

# Computer Vision Papers

## Perception

The Plenoptic Function and the Elements of Early Vision, Adelson and Bergen, 1991:

<http://www.cs.cmu.edu/~efros/courses/LBMV07/Papers/adelson-elements-91.pdf>

On Seeing Stuff: The Perception of Materials by Humans and Machines, Adelson, 2001:

[http://web.mit.edu/persci/people/adelson/pub\\_pdfs/adelson\\_spie\\_01.pdf](http://web.mit.edu/persci/people/adelson/pub_pdfs/adelson_spie_01.pdf)

Throwing Down the Visual Intelligence Gauntlet. in Machine Learning for Computer Vision, Tan, Leibo, and Poggio, 2012:

<http://cbcl.mit.edu/publications/ps/MIT-CSAIL-TR-2012-016.pdf>

## Imaging

Accidental Pinhole and Pinspeck Cameras: Revealing the Scene Outside the Picture, Torralba, 2012:

<http://people.csail.mit.edu/torralba/publications/shadows.pdf>

Recovering Intrinsic Images with a Global Sparsity Prior on Reflectance, Gehler, 2011:

<http://people.tuebingen.mpg.de/mkiefel/projects/intrinsic/nips11intrinsic.pdf>

## Features, Segmentation and Detection

Constrained Convolutional Neural Networks for Weakly Supervised Segmentation by Pathak, Krahenbuhl, Darrell, 2015:

<http://www.robots.ox.ac.uk/~vgg/rg/papers/ccnn.pdf>

CPMC: Automatic Object Segmentation Using Constrained Parametric MinCuts Carreira and Sminchisescu, 2010:

<http://www.maths.lth.se/matematiklth/personal/sminchis/papers/cs-cvpr10.pdf>

HOGgles: Visualizing Object Detection Features, Vondrick et. al., 2013:

[http://people.csail.mit.edu/khosla/papers/iccv2013\\_ihog.pdf](http://people.csail.mit.edu/khosla/papers/iccv2013_ihog.pdf)

Diagnosing Error in Object Detectors, Hoiem, 2012:

[http://dhoiem.cs.illinois.edu/publications/eccv2012\\_detanalysis\\_derek.pdf](http://dhoiem.cs.illinois.edu/publications/eccv2012_detanalysis_derek.pdf)

Regionlets for Generic Object Detection, Wang et. al., 2013:

[http://www.ece.northwestern.edu/~mya671/mypapers/ICCV13\\_Wang\\_Yang\\_Zhu\\_Lin.pdf](http://www.ece.northwestern.edu/~mya671/mypapers/ICCV13_Wang_Yang_Zhu_Lin.pdf)

Scale and Affine Invariant Interest Point Detectors. Mikolajczyk and Schmid, 2004:

[http://www.robots.ox.ac.uk/~vgg/research/affine/det\\_eval\\_files/mikolajczyk\\_ijcv2004.pdf](http://www.robots.ox.ac.uk/~vgg/research/affine/det_eval_files/mikolajczyk_ijcv2004.pdf)

On SIFTs and Their Scales, Hassner, 2012:

<http://www.openu.ac.il/home/hassner/projects/siftscales/OnSiftsAndTheirScales-CVPR12.pdf>

## **Tracking and Flow**

Scene Chronology, Matzen and Snavely, 2014:

<http://www.cs.cornell.edu/projects/chronology/>

Tracking-Learning-Detection, Kalal, Mikolajczyk, and Matas, 2010:

[http://epubs.surrey.ac.uk/713800/1/Kalal-PAMI-2011\(1\).pdf](http://epubs.surrey.ac.uk/713800/1/Kalal-PAMI-2011(1).pdf)

Track to the Future: Spatiotemporal Video Segmentation with Long-Range Motion Cues, Lezema, 2011:

<http://www.di.ens.fr/~josef/publications/lezama11.pdf>

On Space-Time Interest Points, Ivan Laptev, 2005:

[http://www.irisa.fr/vista/Papers/2005\\_ijcv\\_laptev.pdf](http://www.irisa.fr/vista/Papers/2005_ijcv_laptev.pdf)

Event Retrieval in Large Video Collections with Circulant Temporal Encoding, Revaud et. al., 2013:

[http://hal.archives-ouvertes.fr/docs/00/80/17/14/PDF/revaud\\_event.pdf](http://hal.archives-ouvertes.fr/docs/00/80/17/14/PDF/revaud_event.pdf)

