**基于双目视觉的UAV自主感知与避让**

# 一、绪论

写2页左右。

1.1 研究背景及意义

1.2 国内外研究现状

UAV障碍物检测

1.3 论文研究的目的

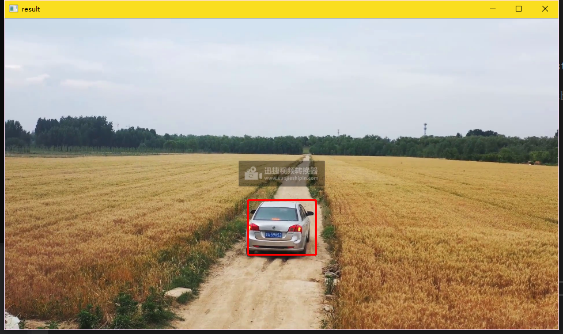
1.4 本论文的主要工作

# 二、场景剖析及硬件选型

## 2.1 场景剖析

无人机飞行障碍物的多样性、复杂性决定了本课题难度。以当前的目标检测算法而言，很难有一种算法能够对所有类型的障碍物都具备检测识别能力。因此，只有针对具体实用场景下的具体障碍类型进行深入研究，才能达到较好的应用效果。

本文设定的研究场景为无人机智能跟拍车辆。大疆无人机（御 Mavic Air等型号）具有智能跟拍车辆的功能，支持对行进车辆的识别跟随能力，无人机可以在顶部俯拍、车尾/侧面跟拍行进中的车辆。在顶部俯拍时，无人机的飞行高度保证了不会与被跟随车辆相撞。在车尾/侧面跟拍时，由于速度不同步，很有可能出现无人机加速接近汽车的情况。这种情况下，被跟随车辆就属于无人机飞行路线中有碰撞风险的障碍物，无人机需要识别、跟踪并规避车辆。

车尾跟拍 侧面跟拍

## 2.2 硬件选型

无人机智能跟拍车辆场景的自主感知与避让，核心功能有3个：障碍物检测（车辆检测）、障碍物追踪（车辆追踪）、危险距离判定（车辆测距）。其中，危险距离判定要求相对高的精度，是硬件选型时主要的关注点。

在经费有限的情况下，无人机较难获得。结合核心功能需求、危险距离判定精度要求等的考量，最终我们选择使用KS861双目摄像头模拟无人机进行实验。



双目摄像头

真实场景中，大疆无人机（御 Mavic Air等型号）在智能跟拍汽车时，无人机和车辆均处于运动状态。模拟场景下，手持双目摄像头前后左右晃动，以模拟无人机的姿态变换；同时，车辆由远及近接近双目摄像头，以模拟碰撞威胁逐步增大的场景。

