# 《计算机视觉》思考题一

1. 什么是计算机视觉？什么是图像处理？两者的区别是什么？

答：

[计算机视觉]: 研究如何让计算机从数字图像中获取深层次理解的技术，将输入的图像转化为对现实世界的理解，并作出相应反馈的技术。

[图像处理]: 主要有图像增强和图像分析两个方面，包括对图片的去噪，锐化，减轻或者消除图像的退化效果等。

[区别]: 图像处理是图像到图像的映射，其评价是主观的。计算机视觉是图像到知识之前的映射，其评价是客观的。图像处理是计算机视觉的基础。

2.你生活中哪些场景使用到了计算机视觉技术？哪些场景使用了图像处理技术？举例说明。

答：

[计算机视觉]: 公司门禁的人脸识别、手机上的花卉识别软件、车辆路口违章检测等。

[图像处理]: 照相馆提供的旧照片复原、制作海报、相机中的滤镜等。

1. 计算机视觉的难点是什么？

答：

主要有如下的两个难点

[1]: 逆问题，图像是三维世界在二维平面上的投影，而根据投影反求出三维空间信息是难解的问题。这也正是计算机视觉需要解决的。

[2]: 从图像到语义概念的映射，这个过程是未知、极其复杂的。

对于第一个难点，目前可以使用先验的知识对问题添加约束，后一个难点可以使用数据驱动的方式，如模式识别、机器学习、大数据等。

4.人类具有很强的人脸识别能力，但是却不能很好的识别指纹。请从进化论的角度解释原因。

答：

脸是生物个体最主要的外在特征之一，并且涉及到对于不同社会阶层的身份识别。而指纹相对比较隐秘，并且差异性相对不明显。

5.判断下列视觉任务中，哪些是人类视觉可以很好的完成的，哪些是计算机视觉可以很好的完成的。

a:你在马路上骑车，看到路边穿梭的行人与车辆，你要判断这些人的行为和意图，从而决定你的行进方向和速度。

答：人类视觉

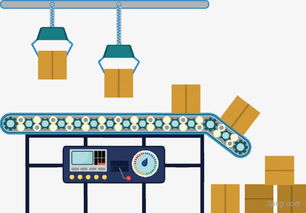
b:根据病理图像判断皮肤是否发生癌变。

答：计算机视觉

c:根据一幅图像中的建筑、风景等信息，判断这幅图像中的场景所在的经纬度。

答：计算机视觉

6.假设某工厂流水线上，使用机器手对传送带上的固定大小的包装箱进行抓取，如下图所示。机械手位于传送带上方的横梁上，机械手末端安装有摄像机，通过摄像机拍摄的图像，判断传送带上是否有包装箱，并估计包装箱的朝向、以及机械手与包装箱之间的距离。请针对这个应用场景，分别设计自顶向下和自底向上的计算机视觉解决方案。请具体说明，这个计算机视觉系统的输出结果是什么，并给出具体的表达形式。



答:

[自顶向下]:

采用基于模型的检查方法

[1] 对包装箱进行三维建模，获得其模型

[2] 当有包装箱经过时，先猜测一个包装箱的方向值d，根据这个d生成个上一把的模型得到当前模型，并与拍照的图片进行比较

[3] 通过比对得到能量值E，若E小于临界值，表示方向值d即为当前包装箱的方向，若大于临界值，则动态修正d的值到d’,回到[2]

[自底向上]:

[1] 模拟大量的集装箱经过机械师的场景，并给出当前是否有包装箱和包装箱朝向的值，以供机械臂进行训练。并给出测试场景，当测试结果识别率达到某一标准，则投入到实际场景应用。

[输出结果]: 识别出集装箱的方向，根据两者之间的距离，输出机械臂在何时以何种速度可以成功抓取当前包装箱的范围。

7. 在上一题的应用场景中，假设只对包装箱上表面粘贴的标签不规范的包装箱进行抓取。假设标签是一个固定大小的矩形，应该粘贴在包装箱表面一个同样大小的矩形框中，如果粘贴位置超出了这个矩形框，那么就认为粘贴不规范。请你设计合理的视觉解决方案，判断一个包装盒的标签是否粘贴规范。

答：

我们使用标签某一方位的坐标来描述当前标签，

特征向量: V = (x, y) // 分别表示坐标的两个方位

类别

Y = 1 // 表示为粘贴规范

Y = 0 // 表示为粘贴不规范

给出大量的训练样本，进行训练

得到先验概率P(V|Y=0) 和 P(V|Y=1)的值

利用贝叶斯公式 P(A|B) = P(A)\*P(B|A) / P(B)

则有:

P(Y=0|V) = P(Y=0) \* P(V|Y=0)/P(V)

P(Y=1|V) = P(Y=1) \* P(V|Y=1)/P(V)