

# 机器视觉与图像处理

## 第1讲 前言与绪论

主讲：汪凯巍 李 鹏

讲义部分资料来自互联网，版权归作者所有

# 机器视觉与图像处理

- 教师：汪凯巍
  - 办公室：教三137
  - 手机：13906527497
  - Email: wangkaiwei@zju.edu.cn
  - 微信： wangkaiwei00
- QQ讨论群：837474927

2

# 机器视觉与图像处理

- 助教：叶耀祖
- 办公室：教三304
  - 手机：13600547003
  - Email: 3140102472@zju.edu.cn
  - 学分：3（48学时）
- 课程网站：<http://opt.zju.edu.cn/eclass/jqsj>

3

## 本讲内容

- 课程性质与考核方式
- 机器视觉的定义
- 机器视觉的价值-有什么用？
- 机器视觉的授课内容

## 课程性质


- 专业选修课
- 每周3小节。
- 背后是火热的人工智能，相关需求暴涨。
- 很多大型科技公司和创业公司以此为发展方向或技术支撑。
- 光电专业在其中大有空间。

## 考核方式

### 课程考核与成绩评定方式

1. 平时：30%（4-5次编程练习+课堂讨论互动）
2. 期中随堂：30%
3. 期末考试：40%（大作业或课程设计）
4. 编程练习和大作业/课程设计提交到课程网站  
<http://opt.zju.edu.cn/eclass/jqsj/>

## 课程目标

1. 每个人都“通过”。
  2. 认识到“机器视觉是非常有用的”。
  3. 掌握机器视觉的“常规操作”：硬件、软件、算法，构建基本机器视觉系统。
  4. 通过学习、讨论和动手，对机器视觉充满好奇和兴趣，甚至开始深入研究或开启一段创业。
- 

## 授课内容（1）

1. **图像的获取**：二维和三维成像的光学系统，摄像机几何模型，光电转换与图像数字化，亮度视觉与颜色视觉的基本概念、模型及视觉特性。
2. **图像分析的基础变换**：图像灰度直方图、图像像素的邻域、像素间的距离度量、二维离散傅立叶变换、二维离散余弦变换。
3. **图像增强和恢复**：图像的空域、频域增强技术，图像退化的基本模型及图像恢复的基本思想，图像恢复的基本技术。直方图均衡化和规定化、平滑滤波器和锐化滤波器、同态滤波、图像退化的数学模型、逆滤波。

## 授课内容（2）

**4. 形态学图像处理：**二值化图像与灰度图像的形态学运算方法，基于形态学运算方法的相关应用。

**5. 图像分割：**边缘检测、微分算子、阈值分割、区域生长、评价测度等。

**6. 立体视觉：**常见立体视觉的获取途径-飞行时间法、结构光投影、双目视觉。双目匹配的局部算法和全局算法，基于三维特征的机器视觉应用

**7. 图像识别基础：**图像目标特征的提取与表达，图像目标特征的描述，图像目标的分类和估计、边界特征及区域特征的表达、测量和描述方法，分类器。基础神经网络。

## 参考资料

□ 数字图像处理与机器视觉（张铮 徐超），人民邮电出版社。有实例和源代码。

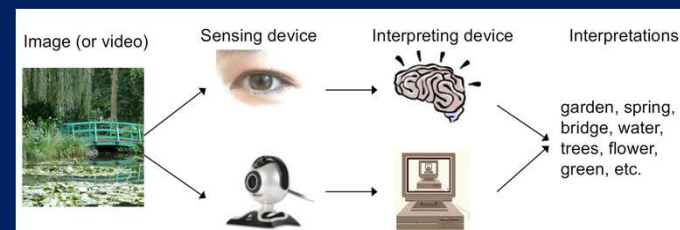


□ 机器视觉（Paul Horn，王亮等译）中国青年出版社，MIT经典教材。理论分析到位。



## 了解机器视觉

## 机器视觉（Machine vision）



让“机器”像人一样识别物品、理解场景

Make computers understand images and video or **any visual data**.

