



# 毕业设计终期答辩

## 基于CMOS的视觉识别与三维重建应用研究

机械与能源工程学院 机械设计制造及其自动化（中意班）

蔡中天 1352050 指导老师：朱传敏

## 目录

01	论文绪论
02	研究方法
03	内容阐述
04	成果展示
05	未来发展

## • 课题背景

- ▶ 三维计算机视觉里有三个关键的任务：图像匹配，**相机定位**，还有**三维重建**。
- ▶ 核心目标：感知周围的三维环境结构／**感知物体的形状**／**感知物体和物体之间的相对位置**

吴毅红：三维计算机视觉的应用从传统工业领域逐渐走向生活

机器视觉

中国人工智能学会

吴毅红

2017-06-07 20:00

（吴毅红：中科院自动化所模式识别国家重点实验室研究员）

## • 课题背景

- ▶ 机器视觉技术已经进入人们的生活。



- ▶ 典型的例子：特斯拉电动轿跑车利用视觉识别探测路况，其中就用到了三维重建技术。



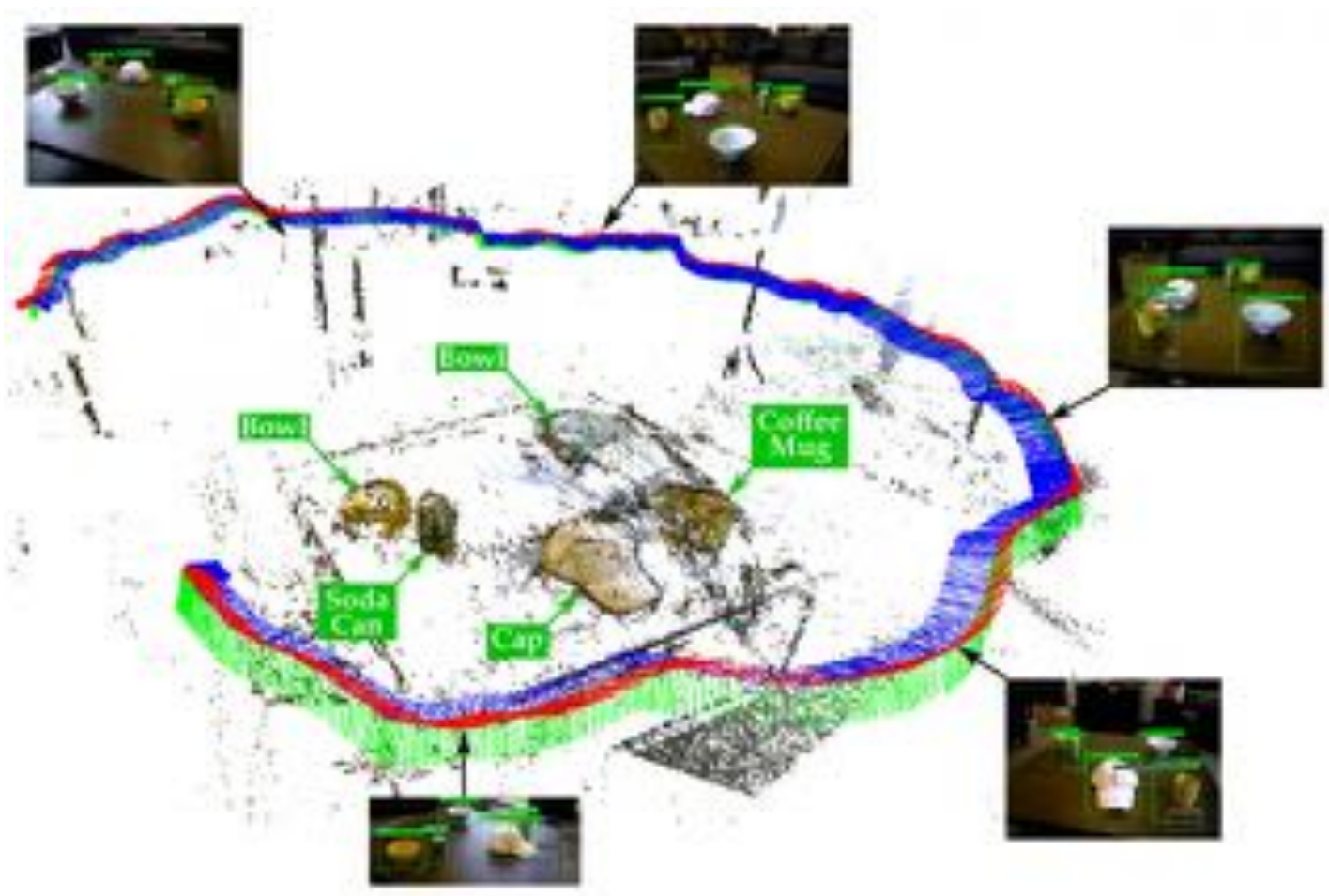
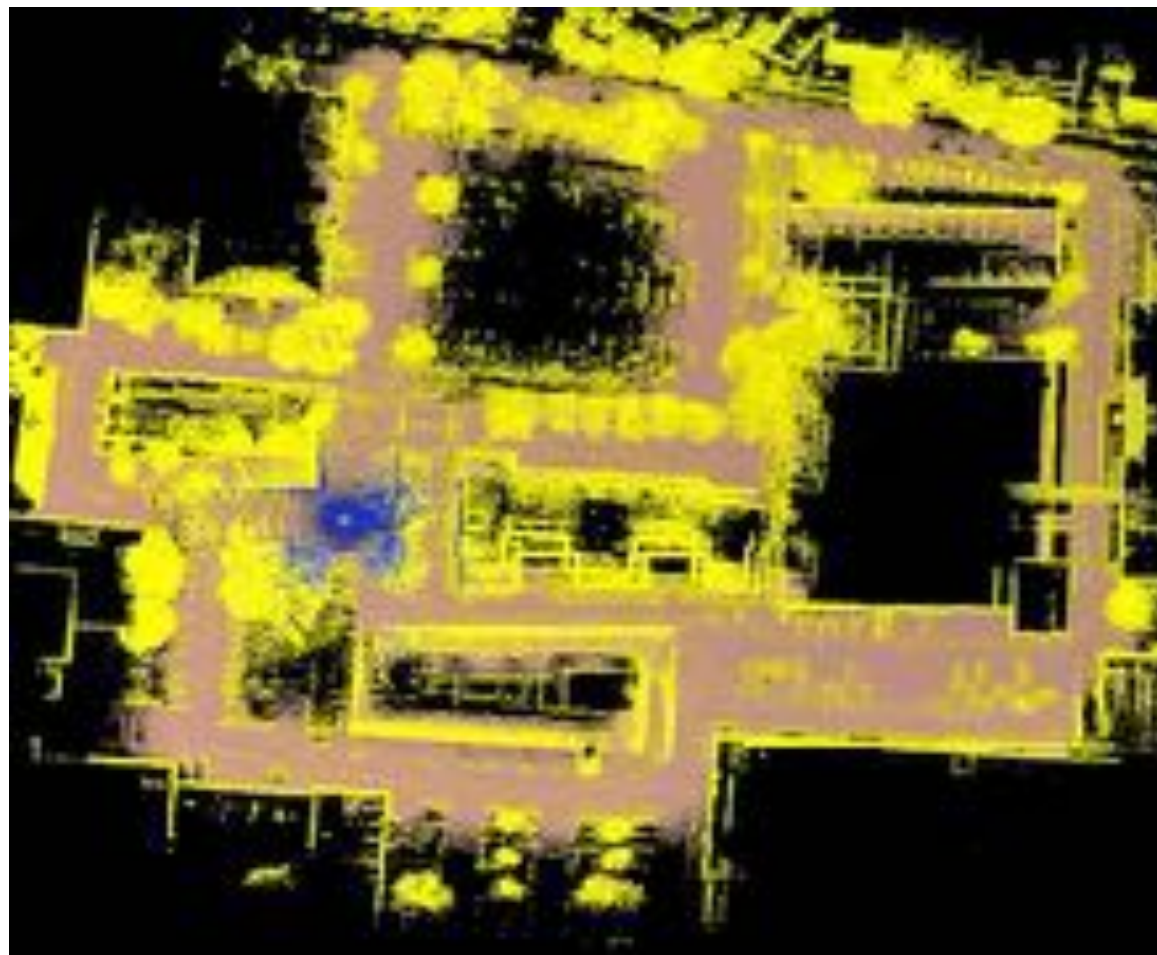
- 课题背景



- ▶ 但是，这项技术并不安全。三维重建技术需要更多的研究。

## • 课题背景

- ▶ 三维计算机视觉和重建技术的热点研究领域及应用场景例举。
- SLAM: Simultaneous localization and mapping （实时定位与地图构建）



类似：地图测绘、工业仿真、三维扫描等等。



## • 课题背景

- ▶ 三维计算机视觉和重建技术的热点研究领域及应用场景例举。
- 工业视觉检测：在自动化生产过程中，机器视觉系统广泛用于**工况监视**、**成品检验**和**质量控制**等领域。

## • 医疗影像诊断

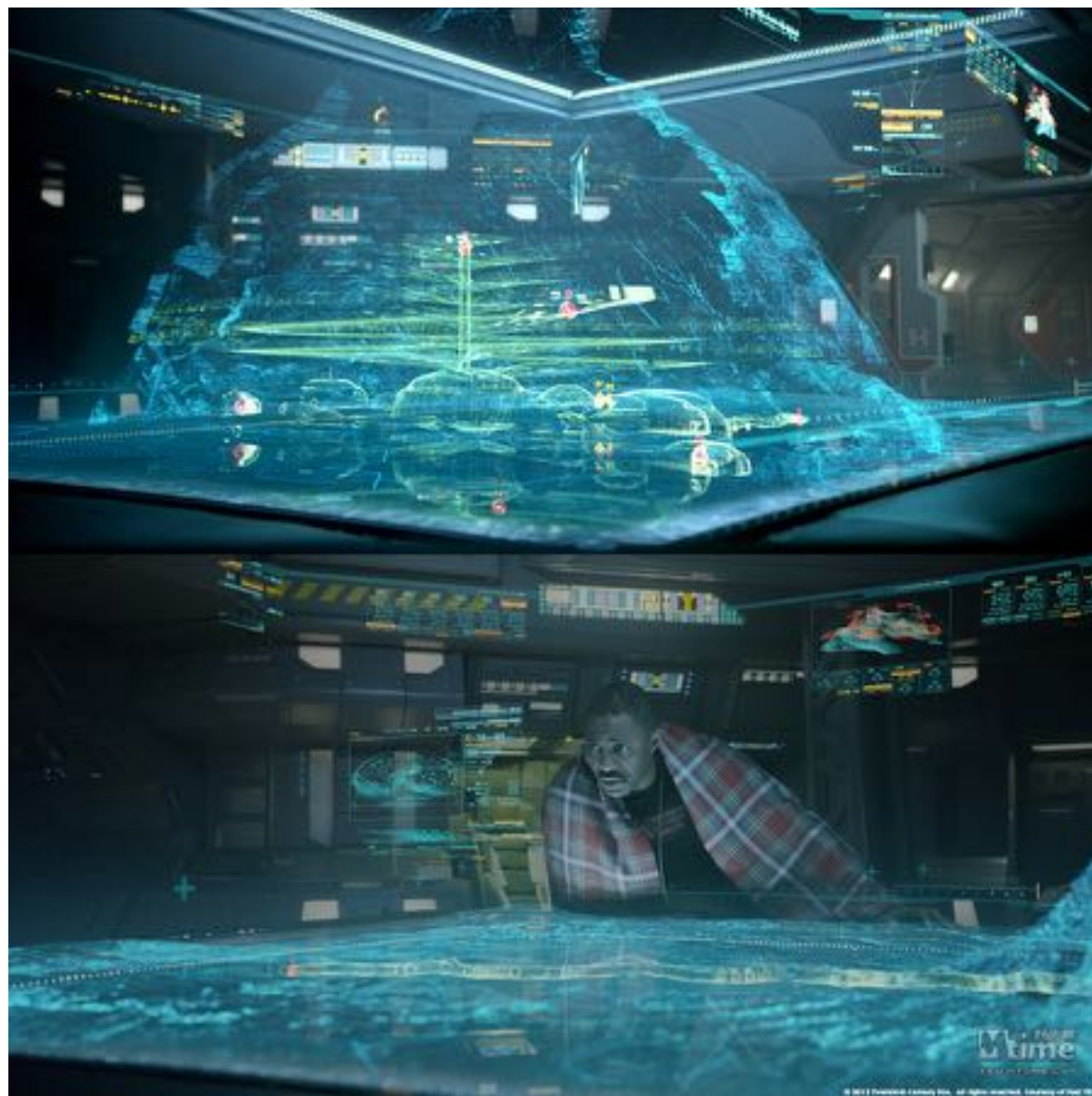
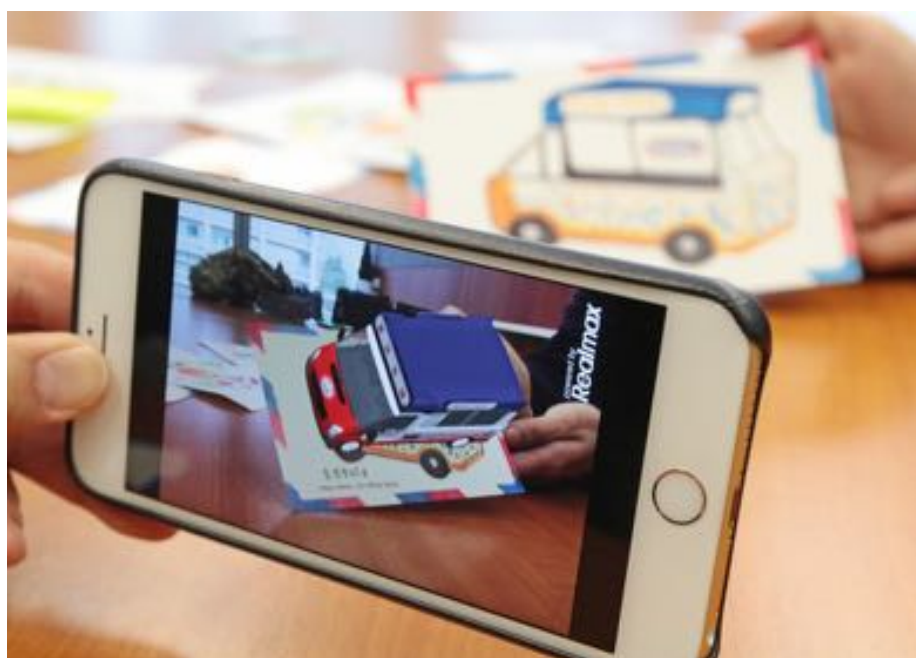




