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

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

Diagnosing Error in Object Detectors, Hoiem, 2012:
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

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

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

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

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

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

InterpoNet, A brain inspired neural network for optical flow dense interpolation, 2017, Zweig and Wolf:
http://arxiv.org/pdf/1611.09803v3.pdf

FlowNet: Learning Optical Flow with Convolutional Networks, Dosovitskiy et al., 2015:
http://lmb.informatik.uni-freiburg.de/Publications/2015/DFIB15/flownet.pdf

Optical Flow Requires Multiple Strategies (but only one network), Schuster, Wolf and Gadot, 2017:
http://arxiv.org/pdf/1611.05607v3.pdf

Optical Flow Estimation using a Spatial Pyramid Network, Ranjan and Black, 2016:
http://arxiv.org/pdf/1611.00850v2.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

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

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/

Aggregating Local Image Descriptors into Compact Codes. Jégou, 2012:
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

YOLO9000: Better, Faster, Stronger, Redmon and Farhadi, 2016: http://arxiv.org/pdf/1612.08242v1.pdf

Mask R-CNN, He et al., 2017: http://arxiv.org/pdf/1703.06870v1.pdf

SSD: Single Shot MultiBox Detector, Liu et al, 2016:
http://arxiv.org/pdf/1512.02325v5.pdf

Speed/accuracy trade-offs for modern convolutional object detectors, Huang et al., 2016: http://arxiv.org/pdf/1611.10012v1.pdf

R-FCN: Object Detection via Region-based Fully Convolutional Networks, Dai et al.: https://arxiv.org/pdf/1605.06409.pdf

Learning Rich Features from RGB-D Images for Object Detection and Segmentation, Gupta et al, 2014:
https://people.eecs.berkeley.edu/~sgupta/pdf/rcnn-depth.pdf

Image-to-Image Translation with Conditional Adversarial Nets, Isola et al, 2016:
https://arxiv.org/pdf/1611.07004v1.pdf
https://phillipi.github.io/pix2pix/

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

BundleFusion: Real-time Globally Consistent 3D Reconstruction using Online Surface Re-integration, Dai et al, 2016:
http://graphics.stanford.edu/projects/bundlefusion/
https://arxiv.org/pdf/1604.01093v1.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

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_learning3dstructure.pdf

Recovering Surface Layout from an Image, Hoiem, 2007:
http://www.ri.cmu.edu/publication_view.html?pub_id=5818

SegICP: Integrated Deep Semantic Segmentation and Pose Estimation, 2017:
http://arxiv.org/pdf/1703.01661v1.pdf

SemanticFusion: Dense 3D Semantic Mapping with Convolutional Neural Networks, 2016: http://arxiv.org/pdf/1609.05130v2.pdf

3D Face Reconstruction by Learning from Synthetic Data, 2016:
http://arxiv.org/pdf/1609.04387v2.pdf

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

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

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

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

3D Face Morphable Models "In-the-Wild", 2017:
http://arxiv.org/pdf/1701.05360v1.pdf

Other Convolutional Networks in Computer Vision

A New Method to Visualize Deep Neural Networks, 2016:
http://arxiv.org/pdf/1603.02518v2.pdf

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

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

Densely Connected Convolutional Networks, 2016:
http://arxiv.org/pdf/1608.06993v3.pdf

Inception-v4, Inception-ResNet and the Impact of Residual Connections on
Learning, 2016: http://arxiv.org/pdf/1602.07261v2.pdf

Rethinking the Inception Architecture for Computer Vision (Inception-v3),
2015: http://arxiv.org/pdf/1512.00567v3.pdf