

# 计算机视觉研究入门全指南-SWUST

原文出处：

<https://sites.google.com/site/mostafasibrahim/research/articles/how-to-start>

## 目录

顶级会议和期刊.....	2
顶级专家作者.....	2
顶级研究组.....	2
博客.....	3
CV 工业界的实验室和创业公司.....	4
如何开始研究.....	4
使用课本和课程.....	4
从论文开始.....	4
从代码开始.....	5
机器学习.....	5
一些推荐论文.....	6
积累经验.....	6
材料.....	7
在线视频和会谈.....	7
课程.....	7
计算机视觉.....	7
计算机视觉和图像处理编码.....	7
人类视觉.....	7
其他.....	7
精彩的应用程序.....	8
令人兴奋的算法.....	8
其它.....	9
工作机会.....	9
数据集.....	9
软件.....	9
有用的网站.....	9
Ad-hocks.....	9
链接.....	9

## 顶级会议和期刊

第一梯队顶级会议: CVPR, ECCV, ICCV, NIPS, IJCAI

高声誉第二梯队的顶级会议: BMVC

著名的第二梯队顶级会议: ICIP, ACCV, ICPR, SIGGRAPH

顶级期刊: PAMI, IJCV

著名期刊: CVIU, IVC

Microsoft Academic Research 列出的 顶级会议

Ranks from Core

Ranks from Arnetminer

source 列出了近几年的会议论文

journal 列出了期刊的影响因子

来自 EigenFactor 的期刊分数

## 顶级专家作者

微软学术 Microsoft Academic authors list

谷歌学术 Google Scholar List

HOG 特征作者 Navneet Dalal

Jitendra Malik.

Gary Bradski OpenCV 创始缔造者

David Lowe SIFT 特征发明人

List of vision people (but not necessarily top authors)

Computer Vision: Algorithms and Applications by Richard Szeliski

## 顶级研究组

Check them here

Check others here

CMU: Robotics everywhere.

LEAR

ImageLab Group

Machine Vision Laboratory at UWE

ALCOR

Centre for Image Processing and Analysis (CIPA)

ImageMetry

VISILAB

GRIMA – Machine Intelligence Group

Vision and Sensing Research Group – University of Canberra

CAVE – Computer Vision Laboratory at Columbia University

Computational Biomedicine Laboratory (CBL), University of Houston

Vision Lab – University of Antwerp.

Visual Geometry Group, Oxford UK (Andrew Zisserman's group)

LEAR, Grenoble, France (Cordelia Schmid's group)

WILLOW, Paris France (Jean Ponce's group)

CVLAB EPFL, Laussane Switzerland (Pascal Fua's group)

Computer vision group ETH, Zurich Switzerland (Luc Van Gool's group)

UCB (Malik, Darrel, Efros)

UMD (Davis, Chellappa, Jacobs, Aloimonos, Doermann)

UIUC (Forsyth, Hoiem, Ahuja, Lazebnik)

UCSD (Kriegman)

UT-Austin (Aggarwal, Grauman)

Stanford (Fei-Fei Li, Savarese)

USC (Nevatia, Medioni)

Brown (Felzenszwalb, Hays, Sudderth)

NYU (Rob Fergus)

UC-Irvine (Ramanan, Fowlkes)

UNC (Tamara Berg, Alex Berg, Jan-Michael Frahm)

Columbia (Belhumeur, Shree Nayar, Shih-Fu Chang)

Laboratory for Computational Intelligence, University of British Columbia, Vancouver (David Lowe's group)

Computer Science Department, University of Toronto, Toronto (Deep Learning fame Hilton, Srivastava, Salakhutdinov)

## 博客

Tomasz Malisiewicz blog  
The Serious Computer Vision Blog  
Research blog of Roman Shapovalov  
Computer Vision Talks  
Steves Computer Vision Blog  
The Computer Vision  
Computer Vision Blog  
Andy's Computer Vision and Machine Learning Blog  
Computer Vision Models  
solem's vision blog  
uncannyvision blog  
Blogs on Computer Vision, Machine Vision and Image Processing  
All About Computer Vision  
Open Computer Vision

## CV 工业界的实验室和创业公司

Microsoft and Google  
IBM Research  
NEC Labs America  
Acute3D (Sophia Antipolis, France) was founded in 2011.  
Bubbli  
ShoppTag  
Oculusai  
Videosurf (video search)  
Willow garage (robotics)  
Sportvision (sports broadcast)  
Intelli-vision (surveillance)  
Gauss Surgical  
Adobe's Advanced Technology Labs  
Dolby