## • 课题背景

- ▶ 三维计算机视觉和重建技术的热点研究领域及应用场景例举。

- AR、VR……

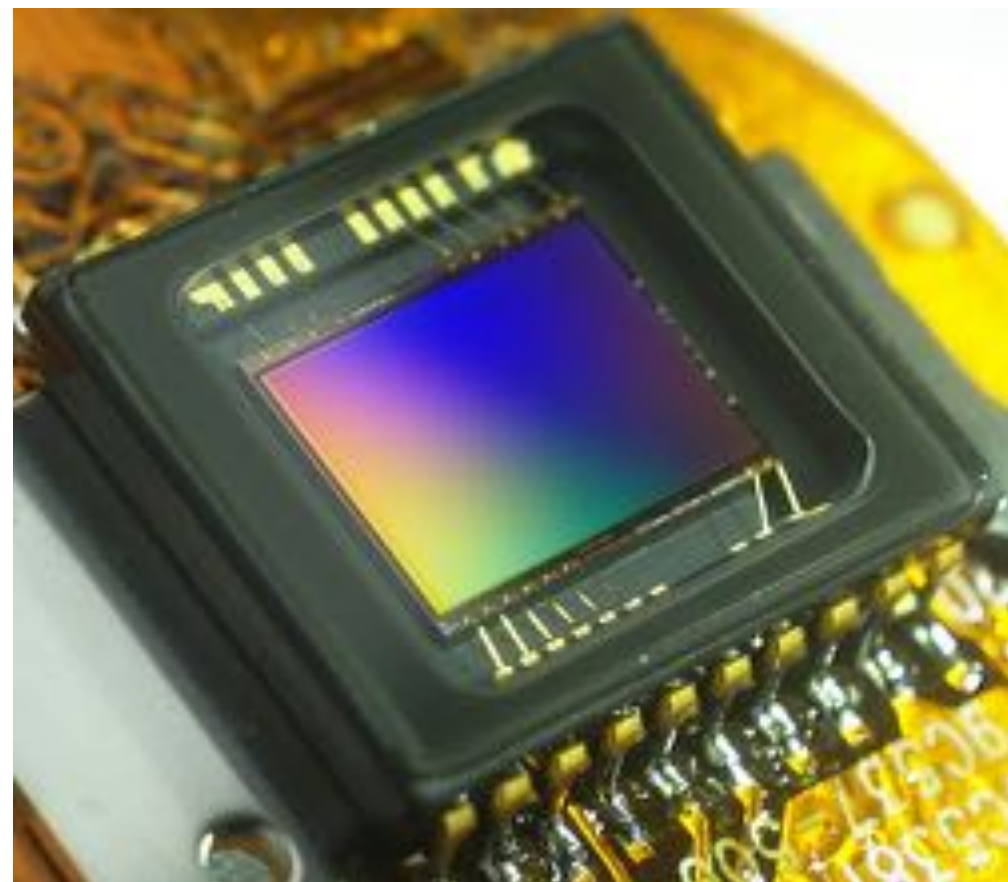




- 课题选题

基于 **CMOS** 的视觉识别与三维重建应用研究

- ▶ WHY CMOS?



- 便宜；常见

• 课题选题

▶ WHY CMOS?

CCD的数据处理速度有先天瓶颈！

表 1.1 CMOS 与 CCD 图像传感器技术细节和性能比较<sup>[12]</sup>

指标 / 项目	CCD	CMOS
像素输出信号	电荷包	电压
传感器输出信号	电压（模拟）	比特（数字）
相机输出信号	比特（数字）	比特（数字）
传感器复杂度	低	高
应用复杂度	高	低
研发成本	高	低
灵敏度	高	中
系统噪声	低	中
响应速度	中	中上
动态范围	大	中
响应均匀性	好	中下
信号处理速度	中上	高
随机窗口读取能力	有限	任意局部画面
电源电压	多电源高电压	单电源低电压

• 重要项对比

CCD Wins!

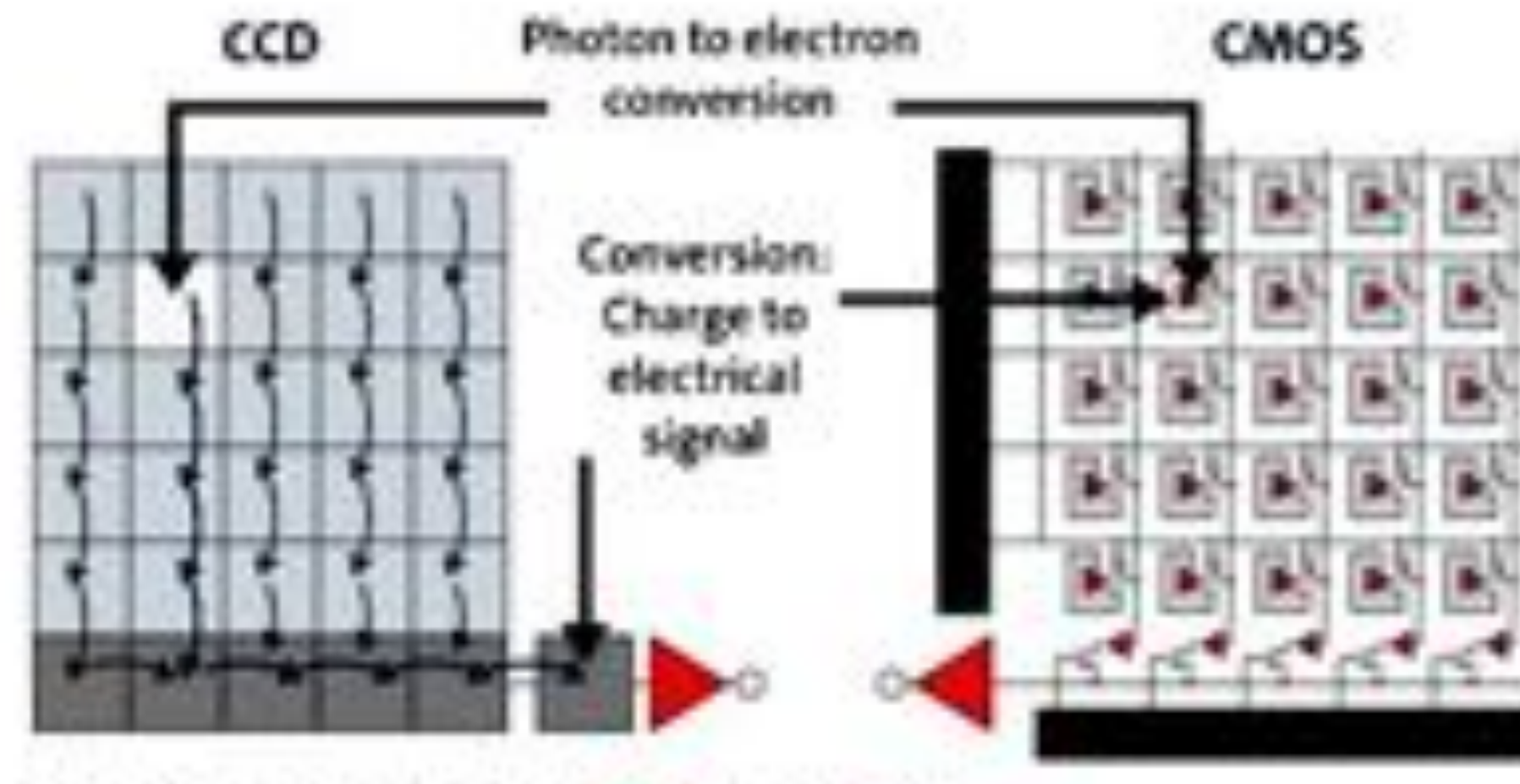
CMOS Wins!





- 课题选题

- ▶ WHY CMOS?



1. 单列一起输出 (先)

2. 批量数模转换 (后)

1. 单个数模转换

2. 自上而下自左而右输出

(两步同时进行)

## • 课题选题

### ▸ WHY CMOS?

### ▸ 随着工艺和图像处理算法的发展，CCD的优势逐渐荡然无存。



### ▸ CCD目前仅生存于部分工业领域和电视行业。



## • 课题选题

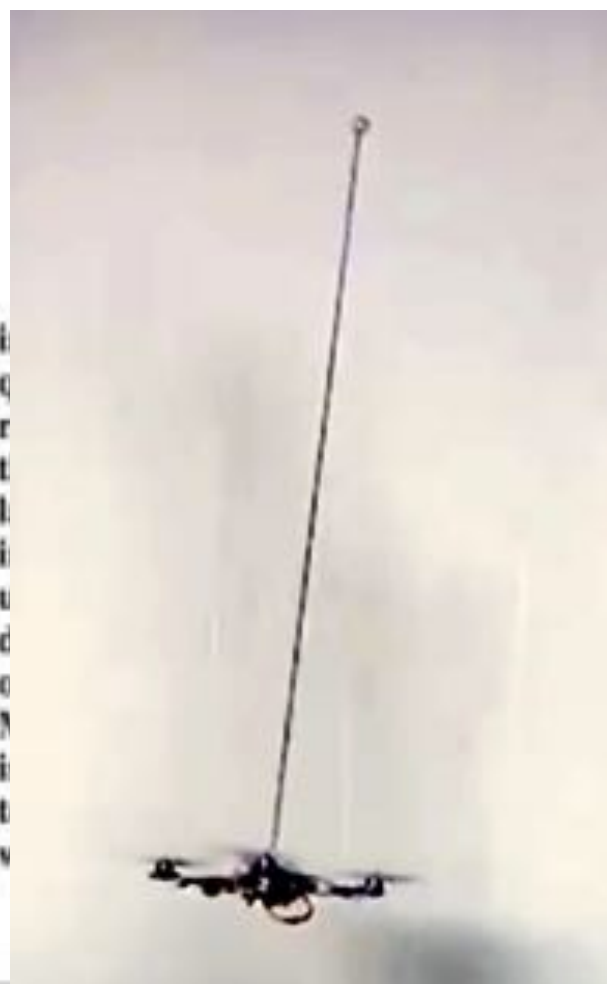
### ▸ WHY CMOS?



### ▸ 在民用领域，CMOS传感器已是绝对霸主。

- 课题选题 —> 来源

- 实现倒立摆的自平衡 A Flying Inv
- 实现倒立摆的抛接！！



Markus H

control problem  
pendulum on  
tic and dynamic  
to find nominal  
ular trajec  
ninal trajectories  
e dynamics is in  
lows for the time-  
he balancing per  
the ETH Zurich  
at potential for t  
ing learning met  
ing systematic er

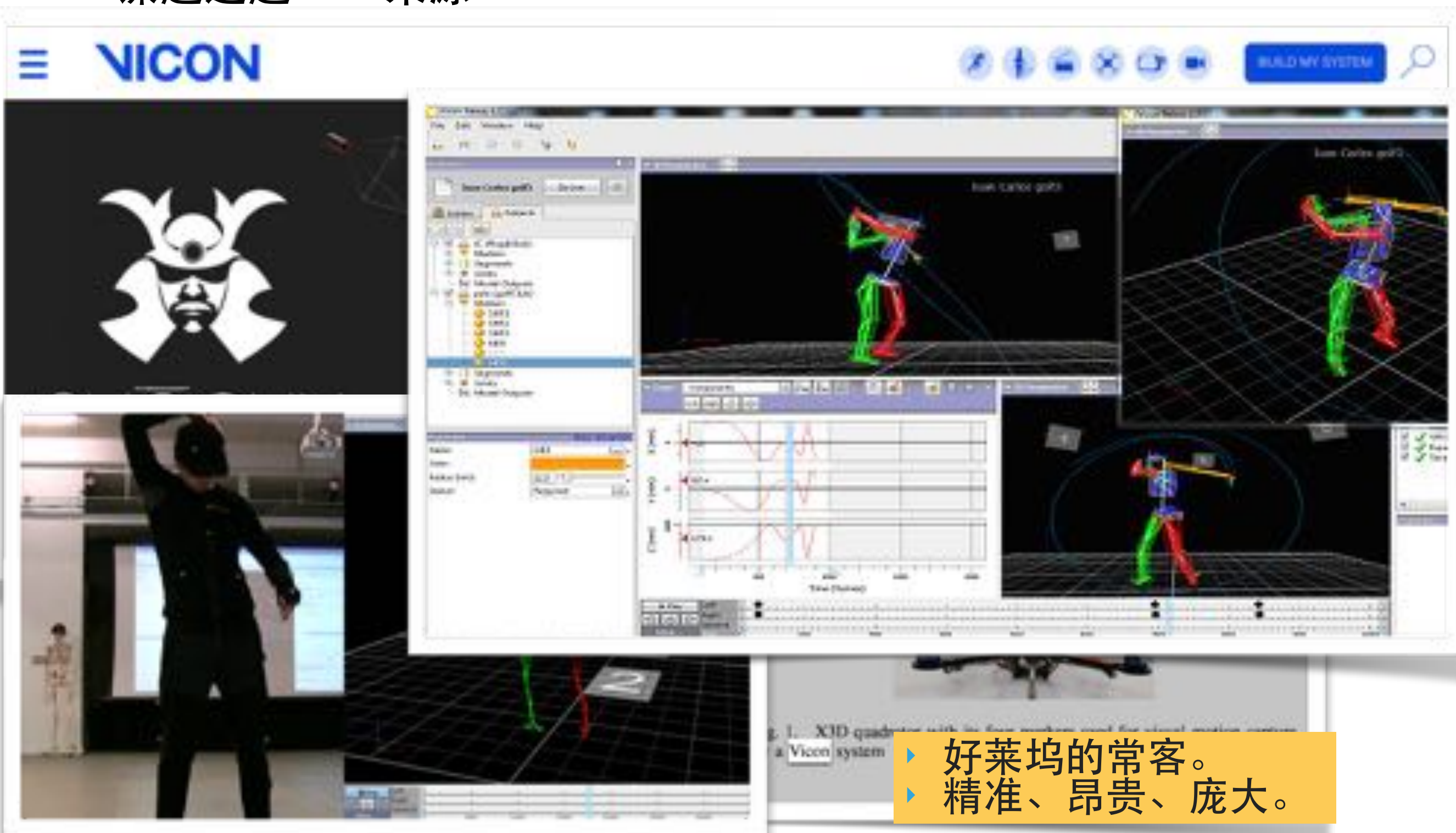


(十分先进的控制算法)

- <http://flyingmachinearena.org>



• 课题选题 —> 来源



## • 课题选题

统目前广泛应用于电影特效行业和人体运动分析等领域。而我们希望做到的，是在减少成本的基本要求下，增加源项目的应用范围：即不再需要在配套有专业 Vicon 系统的实验室环境中，而是在任何地点，不论室内室外，都可完成都可让飞行器完成倒立摆自平衡功能。这即为本文所做工作的最初出发点。

### ▶ 最终确定选题

**基于 CMOS 的视觉识别与三维重建应用研究**  
**Research into The Application of CMOS-Based Vision Recognition and 3-Dimensional Reconstruction**



2	理论与算法研究.....	10
2.1	图像预处理.....	10
2.1.1	灰度图像转换.....	10
2.1.2	二值图像.....	11
2.1.3	图像平滑.....	11
2.2	圆形物体识别与提取.....	14
2.2.1	圆形度法.....	14
2.2.2	霍夫变换法.....	14
2.3	摄像机数学模型及三维坐标线性匹配.....	16
2.4	.....	16
2.5	.....	18
3	代码实.....	21
3.1	.....	21
3.2	核心功能的代码实现.....	23
3.3	软件封装.....	26
4	测试结果与讨论.....	30
4.1	测试方案介绍.....	30
4.2	具体测试步骤与结果.....	30
4.2.1	实验设计.....	30
4.2.2	实验结果.....	31
4.3	结果分析.....	34
4.4	系统可优化性讨论和改进方案.....	34

