

(1) 匹配代价计算

立体匹配代价的计算即计算左右两像素之间的相似程度, 根据代价函数的不同对应的匹配结果也会不同, 常见的匹配代价计算方式包括灰度差的绝对值, 灰度值的平方差等

(2) 代价聚合

匹配代价的聚合通常是以待匹配点为中心, 选择一定的区域作为聚合窗口, 对窗口内的匹配代价进行叠加或平均来增强匹配代价的可靠性, 提高匹配的准确度。目前常见的代价聚合方式有固定窗的代价聚合, 基于交叉的代价聚合[14]等。

(3) 视差计算

视差计算通常采用胜者全拿的策略, 即寻找使匹配代价 $C(p, d)$ 最小的视差d, 如公式2-7所示。

$$D_p = \arg \min_d C(p, d) \quad (2-7)$$

(4) 视差优化

视差优化的步骤是通过对计算所得的视差图进行错误校验, 来得到精度更高的视差图, 通常会采用左右一致性检查, 中值滤波和双边滤波等技术来消除视差图中的某些异常点, 也可以通过半全局匹配[15]进行亚像素增强, 来精化视差。

2.3 立体匹配的约束条件

为减少立体匹配的搜索范围, 提高匹配精度与速度, 我们通常需要参考某些约束条件, 目前常见的约束条件有以下几种:

(1) 唯一性约束

左右两图像中的点一一对应, 左图中的任何一点只能和右图中的一个点唯一对应。

(2) 相似性约束

左右图像中匹配的两个点具有相似的属性, 如灰度, 梯度变化, 几何轮廓, 颜色值等都具有相似性。

(3) 连续性约束

空间中的物体表面一般是连续的, 除物体的边缘部分外, 其他地方的视差值应该是一个缓慢变化的过程, 可以以此来约束某点和邻域像素的视差。

(4) 顺序性约束

立体匹配中某极线上的点, 在对应极线上的匹配点的顺序是不变的。

(5) 视差范围约束

图像的视差值不可能无限大，在匹配时，仅需要搜索视差范围内的点即可。选择合适的视差范围可以提高匹配的精度和效率。

(6) 左右一致性约束

如果以左图为准时，左图的点p在右图的匹配点为点q，那么，以右图为准时，右图的点q在左图的匹配点应为点p。我们通过做左右一致性检查来消除视差图的毛刺。

(7) 极线约束

极线约束的原理已在2.1.2小节对极几何原理中详细介绍了，左图中的待匹配点在右图中的对应点一定在右图的极线上，匹配搜索时仅需要搜索该极线上的点即可。

2.4 立体匹配的难点

立体匹配算法经过不断的发展，已经有了极大的进步，但仍然面临众多挑战。目前立体匹配的难点主要集中在以下几个方面：

(1) 镜面区域

由于双目视觉中的两摄像机是从不同角度拍摄的图像，所以会存在某些镜面区域因为光照问题而导致的灰度信息出现明显变化。如图2-6中光照所产生的镜面反射导致亮度、色调、饱和度等均出现变化。

图2-6 镜面区域

(2) 遮挡区域

双目视觉中，同一场景在左右两幅图中并不完全一样，由于视线阻挡会出现遮挡区域，如图2-7所示，左图的部分区域被遮挡，在右图中无法观测到，自然也无法找到对应点。

图2-7 遮挡区域

(3) 弱纹理和重复纹理区域

弱纹理和重复纹理问题是立体匹配中的经典难题。由于目标场景的纹理特征单一或部分纹理特征非常相似导致匹配误差。如图2-8中，目标场景纹理特征不丰富，很难得到准确的视差值。

图2-8 弱纹理和重复纹理区域

(4) 透明区域

一般情况下，场景中的物体在边缘变化会比较剧烈，我们以此来判断物体边缘。但是也存在某些前景后景相似的情况，如出现透明物体时，这些区域的边缘灰度变化不大，导致所得视差的边缘模糊，甚至误匹配。

图2-9 透明区域

2.5 本章小结

本章节主要介绍了双目立体视觉的相关原理。首先本章详细介绍了双目立体视觉的基本原理，详细介绍了摄像机的成像模型、对极几何和视差理论。随后，本章对立体匹配的经典步骤，约束条件以及实现难点等做了概略说明。

指 标	
疑似剽窃文字表述	
1.	一般由两摄像机从不同角度同时获取目标场景的两幅图像，并基于视差原理恢复出场景的三维几何信息，重建场景中物体的三维轮廓及位置。
2.	空间物体开始，通过透镜到达摄像机的成像平面，从而形成摄像机捕捉到的数字图像。摄像机的成像模型可以通过图像坐标系、世界坐标系和摄像机坐标系之间的
3.	坐标系及其相互关系。 (1) 图像坐标系 图像坐标系可以分为两种：图像物理坐标系和图像像素坐标系，他们的区别主要在于原点的位置不同。
4.	世界坐标系 (3) 摄像机坐标系 摄像机坐标系是联系图像坐标系和世界坐标系的桥梁。
5.	代价聚合 匹配代价的聚合通常是以待匹配点为中心，选择一定的区域作为聚合窗口，对窗口内的匹配代价进行叠加或平均来增强匹配代价

相似文献列表 文字复制比：15.5%(409) 疑似剽窃观点：(0)		
1	121002108-屈小波-卷积神经网络算法的研究 屈小波 - 《大学生论文联合比对库》 - 2016-06-07	10.3% (273) 是否引证：否
2	KITTI计算机视觉算法测评平台新排名:中国创业公司图森领先-资讯-海外网 - 《网络 (http://news.haiwainet.cn) 》 - 2016	4.8% (128) 是否引证：否
3	全球自动驾驶算法排行榜公布 图森互联多项第一 - 《网络 (http://tech.163.com/) 》 - 2016	4.8% (128) 是否引证：否
4	全球自动驾驶算法排行榜公布 图森互联多个项目折桂-中新网 - 《网络 (http://www.chinanews.com) 》 - 2016	4.8% (128) 是否引证：否
5	图森互联自动驾驶算法横扫排行榜 世界第一 - 《网络 (http://www.chexun.com) 》 - 2016	4.8% (128) 是否引证：否
6	全球自动驾驶算法排行榜公布 图森互联多项领跑 - 综合新闻 - 《网络 (http://auto.ubetween.com) 》 - 2016	4.8% (128) 是否引证：否
7	KITTI计算机视觉算法测评平台公布排名 中国创业公司领先_科技时代 - 《网络 (http://tech.sina.com) 》 - 2016	4.5% (119) 是否引证：否

