

# 计算机视觉

维基百科，自由的百科全书

**计算机视觉**是一门研究如何使机器“看”的科学，更进一步的说，就是指用摄影机和计算机代替人眼对目标进行识别、跟踪和测量等机器视觉，并进一步做图像处理，用计算机处理成为更适合人眼观察或传给仪器检测的图像<sup>[1]</sup>。

作为一門科学学科，计算机视觉研究相关的理论和技术，试图建立能够从图像或者多维数据中获取「信息」的人工智能系统。这里所指的信息指香农定义的，可以用来帮助做一个“决定”的信息。因为感知可以看作是从感官信号中提取信息，所以计算机视觉也可以看作是研究如何使人工系统从图像或多维数据中“感知”的科学。

作为一个工程学科，计算机视觉寻求基于相关理论与模型来建立计算机视觉系统。这类系统的组成部分包括：

1. 过程控制（例如工业机器人和无人驾驶汽车）
2. 事件监测（例如图像监测）
3. 信息组织（例如图像数据库和图像序列的索引建立）
4. 物体与环境建模（例如工业检查，医学图像分析和拓扑建模）
5. 交互互动（例如人机互动的输入设备）

计算机视觉同样可以被看作是生物视觉的一个补充。在生物视觉领域中，人类和各种动物的视觉都得到了研究，从而建立了这些视觉系统感知信息过程中所使用的物理模型。另一方面，在计算机视觉中，靠软件和硬件实现的人工智能系统得到了研究与描述。生物视觉与计算机视觉进行的学科间交流为彼此都带来了巨大价值。

计算机视觉包含如下一些分支：画面重建，事件监测，目标跟踪，目标识别，机器学习，索引建立，图像恢复等。



NASA火星探测车的雙攝影機系統

## 目录

计算机视觉的发展现状

相邻领域的异同

计算机视觉的经典问题

识别

运动

场景重建

图像恢复

计算机视觉系统

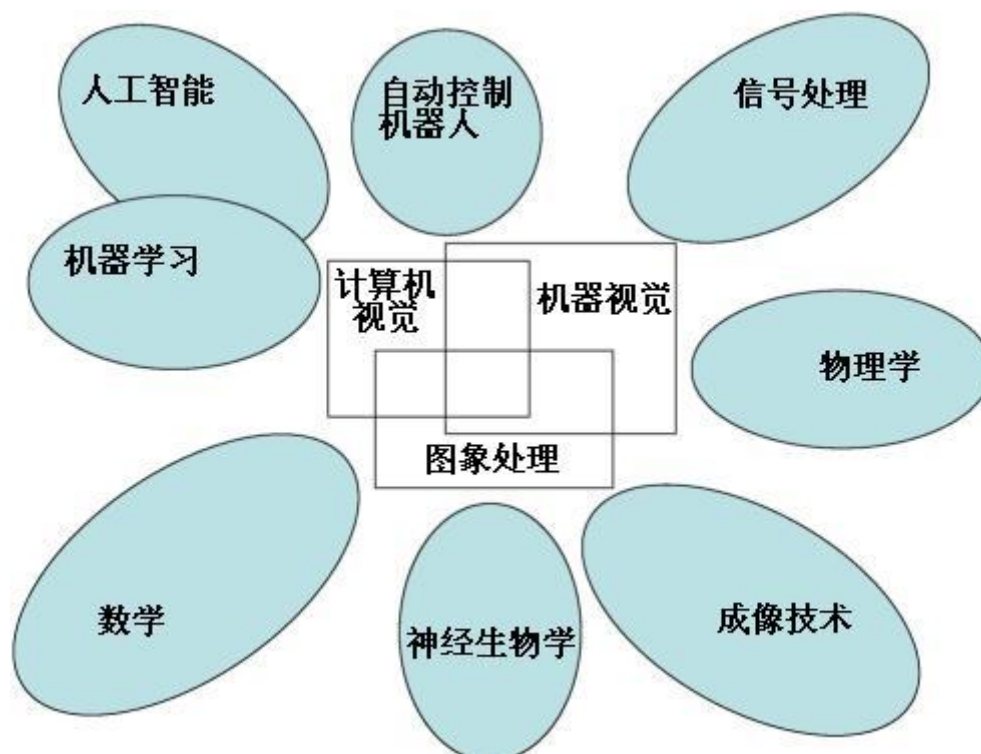
影響視覺系統的要件

参考文献

外部連結

參見

# 计算机视觉的发展现状



计算机视觉与其他领域的关系

计算机视觉领域的突出特点是其多样性与不完善性。

这一领域的先驱可追溯到更早的时候，但是直到20世纪70年代后期，当计算机的性能提高到足以处理诸如图像这样的大规模数据时，计算机视觉才得到了正式的关注和发展。然而这些发展往往起源于其他不同领域的需要，因而何谓“计算机视觉问题”始终没有得到正式定义，很自然地，“计算机视觉问题”应当被如何解决也没有成型的公式。