理论研究



工程实践

实际过程：根据所查找的文献，一边学习理论，一边实现代码。根据需求添加功能。

优点：通过及时的实践加深对理论知识的理解。

## • 图像预处理 —> 灰度&二值(黑白)图像

- A. 保留亮度信息
- B. 保留图像细节
- C. 每个像素8个字节



▶ 原始图像

- A. 保留少数特殊信息
- B. 其他信息基本流失
- C. 每个像素1个字节



▶ 灰度图像 (Grayscale Image)



▶ 二值图像 (Binary Image)



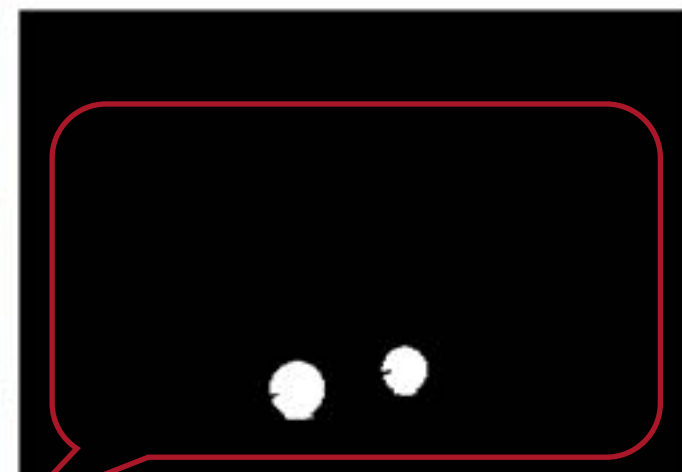
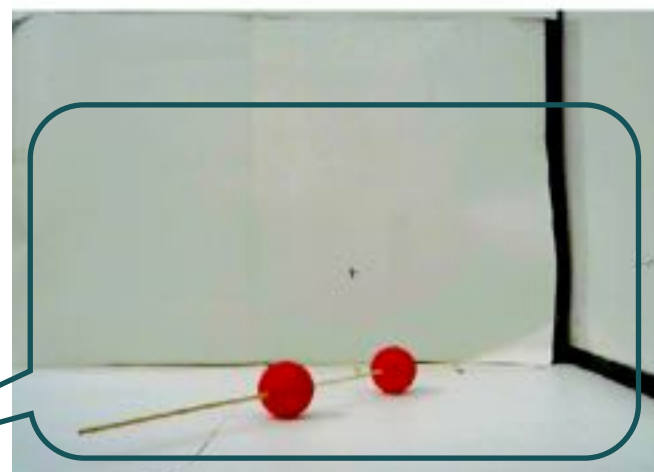
## • 图像预处理 —> 二值图像(黑白)

环境噪声，容易在之后环节造成误识别

### ▶ 随机环境



### ▶ 实验环境



匀质白色背景，特殊颜色目标物

区分效果效果接近完美

```
function [vidOriginal]=CallCamera(numCam,handleAx)
vidOriginal=videoinput('winvideo',numCam,'YUY2_960x720');
camRes=get(vidOriginal,'videoResolution');
nBands=get(vidOriginal,'NumberOfBands');
axes(handleAx);
hImage=imshow(zeros(camRes(2),camRes(1),nBands));
preview(vidOriginal,hImage);
end
function [centers, radii, frame]=ImgPreprocessing(vidOriginal)
frame=getsnapshot(vidOriginal);
I1=frame(:,:,3);
frame=imbinarize(I1,'global');
[centers, radii] = imfindcircles(frame,[25 75]);
end
function showBiIm(handleAx,centers,radii,frame)
axes(handleAx);
imshow(frame);
viscircles(centers, radii,'EdgeColor','b');
end
```



## • 图像预处理 —> 平滑算法

## ▶ 计算机视觉中的重要环节



图 2.1 一张带有典型“椒盐噪声”的图像

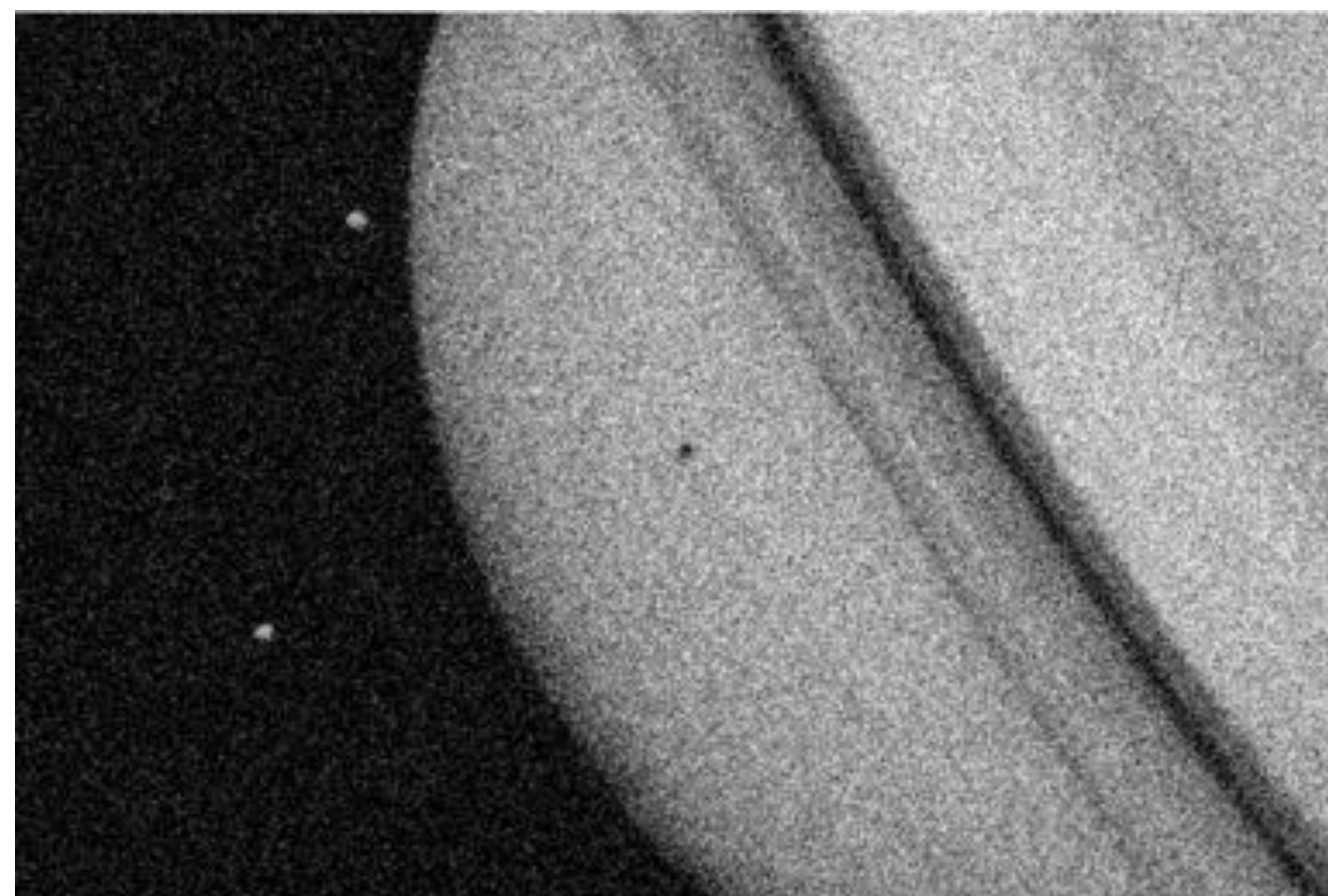


图 2.2 一张带有典型“高斯噪声”的图像

### ▶ 概率密度函数

$$p(z) = \begin{cases} P_a, & z = a \\ P_b, & z = b \\ 0, & \text{else} \end{cases}$$

$$p(z) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left\{ -(z - \mu)^2 / 2\sigma^2 \right\}$$

其中 $z$ 表示图像元素的灰度值； $\mu$ 表示 $z$ 的期望； $\sigma$ 表示 $z$ 的标准差。



## • 图像预处理 —> 平滑算法

## • 计算机视觉中的重要环节



图 2.1 一张带有典型“椒盐噪声”的图像

• 解决方案：  
滤波器

$$y_{ij} = \text{Med}[f(m, n)]$$

均值滤波

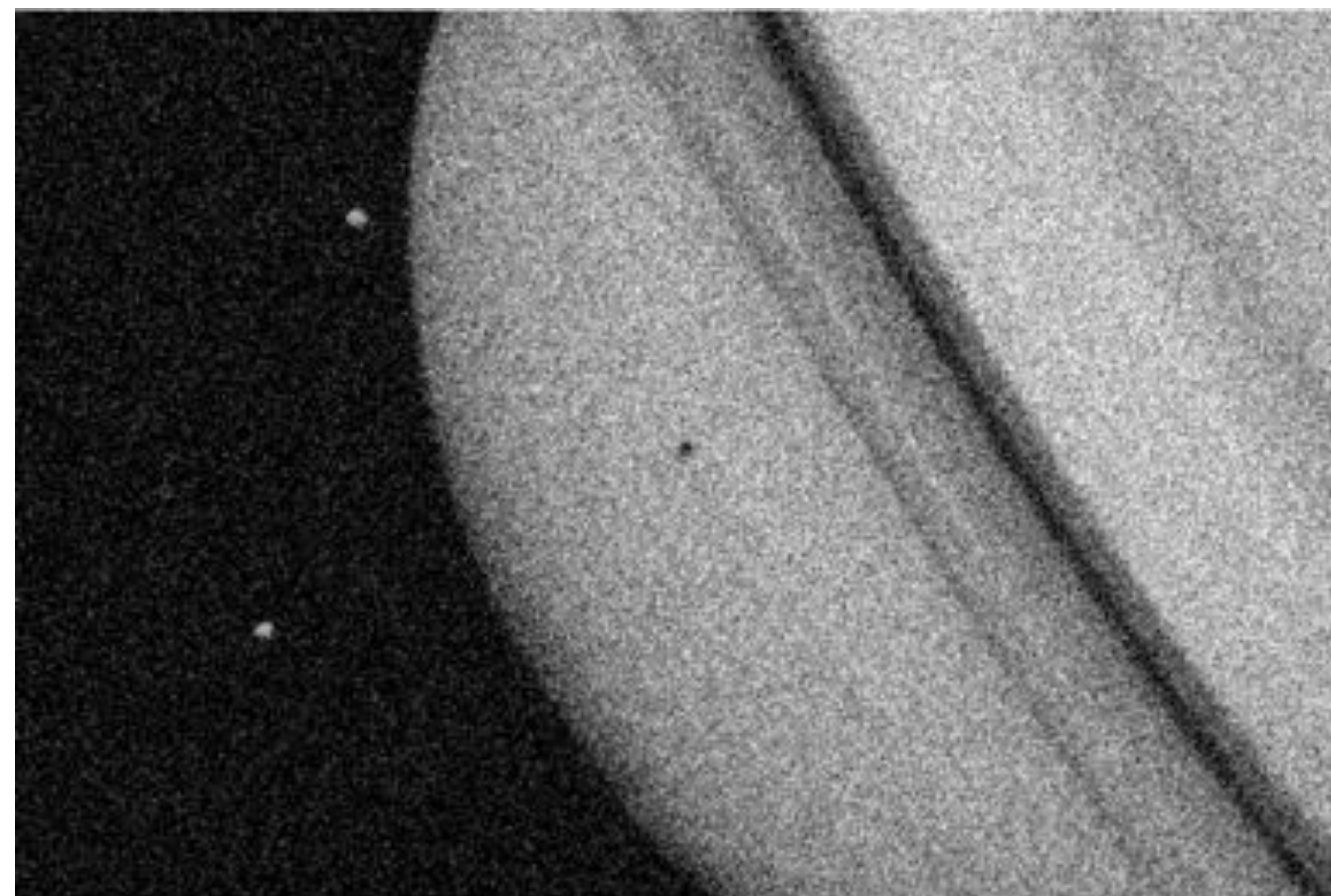


图 2.2 一张带有典型“高斯噪声”的图像

$$G(u, v) = H(u, v)F(u, v)$$

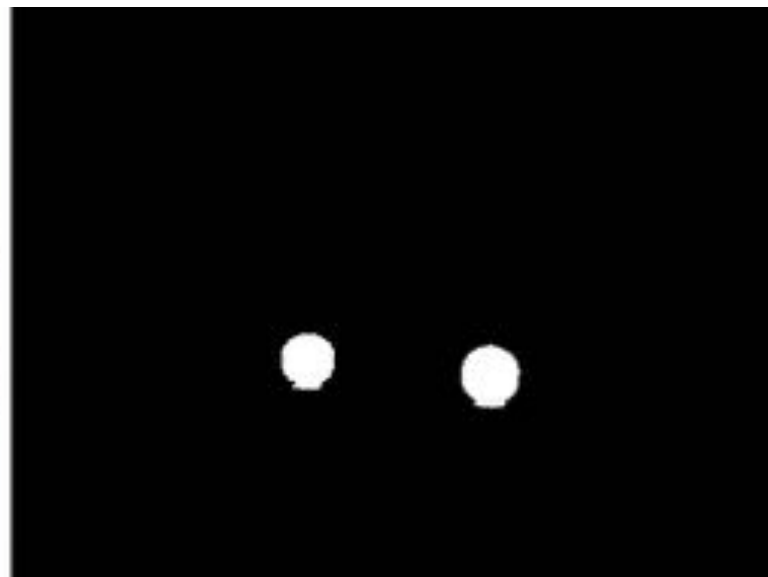
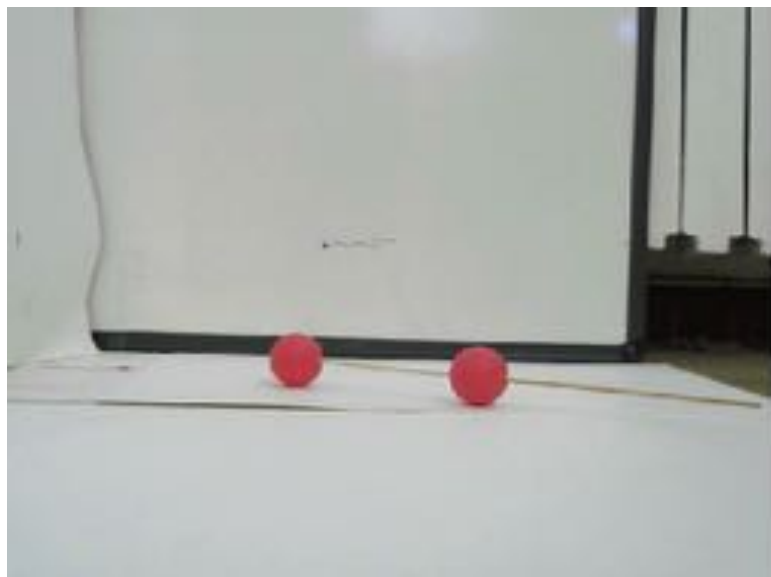
$$H(u, v) = \begin{cases} 1, & D(u, v) \leq D_0 \\ 0, & D(u, v) > D_0 \end{cases}$$

