Xiaofeng Ren

Amazon Go

xiaofenr at amazon dot com

We have launched!!! I am a senior principal scientist at Amazon. Since 2013, I have been the lead scientist at Amazon Go, using computer vision and machine learning to re-invent retail. We have launched our first "Just-Walk-Out" store that automatically figures out purchases without customer effort, completely eliminating check-out (that "unnecessary" and annoying wait). (CV,Google Scholar).

I am interested in all aspects of computer vision, as I believe all are needed to solve it. Prior to Amazon, I worked on using RGB-D (color+depth, a.k.a. Kinect style) cameras, ranging from 3D mapping and modeling to everyday object recognition. I worked on many other vision problems, including image descriptors, boundary detection, image segmentation, figure-ground grouping, object and pose recognition, human body detection and pose estimation, object segmentation and tracking, and optical flow. I had opportunities to work on vision-related problems in robotics and human-computer interaction.

I was a research scientist at Intel Labs during 2008-2013, working closely with faculty and students at University of Washington. Prior to sunny Seattle, I was a research assistant professor at the Toyota Technological Institute at Chicago (TTI-C). I received my Ph.D. from U.C. Berkeley in 2006, under the supervision of Jitendra Malik.

(Old) Updates

- **Code available**: here is a preliminary version of the training code for our NIPS2012 paper on sparse code based contour detection.
- **Code available**: learned sparse coding features for object detection (CVPR13), replacing HOG (30%=>34% on PASCAL 2007).
- **Code available**: I have released a version of contour detection algorithm using sparse code gradients (NIPS2012) , with experimental setups for both BSDS500 (color only) and NYU Depth v2 (RGB-D). Enjoy!!
- **Code available**: I have released a version of our scene labeling algorithms using kernel descriptors (CVPR12), with experimental setups for both Stanford Background (outdoor, color only) and NYU Depth (indoor, RGB-D) datasets.
- **Code available**: I have put together C++ implementations of both our kernel descriptor features (NIPS10) and the more recent hierarchical sparse coding features (NIPS11) in a live webcame demo, here released under BSD license. What's more, the demo is now running on an Android phone!!
- Our book on Consumer Depth cameras (built on the CDC4CCV workshop at ICCV 2011) is now published by Springer.
- Progress on contour detection: move beyond Pb and use sparse codes to compute local oriented gradients. F=**0.74** (up from 0.71 of gPb) on BSDS500, a large step forward (human=0.80). Great RGB-D results: F=0.62 (vs gPb 0.53) on NYU Depth (v2).
- Check out the demo video for our Ubicomp paper on fine-grained kitchen activity recognition and tracking. News article at New Scientist.
- Co-organized the third installment of the 3rd RGB-D workshop at RSS. All papers are online.
- A C++ implementation of our kernel descriptor features (NIPS10, IROS11), ~5 times faster than the matlab version, and with a live demo using webcams. Please try this out and let me know your comments and suggestions.
- Upcoming papers at ISER, Ubicomp and the Robotics and Automation Magazine (RAM).
- Second Workshop on Consumer Depth Cameras for Computer Vision, at ECCV 2012.
- Co-organizing the Second Workshop on Egocentric Vision at this CVPR. We welcome both full papers (Apr 17) and extended abstracts (Apr 24). Please consider submitting your work in progress!
- CVPR paper accepted on scene labeling on both RGB-D (indoor) and image-only (outdoor) scenes. Preprint available.
- Liefeng's NIPS paper on hierarchical orthogonal matching pursuit for learning image features.
- The (Matlab) code for kernel descriptors is now available. Please try it out!!
- Check out our ubicomp final video on interactive mapping on youtube.
- Co-organized the 2nd RGB-D workshop at RSS on advanced perception using depth cameras; it's a great success!! 18 presentations, 7 demos and over 70 attendees. All the papers, videos, slides will be available online.
- Two BMVC papers to appear on material recognition and video segmentation.
- Two IROS papers to appear on depth kernel descriptors and object discovery based on scene changes.
- Our interactive mapping work (online user interaction in real-time mapping) to appear at Ubicomp 2011.
- Our sparse distance learning paper won the best vision paper award at ICRA!!
- Co-organizing the 2nd RGB-D workshop at RSS on advanced perception using depth cameras, and the ICCV Workshop on Consumer Depth Cameras.
- Our RGB-D dataset (ICRA 2011) is now available online covering 300 everyday objects. Send us comments!
- Two papers to appear at CVPR 2011 on kernel descriptors and egocentric video.
- Three papers accepted at ICRA 2011 on object recognition and discovery. Congratulations Kevin and Evan.
- OASIS / Interactive LEGO demo at CES, robust RGB-D recognition using kernel descriptors.
- HeatWave paper as **honorable mention** at CHI, congratulatiosn Eric, Gabe and Sidhant!!
- Kernel Descriptor paper at NIPS 2010, framework for local features beyond SIFT with strong recognition results.
- Colloquium talk at the UW CSE Deparment on RGB-D Perception beyond gestures.
- New RGBDvision channel on youtube (~300,000 clicks already).
- Affiliate assistant professor appointment at the CSE Deparment, University of Washington.
- Co-organized the RGB-D Perception workshop at RSS 2010 (all slides online).
- Co-organized the Egocentric vision workshop at CVPR 2009 (presented the Intel egocentric object dataset).