原文内容

第3章数据集

由于卷积神经网络在图片特征提取方面的优秀表现，本文使用卷积神经网络来初始化匹配代价。高精度的立体视觉数据集对卷积神经网络的构建非常重要，优质的标签数据可以更好的训练模型。KITTI数据集和Middlebury数据集是目前主流的立体视觉数据集。

3.1 立体视觉数据集

3.1.1 KITTI数据集

KITTI立体视觉数据集[12,13]由德国卡尔斯鲁厄理工学院和丰田美国技术研究院联合创办，是目前国际上最大的自动驾驶场景下的计算机视觉算法评测数据集。该数据集可以用于评测立体图像，光流，3D物体检测和3D追踪等计算机视觉技术。

数据集由安装在汽车车顶上两个相隔约54厘米的摄像机拍摄所得并经整理的灰度图像对构成。数据集中包含市区、乡村和高速公路等场景采集的真实图像数据，每张图像中最多达15辆车和30个行人，还有各种程度的遮挡与截断。原始数据集由389对立体图像和光流图，39.2 km视觉测距序列以及超过20万3D标注物体的图像组成。图像分辨率为1240×376。旋转激光扫描仪安装在左侧相机后面，记录地面的真实深度，标记了约30%的图像像素。

数据集中包含训练数据集和测试数据集。测试数据集的地面真实视差并未公开，而是提供了一个在线排行榜供研究人员在测试集上评估他们的算法。每隔三天允许提交一次。错误率是计算那些真实视差和预测视差相差超过三个像素的像素的比例。这就意味着，视差误差为3厘米则转换成物理距离为距离相机2米，而80厘米的误差，就是距离相机10米。KITTI的每个子数据集中还提供了开发工具，用于视差数据的提取，保存以及评估，主要有C++和Matlab两个版本。

目前共发布了两个KITTI立体视觉数据子集：KITTI 2012和KITTI 2015。2012年的数据集包含194个训练图和195个测试图像，2015年的数据集包含200个训练图和200个测试图。较新的数据集引入了一个微妙但重要的变化：运动车辆是密集标记的，并且汽车玻璃也被包括在评估中。

3.1.2 Middlebury数据集

Middlebury立体视觉数据集[3,16,17,18]最早由Scharstein等人在2002年提出，数据集中的图像对来自光照可控的室内场景。数据集通过结构光来测量真实视差，视差的密度和精度都比KITTI数据集更好。该数据集分别在2001年，2003年，2005年，2006年和2014年发布了共五个独立的数据集。

在2005年，2006年和2014年的数据集中，每一个场景都是根据多种光照条件和快门曝光方式拍摄的，其中有一组典型的图像对是同一个场景在四种光照条件和七种快门曝光方式下拍摄的共计28张图片。

该数据集也有一个类似KITTI的在线排行榜，显示了所有提交算法的排名表。参与者只有一次提交结果的机会，且需要通过作者审核。测试数据集包含15幅来自2005年和2014年数据集的图像。测试数据集提供全分辨率，半分辨率和四分之一分辨率的测试图像，全分辨率图像的分辨率约为3000×2000。误差率的计算基于全分辨率，如果算法输出的是半分辨率或四分之一分辨率的视差图，需要在计算错误率之前上采样。

误差率通过计算真实视差与预测视差相差超过两个像素的像素点的百分比得到，使用半分辨率计算视差图时，错误容忍仅为一个像素。在默认情况下，评估服务器上的误差率仅仅计算非遮挡的像素点。而在最终在线报告上的误差率为十五对测试图像的加权平均，权重由数据集的作者设置。

与KITTI数据集相比，Middlebury在计算全分辨率时对实验计算机内存要求更高，图片较少，部分数据子集图像矫正并没有严格对应，即使通过OpenCV库中的标准矫正程序也会在Middlebury中造成最多九个像素的垂直视差错误[18]。

3.2 构造数据集

由于实验条件的限制，本文使用KITTI 2012和KITTI 2015数据集来构造实验数据集。本文从KITTI 2012和KITTI 2015中分别随机选出162张和168张图片构成训练数据集。KITTI 2012和KITTI 2015剩余图片混合后，随机挑出32张构成验证数据集，混合数据集剩余的32张构成测试数据集。最终训练数据集，验证数据集和测试数据集的比例约为10:1:1。

训练数据集仍需要进一步预处理转换成可训练的块数据。本文在训练数据集中每一个非零视差的图像位置处提取一对积极的和一对消极的示例。这种做法是为了确保训练数据集中包含相等数量的正例和负例。以下部分是积极示例和消极示例的具体提取步骤。

用<Pn×nLp,Pn×nRq>表示一对图像块，其中Pn×nLp为左图中以位置p=(x,y)为中心的n×n图像块，Pn×nRq是右图中以位置q为中心的n×n图像块，d表示位置p处的正确视差。一对积极示例的获取是通过将右图像块的中心设置为q=x-d+opos,y
其中opos是从区间[-dataset_pos, dataset_pos]中随机挑选的。其中opos不为零是因为训练网络时，有一定程度的噪声可以降低网络的过拟合。

一对消极示例的获取是通过将右图像块的中心设置为

q=x-d+oneg,y

其中oneg是从区间[dataset_neg_low, dataset_neg_high]或它的原点对应区间[-dataset_neg_high, -dataset_neg_low]中选出的。随机偏移oneg确保了生成的图像块以不同与前者的3D点为中心。

积极示例和消极示例的图像块还需要通过将图像的像素强度值减去它们的平均值再除以标准差来归一化。本文将归一化后的积极示例和消极示例随机混合在一起，以1:1的比例构成一批128对的图像块，这就是我们最终的训练数据集。

构造数据集过程中的超参数包括积极示例的偏移 (dataset_pos)，消极示例的偏移 (dataset_neg_low, dataset_neg_high) 以及图像块的大小 (input_patch_size)。

3.3 本章小结

本章节介绍了立体视觉相关数据集的情况。首先本章详细介绍了目前主流的立体视觉数据集KITTI和Middlebury的具体情况。随后，本章介绍了以KITTI 2012和KITTI 2015构造本文使用的训练、验证和测试数据集的分配策略，并重点介绍了训练数据集的预处理步骤及相关超参数。

指 标	
疑似剽窃文字表述	
1.	包含市区、乡村和高速公路等场景采集的真实图像数据，每张图像中最多达15辆车和30个行人，还有各种程度的遮挡与截断。原始数据集
2.	旋转激光扫描仪安装在左侧相机后面，记录地面的真实深度，标记了约30%的图像像素。
3.	2012年的数据集包含194个训练图和195个测试图像，2015年的数据集包含200个训练图和200个测试图。
4.	一个类似KITTI的在线排行榜，显示了所有提交算法的排名表。参与者只有一次提交结果的机会，
5.	在默认情况下，评估服务器上的误差率仅仅计算非遮挡的像素点。而在最终在线报告上的误差