## 机器视觉在日常生活中的应用

### 案例：相机上的人脸检测



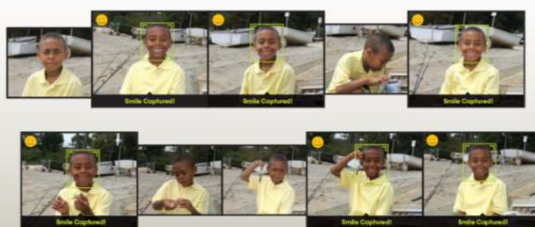
#### Face detection

- Digital cameras detect faces

### 案例：笑脸检测(Smile detection)

#### The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.



[Sony Cyber-shot® T70 Digital Still Camera](#)

### Object recognition (in mobile phones)



Point & Find, Nokia (拍立寻, 拍照获得图片物体信息)

Google Goggles

## 拍照识别名片



## 案例：Login without a password...



Fingerprint scanners on many new laptops, other devices



Face recognition systems now beginning to appear more widely  
<http://www.sensiblevision.com/>



Convenience

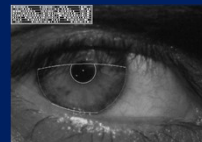
- ☐ Facial recognition
- ☐ Auto unlock Cell Phone
- ☐ 2D & 3D

## 案例：Vision-based biometrics



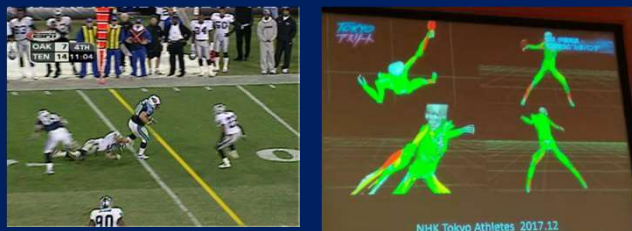
Age 12      Age 30

"How the Afghan Girl was Identified by Her Iris Patterns" Read the [story](#)  
[Wikipedia](#)



## 机器视觉在体育、游戏中的应用

## 案例：Sports



对运动场景分析：可以帮助训练、提高比赛成绩

Sportvision first down line  
Nice [explanation](http://www.howstuffworks.com) on [www.howstuffworks.com](http://www.howstuffworks.com)

<http://www.sportvision.com/video.html>

图片来源：东京大学，Yoshihiko 教授

## 案例：Interactive Games: Kinect

- 使用视觉设备和算法帮助识别玩家的姿态，挥拍的动作（打球），身体的倾斜程度（滑雪）



Fun

- ☐ Play with the computer
- ☐ Just like playing with friends

## 案例：Augmented Reality (AR)



- 为了在特定的位置向视网膜投射虚拟的图像，需要对场景实现准确的理解

Magic Leap, Oculus, Hololens, etc.

## 案例：Special effects: motion capture (特效：动作捕捉)



*Pirates of the Caribbean*, Industrial Light and Magic (工业光魔)  
<https://blog.csdn.net/chinavfx/article/details/53613285>

## 机器视觉在教育行业的应用

### 拍照知答案



要是微积分和物光啥的也能这样就好了

### 拍照识花



### 机器视觉在农业中的应用例子



## 案例：农业采摘机器人



茄子采摘机器人

胡慧明：基于双目视觉的棚室番茄采摘关键技术研究

## 案例：农业采摘机器人

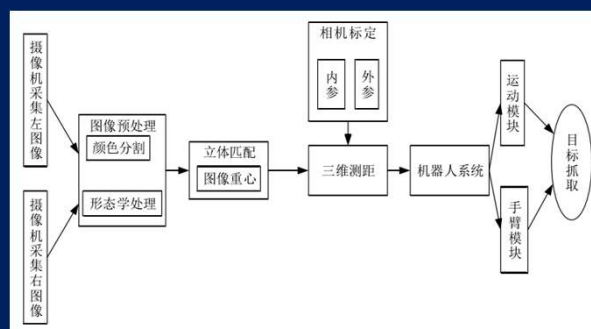


图 1.3 黄瓜采摘机器人

黄瓜采摘机器人

胡慧明：基于双目视觉的棚室番茄采摘关键技术研究

## 案例：农业采摘机器人



技术路线图

胡慧明：基于双目视觉的棚室番茄采摘关键技术研究

## 案例：机器人农场助手

- 禽类生病时，移动缓慢，倾向于避开同伴
- 分析图像的方法可以自动检测出来生病家禽
- 可以隔离开生病家禽，防止疾病传播，降低抗生素使用
- 对规模化农场特别有价值



