Homepage of Mostafa S. Ibrahim

	Search this site

HOME	NEWS	RESEARCH	COMPUTER VISION

VISION ARTICLES	PROGRAMMING COMPETITIONS	MISC

Computer Vision > Vision Articles >

Getting Started in Computer Vision Research

Last update: May 21, 2015.

The *content of this page is mainly collected from the web, especially from **Quora** website. Some of texts were taken "as it is". Searching Google with one of statements hopefully will indicates its source. **Sorry** for not keeping the original source.*

In following notes, there will be too many links to check. Discovering all of them at once may be waste of time. Move to How to Start Section. Focus with it. Through discovering the field, open them from time to time to know about them.

CONTENTS

Top Conferences and Journals

- Top tier Conferences: CVPR, ECCV, ICCV, NIPS, IJCAI
- High Prestigious Second Tier Conferences: BMVC
- Prestigious Second Tier Conferences: ICIP, ACCV, ICPR, SIGGRAPH
- Top tier Journals: PAMI, IJCV
- Less Journals: CVIU, IVC
- Microsoft Academic Research list of top conferences
- Ranks from Core
- Ranks from Arnetminer
- Good source for recent papers in conferences
- List of some journal impact factors
- Journals scores from EigenFactor

Top Authors

- Microsoft Academic authors list
- Google Scholar List
- HOG features by Navneet Dalal
- Jitendra Malik.
- Gary Bradski started OpenCV
- David Lowe proposed the SIFT feature
- List of vision people (but not necessarily top authors)
- Computer Vision: Algorithms and Applications by Richard Szeliski

Top Groups

- 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)
- Centre for Vision Research, York University, Toronto

Blogs

- 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](#)

Industry Labs & Startups

- 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

How to Start in Research

- I like to divide vision topics to 2 types.
- These topics that involve AI learning. E.g. Image Classification, OCR, Video Tracking, etc
  - Most of this document about this type.
  - Learning means we have much data (e.g. 1M image and their labels), we learn the pattern (e.g. classify character in image in range 0–9)
  - For this type, you have to learn too Machine Learning. See Machine Learning Section
- Other topics than involves algorithms not learning based. E.g. 3D Reconstruction, OpticalFlow, Panorma.

Using a textbook/course

- One direct way is to start from books

- Don't stuck in books. Remember, you want to start research. Try to understand the basics and do some coding. Keep Your eyes on recent interesting work for you.

- Try to identify the different research vision problems..see which more exciting for you.

- Then move to next element: "From Papers"

From Papers

- Start with papers from **top tier conferences** and Journals. Other **low rated** conferences may **fake** results and waste your time.
  - CVPR maintains the list of the important conferences and many of their papers.
  - Use the papers to know what are the available tracks..wiki will help too
  - Use Google Scholar to find surveys on a specific problem. Survey save much of your time.
  - Initially consider the last 3 years. Say we are on 2014..consider 2011, then 2012, then 2013. Don't start with 2014.
  - Collect papers such that title seems related. Google them to find if there is **source code**. Try to start with source code papers.
- Start will be tough, as you meet many jargons & tools that you don't know. Be patient. Google them up, ask on forums like Quora or Stackoverflow.
- Try to find a specific point (E.g. 3D reconstruction, point clouds, scene understanding, object recognition, big image data, multiple target tracking, image descriptor theory, etc...). Check wiki or conference tracks to find what interests you.
- Use conferences to know its papers or use Google Scholar
- Target work authoritative researchers. Or work with high # of citations
- Preferred to start with research work that has available running software
- For **learning about engineering**, which is preferred in start, pick a simple and nice paper for you and implement it. Make your target generating same results as the paper. While doing so, many questions will popup and many times you will have to do some assumptions, as typically not everything is mentioned. Also many implementation details, like how to do this point efficiently won't be listed. You will understand issues like performance, experiments, etc. Papers examples: Viola Jones face detection, Christophe Lampert Efficient Subwindow Search, or Brian Fulkerson superpixel neighborhoods. It is a very good idea to implement a paper that has full code available, so that you could check it if you stuck or even finished but with poor implementation.

- For your **research work,** Try to build over existing code rather than implementing from scratch.
- You may contact authors for code, if it is not on public
- If after several trials to understand a paper, it still hard, move to another one. Or move to another field.
- Subset of best award [papers](#)
- A graduate seminar course that depends on [papers](#).