（1）无人机的缺点在于：

①经费不足，无人机购置成本较高，大疆的御Mavic Air2京东售价4999元；

②京东上的无人机多为单目摄像头，而单目测距精度相对较差；

③无人机属于公安管控物品，违规飞行可能给导师和学校带来潜在风险。

（2）双目摄像头替代的优点在于：

①双目摄像头与车辆相互配合，能够模拟无人机的姿态变换与距离拉远拉近；

②双目摄像头价格相对便宜，且不会给导师和学校带来使用风险；

③双目摄像头130w像素，拍照性能优良，采集的视频数据能满足实验要求；

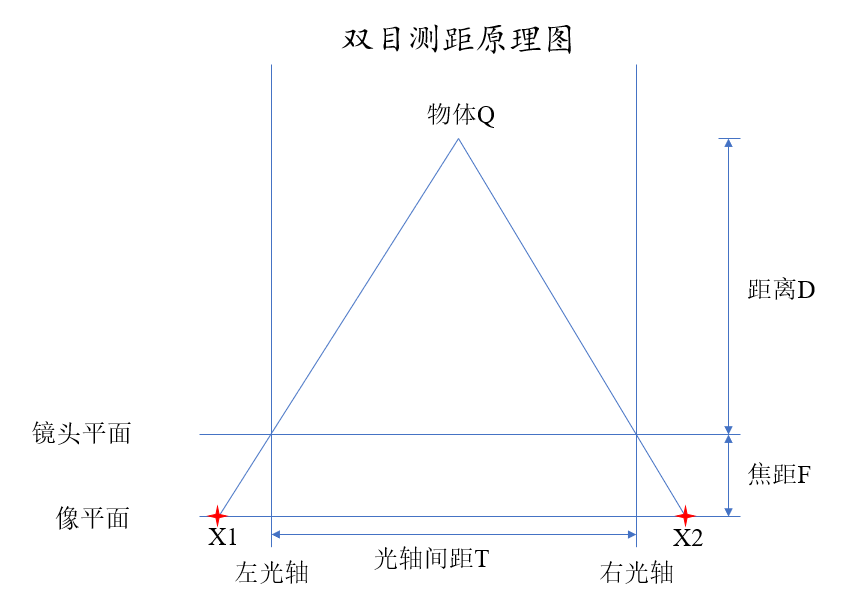
④双目摄像头具备良好的便携性,实验数据可以方便地导出；

⑤使用KS861双目摄像头在50m内的危险距离测量理论误差仅为2.6%。

## 2.3 危险距离测量精度

### 2.3.1 危险距离测量原理

为了保证危险距离测量的精度，我们在一开始就设计了使用双目测距模型。我们人类拥有两只眼睛，物体在左眼视角与右眼视角的内容有差别，这些差别反馈到人脑后我们就能估计出物体距离我们有多远。双目测距模型就是在模仿人类两只眼睛的功能，原理如下图所示：



物体Q在左像平面成像点X1，在右像平面成像点X2。镜头平面与像平面之间的焦距为F，左右光轴的间距为T，需要求解出物体Q与镜头平面的距离D。

原理图中表述的是理想模型：左右两个摄像头焦距等参数完全相同，它们的像平面和镜头平面也相同，它们的左右光轴相互平行。因此，我们可以根据相似三角形的原理，得到距离D的表达式：

, 其中

为物体Q在左右两个摄像头上成像的像素坐标差值，也就是双目视差。结合上述公式，在进行双目危险距离测量时，已知焦距F、光轴间距T和双目视差，即可求解得到物体Q与镜头平面的距离D。当D值小于某个阈值，我们可判定处于高碰撞风险状态。

我们已经得到危险距离D的表达式：

, 其中

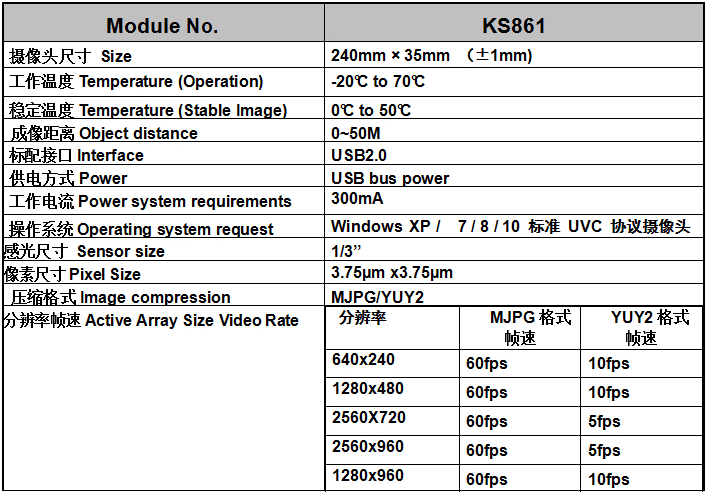
将公式两边同时微分:

上述两式整合，得到危险距离测量的相对误差公式：

表示像元尺寸，F表示焦距，这两个参数在摄像头选定之后即固定。那么在被测距离D一定的情况下，通过增大光轴间距T，可以减小危险距离测量的相对误差。

### 2.3.2 危险距离测量精度计算

KS861双目摄像头的参数如下：



实验条件设定如下：

测量距离：0~50m

光轴间距：0.6m

焦距：12mm

像元尺寸：3.75um

因此，KS861双目摄像头危险距离测量的理论误差为：50 \* 3.75 \* 10-6/(12\*10-3\*600\*10-3) = 2.6%。式中，50表示最大测量距离50m，3.75 \* 10-6表示3.75um的像元尺寸，12\*10-3为摄像头的焦距12mm，600\*10-3为两颗摄像头的基线间隔600mm。2.6%的理论测量误差，完全能够满足我们的实验要求，这也是选型KS861双目摄像头的重要理论依据。

# 三、实验数据采集与矫正

## 3.1 数据采集

数据采集时模拟无人机智能跟拍汽车的场景。汽车与双目摄像头的距离在0~50m之间，每间隔5m设置一个标记点位（图中黄色的橘子），以评判算法估算的危险距离与实际距离之间的误差大小。双目摄像头和车辆均处于运动状态：手持双目摄像头在原地前后左右晃动，以模拟无人机的姿态变换；同时，车辆由远及近接近双目摄像头，以模拟无人机与车辆相对距离减小，碰撞威胁逐步增大的场景。