5. 第4章立体匹配算法设计		总字数：4647
相似文献列表 文字复制比：10.6%(492) 疑似剽窃观点：(0)		
1	基于图像序列的目标三维重建技术研究 齐南(导师：毛征) - 《北京工业大学硕士论文》 - 2013-06-01	2.9% (137) 是否引证：否
2	正在崛起的变革力量—2016中国人工智能产业发展报告 - 微口网 - 《网络 (http://www.vccoo.com) 》 - 2016	2.0% (95) 是否引证：否
3	基于CNN的人脸识别系统研究 廉东泽 - 《大学生论文联合比对库》 - 2016-05-26	2.0% (94) 是否引证：否
4	动态天气图像复原方法研究 芦雪(导师：董慧颖) - 《沈阳理工大学硕士论文》 - 2013-12-01	2.0% (92) 是否引证：否
5	用于啤酒生产线上验瓶机监控系统 查进(导师：徐乐年) - 《山东科技大学硕士论文》 - 2009-05-01	1.7% (77) 是否引证：否
6	基于正则的图像复原算法研究	1.5% (72)

	马明川 - 《大学生论文联合比对库》 - 2014-05-31	是否引证：否
7	102033_马明川_基于正则化的图像复原算法研究_论文定稿_1402041755331 马明川 - 《大学生论文联合比对库》 - 2014-06-09	1.5% (72) 是否引证：否
8	机器视觉技术在啤酒瓶在线检测系统中的应用研究 陈西广(导师：徐乐年) - 《山东科技大学硕士论文》 - 2008-06-01	1.5% (72) 是否引证：否
9	基于车牌识别技术的车辆管理系统的研发 董镭(导师：艾矫燕) - 《广西大学硕士论文》 - 2008-06-01	1.5% (72) 是否引证：否
10	基于双目立体视觉的三维重建方法研究 党乐(导师：贺昱曜) - 《长安大学硕士论文》 - 2009-04-12	1.5% (72) 是否引证：否
11	酒瓶检测中的机器视觉检测技术研究 李丽(导师：彭延军) - 《山东科技大学硕士论文》 - 2009-05-01	1.5% (72) 是否引证：否
12	灰度不均的弱边界血管图像分割方法 吴杰;朱家明;张辉; - 《计算机应用》 - 2016-06-10	1.5% (70) 是否引证：否
13	基于shearlet变换的非线性滤波图像去噪 姜真;毛志龙;朱丽华;丁然; - 《科技经济导刊》 - 2016-06-05	1.5% (70) 是否引证：否
14	基于小波分析的地面三维激光扫描点云数据的滤波方法研究 靳洁(导师：田永瑞) - 《长安大学硕士论文》 - 2013-05-25	1.5% (70) 是否引证：否
15	纪昌青-彩色图像小麦颗粒特征识别系统设计 纪昌青 - 《大学生论文联合比对库》 - 2013-05-24	1.5% (70) 是否引证：否
16	基于空间滤波的图像去噪算法研究 田泽宇 - 《大学生论文联合比对库》 - 2014-05-19	1.5% (70) 是否引证：否
17	软件学院_10212045_张志博 软件学院 - 《大学生论文联合比对库》 - 2014-05-20	1.5% (70) 是否引证：否
18	基于空间滤波的图像去噪算法研究 田泽宇 - 《大学生论文联合比对库》 - 2014-05-21	1.5% (70) 是否引证：否
19	非局部均值去噪 王云霞 - 《大学生论文联合比对库》 - 2013-06-04	1.5% (70) 是否引证：否
20	基于断层扫描数据的心脏三维重建 张锡英;孟繁平;邱兆文; - 《重庆理工大学学报(自然科学)》 - 2016-12-15	1.5% (69) 是否引证：否
21	自适应红外目标跟踪算法 代少升;徐飞; - 《半导体光电》 - 2014-02-15	1.4% (67) 是否引证：否
22	运动模糊雾霾图像增强算法研究 石鑫(导师：李世平) - 《东北大学硕士论文》 - 2013-06-30	1.4% (67) 是否引证：否
23	基于计算机视觉的微小轴承外表面缺陷检测研究 陈廉清(导师：焦生杰) - 《长安大学硕士论文》 - 2008-08-25	1.3% (60) 是否引证：否
24	基于图象处理的汽车仪表自动检测系统 陈艳海(导师：陈建勋) - 《武汉科技大学硕士论文》 - 2006-11-12	1.2% (55) 是否引证：否
25	基于深度卷积神经网络的手势识别研究 陈祖雪(导师：马苗) - 《陕西师范大学硕士论文》 - 2016-05-01	1.1% (53) 是否引证：否
26	蒋乐乐_22120555_基于深度学习的图像分类算法研究 (基于matlab) 蒋乐乐 - 《大学生论文联合比对库》 - 2016-06-06	1.1% (50) 是否引证：否
27	基于SVM的车辆自动分类方法研究与实现 魏娜(导师：龚声蓉) - 《苏州大学硕士论文》 - 2008-10-01	1.0% (47) 是否引证：否
28	基于社会计算和深度学习的社交网络特定内容监控 操晓春;荆丽桦;王蕊;张锐;董振江;熊红凯; - 《计算机科学》 - 2016-09-14 1	0.9% (44) 是否引证：否
29	缩微智能车的环境感知与控制决策算法研究 张祖锋;徐友春;张鹏;朱增辉; - 《第八届中国智能交通年会优秀论文集——智能交通与安全》 - 2013-09-26	0.7% (33) 是否引证：否
30		0.6% (30)

原文内容

第4章立体匹配算法设计

4.1 卷积神经网络

卷积神经网络 (Convolutional Neural Network, CNN) 是当前深度学习技术中极具代表性的网络结构之一。卷积神经网络通过权值共享和局部连接的结构模拟生物神经网络, 不仅避免了传统模式识别算法中复杂的特征提取过程, 而且解决了神经网络中全部采用全连接导致参数量过大难以训练的问题。卷积神经网络的网络结构对平移、比例缩放、倾斜或者其他形式的变形具有高度不变性, 其在图像识别、语音识别和自然语言处理等问题上表现尤为突出。

卷积神经网络的历史最早可以追溯到二十世纪六十年代。1962年Hubel和Wiesel在研究猫的视觉皮层细胞时发现, 每一个视觉神经元只会处理一小块区域的视觉图像, 提出了感受野(Receptive Field)的概念, 即猫的视觉系统是分级的, 这种分级可以看成是逐层迭代、抽象的过程[19]。基于上述提出的感受野概念, 日本学者Fukushima在1984年构建了神经认知机(Neocognitron) [20]。神经认知机是卷积神经网络的第一个实现网络, 该网络中包含两类神经元, 用来抽取特征的S-cells, 和用来抗形变的C-cells, 其中S-cells对应现在主流卷积神经网络中的卷积核滤波操作, 而C-cells则对应激活函数、最大池化(Max-Pooling) 等操作[20]。