From Code

- From code to paper, is to start from some available codes to understand the problem
- Find an Open Source Library and try it. E.g. OpenCV
  - there are many good books
  - Youtube playlists
  - https://www.youtube.com/playlist?v=MfnEtFAWooQ&list=PLo1wvPF7fMxQ_SXibg1azwE
  - https://www.youtube.com/playlist?v=xEnPZ78queI&list=PLDqunwM5dbtlbEuXv1rB7OFB
  - https://www.youtube.com/playlist?v=lwsHuSITs3c&list=PLTgRMOcmRb3PvUZpNTRsdk
  - https://www.youtube.com/playlist?v=cgo0UitHfp8&list=PLvwB65U8V0HHCEyW2UTyOu
- Learn matlab and use it to write initial solution prototype.
- Helpful: Join OpenCV [yahoo group](#) and read comments & messages.
- Pick an interesting toy project for you and work in it

Machine Learning

- ML are the core algorithms to LEARN from the data.
- For vision, specially beginners, you don't learn much. You can also use things as black box
  - It is a hard field by the way. To be guru, you put time.
- The more you grow in the field, learn more about details.
- Initially, You should learn some basics + recent used algorithms.
- Every 4–5 years, there are some algorithms that are more popular in literature
  - E.g. 3 years ago, SVM is very popular one
  - Nowadays (2014, 15..), Deep Learning is the best in performance.
- To establish basics in the field:
  - Finish Andrew NG Machine Learning [course](#) on Coursera.
- To know what are the recent used algorithms for you
  - Either ask someone works in this problem

- Or download top conference papers in range 2–3 years in YOUR problem. Skim it and know what learning algorithms they used.
- Overall should be little who repeat alot. Focus on them
- Then
  - Try to read little more about these algorithms
  - Try to do some coding. Search for popular tools and use them
    - E.g. for SVM (libsvm), CNN(Caffe)
- Now, you can go back to the papers/book and continue reading and you will find topics are easier when ML is involved.
- To be more guru in ML
  - To understand more about how learning happens?
    - See videos of "Learning From Data" for Dr Mostafa.
    - For **Arabic** people: See CS395: Pattern Recognition for Dr Waleed.
    - Bishop book: Pattern Recognition and Machine Learning
  - To understand more algorithmic topics and math behind them
    - See Andrew Nn Standford Machine Learning Course.
    - More on web: Videos and Books

Some papers

- It is hard to say what are good papers to read. Better determine problem and follow references.
- Top publications in vision
  - You don't have to read them all. Just what is releated to your problem
- What are the must–read papers in the field of computer vision for a student in pursuing research in the field?
- Courses in universities..hopefully useful
  - CS395T: Visual Recognition, Fall 2012
  - CMPT888: Human Activity Recognition, Summer 2010
  - CMPT882: Recognition Problems in Computer Vision, Summer 2009

Acquiring Experience

- You typically learn to deal with all of these issues while getting a PhD

- **How do you address all such issues in your research efficiently and reliably ?**
  You basically have to be a member of a research group for several years in order to learn a little bit about all of these issues. If you're in a lab which focuses on object detection, there will be many students around you struggling with the same issues, and talking to fellow students in the middle of the night is the only way I know that you can gain the expertise you inquire about.

- **How do you debug the code and tune the parameters efficiently?**
  The best practices for debugging will be acquired as you look over the shoulders of more advanced students. You should be comfortable with debugging in general before you start debugging machine learning algorithms. Debugging machine learning algorithms is not like debugging quicksort. If you fix all bugs, your algorithm might still not work because of other issues such as lack of data, too low model complexity, etc. To be quite frank, debugging vision/learning algorithms is more like art than science.

  Tuning parameters of an algorithm or software library you did not write is never easy. You should learn how to use validation data correctly, have a good sense of how it takes to run the full training/evaluation pipeline, and be ready to use a cluster of computers for cross–validation.

- **How do you implement a large scale of problem with a personal PC?(For image/video analysis, there may be a large volume of data beyond your RAM, how to deal with it?)**
  In general, you don't implement a large–scale problem on a single PC. One of the most valuable skills I learned in graduate school is how to parallelize computations across a cluster. This is unfortunately one of those make–it–or–break–it skills. While not impossible, it is hard for universities/labs without clusters to compete with universities with medium–to–large sized clusters. This is also one of the reasons why so many Professors are joining organizations like Google and Facebook –– they have the data AND the computational resources to let senior researchers work on more and more ambitious large–scale problems.