1、选择场地，选择起始线，卷尺测量4个标记点位（20m、30m、40m、50m，放小凳子）；预计20min。

2、车辆依次开到点位上，每个点位采集一组视频，从左侧到正中到右侧；（选择出距离合适的放进论文）；预计15min。

3、从50m处缓慢倒车，再缓慢开回50m处，往返3次。录制左侧、正中、右侧的视频。预计15min。

数据采集的要点如下：

①汽车与双目摄像头的距离：0~50m；

②设置7个标记点位（20m、25m、30m、35m、40m、45m、50m），定点采集图像，左侧/正中/右侧；

③双目摄像头和车辆均处于运动状态；

④双目摄像头在原地前后左右晃动，以模拟无人机的姿态变换；

⑤从50m处缓慢倒车，逐渐接近双目摄像头，以模拟无人机与车辆相对距离减小，碰撞威胁逐步增大的场景。



双目摄像头及实验车辆



6个评判点位（橘子标记）

通过上述模拟场景，我们采集得到一段视频作为本文的数据研究对象。该视频时长54s，1080p（1920\*1080），帧率30fps，共计54\*30=1620帧。下图即为模拟无人机智能跟拍功能时，双目摄像头记录到的画面。



实验画面

## 3.2 数据矫正

### 3.2.1 高斯滤波

一般而言，摄像头采集得到的图像会存在各种噪声，比如环境噪声、摄像头电子元件引入的噪声等等。为了减小噪声的影响，并提高后续环节（车辆检测、危险距离测量）的精度，有必要对原始数据进行高斯滤波处理。

#### 3.2.1.1 高斯分布

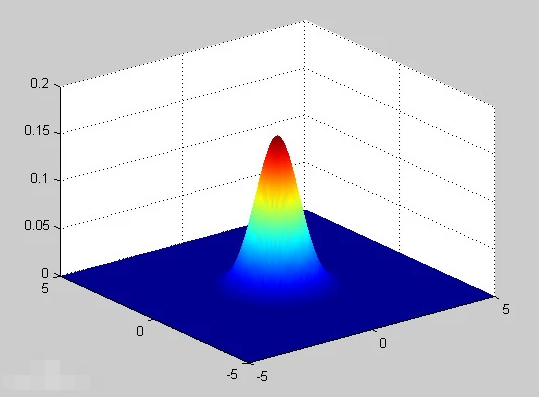
高斯滤波通过数学上的高斯分布函数进行。根据以前的数学知识，我们知道均值为0的一维高斯分布公式为：