儘管如此，人们已開始掌握部分解决具体计算机视觉任务的方法，可惜这些方法通常都僅適用於一群狹隘的目標（如：臉孔、指紋、文字等），因而无法被廣泛地应用於不同場合。

对这些方法的应用通常作为某些解决复杂问题的大规模系统的一个组成部分（例如医学图像的处理，工业制造中的质量控制与测量）。在计算机视觉的大多数实际应用当中，计算机被预设为解决特定的任务，然而基于机器学习的方法正日渐普及，一旦機器學習的研究進一步發展，未來「泛用型」的電腦視覺應用或許可以成真。

人工智能所研究的一个主要问题是：如何让系统具备「计划」和「决策能力」？从而使之完成特定的技术动作（例如：移动一个机器人通过某种特定环境）。这一问题便與计算机视觉問題息息相關。在这里，计算机视觉系统作为一个感知器，为决策提供信息。另外一些研究方向包括模式识别和机器学习（這也隶属于人工智能領域，但与计算机视觉有着重要联系），也由此，计算机视觉時常被看作人工智能與计算机科学的一个分支。

物理是与计算机视觉有着重要联系的另一领域。

计算机视觉关注的目标在于充分理解电磁波——主要是可见光与红外线部分——遇到物体表面被反射所形成的图像，而这一过程便是基于光学物理和固態物理，一些尖端的图像传感器甚至会应用到量子力学理论，来解析影像所表示的真實世界。同时，物理学中的很多测量难题也可以通过计算机视觉得到解决，例如流体运动。也由此，计算机视觉同样可以被看作是物理学的拓展。

另一个具有重要意义的领域是神经生物学，尤其是其中生物视觉系统的部分。

在整个20世纪中，人類對各种动物的眼睛、神经元、以及与视觉刺激相关的脑部组织都進行了广泛研究，这些研究得出了一些有关“天然的”视觉系统如何運作的描述（尽管仍略嫌粗略），这也形成了计算机视觉中的一个子领域——人们试图建立人工系统，使之在不同的复杂程度上模拟生物的视觉運作。同时计算机视觉领域中，一些基于机器学习的方法也有参考部分生物機制。

计算机视觉的另一个相关领域是信号处理。很多有关单元变量信号的处理方法，尤其对是时变信号的处理，都可以很自然的被扩展为计算机视觉中对二元变量信号或者多元变量信号的处理方法。但由于图像数据的特有属性，很多计算机视觉中发展起来的方法，在单元信号的处理方法中却找不到对应版本。这类方法的一个主要特征，便是他们的非线性以及图像信息的多维性，以上二點作为计算机视觉的一部分，在信号处理学中形成了一个特殊的研究方向。

除了上面提到的領域，很多研究课题同样可被当作纯粹的数学问题。例如，计算机视觉中的很多问题，其理论基础便是统计学，最优化理论以及几何学。

如何使既有方法通过各种软硬件实现，或说如何对这些方法加以修改，而使之获得合理的执行速度而又不损失足够精度，是現今電腦視覺領域的主要課題。

## 相邻领域的异同

---

计算机视觉，图像处理，图像分析，机器人视觉和机器视觉是彼此紧密关联的学科。如果你翻开带有上面这些名字的教材，你会发现在技术和应用领域上他们都有着相当大部分的重叠。这表明这些学科的基础理论大致是相同的，甚至让人怀疑他们是同一学科被冠以不同的名称。

然而，各研究机构，学术期刊，会议及公司往往把自己特别的归为其中某一个领域，于是各种各样的用来区分这些学科的特征便被提了出来。下面将给出一种区分方法，尽管并不能说这一区分方法完全准确。

计算机视觉的研究对象主要是映射到单幅或多幅图像上的三维场景，例如三维场景的重建。计算机视觉的研究很大程度上针对图像的内容。

图像处理与图像分析的研究对象主要是二维图像，实现图像的转化，尤其针对像素级的操作，例如提高图像对比度，边缘提取，去噪声和几何变换如图像旋转。这一特征表明无论是图像处理还是图像分析其研究内容都和图像的具体内容无关。