高斯滤波

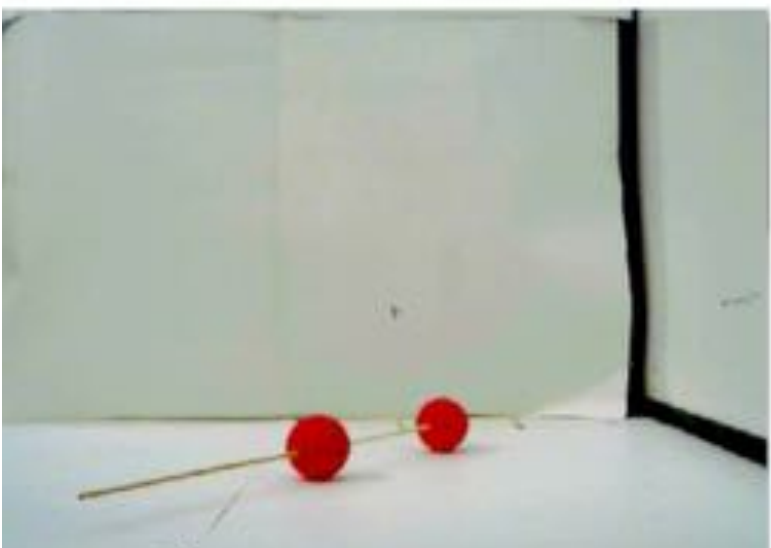
## • 图像预处理 —> 平滑算法

▶ 计算机视觉中的重要环节

### ▶ 不添加平滑



### ▶ 添加平滑和色彩增强



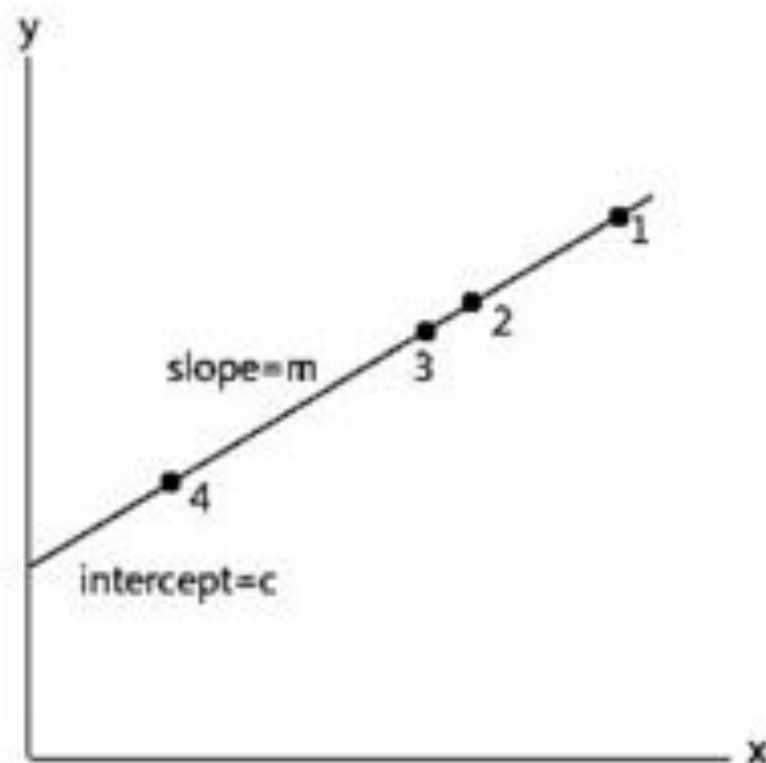
▶ 二值化效果已经平滑且在实验环境中没有多余噪声干扰，在本课题的实验中可以暂时不使用平滑算法，以减轻计算负担。



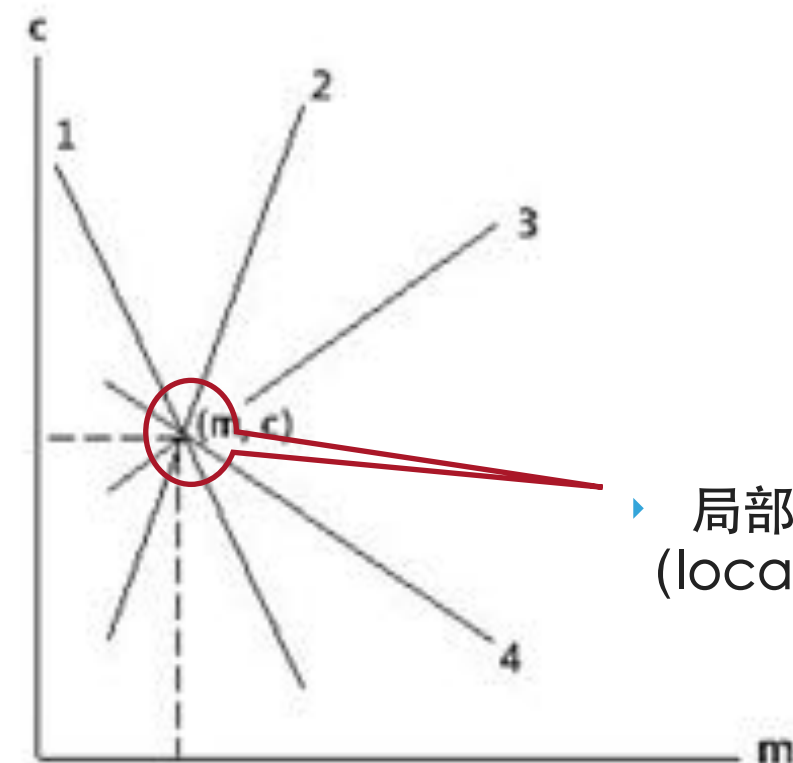
- 类圆识别 → Hough 变换法

- ▶ 示例：直线

- ▶ 几何坐标



- ▶ 参数坐标

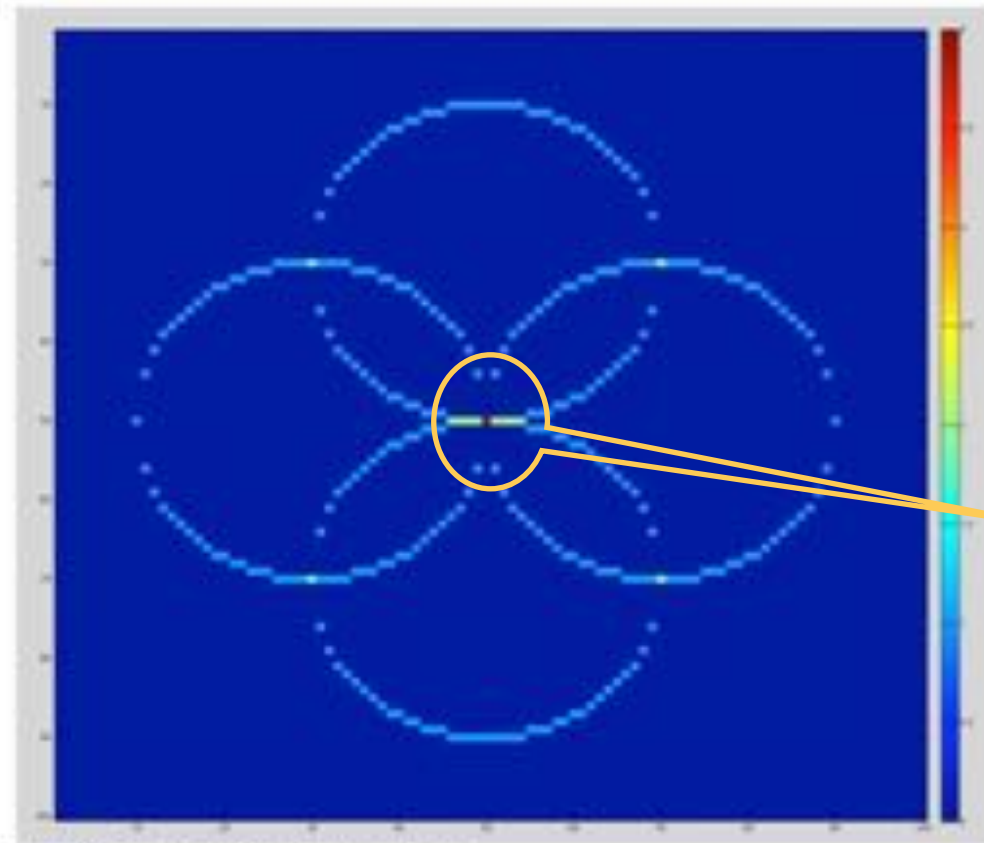
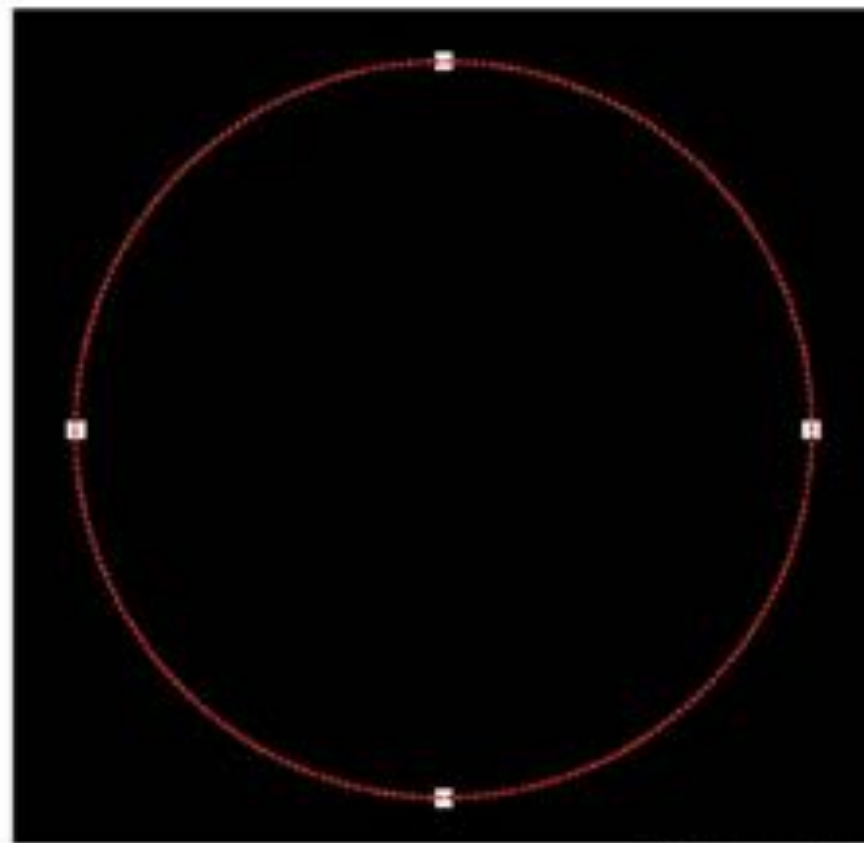


- ▶ 局部最大值  
(local maxima)

图 2.3 对直线  $y = mx + c$  的霍夫变换（左：xy域；右：mc域）

## • 类圆识别 -> Hough 变换法

### ▶ 半径已知的圆



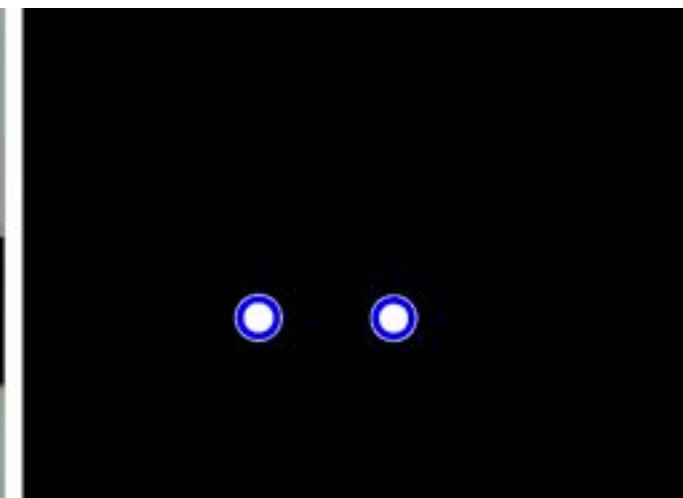
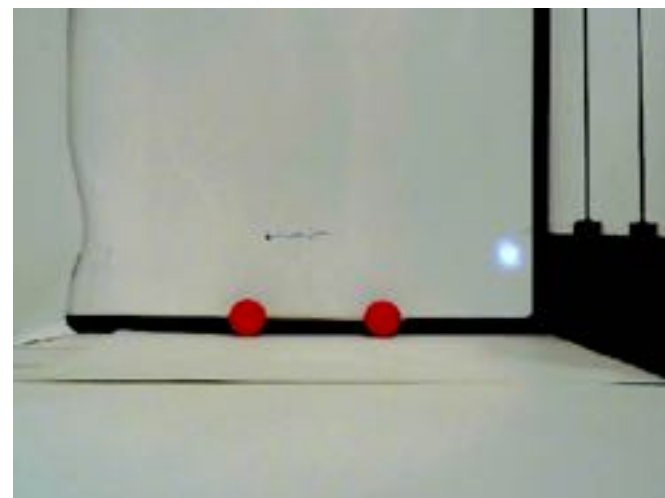
▶ 局部最大值  
(local maxima)