随着后来研究者的不断改进, LeCun等人在1988年结合BP算法实现了LeNet-5模型, 其在数字识别领域的表现非常强大, 在银行支票的手写识别应用中达到了95.1%的识别精度[21]。

但是在接下来近十年的时间里, 卷积神经网络的相关研究趋于停滞, 主要原因有三点: 一是当时互联网尚未广泛普及, 很难取得足够的标签数据以训练网络; 二是神经网络在进行训练时的计算量极其之大, 以当时的硬件计算能力几乎不可能实现; 三是包括SVM在内的浅层机器学习算法渐渐开始崭露头角。直到2006年, Hinton提出深度置信网络 (DBN), 使用一种贪心无监督训练方法来解决并取得了良好结果[22]。DBN (Deep Belief Networks) 的训练方法降低了学习隐藏层参数的难度, 并且该算法的训练时间和网络的大小和深度近乎线性关系[22]。CNN重新回到研究者的视野中, 并取得了长足发展。

随着大数据时代的到来和硬件性能的提升, 研究者们不仅能够取得足够标签数据用于训练, 而且能够借助图形处理单元 (GPU) 的强大运算能力快速的训练和优化模型。2012年Krizhevsky等人凭借AlexNet网络赢得了2012 ImageNet竞赛的冠军 (ImageNet竞赛为1000类图片分类任务, 难度较大), 该网络将分类误差从26%降到15%, 远远超越了之前的方法[23,24]。随着AlexNet的成功, 卷积神经网络开始被广泛应用到各个领域, 例如谷歌将神经网络用于图片搜索引擎中, 亚马逊将其用于商品推荐, Pinterest 将其用于个性化主页推送, 还有打败李世石的AlphaGo也用到了卷积神经网络, 可以说卷积神经网络在众多课题中都取得了突破性的成果。

随着卷积神经网络的深入发展, 研究人员开始尝试在一些早已被充分研究的问题中应用神经网络。Zbontar和LeCun在2015年提出使用卷积神经网络进行立体匹配[25]。他们使用相似和不相似的图像对构建二元分类数据集来进行有监督的训练[25]。卷积神经网络的输出用于初始化立体匹配代价, 并配合一系列的后处理步骤: 基于交叉的代价聚合, 半全局匹配, 左右一致性检查, 亚像素增强, 一个中值滤波器和一个双边滤波器来提高视差图的准确性[25]。他们的立体匹配算法在主流立体视觉评测库KITTI和Middlebury中表现优异, 引起了广泛的关注。目前在KITTI的排行榜中排名前十的算法几乎都用到了卷积神经网络。

4.2 网络结构

鉴于卷积神经网络在图像处理领域的优秀表现, 本文采用Siamese网络来比较图像块之间的相似度。Siamese网络中包含两路结构相似的子网, 它们分别训练, 但共享各个层的参数, 并在最后的非共享连接层进行相似度比较[26]。本文采用的网络结构如图4-1所示。

在图4-1网络结构中, 两个子网由多个卷积层和整流线性单元 (ReLu) 组成。输入的左右图像块提取特征向量后通过计算它们之间的余弦相似性来获得相似度得分, 其中余弦相似性的计算分为两步进行, 归一化和点积。这样每个位置仅需要执行一次归一化操作, 减少了运行时间。余弦相似性的计算如公式4-1所示。

$$\text{similarity} = \frac{A \cdot B}{\|A\| \|B\|} = \frac{1}{\|A\|} \frac{A \cdot B}{\|B\|} = \frac{1}{\|A\|} \text{normal}(B) \quad (4-1)$$

其中A, B分别表示左右特征向量。上式可以拆分为以下两步。

$$\text{normal}(X) = \frac{X}{\|X\|} = \frac{1}{\|X\|} X \quad (4-2)$$

$$\text{similarity} = \text{normal}(A) \cdot \text{normal}(B) \quad (4-3)$$

也就是上文所说的归一化和点积步骤。

图4-1 本文网络结构

本文选择铰链损失 (Hinge loss) 最小化来训练网络。损失函数如公式4-4所示。

$$L_{\text{Hinge}}(s^+, s^-) = \max\{0, m + s^- - s^+\} \quad (4-4)$$

其中 s^+ 和 s^- 表示以同一图像位置为中心的积极和消极示例计算所得的相似度得分。 s^+ 是积极示例的网络输出, s^- 是消极示例的网络输出。余量 m 为正实数。当 s^+ 大于 s^- 至少余量 m 时, 网络损失为零。本文实验中设置余量 m 为0.2。

Siamese网络的输出用于初始化匹配代价, 如公式4-5所示。

$$CCNNp,d=-s(<PL(p), PR(p-d)>) (4-5)$$

其中 $s(<PL(p), PR(p-d)>)$ 是当网络输入图像块 $PL(p)$ 和 $PR(p-d)$ 时的输出。网络输出的相似度得分的负数表示匹配代价。

4.3 后处理步骤

卷积神经网络的原始输出不足以产生准确的视差图，特别是在某些低纹理和重复纹理区域存在明显的误差。通过应用一系列的后处理步骤可以有效的改善视差图的质量。本文使用的后处理步骤主要参考了Zbontar和LeCun的论文[25]，其中包括左右一致性检查，半全局匹配，中值和双边滤波。

4.3.1 左右一致性检查

左右一致性检查是进行误匹配检测最为常见和有效的方式，可以有效的消除遮挡区域减少误匹配，得到精度更高的视差图。令 DL 表示通过以左图作为参考图得到的视差图， DR 表示以右图作为参考图得到的视差图，在某些遮挡区域它们的视差是不同的。我们通过左右一致性检查来检测这些冲突，并应用以下规则来按顺序标记每个位置 p ：

正确：当 $d=DLp$ 时， $d-DRp-d \leq 1$

错误：当 $d \neq DLp$ 时，存在 d 使得 $d-DRp-d \leq 1$

遮挡：除上述两种情况外

对标记为“遮挡”的位置，我们从背景中获取新的视差值。我们向左遍历直到找到一个标记为“正确”的点并将该点的视差值作为遮挡点的视差。对于标记为“错误”的位置，我们以该点为中心的十六个不同方向中寻找最近的标记为“正确”的点，并取这些视差的中值作为错误点的视差。对于标记的“正确”的点，我们不进行操作。经过左右一致性检查后的视差图为 $DINT$ 。

