Q：如果请你来模拟人类视觉系统构建一个机器视觉系统，请指出你认为其中最重要的需要建模的人类视觉系统的3个特点是什么？为什么？

A：首先我们需要先清楚人类视觉系统的组成部分及其运行规律，才能根据其结构对应构建出机器视觉系统。通过对人类视觉系统的深入了解，我认为其中最重要的需要建模的人类视觉系统的三个特点为：

**第一个为眼球对物体图像信息的捕捉。**事实上，物体可能是静止的，也可能是运动的，这使得我们的眼球需要时刻跟随着物体的运动而运动，从而持续捕捉到物体的图像信息，而不应该是一个完全静止的系统。更加复杂的一个场景是，人需要**在高速运动中**持续捕捉物体的信息，如在足球比赛中，需要边快速跑动边观察球的位置。这说明我们的观察器件——眼球，并不总是在一个稳定的场景观察的。（通过查询可知，人类视觉系统中有一个名为“前庭”的结构，可以控制眼肌的运动，维持观察的相对稳定）这与当前cv的防抖动算法相类似。

**第二个是视网膜对图像亮度和颜色的感知。**这主要由两种感光细胞实现：视杆细胞和视锥细胞。视杆细胞感知亮度，而视锥细胞感知颜色。我们需要将尽可能正确的且符合人类识别预期的图像成功感知，并将光刺激转换为相应的电刺激信号，传输到神经细胞。这跟目前cv中对图片进行RGB转化相类似，本质上是转换为计算机能够看懂的亮度与颜色。

**第三个是视网膜的神经细胞对光信息的提取和处理，以及经过分析计算最终形成目标图像。**事实上，我们并不能够将眼球锁捕捉到的所有信息不经处理，原封不动一股脑地就传输到大脑皮层中：这可能会导致大脑来不及处理过多的信息，从而导致大脑过负荷，所需能量过多。而这些神经细胞的功能主要有三点：一是对接收到的光信息进行局部的分类感受；二是对光信息进行局部的滤波处理，其中包括时域、空域与颜色的处理。这为最终在大脑皮层生成图像打下了良好的基础。而这跟目前cv领域的滤波处理，卷积提取并通过一系列的神经网络的计算最终输出最终的目标图像相对应。

上述的三个方面是人类视觉系统中最重要的部分，也是图像光信号经人眼变换为点信号传输到大脑最后完成图像构建的关键步骤。而机器视觉系统的主要目标是：对目的图像进行捕捉，提取其光信息和颜色信息，转换为RGB数据，经过对图像矩阵的信息提取和各种类神经细胞的处理后输出最终的包含关键信息的图像。经过比对，它们完全契合对应。

以上。