一维高斯分布的曲线图为：



均值为0的二维高斯分布公式为：

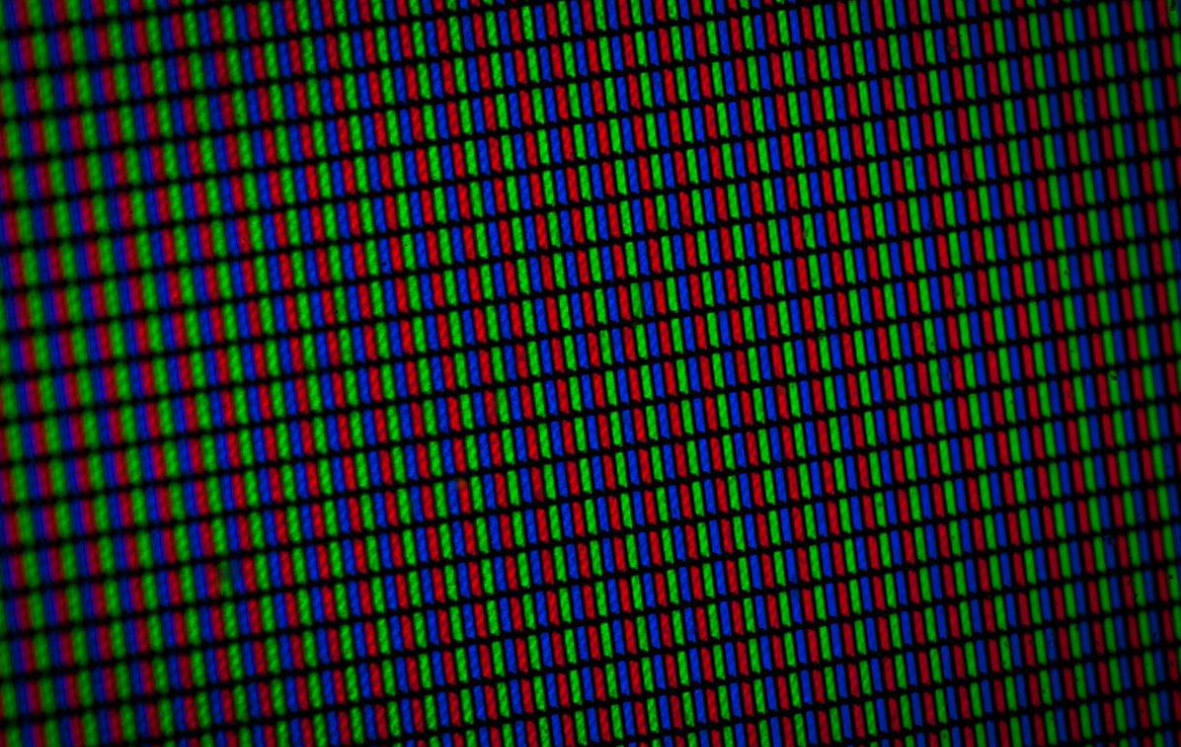
二维高斯分布的曲线图为：



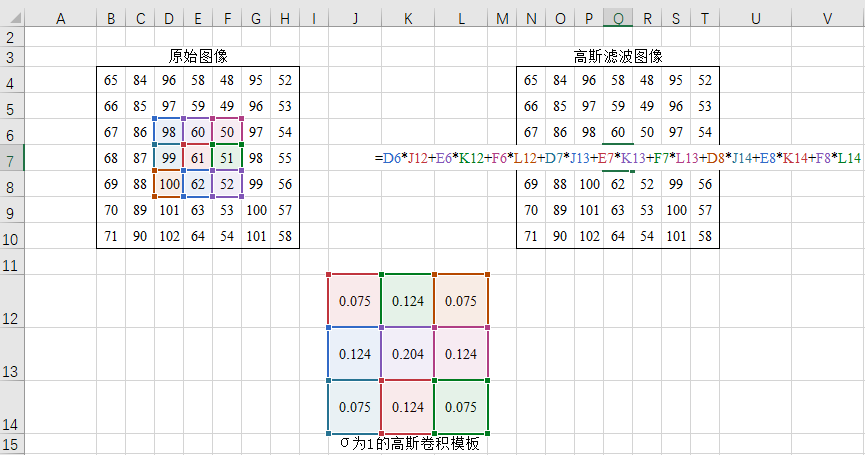
从一维和二维高斯分布的图形可以直观的看到，越靠近中心的部位，值越大，表示所占的权重越大；越靠近边缘的部位，值越小，表示所占的权重越小。

#### 3.2.1.2 数字图像

数字图像由表示红、绿、蓝三种原色的数字表示，并在宽度和高度两个方向堆积。因此，数字图像在计算机内部可以被视为二维矩阵。



高斯滤波的过程就是使用二维高斯函数与原始图像像素值加权求和的过程。图像处理中常用3\*3的二维高斯函数矩阵（边缘部分权值较小，可以忽略不计），中心部分权值占比高，稍远部分权值占比低，例如下图中：



### 3.2.2相机标定

在2.3.1节危险距离测量原理中，存在两个假设条件：

①两颗摄像头的光轴平行；

②两颗摄像头的像平面为同一平面。

但是在摄像头的生产过程中，由于加工装配的误差，上述两个条件很难得到满足。我们通过相机标定来求解相机参数以及两颗摄像头的相对位置关系。

在图像测量过程以及机器视觉应用中，为确定空间物体表面某点的三维几何位置与其在图像中对应点之间的相互关系，必须建立相机成像的几何模型，这些几何模型参数就是相机参数。在大多数条件下这些参数必须通过实验与计算才能得到，这个求解参数的过程就称之为相机标定。无论是在图像测量或者机器视觉应用中，相机参数的标定都是非常关键的环节，其标定结果的精度及算法的稳定性直接影响相机工作产生结果的准确性。

#### 3.2.2.1 坐标系转换

首先介绍一下坐标系转换中用到的4种坐标系：世界坐标系、相机坐标系、图像坐标系、像素坐标系。坐标系转换就是为了将空间的三维世界坐标系转换至图像处理的二维像素坐标系，并在此过程中求得世界坐标系与像素坐标系的转换矩阵。

①世界坐标系（world coordinate）：也称为测量坐标系，是一个三维直角坐标系，以其为基准可以描述相机和待测物体的空间位置。世界坐标系的位置可以根据实际情况自由确定。世界坐标系的最小单位为mm。