  If you are unable to get access to a large cluster, then I would recommend applying for an internship at a place like Google. You will learn so much (at least I did) while you're there. While you won't be able to bring back home any code you write there, you will learn lots of scaling lessons that will impact your life as a student.

  If you have to work on a single machine, you will have to cut up the dataset into smaller chunks and incrementally load the chunks into memory.

Material

Online Videos & Talks

- Online Course: Discrete Inference and Learning in [Artificial Vision](#)
- UCF Computer Vision Video Lectures: [Videos](#)
- EGGN 512 – Computer Vision [Videos](#)
- [Video Lectures](#) includes many in computer vision.
- [Tech Talks](#) For some conferences, like [ICML2011](#), they host video for most (all?) of the talks from the event. Others, like [CVPR2011](#), only have selected videos. This is a great way to learn about a lot of recent work without solely relying on reading papers.
- [CVPR2010](#), they host a lot of videos for the talks. They also have a lot of ML videos for summer schools.
- Wired, IEEE Spectrum, TechCrunch, TED, BigThink, Sixty Symbols, [GISCIA](#), http://www.youtube.com/user/GoogleTechTalks,

Courses

- [Intro to Computer Vision](#) (Stanford; Prof Fei–Fei Li) Fairly standard CV course.
- [Computer Vision](#) (UIUC; Prof Forsyth) Fairly standard CV course.
- [Learning–based Methods in Vision](#) (CMU; Prof Alexei Efros) I learned a lot about texture (texton) recognition and some state of the art methods using fancy ML techniques.
- [Grounding Object Recognition and Scene Understanding](#) (CMU; Prof Antonio Torralba) This is an ongoing class focusing on higher level vision. The first lecture looks promising, but I'm not exactly sure what the rest of the class will be like.
- Machine Vision MIT [Course](#)
- Advances in Computer Vision MIT [Course](#)

Computer Vision

- [Computer Vision: Models, Learning, and Inference](#) – This is a great (free!) preprint that leans heavily towards machine learning. Each section provides background on a set of models or machine learning tools involved, and methods of inference. The beginning is an in–depth overview of the necessary probability and machine learning concepts. I just started going through this book but it has been really useful for getting an overview of things like parts models and shape models.
- [Computer Vision: Algorithms and Applications](#) – by Richard Szeliski. A survey book. This is more traditionally laid out textbook that is referenced in a number of current Intro to CV classes such as Fei–Fei Li's above and the current CV course at my school (JHU).
- Multiple View Geometry in Computer Vision – Richard Hartley and Andrew Zisserman

- Computer Vision a Modern Approach – David Forsyth and Jean Ponce
- Visual Object Recognition: Synthesis Lectures on Artificial Intelligence and Machine Learning – Kristen Grauman and Bastian Leibe
- Introduction to 3D Computer Vision by Trucco and Verri
- Digital Image Processing 3rd Edition by Gonzales and Woods
- Practical Algorithms for Image Analysis
- http://www.computervisiononline.com/books

Computer Vision & Image Processing Coding

- Programming Computer Vision with Python – Jan Erik Solem
- Learning OpenCV – Gray Bradski and Adrian Kaehler
- Fundamentals of Digital Image Processing: A practical approach with Examples in Matlab – Chris Solomon and Toby Breckon

Human Vision

- Vision: A Computational Investigation into the Human Representation and Processing of Visual Information – David Marr
- Steps Towards a Theory of Visual Information: Active Perception, Signal–to–Symbol Conversion and the Interplay Between Sensing and Control – Stefano Soatto
- Basic Vision: An Introduction to Visual Perception – Robert Snowden, Peter Thompson and Tom Troscianko
- Programming Computer Vision with Python

Others

- **CV Papers** is a collection of recent computer vision papers from the top/largest vision conferences.

- Visual Recognition and Machine Learning Summer School, Grenoble, 2012

- I would take a few machine learning courses and also take a few courses in signal processing/ time–frequency analysis/ wavelet analysis.