4.3.2 半全局匹配

半全局匹配是Hirschmuller在2008年提出的，通过对视差图进行平滑约束来改善匹配代价[27]。半全局匹配通过使一个依赖于视差图 D 的能量函数 $E(D)$ 最小化来求解最优视差，能量函数形式如公式4-6所示。

$$ED=p(CCNNp,Dp+q \in NpP1 \cdot 1Dp-Dq=1+q \in NpP2 \cdot 1\{Dp-Dq>1\}) (4-6)$$

公式4-6中的粗体的 p ， q 表示图中的某个像素点， Np 表示以 p 点为中心的图像块， $CCNNp,Dp$ 表示当 p 点的视差取 Dp 时，该点的匹配代价。 $P1$ ， $P2$ 为惩罚系数，他们分别适用于与 p 点相邻像素的视差值和 p 的视差值相差为1和大于1的情况。 $1\{\cdot\}$ 表示如果函数中的参数为真，则返回1，否则返回0。

我们可以通过动态规划在单一方向上的能量函数最小化来实现 $E(D)$ 的最小化。虽然Hirschmuller选择了从十六个方向优化结果，但本文出于匹配速度上的考虑仅从水平和垂直的四个方向上优化取平均数来最小化能量。为最小化方向 r 上的能量，可以得出公式4-7所示的具有递推关系的公式。

$$C_{rp,d}=CCNNp,d+\min(C_{rp-r,d},C_{rp-r,d-1}+P1,C_{rp-r,d+1}+P1,\min(C_{rp-r,k}+P2-\min(C_{rp-r,k}),k)) (4-7)$$

公式4-7由三大项组成，第一项表示Siamese网络计算所得的匹配代价，第二项是四个候选值中的最小值。第一个候选值是之前所有像素视差值取 d 时的最小匹配代价，第二个候选值是之前所有像素视差值取 $d-1$ 时的最小匹配代价与惩罚系数 $P1$ 之和，第三个候选值是之前所有像素视差值取 $d+1$ 时的最小匹配代价与惩罚系数 $P1$ 之和，第四个候选值是之前所有像素视差值取其他值时的最小匹配代价与惩罚系数 $P2$ 之和。第三项是之前所有像素视差值取其他值时的最小匹配代价，这一项是为了防止 $C_{rp,d}$ 的值过大影响视差图的优化。

惩罚系数 $P1$ 和 $P2$ 是为优化图像边缘而设置的，令 $D1=|ILp-ILp-r|$ ， $D2=|IRp-d-IRp-d-r|$ ， $P1$ 和 $P2$ 遵循以下规则：

如果 $D1 < \text{sgm_D}$ ， $D2 < \text{sgm_D}$ ，则 $P1=\text{sgm_P1}$ ， $P2=\text{sgm_P2}$

如果 $D1 \geq \text{sgm_D}$ ， $D2 \geq \text{sgm_D}$ ，则 $P1=\text{sgm_P1sgm_Q2}$ ， $P2=\text{sgm_P2sgm_Q2}$

其他情况，则 $P1=\text{sgm_P1sgm_Q1}$ ， $P2=\text{sgm_P2sgm_Q1}$

在上述规则中超参数 sgm_P1 和 sgm_P2 为视差图中的不连续点设置了惩罚因子。如果 $D1$ 或 $D2$ 中的一个表示高图像梯度，惩罚因子为 sgm_Q1 ，如果 $D1$ 和 $D2$ 同时表示高图像梯度，则通过较大的 sgm_Q2 来惩罚。当计算两个垂直方向时， $P1$ 的值还需要除以 sgm_V 惩罚因子。因为通常情况下垂直方向上的视差变化比水平方向更为频繁。

最终过的匹配代价 $CSGMp,d$ 是所有四个方向取平均值计算得到的，如公式4-8所示。

$$CSGMp,d=14rC_{rp,d} (4-8)$$

本小节涉及到的超参数有基础惩罚系数 sgm_P1 、 sgm_P2 ，惩罚因子 sgm_Q1 、 sgm_Q2 、 sgm_V ，图像梯度阈值 sgm_D 。

4.3.3 滤波

为消除视差图的某些异常点平滑视差图，本文通过中值滤波和双边滤波来进一步的精化视差。中值滤波的基本思想是用像素点邻域内灰度值的中值来代替该像素点的灰度值，在去除脉冲噪声、椒盐噪声的同时又保留了图像边缘细节。双边滤波是一种结合图像的空间邻近度和像素值相似度的折衷处理，同时考虑空域信息和灰度相似性，达到保边去噪的目的，具有简单、非迭代、局部的特点。

滤波过程中涉及到的超参数有中值滤波孔径线性尺寸 $ksize$ ，双边滤波邻域范围直径 d ，颜色空间的标准差 sigmaColor 和坐标空间的标准差 sigmaSpace 。

4.4 本章小结

本章节首先简单的介绍了卷积神经网络的发展历程，随后详细描述了本文使用的卷积神经网络结构和匹配代价的计算，最后本文详细介绍了立体匹配的后处理算法，重点介绍了左右一致性检查和半全局匹配算法，并简单说明了中值和双边滤波作用

和涉及到的超参数。

指 标
疑似剽窃文字表述
1. 神经网络在进行训练时的计算量极其之大，以当时的硬件计算能力几乎不可能实现；三是包括SVM在内的浅层机器学习算法渐渐
2. 滤波和双边滤波来进一步的精化视差。中值滤波的基本思想是用像素点邻域内灰度值的中值来代替该像素点的灰度值，在去除脉冲噪声、椒盐噪声的同时又保留了图像边缘细节。双边滤波是一种结合图像的空间邻近度和像素值相似度的折衷处理，同时考虑空域信息和灰度相似性，达到保边去噪的目的，具有简单、非迭代、局部的特点。

6. 第5章立体视觉实现与运行	总字数：4549
相似文献列表 文字复制比：2.2%(101) 疑似剽窃观点：(0)	
1 1120121899_孙宇超_视频中动作的自动检测_(评优) 孙宇超 - 《大学生论文联合比对库》 - 2016-05-30	2.2% (101) 是否引证：否
原文内容	

第5章立体视觉实现与运行

双目立体视觉的实现过程可分为离线标定、双目矫正、立体匹配和三维重建四个步骤。离线标定，双目矫正和三维重建均属于图像矫正部分，可借助OpenCV和Matlab等工具完成，不属于本文的核心研究内容。所以本章简要说明了图像矫正的流程，重点介绍立体匹配算法的实现过程。