图 2.4 对固定半径圆形的霍夫变换

## • 类圆识别 → Hough 变换法

### ▶ 半径未知的圆

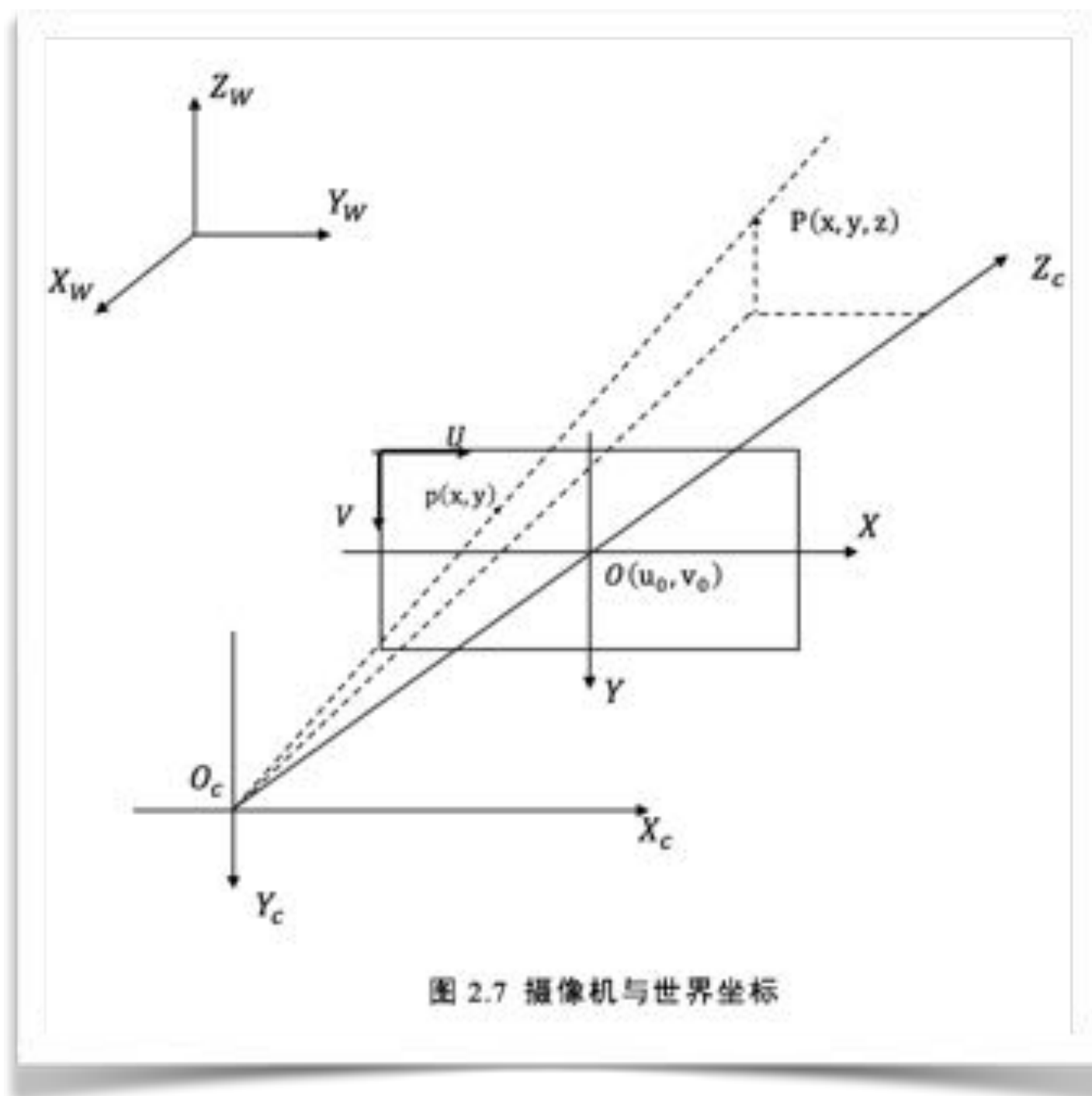
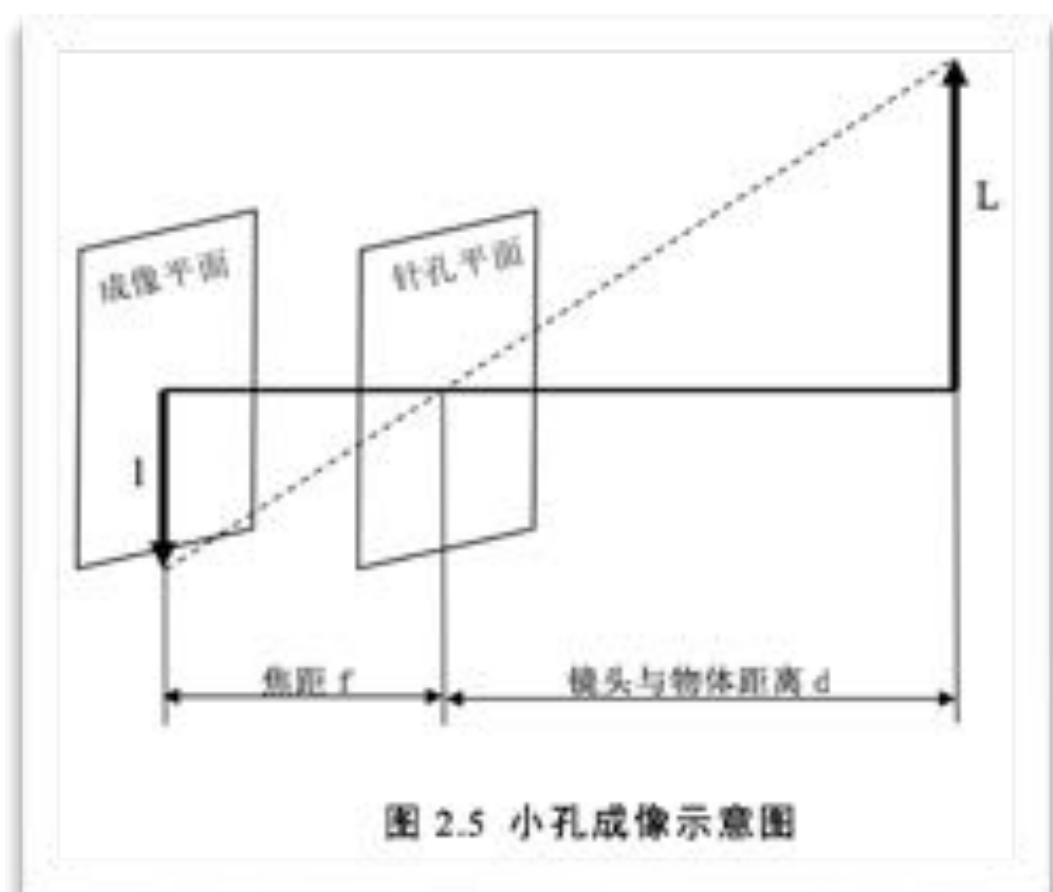
```
function [centers, radii, frame] = ImgPreprocessing(vidOriginal)
    frame = getsnapshot(vidOriginal);
    I1 = frame(:, :, 3);
    frame = imbinarize(I1, 'global');
    [centers, radii] = imfindcircles(frame, [25 75]);
end
function showBiIm(handleAx, centers, radii, frame)
    axes(handleAx);
    imshow(frame);
    viscircles(centers, radii, 'EdgeColor', 'b');
end
```



位置参数有三个： $a$ 、 $b$ 和 $r$ 。但是在计算机处理中，在参数坐标的维度超过两维后，在其中找局部最大值的过程的时间消耗和所需存储空间被证明会急剧增加。因此在实际处理中，一般对于未知半径圆形的识别也不使用三维参数空间，而是增加一步半径估计，转化为一个二维参数空间问题，这种方法现在被广泛使用，在 Matlab 中的 imfindcircles 函数即使用了 降维的霍夫变换 来识别定位圆形特征。



## • 三维匹配



• 三维匹配

本质

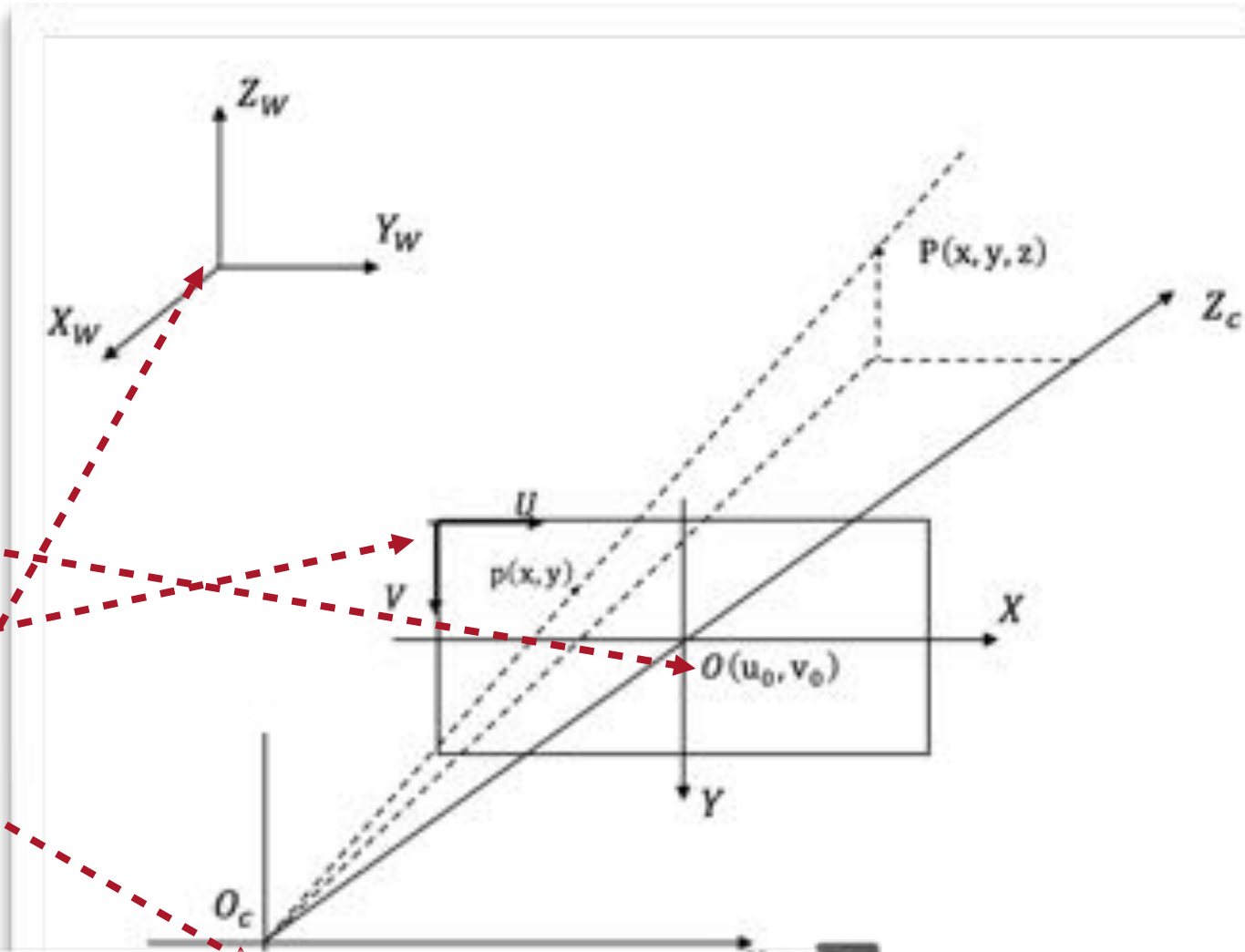
寻找它们之间的关系

(1) 图像的物理坐标系 ( $OXY$ )

(2) 图像的像素坐标系 ( $O_1UV$ )

(3) 摄像机坐标系 ( $O_cX_cY_cZ_c$ )

(4) 世界坐标系 ( $O_wX_wY_wZ_w$ )



由此得出摄像机像素坐标与世界坐标之间的关系：

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{dx} & 0 & u_0 \\ 0 & \frac{1}{dy} & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 & 0 \\ 0 & f_y & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \quad (2.17)$$

## • 三维匹配

- ▶ 一个目标物的情况
- ▶ 两台相机 → 两条三维直线
- ▶ 目标：找到他们的交点P

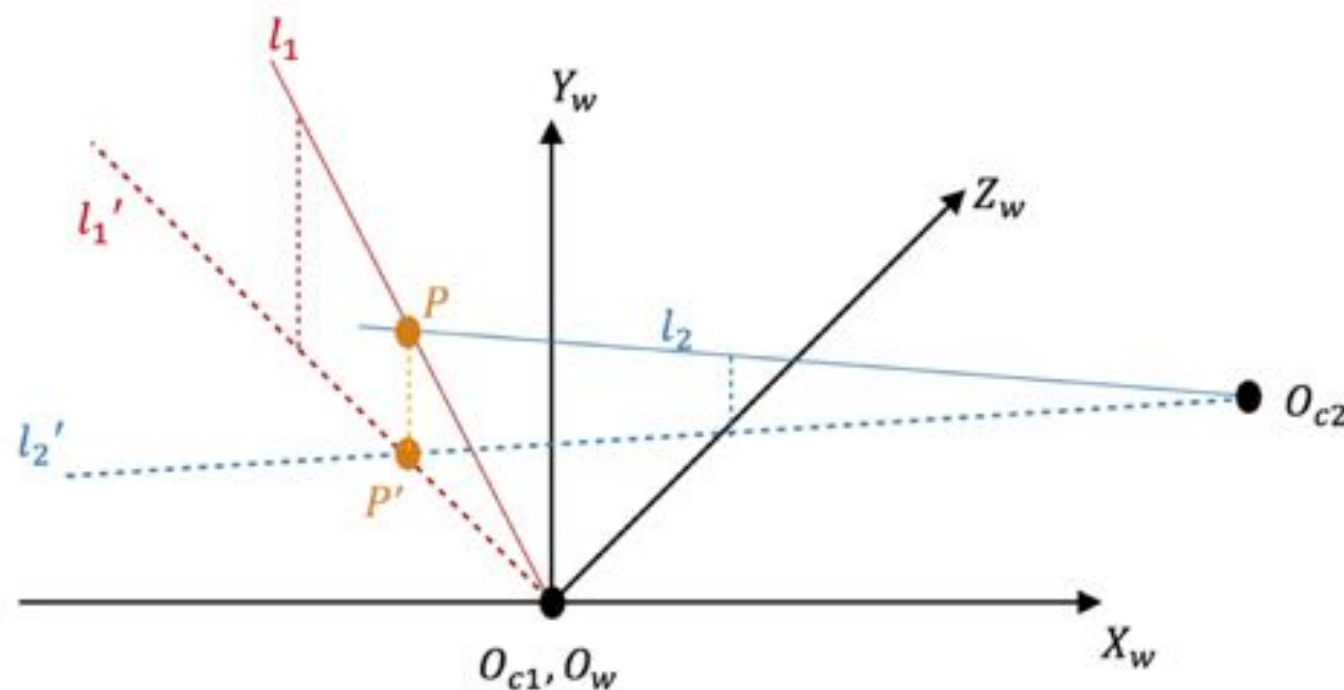


图 3.7 三维匹配关系示意图

```

function [xmin,ymin,zmin]=findIntersection(centers1,centers2,p,f)
% line from camera 1
positionX1=centers1(1)-480;positionY1=360-centers1(2);
point1_0=[0,0,0];point1_1=[positionX1*p,positionY1*p,f];
n1=point1_1 - point1_0;
% line from camera 2
positionZ2=centers2(1)-480;positionY2=360-centers2(2);
point2_0=[500,0,500];point2_1=[500-f,positionY2*p,500+positionZ2*p];
n2=point2_1 - point2_0;
% 求两线交点
project1=[n1(1),n1(3)];
project2=[n2(1),n2(3)];
xzmin=( [project1;project2]*[0 -1;1 0])\[0;500*(project2(2)-project2(1))];
xmin=xzmin(1);zmin=xzmin(2);
ymin=zmin*positionY1*p/f;
end