## 如何开始研究

- 我喜欢我喜欢把计算机视觉问题分为两种类型
- 一些研究方向设计到人工智能基于学习的方法。比如图像分类，OCR，视频跟踪等
  - ◇ 大多数你能看到的论文都是这种方向的。
  - ◇ 学习意味着我们有很多数据 (e.g. 比如 ImageNet 100 万图像和他们的标签),然后学习这种模式 (e.g. 比如分类图像中的字符)
  - ◇ 对这种类型的方向，你必须学习很多机器学习的知识
- 其他研究方向涉及到不需要学习的算法，比如 3D 重建，光流计算，全景拼接

## 使用课本和课程

- 一种直接的方法是从书本开始

- 不要被困在书本里。请记住，你想开始研究。尝试了解基础知识并进行一些编码。保持你的眼睛定睛在对你来说最近有趣的工作上。
- 尝试找出不同的研究视觉问题..看哪个更令你兴奋。
- 然后你要进入下一个阶段：“从论文开始”

## 从论文开始

- 从顶级会议和期刊的论文开始。其他低级别的会议可能会有虚假结果并浪费您的时间。
  - ✧ CVPR 保留重要会议和许多论文的清单。
  - ✧ 使用文件知道什么是可用的轨道.. Wiki 也会有帮助
  - ✧ 使用 Google Scholar 查找特定问题的综述。综述可以节省大部分时间。
  - ✧ 考虑最近过去 3 年的论文。假设我们在 2014 年，考虑 2011 年，然后是 2012 年，然后是 2013 年。不要从 2014 年开始。
  - ✧ 收集文件，使标题看起来相关。搜索他们找到是否有源代码。尝试从源代码文件开始。
- 开始将是艰难的，因为你遇到了许多你不知道的术语和工具。耐心一点。谷歌搜索他们，在论坛上提问，如 Quora 或 Stackoverflow。
- 尝试找到一个特定研究方向（例如 3D 重建，点云，场景理解，物体识别，大图像数据，多目标跟踪，图像描述符理论等）。查看 wiki 或会议论文目录以查找您感兴趣的内容。
- 使用会议来了解某方向论文或使用 Google 学术搜索
- 关注那些研究工作更权威的的研究人员。关注高引用次数文献。
- 首选从有运行软件的研究工作开始，节省你的时间。
- 为了学习一些工程实现方向，请为您选择一个简单而漂亮的论文然后实现它。复现论文的结果。在这样做的时候，会有很多问题弹出，很多时候你将不得不做一些假设，因为你所看到的论文中通常并不是所有的都提到了。还有许多实现细节，比如如何有效地实现这一点不会被列出。您将了解诸如性能，实验等问题。可选择的论文比如：Viola Jones face detection, Christophe Lampert Efficient Subwindow Search, or Brian Fulkerson superpixel neighborhoods 等。实现具有完整代码的论文是一个非常好的主意，以便您可以检查自己的实现有什么问题。
- 对于你自己的研究工作，要尝试使用现有开源代码，而不是一切都从头开始，不要重复造轮子！
- 如果论文没有公开代码，你可以尝试联系作者是否可以得到代码。
- 如果理解几次尝试理解一篇论文仍然很难，就转到另一片论文。或者换一个方向。（这是你在寻找研究方向的时候）
- 这也许对你有用，最佳获奖论文集
- 研究生研讨课程取决于论文。

## 从代码开始

- 从代码到纸张，是从一些可用的代码开始理解你所研究的问题
- 找一个开源库，然后尝试它，比如 OpenCV
  - ✧ 有很多不错的书关于 OpenCV
  - ✧ Youtube 上也有不少视频：
  - ✧ [https://www.youtube.com/playlist?v=MfnEtFAWooQ&list=PLo1wvPF7fMxQ\\_SXibg1azwBfmTFn02B9O](https://www.youtube.com/playlist?v=MfnEtFAWooQ&list=PLo1wvPF7fMxQ_SXibg1azwBfmTFn02B9O)
  - ✧ <https://www.youtube.com/playlist?v=xEnPZ78quel&list=PLDqunwM5dbtIbEuXv1rB7OFBoRzEF8GH6>
  - ✧ [https://www.youtube.com/playlist?v=lwsHuSITs3c&list=PLTgRMOcmRb3PvUZpNTRsdKzVuZ4z\\_s444](https://www.youtube.com/playlist?v=lwsHuSITs3c&list=PLTgRMOcmRb3PvUZpNTRsdKzVuZ4z_s444)
  - ✧ <https://www.youtube.com/playlist?v=cgo0UitHfp8&list=PLvwB65U8V0HHCEyW2UTyOjym5FsdqfbHQ>
- 学习 Matlab 并使用它来编写初始解决方案原型（因为它往往比较快的能够开发出原型）