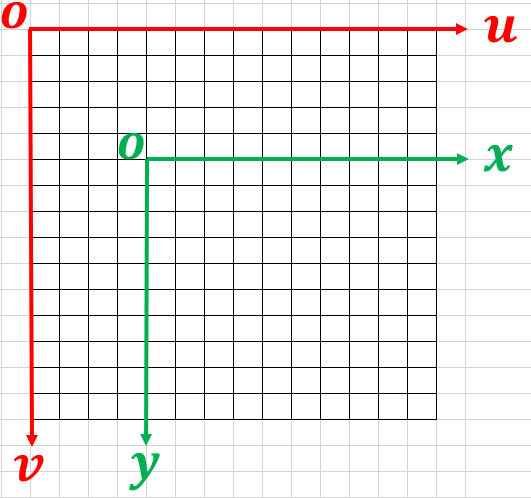
②相机坐标系（camera coordinate）：也是一个三维直角坐标系，原点位于镜头光心处，、轴分别与像面的两边平行，轴为镜头光轴，与像平面垂直。相机坐标系的最小单位为mm。

③图像坐标系（image coordinate）：是像平面上的二维直角坐标系。图像坐标系的原点为镜头光轴与像平面的交点（也称主点，principal point），它的轴与相机坐标系的轴平行，它的轴与相机坐标系的轴平行。图像坐标系的最小单位为mm。

④像素坐标系（pixel coordinate）：是图像处理工作中常用的二维直角坐标系，反映了相机CCD/CMOS芯片中像素的排列情况。它的原点位于图像左上角，横坐标表示像素所在的列，纵坐标表示像素所在的行。像素坐标系与图像坐标系可以简单理解为平移关系，它们同处于像平面。像素坐标系的轴与图像坐标系的轴平行，像素坐标系的轴与图像坐标系的轴平行。像素坐标系的最小单位为像素。

3.2.2.1.1图像坐标系转换为像素坐标系

像素坐标系与图像坐标系是平移关系：

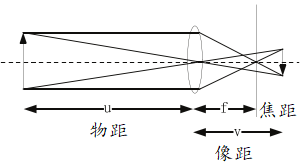


它们之间的转换矩阵关系为：

、表示像元尺寸，即单一像素在轴和轴方向上的物理长度（其单位可理解为 um/像素）。、为像素坐标系原点在图像坐标系中的位置坐标。

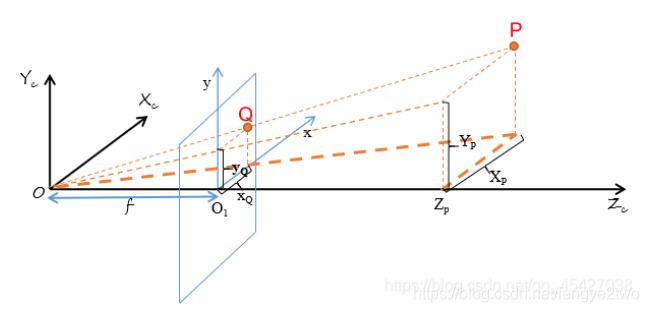
3.2.2.1.2相机坐标系转换为图像坐标系

相机成像

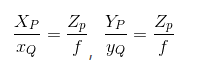


相机的镜头相当于一个凸透镜，感光元件就处在这个凸透镜的焦点附近，所成的像是一个倒像。在实际中，当物距足够远时(远大于两倍焦距)，凸透镜成像可以看作是在焦距处的小孔成像，即所谓的“针孔模型”。通过针孔模型，我们可以得到三维坐标到二维坐标的的对应关系。

小孔成像

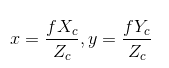


Q点位于图像坐标系，P点位于相机坐标系。根据三角形相似原理，上图可以得到如下比例关系：



变换矩阵

将Xp、Yp、Zp用相机坐标系的Xc、Yc、Zc表示，xQ、yQ用图像坐标系的x、y表示，并变换x、y到左边：



所以，相机坐标系与图像坐标系之间的矩阵变换关系为：

3.2.2.1.3 世界坐标系转换为相机坐标系

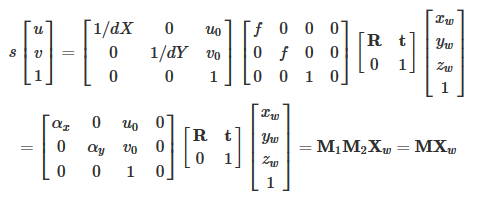
世界坐标系和相机坐标系都是三维坐标，任意两个三维坐标之间均可以通过旋转和平移进行转换：

其中R为3\*3的旋转矩阵，T为3\*1的平移矢量，为相机坐标系的齐次坐标，为世界坐标系的齐次坐标。

由于世界坐标系和相机坐标系之间的变换矩阵与相机无关，因此称为外部参数矩阵。

3.2.2.1.4 世界坐标系转换为像素坐标系

将前面的几个变换矩阵进行连乘，即可得到世界坐标系与像素坐标系的转换关系（s为zc）：



其中，ax=f/dx、ay=f/dY，称为u、v轴的尺度因子，M1称为相机的内部参数矩阵，M2称为相机的外部参数矩阵，M称为投影矩阵。

#### 3.2.2.2 单目相机标定

单目相机标定能够求得摄像头的内部参数，即M1矩阵。我们使用张友正相机标定法来完成，OpenCV库中的接口函数：calibrateCamera已封装实现张友正相机标定的过程。我们只需要按照要求填写入参即可。