5.1 图像矫正

在进行立体匹配算法研究时，我们的数据集中都是经过矫正的图像，可以直接用来实验。但是在实际应用中，我们的摄像机并不能如理想的双目视觉模型那样处于绝对水平的位置上，而且普通的摄像机往往存在一定程度的光学畸变，所以我们需要对图像进行离线标定，并在实际运行过程中进行双目矫正。

5.1.1 双目摄像机选取

本文使用的双目摄像机是由两个单目摄像机组成的，如图5-1所示，它们的配置如表5-1所示。

图5-1 双目摄像机

表5-1 摄像机配置参数

USB接口 2.0高速传输视角 50度

最大分辨率 640×480 摄像头尺寸 30 mm×25 mm

输出格式 YUV 焦距 2.8 mm

5.1.2 摄像机离线标定

摄像机的离线标定通常有两种方式可以实现：第一种是直接利用OpenCV图像处理库[28]的内置函数进行标定；第二种是通过Matlab的立体标定工具Stereo Camera Calibrator进行离线标定。由于第一种方法通常会产生过于夸张的矫正效果，所以文本采用第二种方式。

摄像机离线标定通常采用棋盘标定法实现。本文使用的棋盘如图5-2所示。

图5-2 棋盘图

棋盘由黑白相间的正方形组成，为达到最好的标定效果，棋盘的长和宽的方格数必须要有一个为偶数，一个为奇数。本文的棋盘方格数为6×9，实际打印后每个正方形的边长为29mm，棋盘被固定在一硬纸箱底部。本文选取了15组不同的图像对进行标定，图像对左右图如5-3所示。

Matlab的标定和矫正是一个自动化的过程，矫正过程中会自动修剪，只保留左右图像的共同区域。软件界面如图5-4所示，矫正前后的图像对如图5-5所示，矫正误差如图5-6所示，平均矫正误差为0.19个像素，基本满足实际应用需求。

图 5-3 左起依次为：棋盘图像左图，棋盘图像右图

图5-4 Matlab Stereo Camera Calibrator操作界面

图5-5 矫正前后图像对:从上到下依次为矫正前，矫正后

图5-6 矫正误差

5.1.3 双目矫正与三维重建

经过Matlab离线标定所得的矫正参数，可以通过OpenCV的内置函数进行双目矫正和三维重建[28]。本文的矫正参数如图5-7所示，通过双目矫正的图像基本满足左右图的匹配点基本在同一水平线上。

图5-7 矫正参数

5.2 立体匹配

5.2.1 实验环境

随着近年来卷积神经网络的发展，涌现出了大量的深度学习框架，目前常见的有Caffe, TensorFlow, MXNet, Torch, Theano等五个主流框架[29,30,31]，这些框架的优缺点比较如表5-2所示。

表5-2 深度学习框架比较

框架开发语言速度预训练模型 RNN/CNN 多GPU

Caffe C++/Python 快支持 CNN 支持

TensorFlow Python/C++ 中等支持(Inception) 支持(最好) 支持(最好)

MXNet C++/Python/R 快支持支持支持

Torch Lua/C 快支持支持支持

Theano Python/C++ 中等支持(Lasagne) 支持支持

实验主要使用Python作为开发语言，需要GPU加速运算，综合考虑立体匹配算法的需求，本文选择TensorFlow作为模型训练框架。本文实验环境为TensorFlow框架，CentOs系统，硬件: CPU(8 cores): [Intel\(R\) Core\(TM\) i7-4790 CPU @ 3.60GHz](#) , Mem: 15990MB , GPU : [NVIDIA Corporation GM200 \[GeForce GTX TITAN X\]](#) , NVIDIA Corporation GK208 [GeForce GT 730]。

5.2.2 模型训练

本文使用3.2小节所说的训练数据集训练神经网络。训练期间，每次输入一批128对数量相等的积极和消极示例图像对。

本文使用动量项为0.9的动量优化器来减少较链损失。本文共训练了20次迭代，初始学习率为0.001，在一次迭代后改为0.0002，第10次迭代后改为0.00002。训练中为了防止过拟合，本文还引入了权值衰减，衰减系数为0.005，初始学习率和学习率的衰减计划是通过交叉验证优化得到的。本文网络的卷积层参数如表5-3所示。

表5-3 子网卷积层参数

网络层卷积核，特征映射

子网卷积层+ReLU 3×3，64

子网卷积层+ReLU 3×3，64

子网卷积层+ReLU 3×3，64

子网卷积 3×3，64

5.2.3 立体匹配细节

立体匹配的初始视差图 $D(p)$ 采用胜者全拿的策略得到，即对每个像素点 p 寻找使匹配代价 $CCNN_{p,d}$ 最小的视差值 d ，如公式5-1所示。

$$D_p = \arg\min_d CCNN_{p,d} \quad (5-1)$$

其中匹配代价通过Siamese网络的输出初始化得到。考虑到计算量较大，本文通过以下三个细节提高算法效率：

计算匹配代价时执行2.3节所说的视差范围约束，即每个像素点仅计算视差范围内的匹配代价。因为本文使用的数据集集中的最大视差值为228，所以本文视差范围为228。

每个像素点的子网部分仅计算一次，而不是对每个视差值重新计算，即在计算匹配代价前先将每个像素点的子网部分计算一次，子网输出保存后用于后续的匹配代价计算

在运行期间，本文希望输入的是一整张图片，但由于GPU内存的限制，本文每次仅输入一行图像数据（约1220个像素点）进入网络，并同时计算以左图作为参考图的匹配代价和以右图作为参考图的匹配代价，这样子网部分可以重复使用，节省运算时间。

运行过程中，每个输入Siamese网络的图像对都需要通过和训练数据一样的归一化处理，即将图像的像素强度值减去它们的平均值再除以标准差。

验证数据集和测试数据集中的图片已经过双目矫正，可以直接计算视差图，但实际摄像头采集的图片存在一定程度的畸变，所以需要进行图像矫正，矫正所得的参数还可用于三维恢复。