- Helpful: Join OpenCV yahoo group and read comments & messages.
- 选择一个有意思的 toy 项目并实现它

## 机器学习

- 机器学习是从数据中学习的核心算法。
- 对于计算机视觉而言，特别是初学者，最开始的时候你不需要学习太多机器学习。你可以像黑箱一样使用他们就够了
  - ◇ 顺便说一下，这是一个艰难的领域。要成为专家，你需要付出大量时间。
- 你想要在这个领域成长够多，你就要关注更多的细节。
- 最开始，您只需要学习一些基础知识+最近使用的算法。
- 每 4-5 年，都有一些算法在文献中流行
  - ◇ 例如 3 年前（2012 年之前，52CV 君注），SVM 非常受欢迎
  - ◇ 如今（2014 年，2015 年），深度学习往往有最好的表现。
- 建立该领域的基础知识：
  - ◇ 在 Coursera 上完成 Andrew NG 机器学习课程。
- 了解最近使用的算法是什么
  - ◇ 尝试阅读有关这些算法的更多信息
  - ◇ 尝试做一些编码。搜索流行的工具并使用它们
  - ◇ 例如对于 SVM（libsvm），CNN（Caffe）
  - ◇ 要么询问一些专业人士
  - ◇ 或在您的问题中下载 2-3 年范围内的顶级会议论文。浏览它们并且知道他们使用了什么学习算法。
  - ◇ 总的来说应该是很少重复的人。多关注他们
  - ◇ 然后
- 现在，您可以回到前面论文/书籍并继续阅读，当涉及 ML 时，您会发现主题更加容易。
- 更加加深
  - ◇ 请参阅 Andrew Ng Stanford Machine Learning Course
  - ◇ 其他网络上的视频和书籍
  - ◇ 请参阅 Mostafa 博士的“Learning From Data”视频。
  - ◇ 学习 Waleed 博士的 CS395: Pattern Recognition。
  - ◇ 教科书：Pattern Recognition and Machine Learning
  - ◇ 要更多地了解学习如何发生？
  - ◇ 了解更多算法主题和背后的数学

## 一些推荐论文

- 很难说什么是好的论文。也许就是更好地确定问题并作为参考。
- 视觉中的顶级刊物
- What are the must-read papers 什么是计算机视觉领域的必读文章？让学生在这个领域进行研究？
- 非常有用的大学课程
  - ◇ CS395T：视觉识别，2012 年秋季
  - ◇ CMPT888：2010 年夏季人类活动识别
  - ◇ CMPT882：计算机视觉中的识别问题，2009 年夏季

## 积累经验

- 在获得博士学位时，您通常会学会处理所有这些问题
- 您如何高效可靠地解决研究中的所有问题？为了了解所有这些问题，您基本上必须成为研究小组的成员几年。如果你在一个专注于物体检测的实验室里，你周围会有很多学生在解决相同的问题，在深夜与同学交谈是我知

道你可以获得专业知识的唯一途径了解：多交流打听。

- 您如何调试代码并有效调整参数？最佳实践是看更高级学生的优秀代码。在开始调试机器学习算法之前，您应该总体上熟悉调试。调试机器学习算法不像调试快速排序。如果你修正了所有的错误，你的算法可能仍然不起作用，可能是因为其他问题，比如缺乏数据，模型复杂度太低等等。坦率地说，调试视觉/学习算法更像是艺术而不是科学。
- 调整您未编写的算法或软件库的参数并非易事。您应该学会如何正确使用验证数据，了解如何运行完整的训练/评估流程，并准备好进行交叉验证。
- 你如何用个人电脑实现大规模的问题？（对于图像/视频分析，可能会有大量的数据超出你的内存，如何处理它？）一般来说，你不会实现一个大的在一台 PC 上出现问题。我在研究生院学到的最有价值的技能之一就是如何在群集中并行计算。没有群集的大学/实验室很难与拥有大中型集群的大学竞争。这也是许多教授加入 Google 和 Facebook 等组织的原因之一——他们拥有数据和计算资源，可以让高级研究人员处理越来越多的大型问题。
- 如果您无法访问大型集群，那么我会建议您在 Google 这样的地方申请实习。你会在那里学到很多东西（至少我是）。虽然你无法将自己编写的任何代码带回家，但是你会学到很多课程，这些课程会影响你作为学生的生活。如果你必须在一台机器上工作，你将不得不将数据集切割成更小的块，并逐渐将块加载到内存中