```



## • 三维匹配 —> 调整摄像头位姿

- ▶ 通过目标物的像素坐标数据进行调零。

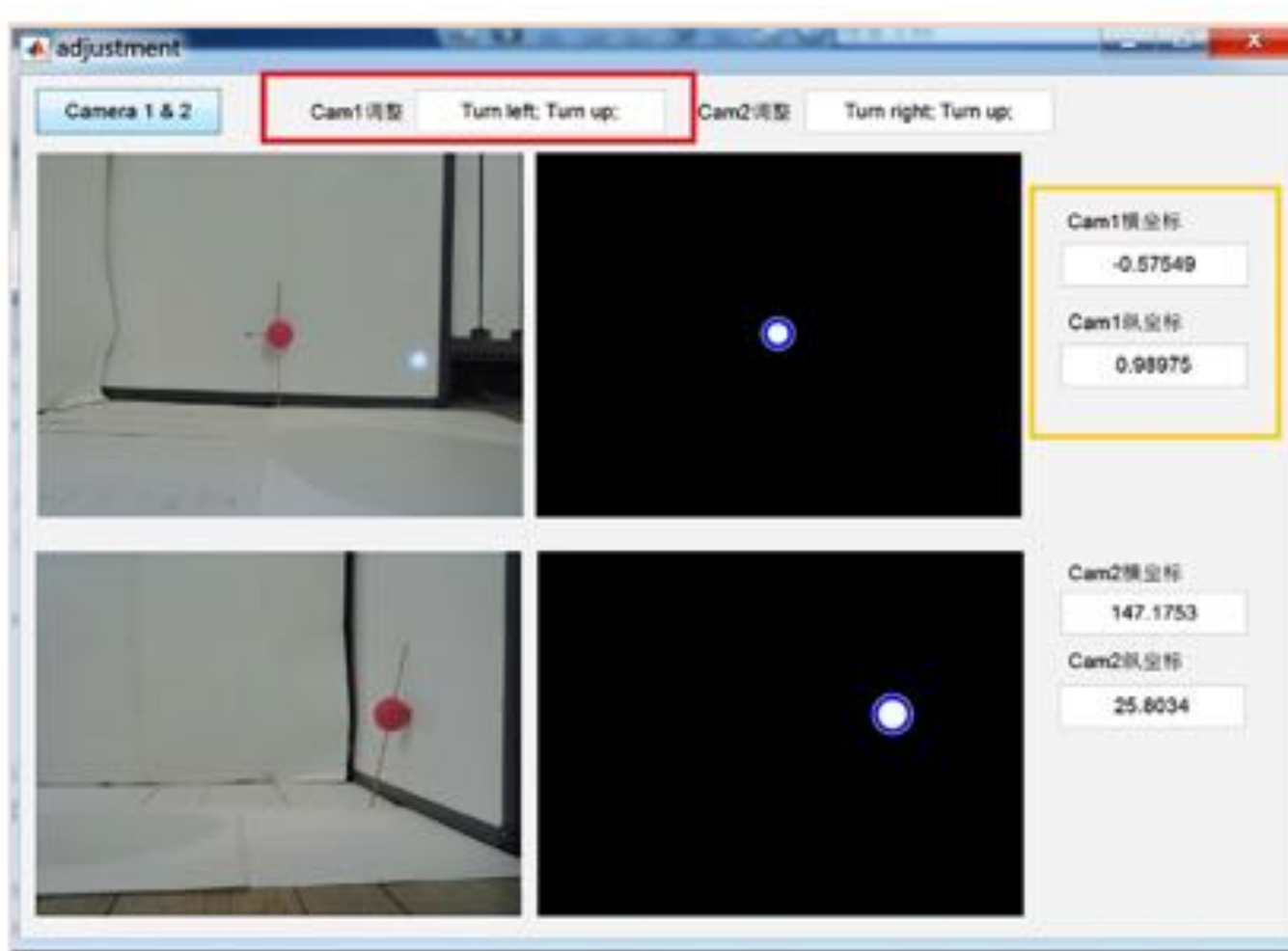


图 3.10 俯仰角和平偏角的调整

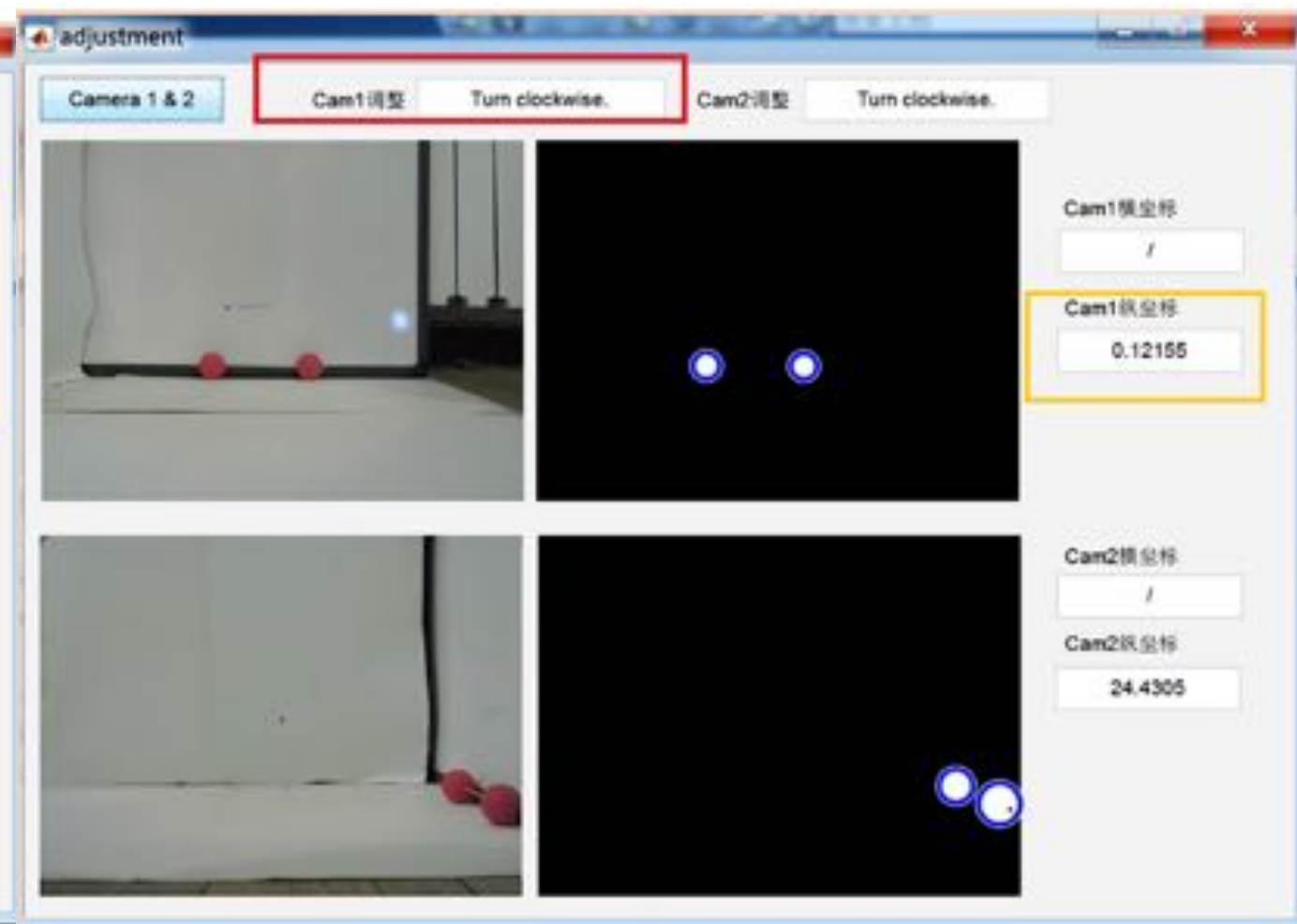


图 3.11 滚转角的调整

## • 实验设计

### 怎么得到目标球的真实位置?

- ▶ 测试目的：系统的准确性和时效性。
- ▶ 针对目标物真实世界坐标值的不可获得性，专门设计实验。

### ▶ 实验环境

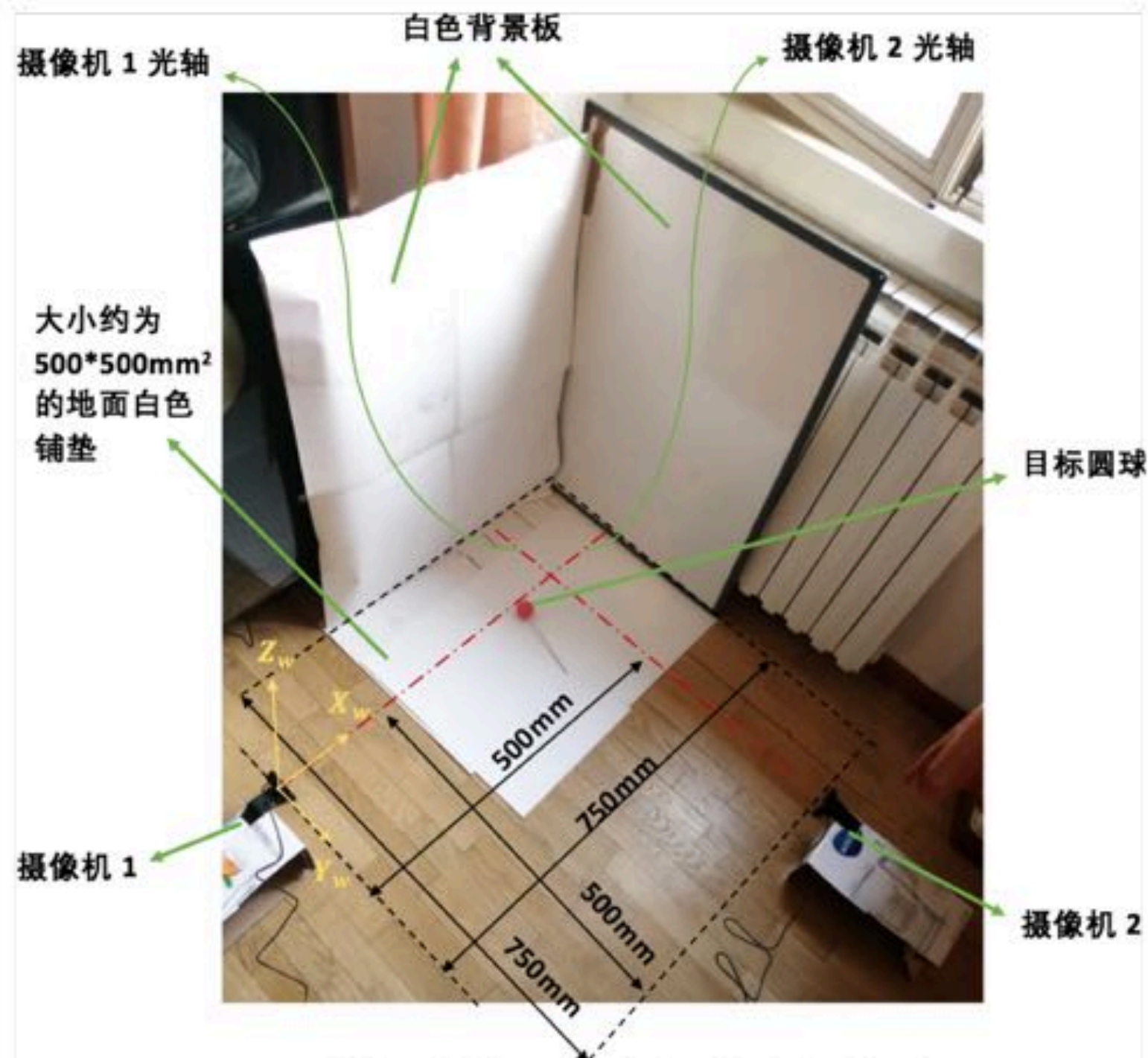
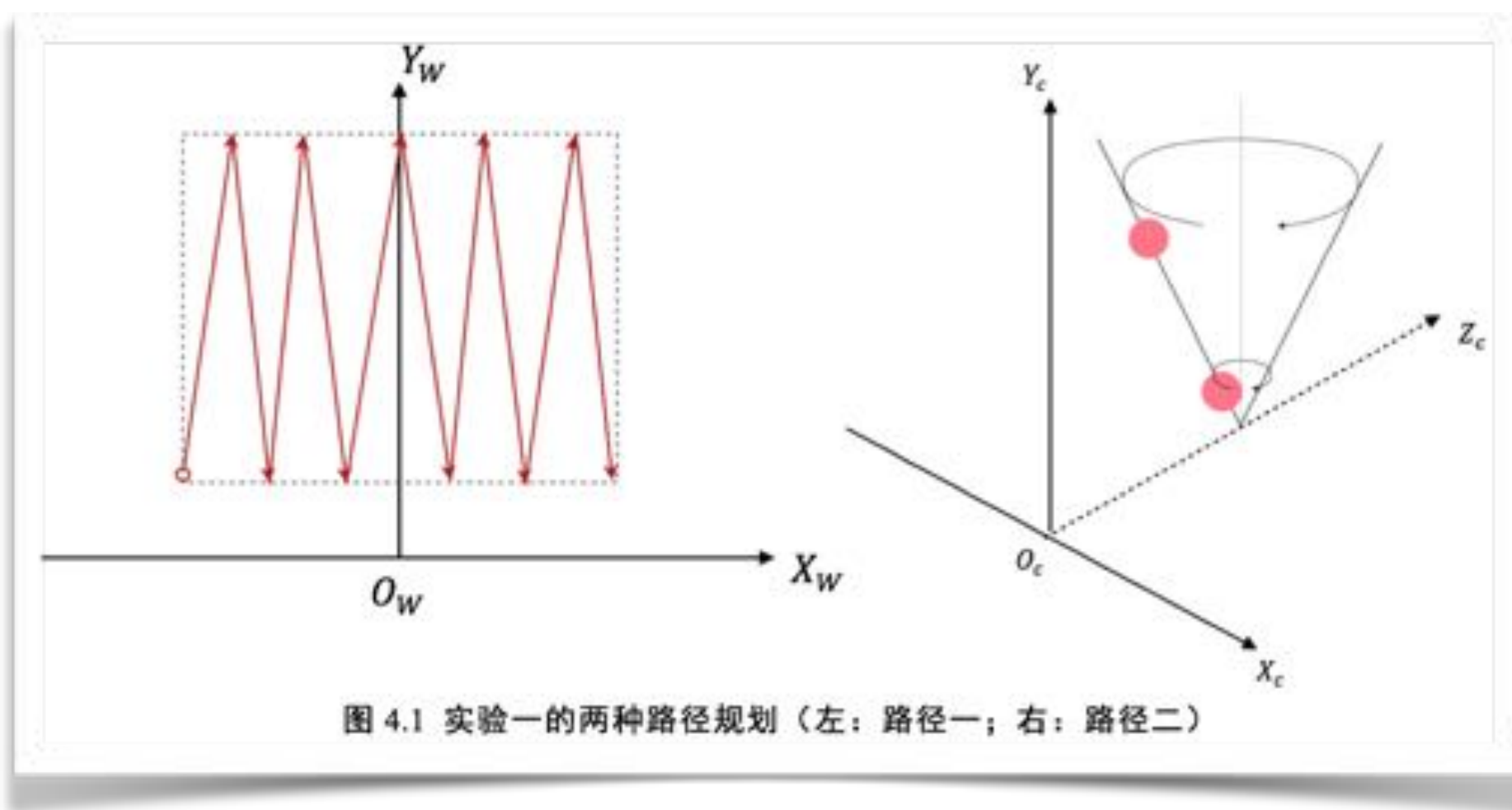