机器视觉主要是指工业领域的视觉研究，例如自主机器人的视觉，用于检测和测量的视觉。这表明在这一领域通过软硬件，图像感知与控制理论往往与图像处理得到紧密结合来实现高效的机器人控制或各种实时操作。

模式识别使用各种方法从信号中提取信息，主要运用统计学的理论。此领域的一个主要方向便是从图像数据中提取信息。

还有一个领域被称为成像技术。这一领域最初的研究内容主要是制作图像，但有时也涉及到图像分析和处理。例如，医学成像就包含大量的医学领域的图像分析。

对于所有这些领域，一个可能的过程是你在计算机视觉的实验室工作，工作中从事着图象处理，最终解决了机器视觉领域的问题，然后把自己的成果发表在了模式识别的会议上。

## 计算机视觉的经典问题

---

几乎在每个计算机视觉技术的具体应用都要解决一系列相同的问题。这些经典的问题包括：

### 识别

一个计算机视觉，图像处理和机器视觉所共有的经典问题便是判定一组图像数据中是否包含某个特定的物体，图像特征或运动状态。这一问题通常可以通过机器自动解决，但是到目前为止，还没有某个单一的方法能够广泛的对各种情况进行判定：在任意环境中识别任意物体。现有技术能够也只能很好地解决特定目标的识别，比如简单几何图形识别，人脸识别，印刷或手写文件识别或者车辆识别。而且这些识别需要在特定的环境中，具有指定的光照，背景和姿态要求。

广义的识别在不同的场合又演化成了几个略有差异的概念：

- **识别（狭义的）**：对一个或多个经过预先定义或学习的物体或物类进行辨识，通常在辨识过程中还要提供他们的二维位置或三维姿态。
- **鉴别**：识别辨认单一物体本身。例如：某一人脸的识别，某一指纹的识别。
- **监测**：从图像中发现特定的情况内容。例如：医学中对细胞或组织不正常技能的发现，交通监视仪器对过往车辆的发现。监测往往是通过简单的图象处理发现图像中的特殊区域，为后继更复杂的操作提供起点。

识别的几个具体应用方向：

- **基于内容的图像提取**：在巨大的图像集合中寻找包含指定内容的所有图片。被指定的内容可以是多种形式，比如一个红色的大致是圆形的图案，或者一辆自行车。在这里对后一种内容的寻找显然要比前一种更复杂，因为前一种描述的是一个低级直观的视觉特征，而后者则涉及一个抽象概念（也可以说是高级的视觉特征），即‘自行车’，显然的一点就是自行车的外观并不是固定的。
- **姿态评估**：对某一物体相对于摄像机的位置或者方向的评估。例如：对机器臂姿态和位置的评估。
- **光学字符识别**对图像中的印刷或手写文字进行识别鉴别，通常的输出是将之转化成易于编辑的文档形式。

## 运动

基于序列图像的对物体运动的监测包含多种类型，诸如：

- **自体运动**：监测摄像机的三维刚性运动。
- **图像跟踪**：跟踪运动的物体。

## 场景重建

给定一个场景的二或多幅图像或者一段录像，场景重建寻求为该场景建立一个三维模型。最简单的情况便是生成一组三维空间中的点。更复杂的情况下会建立起完整的三维表面模型。

## 图像恢复

图像恢复的目标在于移除图像中的噪声，例如仪器噪声，模糊等

# 计算机视觉系统

计算机视觉系统的结构形式很大程度上依赖于其具体应用方向。有些是独立工作的，用于解决具体的测量或检测问题；也有些作为某个大型复杂系统的组成部分出现，比如和机械控制系统，数据库系统，人机接口设备协同工作。计算机视觉系统的具体实现方法同时也由其功能决定——是预先固定的抑或是在运行过程中自动学习调整。尽管如此，有些功能却几乎是每个计算机系统都需要具备的：

- **图像获取**：一幅数字图像是由一个或多个图像传感器产生，这里的传感器可以是各种光敏摄像机，包括遥感设备，X射线断层摄影仪，雷达，超声波接收器等。取决于不同的传感器，产生的图片可以是普通的二维图像，三维图组或者一个图像序列。图片的像素值往往对应于光在一个或多个光谱段上的强度（灰度图或彩色图），但也可以是相关的各种物理数据，如声波，电磁波或核磁共振的深度，吸收度或反射度。