素材来源：British Science Museum



### 案例：蚕卵自动分选

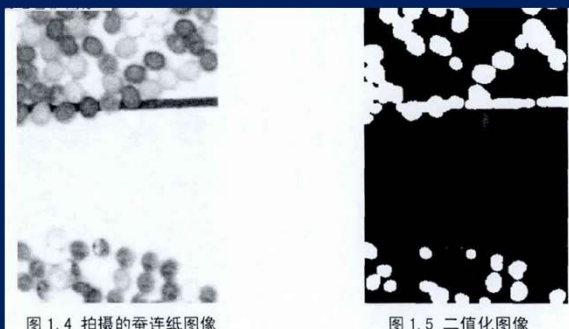


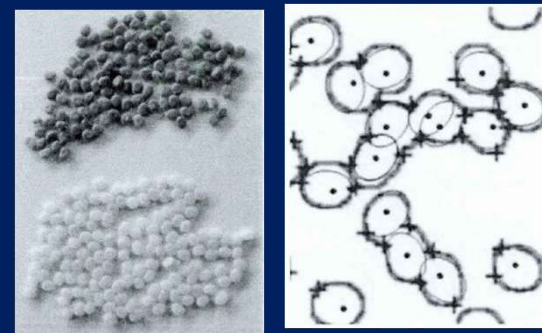
图 1.4 拍摄的蚕连纸图像

图 1.5 二值化图像

雄蚕吃得少，产丝多

华娇：雌雄蚕卵激光自动分选仪及散卵式光电自动分选仪的研究

### 案例：蚕卵自动分选

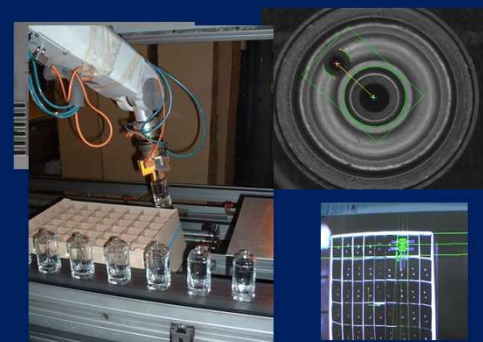


机器视觉获得蚕卵中心位置，激光灭活雌卵

华娇：雌雄蚕卵激光自动分选仪及散卵式光电自动分选仪的研究

## 机器视觉在工业生产中的应用

### 案例：工业机器人的视觉反馈



- 定位，为后续工序或者机器人提供引导
- 为机器人的动作控制提供视觉反馈。其功能为识别工件，确定工件的位置和方向

## 案例：工业机器人的视觉反馈



Vision-guided robots position nut runners on wheels  
视觉引导机器人安装螺钉

## 案例：工业机器人的视觉反馈



机器人无序抓取

无序抓取，图像来源：  
<http://www.app17.com/c136907/products/d8102699.html>

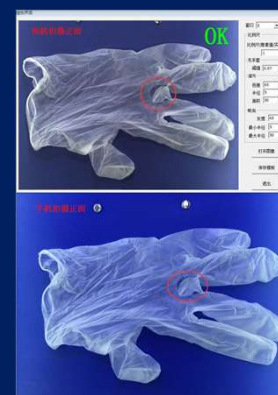
## 案例：工业机器人的视觉反馈



机器人无序抓取

无序抓取，图像来源：  
<http://www.app17.com/c136907/products/d8102699.html>

## 案例:乳胶手套



- 江苏某乳胶手套生产商
- 年产20亿只
- 脱模过程有破损、油污



### 案例:单晶硅片



- 浙江某单晶硅片厂商
- 细分行业龙头
- 大量的人员用于检验硅片的划伤、崩边、裂纹,低效率

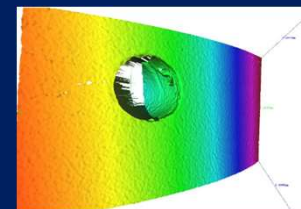
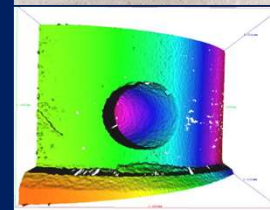


### 案例:航空发动机



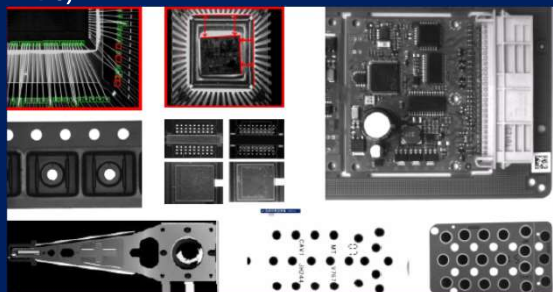
西安某厂商:

- 需要测量航空发动机组件的小孔直径以及与大圆柱体轴线距离。
- 现在人工用塞规,耗时,无法应对规模化生产
- 期待通过光电非接触的方式自动化测量



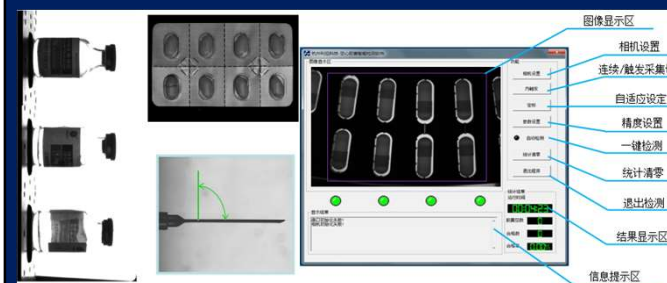
### 案例:PCB板

- 某深圳AOI公司
- 面向电子行业,焊点质量判断(漏焊、虚焊),需求量巨大(手机、PC)



### 案例:胶囊

- 医疗及医疗器械行业:
- 药品的封装、胶囊的检测(中国每年1000亿粒)

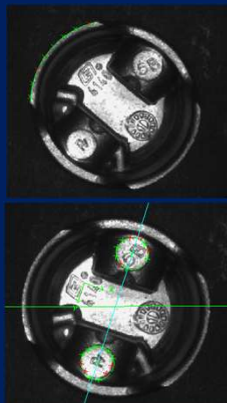


杭州利珀科技

<http://www.hzleaper.com/?p=112>

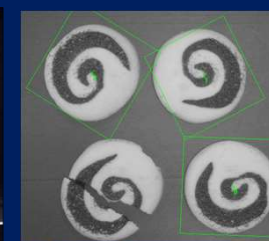
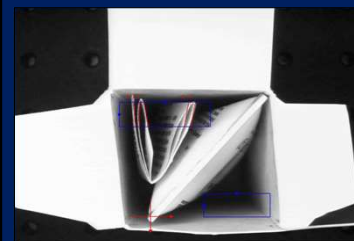
### 案例:火花塞

- 测量  
(如长度、直径、角度测量)



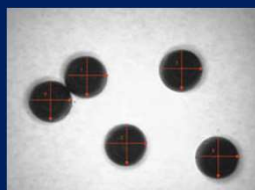
### 案例: 有无、残次品检测

- 检测 (如有无、残次品检测)



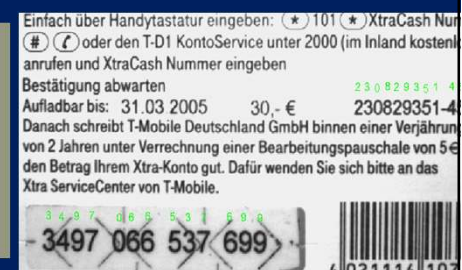
### 案例:工业计数

- 检测: 数量统计



### 案例:编码和数字识别

- 识别:如编码, 数字



## 案例:带材

- 杭州赤霄: 带材、膜、纸表面缺陷检测



放弃出国要创业的校友: 赤霄科技王曦

<https://chixiaotech.com/>

49

## 案例: Optical character recognition (OCR)



Digit recognition, AT&T labs

<http://www.research.att.com/~yann/>



License plate readers

[http://en.wikipedia.org/wiki/Automatic\\_number\\_plate\\_recognition](http://en.wikipedia.org/wiki/Automatic_number_plate_recognition)

## 机器视觉在智能交通的应用

Slide content courtesy of Amnon Shashua

## 案例: Smart cars



- Mobileye
  - Market Capitalization: 11 Billion dollars
  - Bought by Intel for 15 Billion dollars



## 案例：Google cars



Oct 9, 2010. "Google Cars Drive Themselves, in Traffic". [The New York Times](#). John Markoff  
 June 24, 2011. "Nevada state law paves the way for driverless cars". [Financial Post](#). Christine Dobby  
 Aug 9, 2011. "Human error blamed after Google's driverless car sparks five-vehicle crash". [The Star](#) (Toronto)