## 材料

### 在线视频和会谈

- 在线课程：离散推理和人工视觉学习
- UCF 计算机视觉视频讲座：视频
- EGGN 512 – 计算机视觉视频
- 视频讲座包括许多计算机视觉。
- 技术会谈 对于一些会议，如 ICML2011，他们主持视频中的大部分（全部）会谈。其他人，如 CVPR2011，只有选定的视频。这是了解大量近期工作而不依赖阅读报告的好方法。
- CVPR2010，他们为会谈主持了很多视频。他们也有很多夏季学校的 ML 视频。
- Wired，IEEE Spectrum，TechCrunch，TED，BigThink，Sixty Symbols，GISCIA，<http://www.youtube.com/user/GoogleTechTalks>，

### 课程

- 计算机视觉简介（斯坦福大学；李飞飞教授）相当标准的 CV 课程。
- 计算机视觉（UIUC；Forsyth 教授）相当标准的 CV 课程。
- 视觉中的基于学习的方法（CMU；Alexei Efros 教授）我学习了很多关于纹理（纹理）识别和一些使用花式 ML 技术的最先进的方法。
- 基础物体识别和场景理解（CMU；Antonio Torralba 教授）这是一个持续不断的课程，侧重于更高层次的视觉。第一场讲座看起来很有前途，但我不确定班上的其他人会是什么样子。
- 机器视觉 MIT 课程
- 计算机视觉麻省理工学院课程进展

### 计算机视觉

- 计算机视觉：模型，学习和推理 – 这是一个很好的（免费的！）预印本，主要倾向于机器学习。每个部分都提供了一套涉及的模型或机器学习工具的背景以及推理方法。开始是对必要概率和机器学习概念的深入概述。我刚开始阅读本书，但对于获取零件模型和形状模型等概述非常有用。
- 计算机视觉：算法和应用 – Richard Szeliski。一本调查书。这是更传统的教科书，在许多目前的 CV 课程中都有引用，如李飞飞的上述内容以及我校目前的 CV 课程（JHU）。
- 计算机视觉中的多视图几何 – Richard Hartley 和 Andrew Zisserman
- 计算机视觉现代方法 – David Forsyth 和 Jean Ponce
- 视觉对象识别：人工智能和机器学习综合讲座 – Kristen Grauman 和 Bastian Leibe
- 由 Trucco 和 Verri 介绍 3D 计算机视觉
- Digital Image Processing 3rd Edition by Gonzales and Woods

- 图像分析的实用算法
- <http://www.computervisiononline.com/books>

### 计算机视觉和图像处理编码

- 用 Python 编程计算机视觉 – Jan Erik Solem
- 学习 OpenCV – Gray Bradski 和 Adrian Kaehler
- 数字图像处理基础：Matlab 中的实例 – Chris Solomon 和 Toby Breckon

### 人类视觉

- 视觉：视觉信息的人类表现和处理的计算调查 – David Marr
- 迈向视觉信息理论的步骤：主动感知，信号 – 符号转换以及传感与控制之间的相互作用 – Stefano Soatto
- 基本视觉：视觉感知介绍 – 罗伯特斯诺登，彼得汤普森和汤姆 Troscianko
- 用 Python 编程计算机视觉

### 其他

- CV 论文是来自视觉会议的近期计算机视觉论文集。
- 视觉识别和机器学习暑期学校，格勒诺布尔，2012
- 我会参加一些机器学习课程，并参加信号处理/时频分析/小波分析的一些课程。

