

1.计算机视觉和 OpenCV 介绍

1.机器是如何“看”的

我们人类可以通过眼睛看到五颜六色的世界，是因为人眼的视觉细胞中存在分别对红、绿、蓝敏感的 3 种细胞。其中的光感色素根据光线的不同进行不同比例的分解，从而让我们识别到各种颜色。

对人工智能而言，学会“看”也是非常关键的一步。那么机器人是如何看到这个世界的呢？这就涉及到人工智能方向重要的分支--机器视觉。

机器视觉即用机器人代替人眼来做测量和判断，通过机器视觉产品（即图像摄取装置，分 CMOS 和 CCD 两种）将被摄取的目标转换成图像信号，传送给专用的图像处理系统，根据像素分布和亮度、颜色等信息，转变成数字化信号。

图像系统对这些信号进行各种运算来抽取目标的特征，进而根据判别的结果来控制现场的设备动作。

2.机器视觉技术的常见应用

随着人工智能的推进、5G 时代的到来，中国正成为世界机器视觉发展最为活跃的地区之一。机器视觉技术的应用范围涵盖工业、农业、军事、医药、航天、交通、科研、安防等多个行业。下面我们来看看这项技术常见的应用领域：

1) 智慧交通

城市交通拥堵情况日益严重，如何快速有效地检测拥堵状态对于解决这个问题具有极其重要的意义。机器视觉在交通领域上的应用，越来越常见，也发挥着越来越重要的作用。

例如，2016 年杭州萧山试点的“**城市大脑**”人工智能系统便是一个典型应用。通过对监控摄像头、红绿灯每天产生的海量数据进行自我管理，对红绿灯进行统筹调节，使得车辆在高架路上的通行时间平均缩短 5 分钟，在普通道路上的车速提高 15%-20%。

2) 智能家居

机器视觉在智能家居方面的应用，与我们的生活密切相关。人工智能技术的运用，使得我们的家电越来越“聪明”。而视觉模块就相当于智能家电的“**眼睛**”，是感知层最为重要的核心器件之一。

普通家电主要是通过遥控器、APP 进行操控。在增加了摄像头、语音识别、语音合成等感知模块后，家电产品便能够“看懂”、“听懂”、“说出”，和用户进行交流，了解用户需求。

例如，人脸识别门禁系统能够与室内的空调、热水器、电灯等家电结合起来，实现对屋主与陌生人进行识别、区分，从而控制大门及家电的开关。

3.图像识别介绍

图像识别技术是人工智能的重要领域。它是一种利用计算机对图像进行处理、分析，以识别各种不同模式的目标和对象的技术。

正如我们人眼识别图像，往往是从变化最大或突然改变的地方开始，从一个特征到下一个特征。大脑控制眼睛采集图像的主要特征，排除冗杂的非必要信息，再将主要特征的信息整合成完整的视觉映象。

计算机的图像识别与人眼识别图像的原理相同，识别过程大同小异，大致可分为四个步骤：

- 1) 信息获取：通过传感器将光信号、声音信号等转换为电信号，即对基本信息进行获取。
- 2) 图像预处理：主要指采用平滑、去噪等手段对图像进行处理，突出图像的主要特征。
- 3) 特征抽取和选择：在模式识别中，抽取和选择图像特征。这是图像识别过程中的关键步骤之一。
- 4) 图像分类：根据训练结果制定识别规则，即设计分类器，得到特征的主要种类，从而使图像识别的辨识度不断提高。

4. 图像识别技术的常见应用

在 AI 领域，图像识别技术占据着极为重要的地位。随着计算机技术与信息技术的不断发展，图像识别技术的应用范围也在不断拓宽。

1) 遥感图像识别

航空遥感和卫星遥感图像通常采用图像识别技术进行加工，提取有用信息。该技术主要用于地形地质探测、气象卫星云图处理、环境污染检测等。

2) 机器视觉领域

图像处理作为热门研究课题--人工智能的重要组成部分，其应用非常广泛，并且与我们的生活紧密相连。例如快递无人车、家庭服务智能机器人、扫地机器人、玩具机器人等。

5.OpenCV 介绍

OpenCV（Open Source Capture Vision）是一个免费的计算机视觉库。它可以处理图像和视频的各种任务，比如显示摄像头采集到的画面以及让机器人识别现实生活中的物体。



虽然 Python 自带图像处理库 PIL，但是其功能比 OpenCV 逊色很多。OpenCV 提供完整的 Python 接口，而且在我们提供的镜像系统中已经集成 Python3.5 和 opencv-python 库文件，大家可以直接使用这个强大的计算机视觉库。

6.图像在计算机中的存储形式

在识别到图像后，计算机是如何存储不同图像的呢？

一般来说，图像是由一个个像素点构成，而每个像素点又可以用范围在 0-255 的 R、G、B 分量表示。OpenCV 则将每个像素用一个三元数组存储起来，这便可以记录图像的全部信息。另外，我们需要注意 OpenCV 在记录 RGB 图像三个颜色通道的数据时，记录顺序会变为 BGR。

此外，其它标准（例如 HSV）的图像，也是通过多元数组的形式进行存储。OpenCV 图像是.array 类型的二维或三维数组，8 位的灰度图像（只有黑，白色的图像）是一个二维数组，24 位的 BGR 图像是一个三维数组。

比如，对于一个 BGR 图像来说，“image[0,0,0]”元素的第一个值表示像素的 Y 轴坐标或者行号（0 代表顶部）；第二个值代表像素的 X 轴坐标或列号（0 表示最左边）；第三个

值表示颜色通道。

这些记录图像的数组和普通的 Python 数组一样可以单独访问，从而可以获得某个颜色通道的数据，或截取图像某个区域的图像。