图 3.2 实验场地搭建示意、部分尺寸和世界坐标系位置

## • 实验设计 —> 实验1:

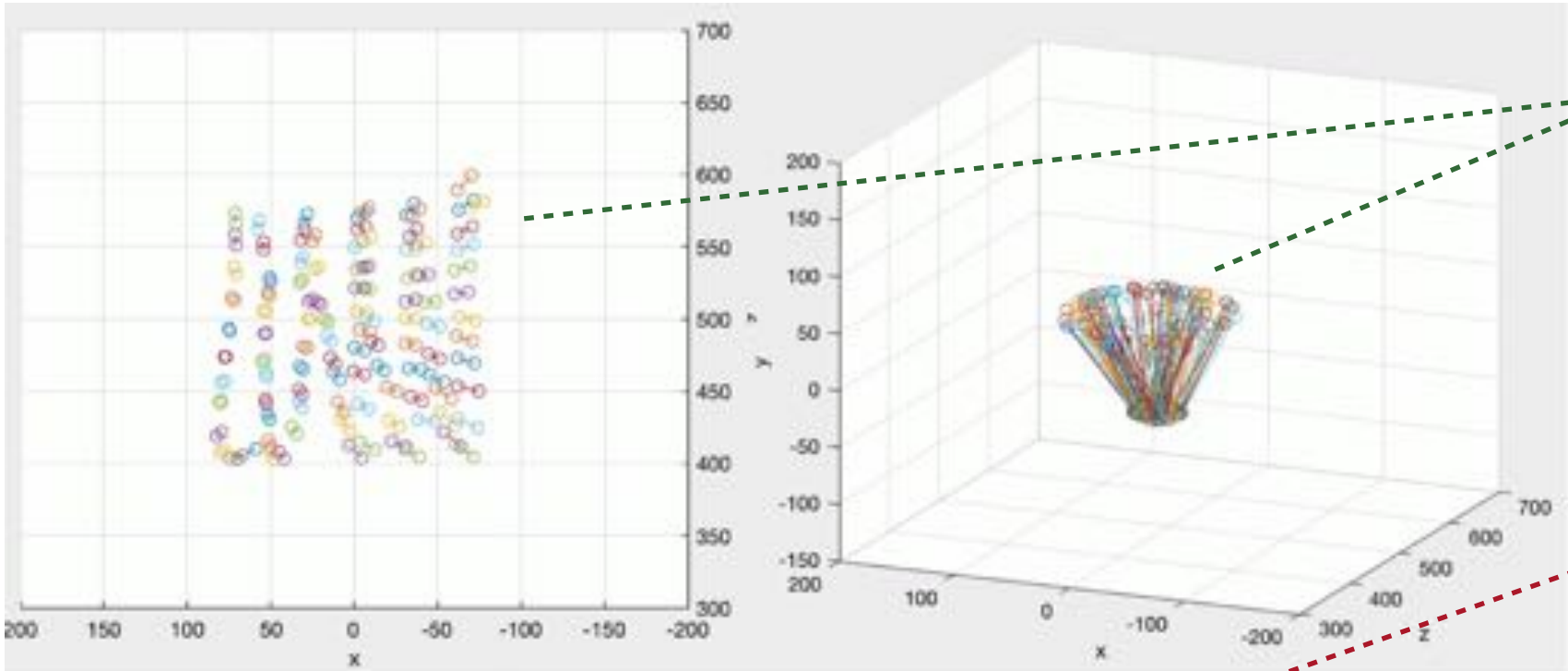
- ▶ 已知：两球之间的真实距离为100mm；
- ▶ 测量并统计：所测得的两球间距。
- ▶ 路径一：两球的轨迹到达拍摄画面的极限
- ▶ 路径二：目标基本保持在画面中央位置





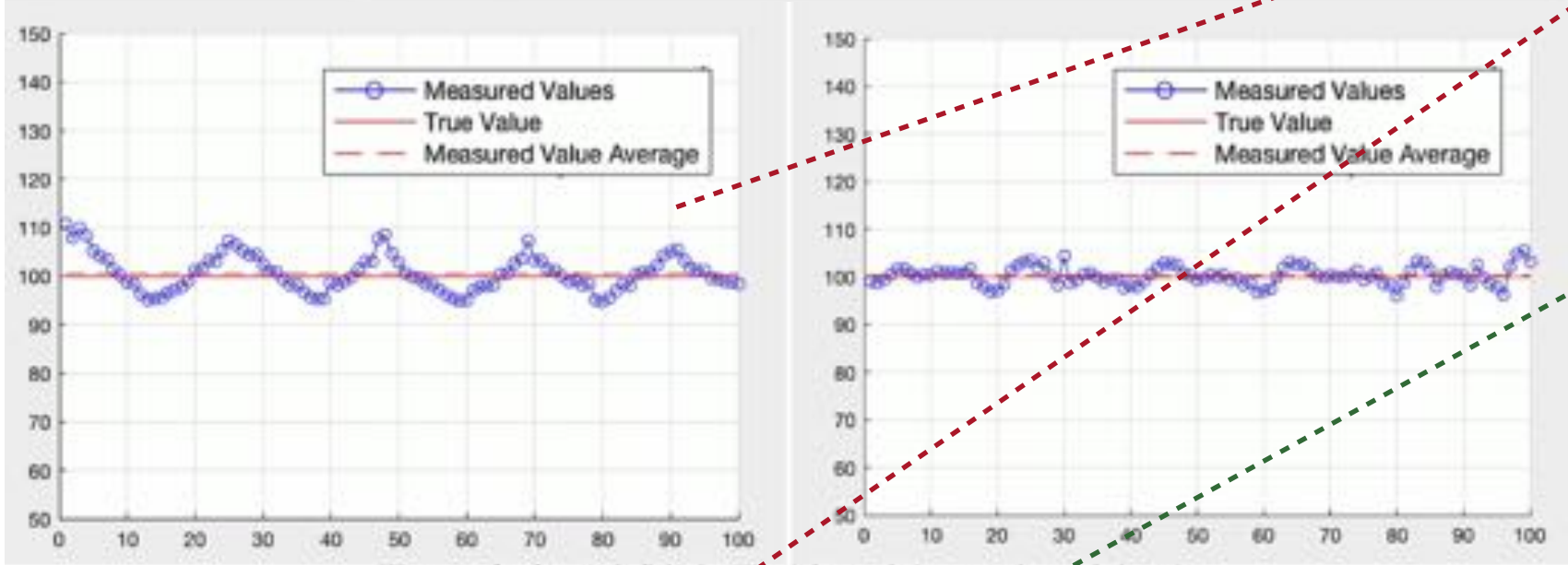
实验结果与分析 —> 实验1:

▶ 路径记录  
(左: 路径1;  
右: 路径2)



▶ 与规划一致，结果较好

▶ 测试数据图  
(左: 路径1;  
右: 路径2)



▶ 波动性大  
▶ 误差大

▶ 均值与真实值差距小，结果较好

平均值 $\bar{D}_{1-1} = 100.5924$ ; 样本方差 $S_{1-1}^2 = 13.9918$ 。

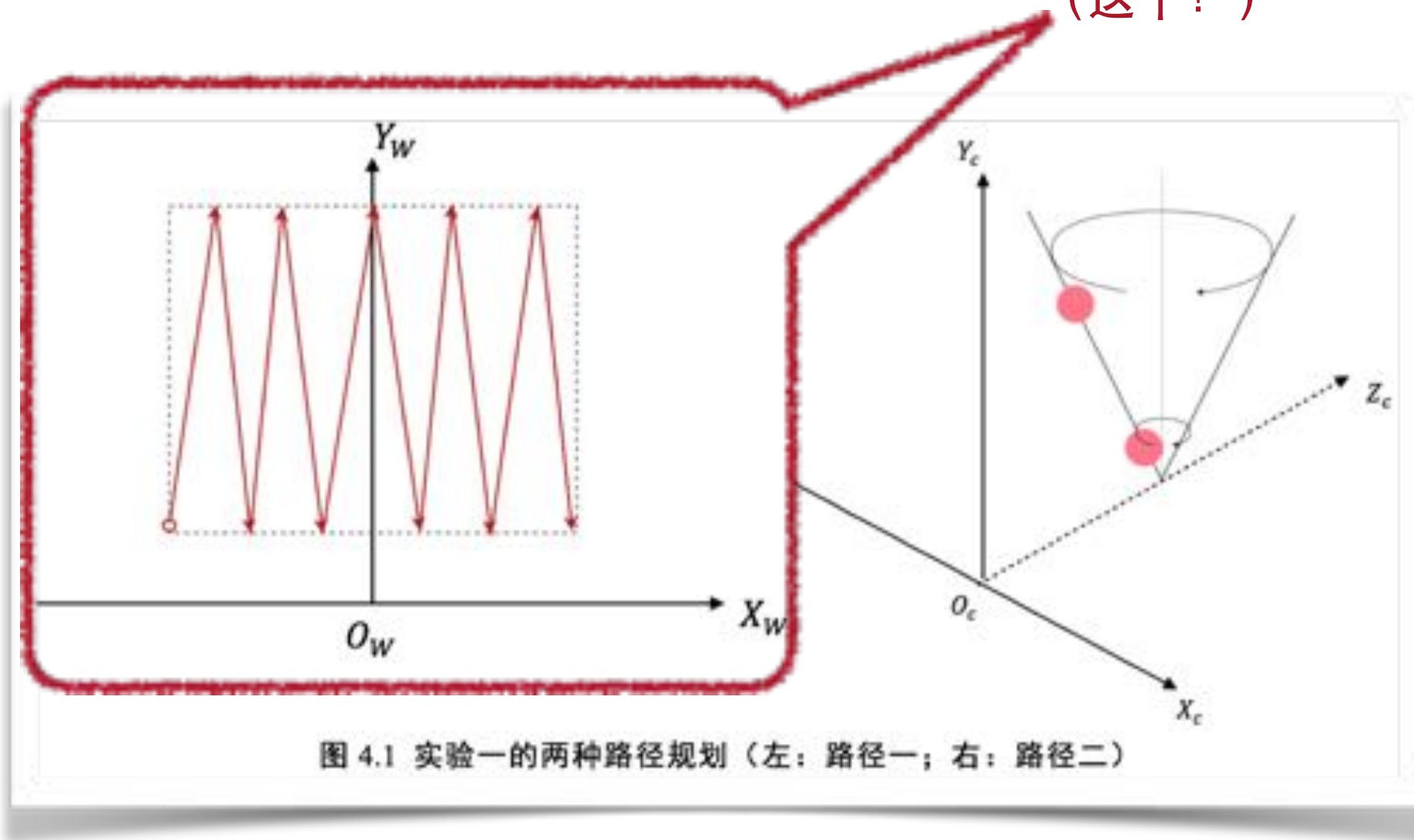
平均值 $\bar{D}_{1-2} = 100.2735$ ; 样本方差 $S_{1-2}^2 = 3.6755$ 。

镜头畸变!

## • 实验设计 → 实验2:

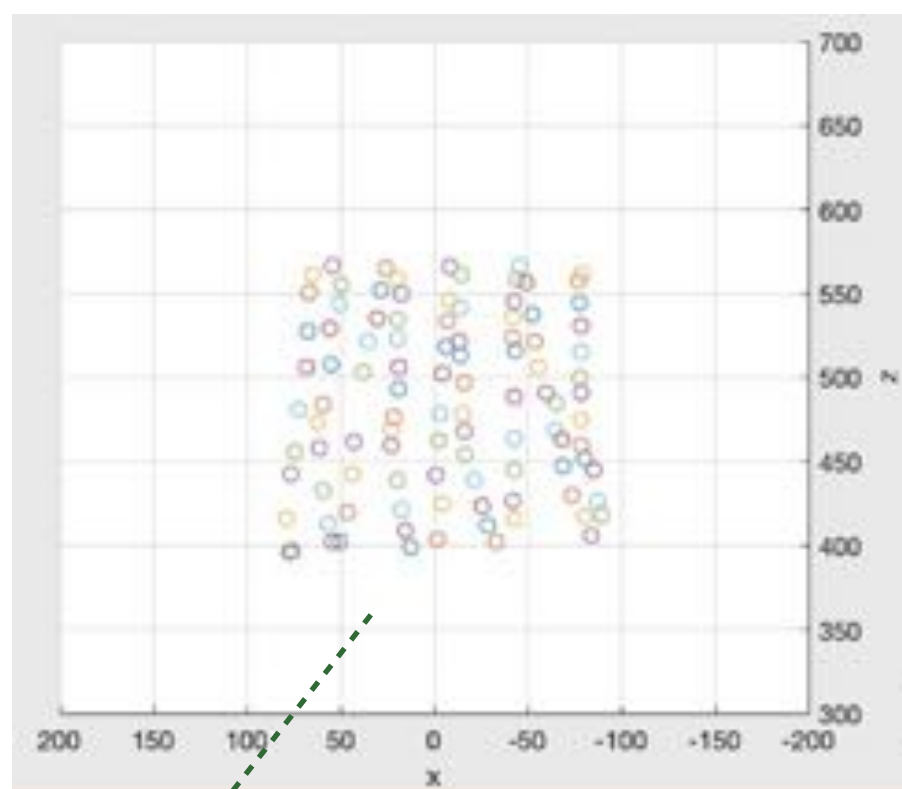
- ▶ 已知：地平面在世界坐标系中的纵坐标为 -60mm（测量摄像头高度）；
- ▶ 测量并统计：所测得的目标物纵坐标值。
- ▶ 路径：同实验1中的路径一

(这个!)