### 精彩的应用程序

- 永不停止图像学习 (NEIL)
  - ◇ 这是一个计算机程序运行 24X7 浏览互联网从互联网数据提取视觉信息。它得到了谷歌和国防部海军研究办公室的支持。
  - ◇ 它目前识别对象 – 对象关系，对象 – 属性关系，场景 – 对象关系，场景 – 属性关系
- 人脸检测
- 网球追踪
- 与深度相机的身体姿势估计
- 微软展示的 3D 扫描技术，Heads Turn
- 颜色变化显示人血流量
- 只有公共 Flickr 照片才能在 3D 中重建整个城市
- 自主物体，例如自驾车
- Predator 对象跟踪
- Kinect Fusion – 从移动 Kinect 实时 3D 模型构建
- Veebot，一个采集血液样本的机器人
- Harp：检测激光的中断以播放音符（简单，强大）。Piano。
- Google 照片搜索
- Physical security
- PTAM 是 AR 的重要应用
- 谷歌眼镜
- 谷歌街景：在街道层面捕捉世界
- Word Lens：基于增强现实相机的语言翻译应用程序。手机摄像头可以识别一种语言的文本，并显示用另一种语言翻译的文字。我发现关于这个应用程序的最好的东西是翻译是在没有连接到互联网的情况下实时执行的！
- CarSafe：该应用程序使用计算机视觉和机器学习算法来监视和检测驾驶员是否疲倦或分心，同时使用两台独立的摄像机跟踪道路状况。本文提供了一些细节和结果：CarSafe：驾驶员安全应用程序，可在智能手机上使用双摄像头检测危险驾驶行为
- iOnRoad：这是一款使用 Qualcomm FastCV 移动优化计算机视觉库的移动驾驶辅助系统应用程序。它使用智能手机的本机相机和传感器来执行各种功能。该应用程序具有先进的功能，如前方碰撞警告，车道偏离警告，车头监控和汽车定位器。
- Jumio：用于在线和移动签出的实时信用卡扫描和验证应用程序。他们还在许多国家提供护照和执照的身份证明。

## 令人兴奋的算法

- HOG 特征+线性 SVM 对物体检测非常有用。
  - ◇ 基于部件的 HOG + SVM
  - ◇ 基于范例的 HOG + SVM
- RANSAC (RANdom SAmple Consensus) – 简单/强大/鲁棒
  - ◇ 高维数据是存在低维结构内。
  - ◇ 最优随机 RANSAC
  - ◇ 与 PROSAC 匹配 – 渐进样本共识
- 霍夫变换算法
- 基于 KD 森林的近似最近邻算法
- 马尔可夫随机场
- 2D 图像拼接, 图像挖掘, 带有 SIFT 算法的纹理对象的三维重建
- SURF
- Viola-Jones: 人脸检测
- 形状上下文
- 可变形零件模型
- 同时定位和映射 Simultaneous localization and mapping

## 其它

### 工作机会

- CVPR 招聘职位
- <http://www.computervisiononline.com/jobs>
- 加入 LinkedIn 并查看图像处理或计算机视觉兴趣小组。
- Adobe 的高级技术实验室 <http://www.adobe.com/technology/> ...

### 数据集

- [点击这里](#)
- 数据集汇总
- 跟踪视频
- 网络上有太多..... Google。

### 软件

<http://www.computervisiononline.com/software>  
<http://www.computer-vision-software.com/blog/>

### 有用的网站

- 谷歌学术
  - ◇ 顶级刊物
  - ◇ Google 学术搜索可以告诉你更多关于研究人员的信息。
  - ◇ Google Scholar 可以告诉你更多关于论文的信息
- 微软学术研究
  - ◇ 您可以查看某个领域排序的顶级关键人物
  - ◇ 您可以在一个领域获得顶级会议和期刊
  - ◇ 你可以知道关于人的引用来了解工作质量。如果某人有 100 个和 100 个引用, 看起来每个作品都被 1 个人使用。另一方面, 如果引用 10000, 则平均被 100 个作品引用。第二个有更强大的参考价值。
- <http://www.scopus.com/>
- [http://wokinfo.com/products\\_tools/analytical/jcr/](http://wokinfo.com/products_tools/analytical/jcr/)
- <http://www.computervisiononline.com>
- <http://www.computervisioncentral.com/>
- <http://computervision.wikia.com>



## Ad-hocks

- ICCV Marr 奖
- 计算机视觉和商业应用
- ImageNet 挑战
- PASCAL 挑战
- Imageworld 用于发布计算机视觉，图像分析和医学图像分析领域的全球事件和学术工作机会。
- 机器人比赛
- 什么是 Deep Learning 仍然无法解决的一些计算机视觉任务？
- Awesome Computer Vision
- Awesome Deep Vision
- Emails Digest in Vision

## 链接

- 学习计算机视觉需要了解哪些数学知识？