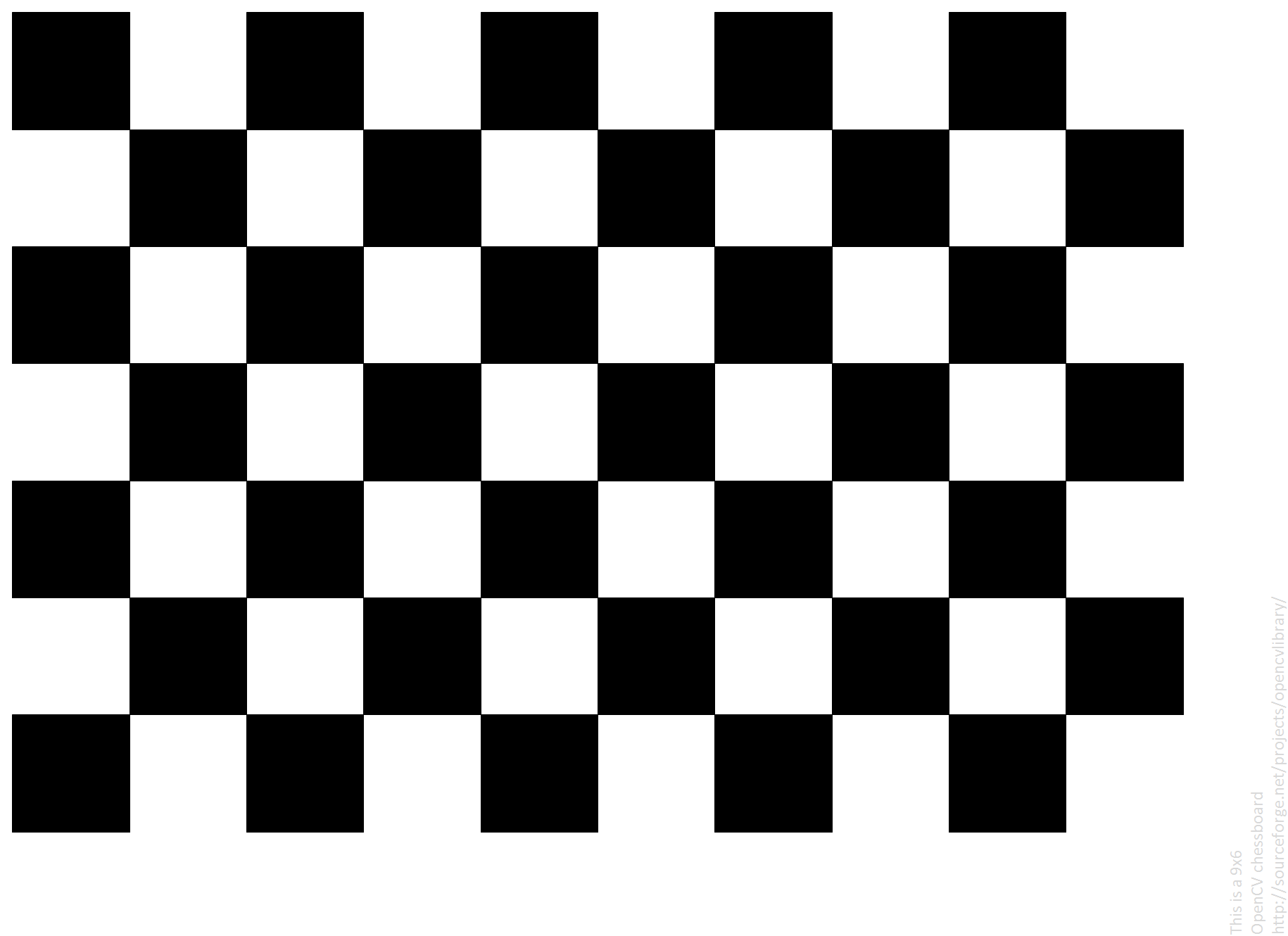
单目相机标定的主要流程如下：

1、棋盘图准备

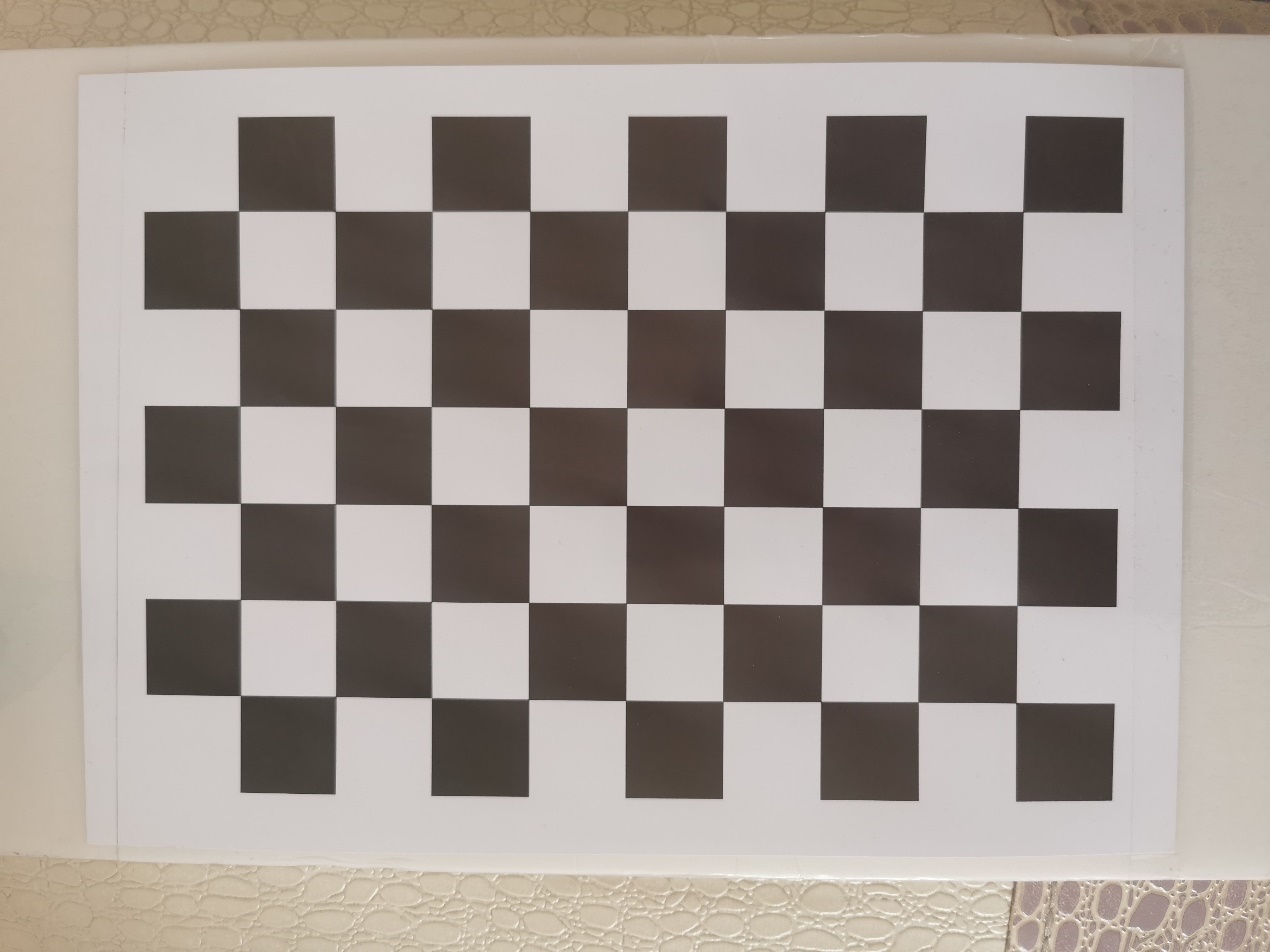
在网上找到黑白方格构成的标定板图样：棋盘图

水平方向有10个黑白相间的格子，9个角点；

竖直方向有7个黑白相间的格子，6个角点。



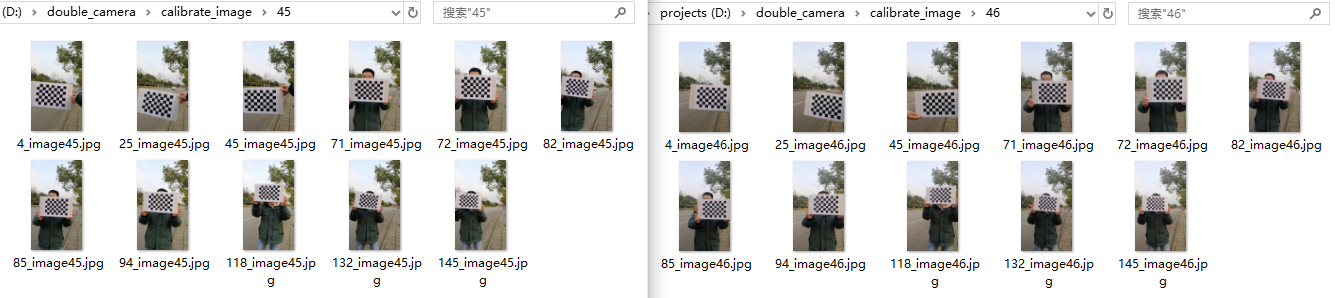
2、打印棋盘图



将棋盘图打印在A4纸上，并用透明胶带粘贴固定在硬纸板上。使用直尺测量黑白格子的实际尺寸为26mm\*26mm。

3、双目摄像头拍摄棋盘图

手持棋盘图，保证左右摄像头中都有完整的棋盘图画面。在不同的距离、角度，拍摄10组棋盘图的照片。



4、提取二维角点

使用cv::findChessboardCorners函数提取角点，这里的角点指黑白格子交叉处的内角点。为了提高精度，使用cv::cornerSubPix函数进一步提取亚像素级的角点imageCorners2D。最后将角点提取结果绘制在画面上，以确认每个角点都被正确检测出来。



5、计算角点三维空间位置

第二步测量到标定板上黑白格子的实际尺寸为26mm\*26mm，我们可以借此得到每个角点的三维空间位置信息realCorners3D（假定z轴尺寸均为0）：第i行第j列角点的三维坐标为（j \* 26, i \* 26, 0）。

6、内参矩阵求解

至此，我们得到了成像画面上的二维角点以及实际空间的三维角点信息，它们是一一对应的。通过calibrateCamera函数即可求解得到内参数矩阵cameraMatrix、畸变矩阵distCoeffs、旋转向量rvecs和位移向量tvecs。