- **预处理**：在对图像实施具体的计算机视觉方法来提取某种特定的信息前，一种或一些预处理往往被采用来使图像满足后继方法的要求。例如：
  - 二次取样保证图像坐标的正确
  - 平滑去噪来滤除传感器引入的设备噪声
  - 提高对比度来保证实现相关信息可以被检测到
  - 调整尺度空间使图像结构适合局部应用
- **特征提取**：从图像中提取各种复杂度的特征。例如：
  - 线，边缘提取和脊偵測
  - 局部化的特征点检测如边角检测，斑点检测

更复杂的特征可能与图像中的纹理形状或运动有关。

- **检测/分割**：在图像处理过程中，有时会对图像进行分割来提取有价值的用于后继处理的部分，例如
  - 筛选特征点
  - 分割一或多幅图片中含有特定目标的部分
- **高级处理**：到了这一步，数据往往具有很小的数量，例如图像中经先前处理被认为含有目标物体的部分。这时的处理包括：
  - 验证得到的数据是否符合前提要求
  - 估测特定系数，比如目标的姿态，体积
  - 对目标进行分类

## 影響視覺系統的要件

---

- 光源佈局影響大需審慎考量。
- 正確的選擇鏡組，考量倍率、空間、尺寸、失真。
- 選擇合適的攝影機（CCD），考量功能、規格、穩定性、耐用。
- 視覺軟體開發需靠經驗累積，多嘗試、思考問題的解決途徑。
- 以創造精度的不斷提升，縮短處理時間為最終目標。

## 參考文獻

---

1. 黄亚勤。基于视线跟踪技术的眼控鼠标研究与实现[D].西华大学, 2011.

## 外部連結

---

- Machine Perception of Three-Dimensional Solids - the paper mentioned by Joseph Mundy in the video (<https://web.archive.org/web/20070329000502/http://www.packet.cc/files/mach-per-3D-solids.html>)
- COnline: The Evolving, Distributed, Non-Proprietary, On-Line Compendium of Computer Vision (<http://homepages.inf.ed.ac.uk/rbf/COnline/>)
- Introduction to computer vision ([https://web.archive.org/web/20070714063452/http://palantir.swarthmore.edu/maxwell/classes/e27/F03/E27\\_F03\\_Lectures.pdf](https://web.archive.org/web/20070714063452/http://palantir.swarthmore.edu/maxwell/classes/e27/F03/E27_F03_Lectures.pdf)) (464KB pdf file)
- CMU's Computer Vision Homepage (<http://www.cs.cmu.edu/~cil/vision.html>)
- Fudan University's Computer Vision Lab (<http://www.cv.fudan.edu.cn>)
- Keith Price's Annotated Computer Vision Bibliography (<http://iris.usc.edu/Vision-Notes/bibliography/contents.html>) and the Official Mirror Site Keith Price's Annotated Computer Vision Bibliography (<http://www.visionbib.com/bibliography/contents.html>)
- HIPR2 image processing teaching package (<http://homepages.inf.ed.ac.uk/rbf/HIPR2/>)

- [USC Iris computer vision conference list \(http://iris.usc.edu/Information/Iris-Conferences.html\)](http://iris.usc.edu/Information/Iris-Conferences.html)
- [How to come up with new research ideas in computer vision? \(in Chinese\) \(http://jbhuang0604.blogspot.com/2010/04/blog-post.html\)](http://jbhuang0604.blogspot.com/2010/04/blog-post.html)
- [People in Computer Vision \(http://jbhuang0604.blogspot.com/2010/05/people-in-computer-vision\\_5496.html\)](http://jbhuang0604.blogspot.com/2010/05/people-in-computer-vision_5496.html)
- [The Computer Vision Genealogy Project \(http://jbhuang0604.blogspot.com/2010/05/computer-vision-genealogy-project.html\)](http://jbhuang0604.blogspot.com/2010/05/computer-vision-genealogy-project.html)

## 參見

---

- [機器視覺](#)
- [人工智能与模式识别](#)
- [图像处理](#)
- [自動光學檢查](#)
- [开放源代码计算机视觉库：OpenCV](#)

---

取自“<https://zh.wikipedia.org/w/index.php?title=计算机视觉&oldid=54797639>”

---

**本页面最后修订于2019年6月12日 (星期三) 19:19。**

本站的全部文字在知识共享 署名-相同方式共享 3.0协议之条款下提供，附加条款亦可能应用。（请参阅[使用条款](#)）  
Wikipedia®和维基百科标志是维基媒体基金会的注册商标；维基™是维基媒体基金会的商标。  
维基媒体基金会是按美国国内稅收法501(c)(3)登记的非营利慈善机构。