## 案例：智慧交通



自动识别车牌，抓拍违章，分析交通流量和拥堵情况，分配路口红绿灯时间

## 机器视觉在安防中的应用

## 案例：身份识别、刑侦



自动身份识别  
代替纸质票据

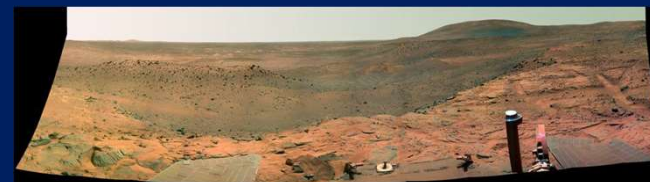
Access

- ☐ Authenticate identity
- ☐ Save time and work load



## 机器视觉在航空航天中的应用

### 案例：Vision in space

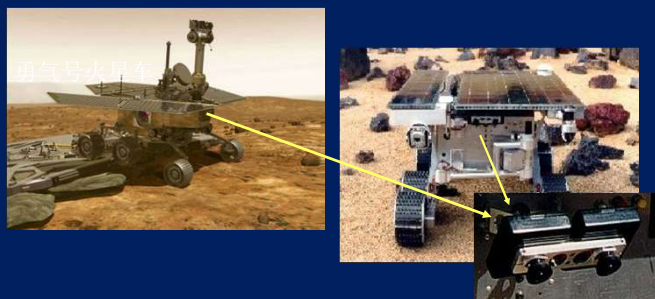


NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

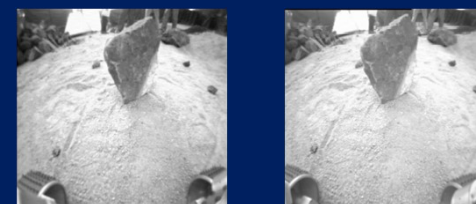
#### Vision systems (JPL) used for several tasks

- Panorama stitching (全景拼接)
- 3D terrain modeling (三维地形建模)
- Obstacle detection, position tracking (障碍物探测、位置探测)
- For more, read "[Computer Vision on Mars](#)" by Matthies et al.

### 自动驾驶



火星车前部图，中部的两个小突出是两个黑白CCD摄像机



Rocky7视觉系统获取的立体图像对



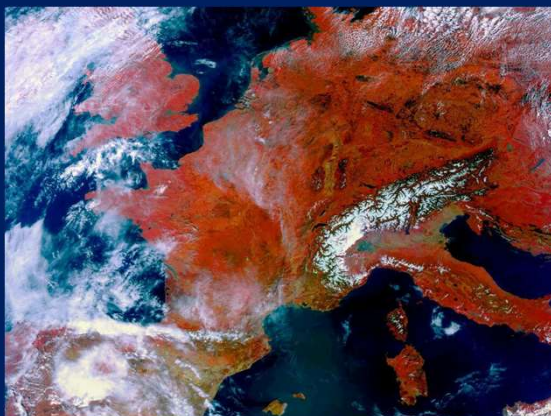
(a) 深度图像



(b) 障碍物探测示意图

Rocky7 视觉系统对场景的深度恢复

## 遥感图片



分析地貌变化、植被、作物产量...

## 机器视觉在医疗行业的应用

### 医疗：医学影像的分析判断、为医生的影像诊断提供辅助

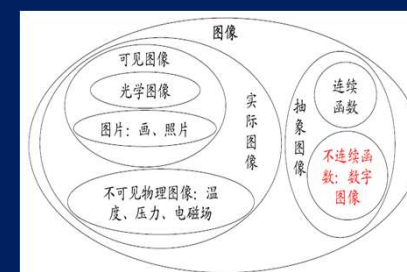


- 《自然》：用近13万张痣、皮疹和其他皮肤病变的图像训练机器识别其中的皮肤癌症状
- 与21位皮肤科医生的诊断结果进行对比发现，诊断准确率与人类医生不相上下，在91%以上

63

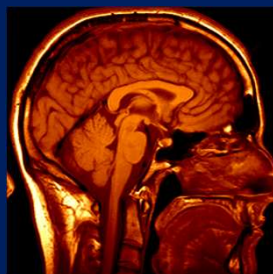
## 机器视觉 (Machine vision)-广义的图像

图像的基本概念：场景画面的一种表示，物理量（一般为颜色）的二维分布 $f(x,y)$   
 数字图像：数字矩阵表示的二维图像。（空间上离散化非连续，幅值上量化）