#### 3.2.2.3 双目相机标定

在单目标定的基础上，对双目相机进行标定，使用OpenCV的接口函数：stereoCalibrate即可求得相机的外部参数矩阵M2。输入左右摄像头的二维角点imageCorners2D、内参数矩阵cameraMatrix、畸变矩阵distCoeffs，以及三维角点realCorners3D，求得左右摄像头之间的旋转向量R和平移向量T。R是3\*3向量，T是3\*1向量，它俩组合在一起形成的3\*4向量即为相机的外部参数矩阵M2。

#### 3.2.2.4 图像矫正

通过双目相机标定，获得了图像的矫正参数，并用它们来矫正图像，OpenCV的接口函数：stereoRectify、initUndistortRectifyMap和remap

<https://blog.csdn.net/fengye2two/article/details/80686409>

<https://blog.csdn.net/weixin_43843780/article/details/89294131>

<https://www.cnblogs.com/zyly/p/9373991.html>

<https://blog.csdn.net/onthewaysuccess/article/details/40736177>

<https://blog.csdn.net/a083614/article/details/78579163>

# 四、障碍物检测

写5页左右。

现有检测算法有哪几种，各自对比的优缺点。YOLO检测的理论基础。

YOLOv4，我们调用别人的接口来识别车辆，不需要额外训练。

<https://blog.csdn.net/baobei0112/article/details/105831613/>

https://blog.csdn.net/qq\_38316300/article/details/105759305

# 五、障碍物跟踪

写5页左右。

现有追踪算法有哪些，各自对比的优缺点。KCF检测的理论基础，数学推理过程。

KCF，跟踪算法。物体识别准确率比YOLO低，但是识别速度比YOLO快很多。

<https://zhuanlan.zhihu.com/p/48249974>

# 六、双目测距

写5页左右。

测量车辆与摄像头的距离，并根据距离进行危险程度判断。

左右视图的匹配：

角点检测：<https://mp.csdn.net/editor/html/79119736>，简化处理，我们用YOLO识别到的矩形框中心点近似代替角点，计算左右图像的视差，进而求出距离

图像形态学处理：<https://blog.csdn.net/fengye2two/article/details/79188087>

选择红色的车辆，利用形态学处理可以分割出车辆的位置

# 七、实验过程及结果

写10页左右。

主要3个指标：

1、车辆识别准确率：每帧图像都使用YOLO检测的矩形框为基准；

2、识别速度（fps）：KCF与YOLO处理图像的不同间隔所对应的识别速度；

3、车辆距离测量误差率：卷尺测量的距离为基准，求得误差率。

# 八、避让路径规划

写2页左右。

现有路径避让算法，我们的场景哪种算法比较合适。

# 九、总结与展望

写2页左右。

总结全文，课题展望。