DeepFlow: Large Displacement Optical Flow with Deep Matching, Weinzaepfel et. al., 2013:

[http://hal.archives-ouvertes.fr/docs/00/87/35/92/PDF/DeepFlow\\_iccv2013.pdf](http://hal.archives-ouvertes.fr/docs/00/87/35/92/PDF/DeepFlow_iccv2013.pdf)

## **Recognition and Detection**

What Makes for Effective Detection Proposals?, J. Hosang et. al., 2015:

<http://arxiv.org/pdf/1502.05082.pdf>

Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation, Girshick, Donahue, Darrell, Malik, 2014:

<http://arxiv.org/pdf/1311.2524v5.pdf>

Matching Local Self-Similarities across Images and Videos, Shechtman, Irani, 2007:

[http://www.wisdom.weizmann.ac.il/~vision/VideoAnalysis/Demos/SelfSimilarities/SelfSimilarities\\_ShechtmanIrani07.pdf](http://www.wisdom.weizmann.ac.il/~vision/VideoAnalysis/Demos/SelfSimilarities/SelfSimilarities_ShechtmanIrani07.pdf)

What is an Object?, Alexe, Deselaers, Ferrari, 2010:

[ftp://ftp.vision.ee.ethz.ch/publications/proceedings/eth\\_biwi\\_00728.pdf](ftp://ftp.vision.ee.ethz.ch/publications/proceedings/eth_biwi_00728.pdf)

Object Detectors Emerge in Deep Scene CNNs, Bolei et. al, 2015:

[http://www.robots.ox.ac.uk/~vgg/rg/papers/zhou\\_iclr15.pdf](http://www.robots.ox.ac.uk/~vgg/rg/papers/zhou_iclr15.pdf)

Cascade Object Detection with Deformable Part Models, Felzenszwalb, 2010:

<http://cs.brown.edu/~pff/papers/cascade.pdf>

You only look once: Unified, real-time object detection, Redmon et. al., 2016:  
<http://pjreddie.com/media/files/papers/yolo.pdf>

The Devil is in the Details: an Evaluation of Recent Feature Encoding Methods, Chatfield, 2011:  
[http://www.robots.ox.ac.uk/~vgg/research/encoding\\_eval/](http://www.robots.ox.ac.uk/~vgg/research/encoding_eval/)

Aggregating Local Image Descriptors into Compact Codes. Jégou, 2012:  
[http://hal.inria.fr/docs/00/63/30/13/PDF/jegou\\_aggregate.pdf](http://hal.inria.fr/docs/00/63/30/13/PDF/jegou_aggregate.pdf)

BING: Binarized Normed Gradients for Objectness Estimation at 300fps, Cheng et. al., 2014:  
<http://mmcheng.net/mftp/Papers/ObjectnessBING.pdf>

In Defense of Nearest Neighbor Based Image Classification, Boiman, 2008:  
[http://www.wisdom.weizmann.ac.il/~irani/PAPERS/InDefenceOfNN\\_CVPR08.pdf](http://www.wisdom.weizmann.ac.il/~irani/PAPERS/InDefenceOfNN_CVPR08.pdf)

## **3D Reconstruction and Recognition**

DynamicFusion: Reconstruction and Tracking of Non-rigid Scenes in Real-Time, Newcombe et. al., 2015:  
<http://grail.cs.washington.edu/projects/dynamicfusion/papers/DynamicFusion.pdf>

Structured Indoor Modeling, Ikehata et. al., 2015:  
[http://www.cse.wustl.edu/~sikehata/sim/iccv2015\\_ikehata.pdf](http://www.cse.wustl.edu/~sikehata/sim/iccv2015_ikehata.pdf)

Sliding Shapes for 3D Object Detection in Depth Images, Shuran Song Jianxiong Xiao, 2014:  
<http://slidingshapes.cs.princeton.edu/paper.pdf>

Real-Time Human Pose Recognition in Parts from Single Depth Images, Shotton et al. 2011:  
<http://research.microsoft.com/pubs/145347/BodyPartRecognition.pdf>

MultiView Stereo Revisited, Goesele, 2006:  
<http://www.gcc.tu-darmstadt.de/media/gcc/papers/Goesele-2006-MSR.pdf>

KinectFusion: Realtime 3D Reconstruction and Interaction Using a Moving Depth Camera, Izadi, 2011:  
<http://research.microsoft.com/pubs/155416/kinectfusion-uist-comp.pdf>