- 广义的图像除了可见光图像外，还包括其它参数表征的“图像”。

## 案例：Medical imaging

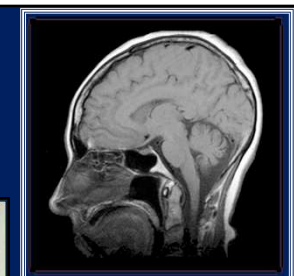


3D imaging  
MRI, CT

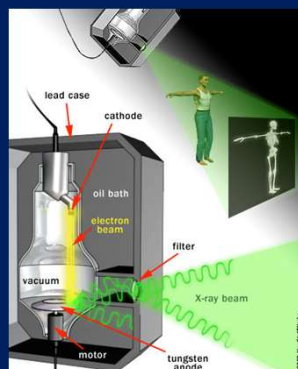


Image guided surgery  
[Günson et al., MIT](#)

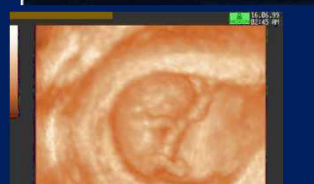
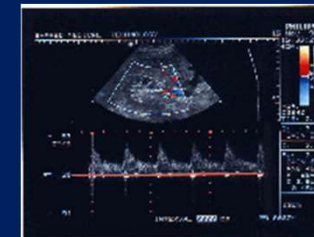
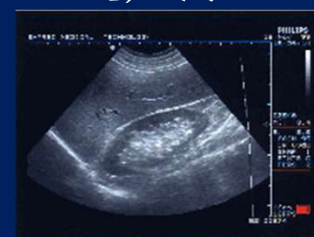
## MRI (核磁共振成像)



## X-Ray



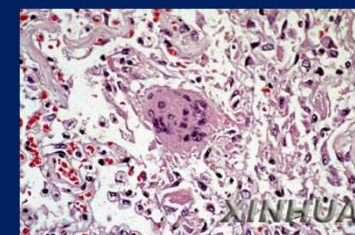
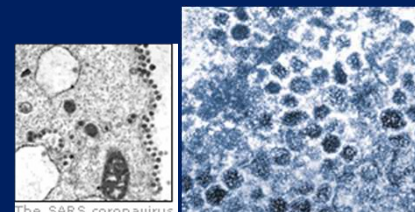
## 超声图像



## 红外图像

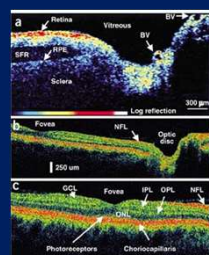


## 医学图像

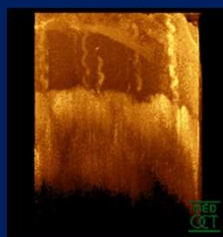


## 医学图像

甚至表面以下组织的断层成像  
OCT-Optical coherence topography



OCT tomogram of an eye



OCT tomogram of a fingertip

视觉感知与辅助研究组的相关工作



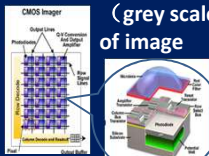
## 光的属性

Simple representation of Electro- Magnetic wave

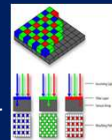
$$e_c(t) = A_c \cos(\omega_c t + \phi_c)$$

Amplitude Frequency-  
wavelength

Intensity  
(grey scale)  
of image



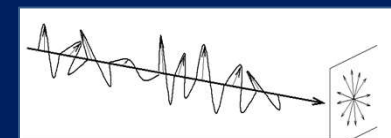
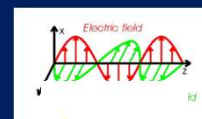
Color(spectra) of  
image



Can not be  
detected  
directly  
(change  
extremely  
fast)

## 光的偏振属性--Polarization

Light is a transverse wave, vibrates in  
direction perpendicular to traveling  
direction.

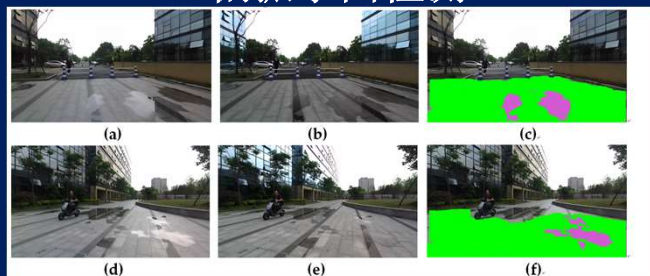


**Non-polarized light**

"Ordinary" light, or natural light, is in general non-polarized: the electric field is on average oriented in an arbitrary direction.