表5-4 超参数

超参数值超参数值

input_patch_size 9×9 ksize 5

dataset_pos 1 sgm_P1 3.5

dataset_neg_low 4 sgm_P2 223.0

dataset_neg_high 10 sgm_Q1 2.0

d 40 sgm_Q2 4.0

sigmaColor 5 sgm_V 1.75

sigmaSpace 5 sgm_D 0.02

本文通过公式5-1所说的胜者全拿（WTA）策略得到以左图作为参考得到的视差图 DL 和以右图作为参考得到的视差图 DR 视差图

。DL和DR通过左右一致性检查得到视差图DINT。DINT 通过公式5-2所示的亚像素增强获得新的视差图DSEp。最后通过一个中值滤波和双边滤波得到了我们的最终视差图。立体匹配流程图如图5-8所示。

$$DSEp=d-C+-C-2(C+-2C+C-) \quad (5-2)$$

式中 $d=DINT(p)$ ， $C-=CSGM(p,d-1)$ ， $C=CSGM(p,d)$ ， $C+=CSGM(p,d+1)$ 。

图5-8 立体匹配流程图

本文采用的超参数如表5-4所示，通过不同参数在验证数据集中的不同表现手动调整和简单的脚本优化得到。在测试数据集中，本文的立体匹配算法的视差误差仅为8.0%，基本满足实际应用需求，视差图如图5-9所示，误差计算方法与KITTI立体视觉评测方式相同，如公式5-3所示。

$$error=N1N2 \quad (5-2)$$

其中N1表示视差绝对误差超过3个像素的点的个数，N2表示参考视差图中有视差值的点的个数。

图5-9 从上到下依次为：左图，右图，视差图

5.3 服务器搭建与客户端设计

5.3.1 服务端搭建

本文的立体匹配代价计算需要GPU加速计算，由于实验中使用的GPU服务器为内网服务器，所以本文需要搭建Ngrok服务实现内网穿透并通过Flask搭建Web服务器。

Ngrok是目前常见的反向代理隧道工具，通过在内网主机和公网转接服务器之间建立安全隧道，让运行内网主机上的服务可以暴露给公网，即内网穿透。本文实验使用的是国内免费Ngrok服务器ITTun.com，Ngrok的配置信息如图5-10所示。

随着近年来互联网的发展，涌现出众多优秀的Web服务器开源框架，其中较为成熟且应用较广的有Nginx，Apache，Django，Flask等。Flask是一个轻量级的Python服务器框架，结构简单灵活可扩展性强，结合本文程序主要使用Python语言开发，所以本文选择Flask框架搭建Web服务器，便于后续集成。

图5-10 Ngrok配置信息

5.3.2 客户端设计

本文实验中的客户端是基于Qt框架搭建的。Qt框架是一个面向对象的跨平台图形用户界面库，开发套件功能强大，应用广泛。本文通过Qt框架在Python的接口PyQt进行开发。初始界面如图5-11，包含“Start”、“Capture”、“UpLoad”、“Disparity”四个按钮，“Left_Camera”，“Right_Camera”两个图像显示屏以及一个文本提示框。

图5-11 初始页面

实验中点击“Start”按钮启动连接的双目摄像头，开启视频录制，如图5-12所示。当出现视频中目标物体后，可以点击“Capture”按钮截取目标图片，同时进行双目矫正，如图5-13所示。确认图片后，可带点击“UpLoad”按钮上传图片至服务端，上传成功提示框显示“UpLoad Success”，否则显示“UpLoad Failed”，如图5-14所示。待服务端计算完毕后，可点击“Disparity”按钮，获取视差图片，获取的视差图会开启另一窗口显示，如图5-15所示，显示的视差图是由初始视差图归一化后经过OpenCV的HSV伪彩色处理得到[28]，点击视差图的某一位置，可以获取该位置的深度信息（以毫米为单位），如图5-16为先后点击图中两个水瓶的位置显示的情况，左侧水瓶约为1300毫米，右侧水平约为1700毫米，深度信息与实际基本相同。

图5-12 开启视频录制

图5-13 截取目标图片并矫正

图5-14 上传图片

图5-15 所得的视差图

图5-16 提示框显示深度信息

5.4 本章小结

本章节介绍了整个立体视觉的实现过程。首先本章介绍了实验中双目摄像机的离线标定、双目矫正和三维重建过程及相关参数，随后本章详细描述了立体匹配算法的实现细节，从环境搭建，模型训练到算法的实现细节。最后，本章简单介绍了服务端的搭建过程和客户端的UI设计。

7. 第6章总结与展望		总字数：987
相似文献列表 文字复制比：17.7%(175) 疑似剽窃观点：(0)		
1	2012040306013李桦 李桦 - 《大学生论文联合比对库》 - 2016-06-07	7.0% (69) 是否引证：否
2	关银凤 论文初稿2 关银凤 - 《大学生论文联合比对库》 - 2016-06-04	6.6% (65) 是否引证：否
3	关银凤 论文初稿3	6.6% (65)

关银凤 - 《大学生论文联合比对库》 - 2016-06-06		是否引证：否
4	基于PHP的考研经验交流网站的设计与实现 张敬宇 - 《大学生论文联合比对库》 - 2015-06-15	5.1% (50)
5	双目视觉中立体匹配算法的研究与实现 蒋昂(导师：汤春明;王景峰) - 《天津工业大学硕士论文》 - 2015-12-01	3.6% (36)
6	旅游发展过程中的政府作用演化与动因研究 赵功凯(导师：梁明珠) - 《暨南大学硕士论文》 - 2009-05-08	3.5% (35)
7	低维MoS ₂ 的制备及光学特性研究 李雪(导师：王晓华;李金华) - 《长春理工大学博士论文》 - 2016-03-01	3.1% (31)
8	反重力铸造充型过程的水力学特征和氧化膜卷入机制 刘闪光(导师：孙剑飞) - 《哈尔滨工业大学博士论文》 - 2015-12-01	3.1% (31)
9	大型重载液体静压转台承载特性及流固热耦合规律研究 王少力(导师：熊万里) - 《湖南大学博士论文》 - 2015-12-28	3.1% (31)
		是否引证：否

原文内容

第6章总结与展望

6.1 论文总结

近年来，随着无人驾驶，机器人导航等技术的发展，计算机立体视觉作为其中重要的感知识别单元引起了研究者们广泛的关注。如何从二维的平面图像中提取出三维的深度信息成为一个重要的命题。双目立体匹配技术是双目立体视觉中的核心步骤，也是本文的研究重点。

本文在现有研究基础上，对立体匹配算法进行了深入研究，为提高匹配算法的精度，主要做了以下工作：

- 1) 本文对双目立体视觉的相关原理进行了深入研究，研究了摄像机成像原理，视差理论，立体匹配的经典步骤和约束条件。
- 2) 本文通过训练卷积神经网络进行图像块的相似性度量，在测试数据集中仅有8.02%的视差误差，已基本达到实际应用的需求。
- 3) 本文完成了基于Qt的双目视觉应用程序设计，用户可以连接双目摄像机，基本实现了双目测距。

6.2 论文研究的不足与展望

本文是在短时间内完成了双目立体视觉系统的设计和搭建，所以立体匹配算法中的许多超参数仍需要进一步优化，网络模型也有待于进一步的训练。实验中的许多模块仅仅出于调试的需求，并未考虑大规模的开发应用，系统容量十分有限。由于GPU内存的限制，本文双目视觉的计算时间超过了实际的应用的需求，后续研究中可以考虑应用多GPU或TPU来加速运算。