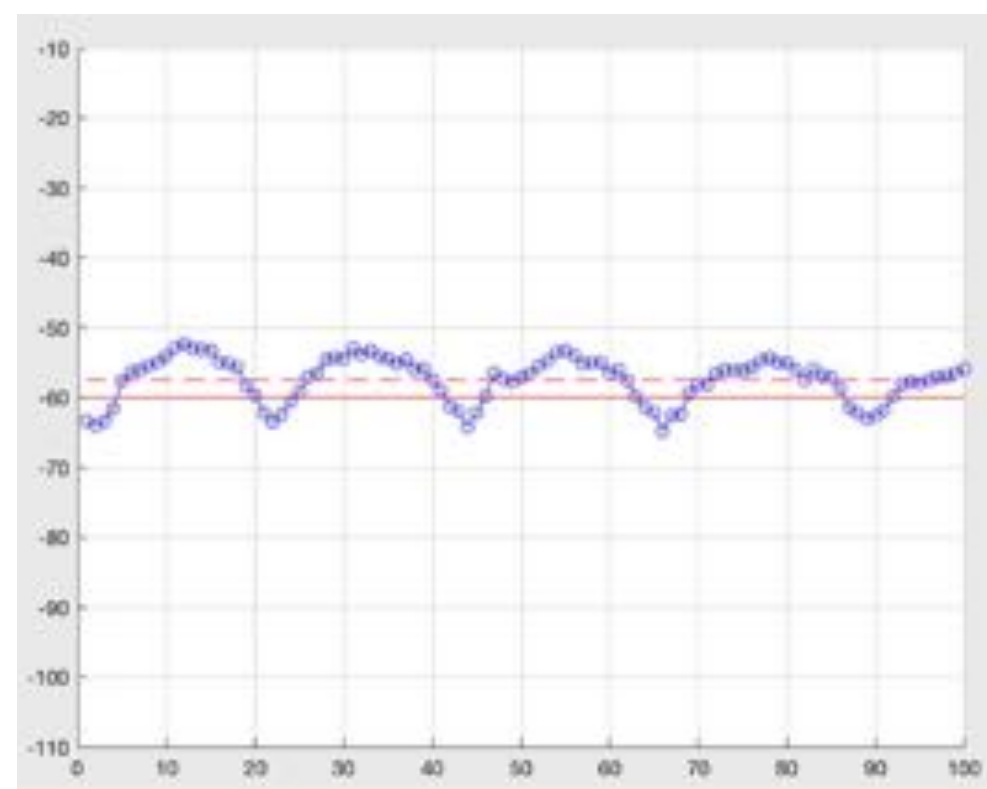
## 实验结果与分析 —> 实验2:

### ▶ 路径记录



▶ 与规划一致，  
结果较好

### ▶ 测试数据图



平均值  $\bar{Y}_{2-1} = -57.4748$ ; 样本方差  $S_{2-1}^2 = 10.1264$ 。

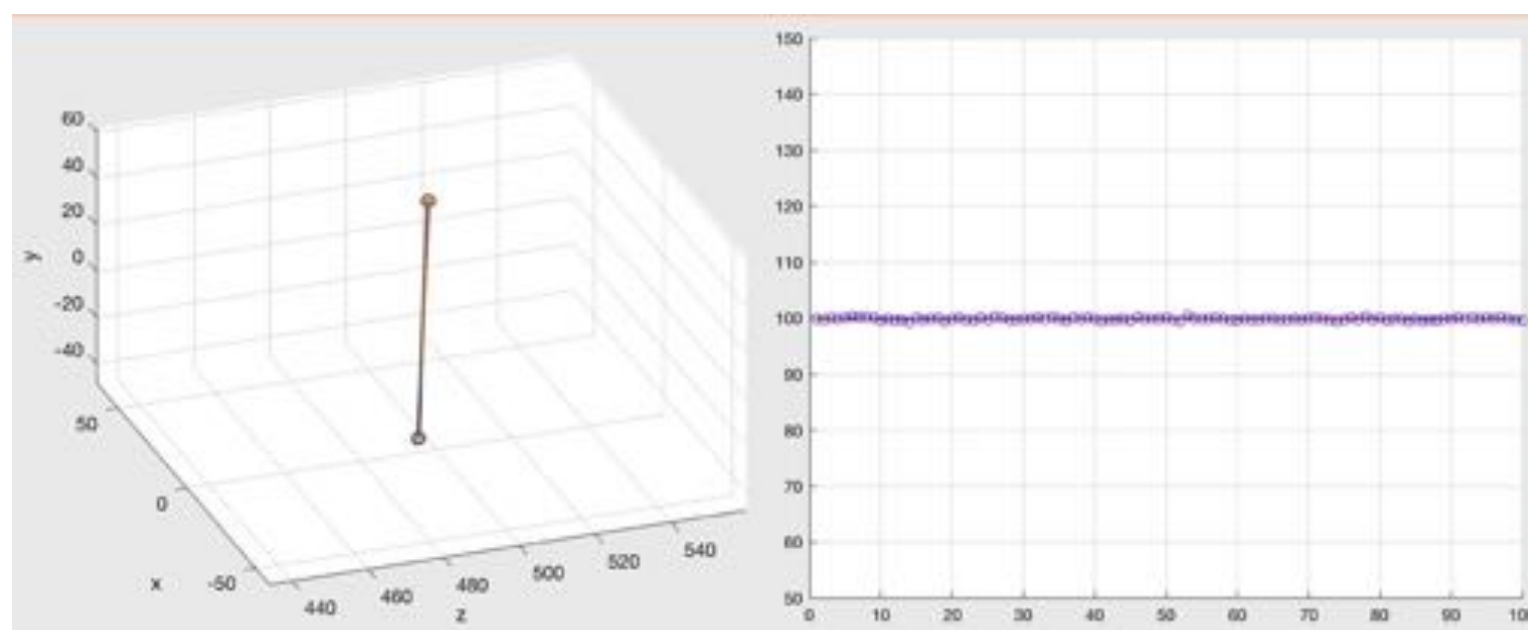
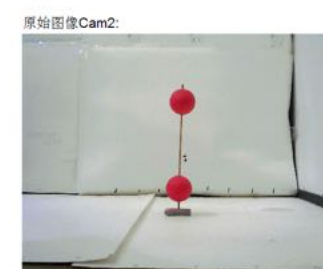
▶ 均值相较实验1，与真实  
值差值大得多，且数据波  
动也较大。

调零误差!



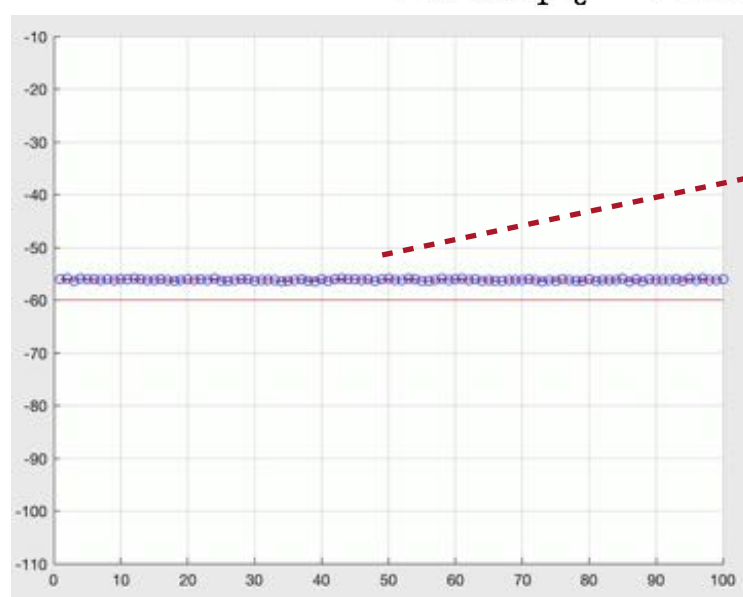
## 实验结果与分析 —> 控制组:

- 设置控制组：目标静止状态下的测量结果，以观察设备和算法本身导致的系统误差



慢!

- 根据每记录100个数据点所花费的时间，可以得到系统的刷新频率约为89.5帧/分钟，即约1.5Hz。



- 控制组测量值仍与真实值存在较大误差。

样本方差相较于之前的实验结果几乎可以忽略，说明系统误差小，算法较为稳定。

平均值  $\bar{Y}_{2-c} = -56.1725$ ; 样本方差  $S_{2-c}^2 = 0.0225$ 。

### 实验1控制组

### 实验2控制组

- 图形用户界面 (Graphical User Interface, GUI)

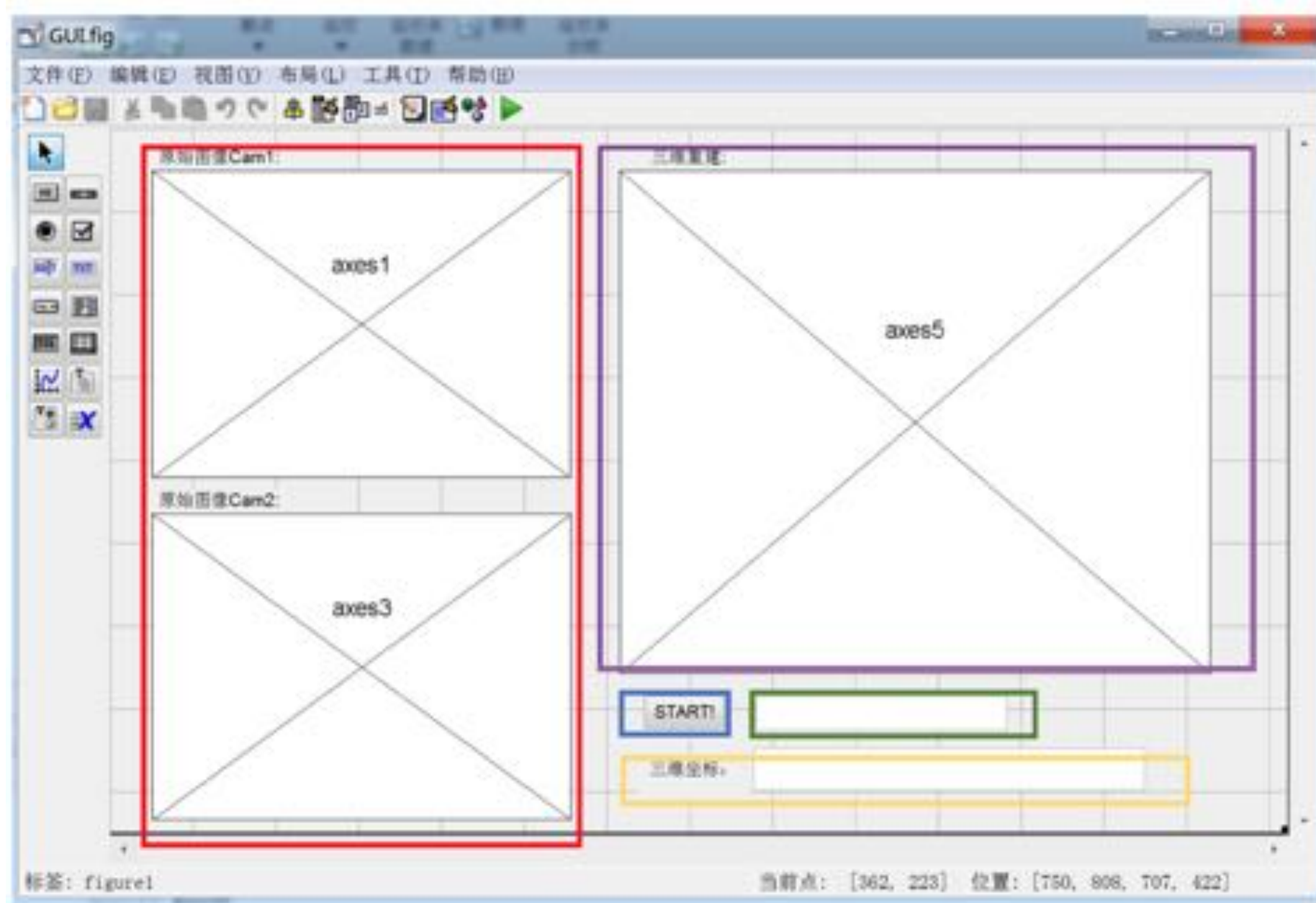


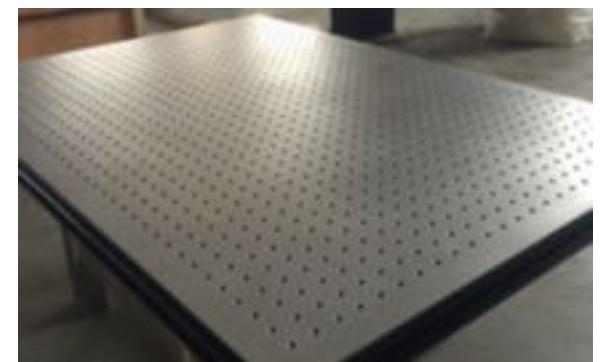
图 3.13 “GUI.fig”图形用户界面设计图

## • 该课题的改进方案

- ▶ 基于个人的经验总结，介于时间和设备的限制，还没有进行实践。

▶ 应该是这样！

▶ 现在是这样！



1. 使用云台和实验台之类的光学实验设备。
2. 使用张氏标定法之类的光学实验常见摄像头标定方法。
3. 调高分辨率
4. 使用高性能电脑
5. ....



- ▶ 一次性标定位置、畸变、焦距等所有内外参数；
- ▶ 现成Matlab Toolbox。



## 该课题的**未来研究方向**

- ▶ 进一步提高实效性（测量频率）：算法复杂度、计算机性能…
- ▶ 进一步提高准确性：设备、标定方法…
- ▶ 扩宽应用场景：更复杂的目标物；更大更随机的环境噪音；刮风、下雨怎么办？（系统稳定性）……

学界、业界的热点

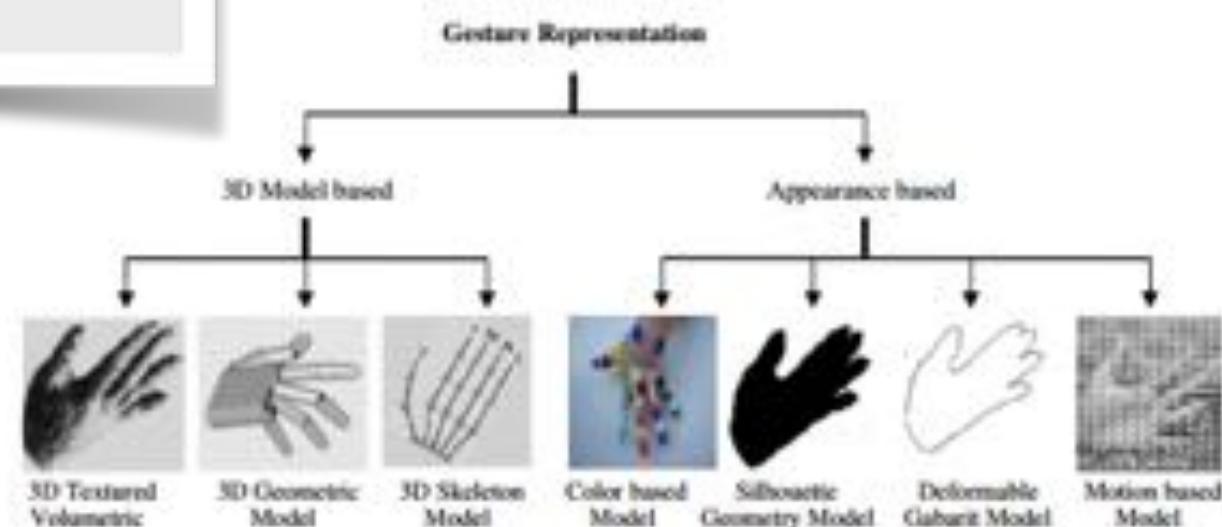
### 大疆Spark无人机正式发布！手势操控让遥控器成为历史

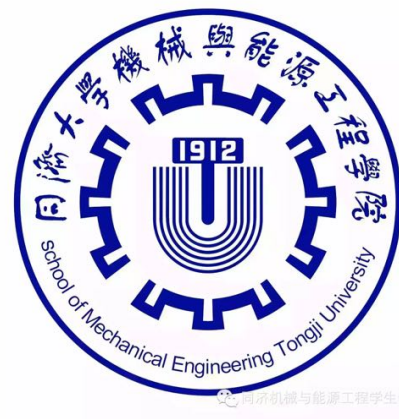
本文作者：金红

2017-05-25 00:30

导语：大疆最mini最智能无人机诞生！

- ▶ 对算法、硬件都是挑战。
- ▶ 配合学习算法、神经网络…？





# 毕业设计终期答辩

结束

## 基于CMOS的视觉识别与三维重建应用研究

机械与能源工程学院 机械设计制造及其自动化（中意班）

蔡中天 1352050 指导老师：朱传敏