Building Rome in a Day, Agarwal, 2009:  
[http://grail.cs.washington.edu/rome/rome\\_paper.pdf](http://grail.cs.washington.edu/rome/rome_paper.pdf)

Depth estimation from image structure, Torralba, 2001:  
<http://cvcl.mit.edu/Papers/Torralba-Oliva02.pdf>

Make3D: Learning 3D Scene Structure from a Single Still Image, Saxena, 2008:

[http://ai.stanford.edu/~asaxena/reconstruction3d/saxena\\_make3d\\_learning3dst ructure.pdf](http://ai.stanford.edu/~asaxena/reconstruction3d/saxena_make3d_learning3dst ructure.pdf)

Recovering Surface Layout from an Image, Hoiem, 2007:  
[http://www.ri.cmu.edu/publication\\_view.html?pub\\_id=5818](http://www.ri.cmu.edu/publication_view.html?pub_id=5818)

## **Faces, People and Crowds**

Modeling Video Evolution for Action Recognition, Fernando et. al., 2015:  
<http://www.robots.ox.ac.uk/~vgg/rg/papers/videoDarwin.pdf>

Deep Convolutional Network Cascade for Facial Point Detection, Yi Sun, Xiaogang Wang, Xiaoou Tang, 2013:  
<http://www.ee.cuhk.edu.hk/~xgwang/papers/sunWTCvpr13.pdf>

Dense Trajectories and Motion boundary histograms for Action Recognition, Wang et al., 2013:  
<https://hal.inria.fr/hal-00803241/document>

Recurrent Network Models for Human Dynamics, Fragkiadaki et. al., 2015,  
[http://www.cv-foundation.org/openaccess/content\\_iccv\\_2015/papers/Fragkiadaki\\_Recurrent\\_Network\\_Models\\_ICCV\\_2015\\_paper.pdf](http://www.cv-foundation.org/openaccess/content_iccv_2015/papers/Fragkiadaki_Recurrent_Network_Models_ICCV_2015_paper.pdf)

Cumulative Attribute Space for Age and Crowd Density Estimation, Chen, Gong, and Xiang, 2013:  
[http://www.robots.ox.ac.uk/~vgg/rg/papers/Chen\\_Cumulative\\_Attribute\\_Space\\_CVPR13.pdf](http://www.robots.ox.ac.uk/~vgg/rg/papers/Chen_Cumulative_Attribute_Space_CVPR13.pdf)

Face Recognition by Humans: Nineteen Results All Computer Vision Researchers Should Know About, Sinha, 2006:  
[http://web.mit.edu/sinhalab/Papers/19results\\_sinha\\_etal.pdf](http://web.mit.edu/sinhalab/Papers/19results_sinha_etal.pdf)

Histograms of Oriented Gradients for Human Detection, Dalal and Triggs, 2005:  
<http://lear.inrialpes.fr/people/triggs/pubs/Dalal-cvpr05.pdf>

Action Recognition with Improved Trajectories, Wang and Schmid, 2013:  
[http://hal.archives-ouvertes.fr/docs/00/87/35/20/PDF/wang\\_iccv13.pdf](http://hal.archives-ouvertes.fr/docs/00/87/35/20/PDF/wang_iccv13.pdf)

## **Other Convolutional Networks in Computer Vision**

Deep Residual Learning for Image Recognition, He, Zhang, Ren, Sun, 2015:  
<http://www.robots.ox.ac.uk/~vgg/rg/papers/deepres.pdf>

Visualizing and Understanding Convolutional Networks, Zeiler and Fergus, 2014,  
<http://arxiv.org/abs/1311.2901>

OverFeat: Integrated Recognition, Localization and Detection using Convolutional Networks, Sermanet et. al., 2014:  
<http://arxiv.org/pdf/1312.6229v4.pdf>

CNN Features off the Shelf: an Astounding Baseline for Recognition, Razavian et al, 2014:

<http://arxiv.org/pdf/1403.6382.pdf>

Learning to Generate Chairs with Convolutional Neural Networks, Dosovitskiy et. al., 2015:

[http://www.robots.ox.ac.uk/~vgg/rg/papers/Dosovitskiy\\_Learning\\_to\\_Generate\\_2015\\_CVPR\\_paper.pdf](http://www.robots.ox.ac.uk/~vgg/rg/papers/Dosovitskiy_Learning_to_Generate_2015_CVPR_paper.pdf)

Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks, Ren, He, Girshick, and Sun, 2016:

<http://arxiv.org/pdf/1506.01497v3.pdf>