参考文献

- [1] 周星，高志军.立体视觉技术的应用与发展[A].工程图学学报,2010年第4期，No.4.
- [2] Marr D C.A Computational Investigation into the Human Representation and Processing of Visual Information [M] . San Francisco : W . H . Freeman and company , 1982 .
- [3] Scharstein D, Szeliski R, Zabih R, et al. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms[J]. International Journal of Computer Vision, 2001, 47(1): 131-140.
- [4] Kong D, Tao H. A method for learning matching errors for stereo computation.[C]. british machine vision conference, 2004.
- [5] Zhang L, Seitz S M. Estimating Optimal Parameters for MRF Stereo from a Single Image Pair[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2007, 29(2): 331-342.
- [6] Scharstein D, Pal C. Learning conditional random fields for stereo[C]. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2007.
- [7] Peris M, Martull S, Maki A, et al. Towards a simulation driven stereo vision system[C]//Pattern Recognition (ICPR), 2012 21st International Conference on. IEEE, 2012: 1038-1042.
- [8] Li Y, Huttenlocher D P. Learning for stereo vision using the structured support vector machine[C]. computer vision and pattern recognition, 2008: 1-8.
- [9] Haeusler R, Nair R, Kondermann D, et al. Ensemble Learning for Confidence Measures in Stereo Vision[C]. computer vision and pattern recognition, 2013: 305-312.
- [10] Spyropoulos A, Komodakis N, Mordohai P, et al. Learning to Detect Ground Control Points for Improving the Accuracy of Stereo Matching[C]. computer vision and pattern recognition, 2014: 1621-1628.
- [11] Zbontar J, Lecun Y. Computing the stereo matching cost with a convolutional neural network[C]. computer vision and pattern recognition, 2015: 1592-1599.

- [12] Geiger A, Lenz P, Stiller C, et al. Vision meets robotics: The KITTI dataset[J]. The International Journal of Robotics Research, 2013, 32(11): 1231-1237.
- [13] Guney F, Geiger A. Displets: Resolving stereo ambiguities using object knowledge[C]. computer vision and pattern recognition, 2015: 4165-4175.
- [14] Zhang K, Lu J, Lafruit G, et al. Cross-Based Local Stereo Matching Using Orthogonal Integral Images[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(7): 1073-1079.
- [15] Hirschmuller H. Stereo Processing by Semiglobal Matching and Mutual Information[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(2): 328-341.
- [16] Scharstein D, Pal C. Learning conditional random fields for stereo[C]//Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on. IEEE, 2007: 1-8.
- [17] Scharstein D, Szeliski R. High-accuracy stereo depth maps using structured light[C]//Computer Vision and Pattern Recognition, 2003. Proceedings. 2003 IEEE Computer Society Conference on. IEEE, 2003, 1: I-I.
- [18] Scharstein D, Hirschmüller H, Kitajima Y, et al. High-resolution stereo datasets with subpixel-accurate ground truth[C]//German Conference on Pattern Recognition. Springer International Publishing, 2014: 31-42.
- [19] Hubel D H, Wiesel T N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex[J]. The Journal of physiology, 1962, 160(1): 106-154.
- [20] Fukushima K, Miyake S. Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position[J]. Pattern recognition, 1982, 15(6): 455-469.
- [21] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [22] Hinton G E, Osindero S, Teh Y, et al. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554.
- [23] Deng J, Dong W, Socher R, et al. Imagenet: A large-scale hierarchical image database[C]//Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009: 248-255.
- [24] Krizhevsky A, Sutskever I, Hinton G E. Imagenet classification with deep convolutional neural networks[C]//Advances in neural information processing systems. 2012: 1097-1105.
- [25] Zbontar J, LeCun Y. Stereo matching by training a convolutional neural network to compare image patches[J]. Journal of Machine Learning Research, 2016, 17(1-32): 2.
- [26] Bromley J, Bentz J W, Bottou L, et al. Signature Verification Using A "Siamese" Time Delay Neural Network[J]. IJPRAI, 1993, 7(4): 669-688.
- [27] Hirschmuller H. Stereo processing by semiglobal matching and mutual information[J]. IEEE Transactions on pattern analysis and machine intelligence, 2008, 30(2): 328-341.
- [28] 于仕琪, 刘瑞祯. 学习 OpenCV (中文版)[J]. 2009.
- [29] Abadi M, Agarwal A, Barham P, et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015[J]. Software available from tensorflow. org.
- [30] Jia Y, Shelhamer E, Donahue J, et al. Caffe: Convolutional architecture for fast feature embedding[C]//Proceedings of the ACM International Conference on Multimedia. ACM, 2014: 675-678.
- [31] Chen T, Li M, Li Y, et al. Mxnet: A flexible and efficient machine learning library for heterogeneous distributed systems[J]. arXiv preprint arXiv:1512.01274, 2015.

致谢

经过几个月来的忙碌工作，我的毕业论文设计也要接近尾声了。作为一个本科学生的毕业设计，由于经验的缺乏，难免会有许多考虑不全之处，在这里，我要由衷的感谢指导老师宣琦副教授，正是在他的悉心指导和帮助下，我进入了深度学习这个全新的领域，学会了如何利用这项技术解决现实问题，他渊博的专业知识，严谨的治学态度，精益求精的工作作风和诲人不倦的高尚师德，给我留下了深刻的印象，使我受益匪浅。在此，谨向导师致以崇高的敬意和衷心的感谢。

离别在即，站在人生的又一转折点上，心中难免思绪万千，一种感恩之情油然而生。感谢所有鼓励支持过我的学院领导、任课老师和学长学姐，是他们给我带来了丰富多彩的大学生活，在他们的帮助下，我参加了各项文体科技竞赛，得到了一次次锻炼自己的机会，包括最后的毕业设计，也是在大家的共同努力下完成的。作为信息人，我不会忘记演讲台上的激昂青春，不会忘记实验室里辛勤汗水，这一次次的经历将会在日后的工作和生活中产生巨大的影响。风雨兼程，感谢一路有你们。别了我的大学，让我们江湖再见。

指 标

疑似剽窃文字表述

1. 毕业论文设计也要接近尾声了。作为一个本科学生的毕业设计，由于经验的缺乏，难免会有许多考虑不全之处，在这里，我要由衷的感谢指导老师宣琦副教授，

说明：1.指标是由系统根据《学术论文不端行为的界定标准》自动生成的。

2. **红色**文字表示文字复制部分；**黄色**文字表示引用部分。

3.本报告单仅对您所选择比对资源范围内检测结果负责。

4.Email：amlc@cnki.net

<http://e.weibo.com/u/3194559873>

http://t.qq.com/CNKI_kycx

“中国知网”大学生论文检测系统