74

## 偏振水面检测

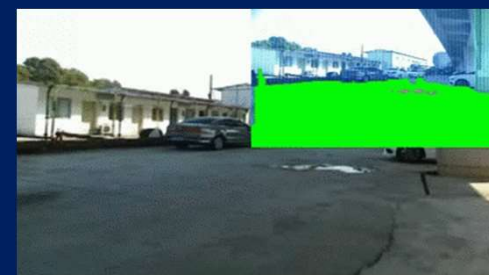


Left images Right images

Detection results with the pRGB-D  
sensor

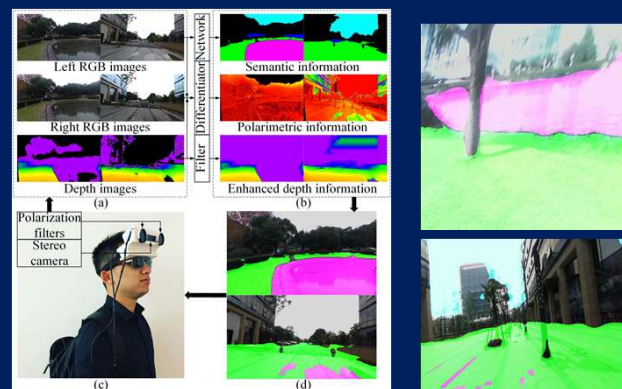
- ❑ The intensity difference of the two cameras show the degree of polarization .
- ❑ The disparity of the stereo pair also generate depth of the scene.
- ❑ Water hazard detection with traversable area detection.

## Water hazard detection: field test



Traversability Awareness + Water Hazards Detection  
→ Sunny/Cloudy/Rainy Adaptability

## Unified detection of traversable area, large water areas and small water puddles

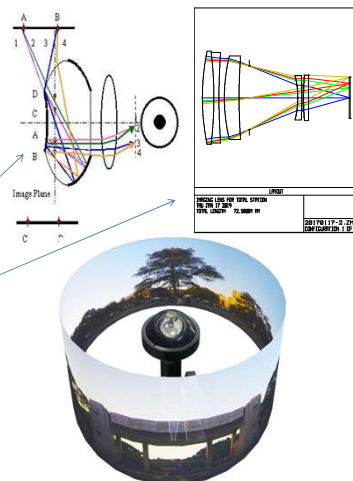


## 全景感知系统

### 硬件

#### Panoramic annular camera

- Specially designed optical path.
- A mixed path of both refraction and reflections.
- Conventional lens feature refraction path.



Fusion with  
Lidar

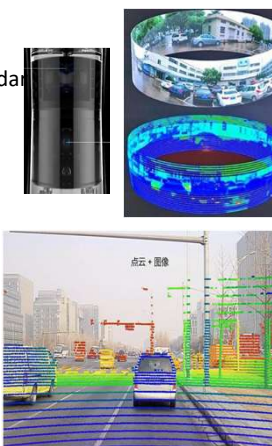


Features similar field of  
view



Both depth point  
cloud and texture  
can be acquired

Lidar





# 单相机全景成像，无盲区



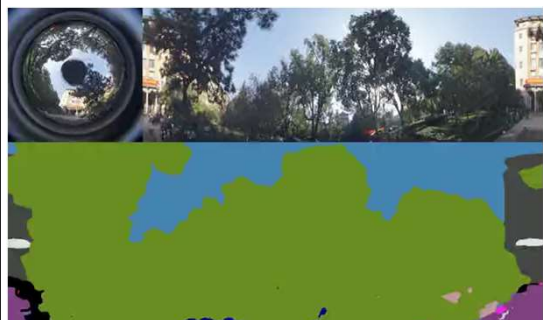
Omnidirectional view of the car( passengers ,road sign, pedestrians, blind zone)