Exciting Applications

- Never Ending Image Learning(NEIL)
  - It is a computer program runs 24X7 to browse the internet to extract visual information from the internet data. It is being supported by Google and Department of Defense's Office of Naval Research.
  - It currently identifies object –object relationships, object–attribute relationships, scene–object relationships, scene –attribute relationships

- [Face detection](#)
- [Tennis ball tracking](#)
- Body pose [estimation](#) with [depth camera](#)
- [Heads Turn](#) as Microsoft Shows Off 3–D Scanning Techniques
- Color changes reveals person [blood flow](#)
- Reconstruct an [entire city](#) in 3D only by public Flickr photographs
- Autonomous objects, e.g. self driving cars
- [Predator](#) Object Tracking
- [Kinect Fusion](#) – realtime 3D model construction from a moving Kinect
- [Veebot](#), a robot that takes blood samples
- [Harp](#): Detecting the interruption of a laser to play a note (simple, powerful). [Piano](#).
- Google Photo [Search](#)
- [Physical security](#)
- [PTAM](#) is great application of AR
- Google Glass
- Google [Street View](#) : Capturing the world at street level
- [Word Lens](#) : An augmented reality camera–based language translation application. The mobile camera can identify text in one language and show the words translated in another language. The best thing I found about this application is that translation is performed in real–time without connection to the internet!
- [CarSafe](#) : This application uses computer vision and machine learning algorithms to monitor and detect whether the driver is tired or distracted, and at the same time track road conditions using two separate cameras. Some details and results are provided in the paper here: [CarSafe: A Driver Safety App that Detects Dangerous Driving Behavior using Dual–Cameras on Smartphones](#)
- [iOnRoad](#) : This is a mobile driver assistance system application using Qualcomm's FastCV mobile–optimized computer vision library. It uses the smartphone's native camera and sensors to perform various functions. The application has advanced features such as forward collision warning, lane departure warning, headway monitoring and car locator.
- [Jumio](#) : A real–time credit card scanning and validation application for online and mobile checkouts. They also provide ID verification of passports and licences in many countries.

Exciting Algorithm

- [HOG](#) features + linear SVMs are quite powerful for object detection.
  - [Part–based HOG+SVM](#)
  - [Exemplar–based HOG+SVM](#)
- [RANSAC](#) (RANdom SAmple Consensus) – Simple / Powerful / Robust

- There is low dimensional structure in your high dimensional data. Go find it.
  - [Optimal](#) Randomized RANSAC
  - [Matching with PROSAC](#) ─ Progressive Sample Consensus
- [Hough transform algorithm](#)
- Approximate Nearest Neighbor [algorithm](#) based on KD–Forests
- Markov random fields
- 2D image stitching, image mining, 3D reconstruction of textured objects with [SIFT](#) like algorithms
- [SURF](#)
- Viola–Jones: [face detector](#)
- Shape [Contexts](#)
- [Deformable](#) Part Models
- Simultaneous localization and mapping

Others

Jobs
- CVPR Jobs Postings
- http://www.computervisiononline.com/jobs
- Join LinkedIn and look at the Image Processing or Computer Vision interest groups.
- Adobe's Advanced Technology Labs [http://www.adobe.com/technology/...](#)

Datasets
- Check from here
- Collection
- Tracking Videos
- there are too many on web..Google.

Software
- My list
- http://www.computervisiononline.com/software
- http://www.computer–vision–software.com/blog/

Deadlines
- [Events](#)
- [Calendar](#)

Helpful sites
- Google Scholar
  - [Top publications](#)
  - Google Scholar could tell you [more about persons](#).

- Google Scholar could tell you more about papers
- Microsoft Academic Research
  - You could get sorted list of top key figures in a field
  - You get top conferences and journals in a field
  - You could know about citation of person to know quality of work. If someone has 100 and 100 citation, it seems every work is used by 1 person. On the other hand, if cited by 10000, it is on average cited by 100 work. Second one is stronger advisor.
- http://www.scopus.com/
- http://wokinfo.com/products_tools/analytical/jcr/
- http://www.computervisiononline.com
- http://www.computervisioncentral.com/
- http://computervision.wikia.com

Ad–hocks

- ICCV Marr Prize
- Computer vision and commercial applications
- ImageNet challenge
- PASCAL challenge
- Imageworld is used to announce worldwide events and academic vacancies within the field of Computer Vision, Image Analysis, and Medical Image Analysis.
- The Great Robot Race

- What are some computer Vision tasks that Deep Learning still does not tackle well?

- Awesome Computer Vision
- Awesome Deep Vision
- Emails Digest in Vision

Links

- What mathematics does one need to know to learn about computer vision?

Comments

You do not have permission to add comments.

Sign in | Recent Site Activity | Report Abuse | Print Page | Powered By **Google Sites**