# 单相机全景成像，无盲区



Video meeting, in collaboration with HUAWEI

# 单相机全景成像，无盲区

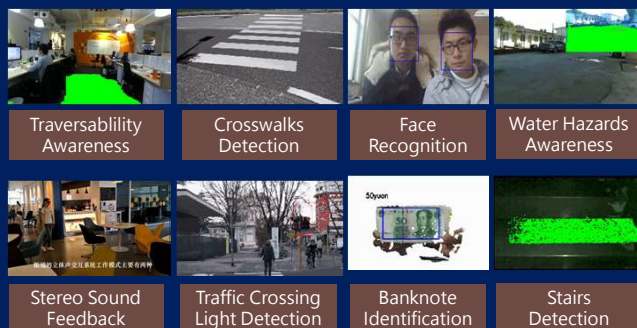


# 深度估计 Depth Estimation

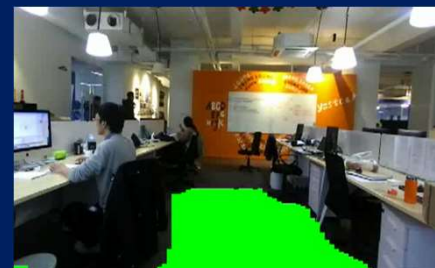


Image Depth Estimation based on PAL

## 其他相关研究



## 可通行路面预测



Passive Stereo Matching + Active Speckle Projecting  
→Indoor/Outdoor & Day/Night Adaptability

Cheng R, Wang K, Yang K, et al. A ground and obstacle detection algorithm for the visually impaired[C]. ICBIISP/IET, 2015: 1-6.  
Yang K, Wang K, Cheng R, et al. A new approach of point cloud processing and scene segmentation for guiding the visually impaired[C]. ICBIISP/IET, 2015: 1-6.  
Yang K, Wang K, Cheng R, et al. Detecting Traversable Area and Water Hazards for the Visually Impaired with a pRGB-D Sensor[J]. Sensors, 2017, 17(8): 1890.  
Huang X, Bai J, Wang K, et al. Target enhanced 3D reconstruction based on polarization-coded structured light[J]. Optics Express, 2017, 25(2): 1173-1184.

## 交通灯和斑马线识别



Pedestrian Crossing Light detection

Crosswalks

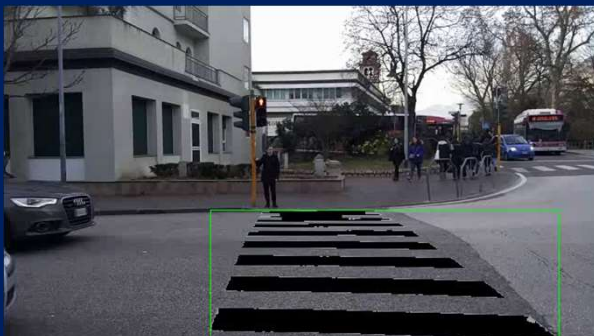
Detection  
Neural network is implemented to help improve the detection accuracy and robustness

Cheng R, Wang K, Yang K, et al. Crosswalks detection for the visually impaired[J]. Journal of Electronic Imaging, 2017, in submission.  
Cheng R, Wang K, Yang K, et al. Real-time Pedestrian Crossing Lights Detection Algorithm for the Visually Impaired[J]. Multimedia Tools and Applications, 2017, in submission.  
Long N, Wang K, Hu W, et al. Low Power Millimeter Wave Radar System for the Visually Impaired[J]. Frontiers of Information Technology & Electronic Engineering, 2017, in submission.

## 交通灯检测



## 斑马线检测



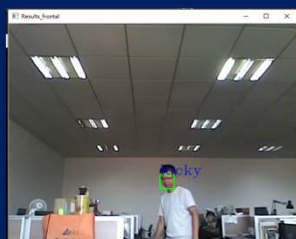
## 人机交互与功能代偿



More than 100 Traversable Directions with bone-conduction headset

Specifically coded sound generated to train the visually impaired in order to find multi obstacles in 1 second

## 人脸检测和钞票识别



Unconstrained Face Detection and Recognition



Fast Banknote Identification

Zhao X, Wang K, Yang K, et al. Unconstrained face detection and recognition based on RGB-D camera for the visually impaired[J]. Proceedings of the SPIE, Volume 10225, id. 1022509 6 pp.(2017), 2017, 225.

## 场景文字识别-Scene text recognition



(a)

(b)

A bus and a pedestrian are identified in (a) and text information on the bus is recognized in (b). From COCO2017 datasets.





## 讨论

1. 想象“未来的、顶级的”机器视觉可以做什么？  
如何改变我们未来的世界？
2. 机器视觉跟计算机视觉是等同的吗？机器视觉就是图像处理？有何分别？
3. 机器视觉还有没有没解决的问题？
4. 我们专业玩机器视觉竞争力在哪里？