Publications

- **Depth Enhancement via Low-rank Matrix Completion**. [pdf] [project]
  Si Lu, Xiaofeng Ren, Feng Liu, in CVPR 2014.
  Ever unhappy with poor-quality depth data? Si and Feng has a solution!

- **Change Their Perception: RGB-D for 3-D Modeling and Recognition**. [read online]
  Xiaofeng Ren, Dieter Fox, and Kurt Konolige, in IEEE Robotics and Automation Magazine (RAM), 2013.
  Summary of our RGB-D work on 3D modeling and recognition for a broad audience

- **Histograms of Sparse Codes for Object Detection**. [abstract] [pdf] [code]
  Xiaofeng Ren and Deva Ramanan, at CVPR 2013.
  Move beyond HOG! Use learned sparse code dictionaries to significantly improve object detection accuracy

- **Multipath Sparse Coding Using Hierarchical Matching Pursuit**. [abstract] [pdf] [code]
  Liefeng Bo, Xiaofeng Ren and Dieter Fox, at CVPR 2013.
  Extend hierarchical matching pursuit (NIPS11) to a reconfigurable architecture that captures structures of varying scale and deformation

- **RGB-D Flow: Dense 3-D Motion Estimation Using Color and Depth**. [abstract] [pdf] [code]
  Evan Herbst, Xiaofeng Ren and Dieter Fox, at IROS 2013.
  General motion estimation using Kinect; extend variational optical flow to RGB-D data

- **Discriminatively Trained Sparse Code Gradients for Contour Detection**. [abstract] [pdf] [test code] [training code]
  Xiaofeng Ren and Liefeng Bo, at NIPS 2012.
  Pushing the limit of local features for contour detection; a large step forward (0.71=>0.74 on BSDS500, human=0.8)

- **SensorSift: Balancing Sensor Data Privacy and Utility in Automated Face Understanding**. [abstract] [pdf]
  Miro Enev, Jaeyeon Jung, Liefeng Bo, Xiaofeng Ren and Tadayoshi Khono, at the Annual Computer Security Applications Conference (ACSAC), 2012.

- **Ontology Guided Approach to Retrieving Disease Manifestation Images for Health Image Base Construction**. (best paper) [abstract]
  Yang Chen, Xiaofeng Ren, Guo-Qiang Zhang, Rong Xu, at Annual IEEE Healthcare Informatics, Imaging, and Systems Biology Conference (HISB), 2012.

- **Fine-Grained Kitchen Activity Recognition using RGB-D**. [abstract] [pdf] [video(wmv)] [news]
  Jinna Lei, Xiaofeng Ren, and Dieter Fox, at Ubicomp 2012.
  Prototype system for tracking objects and recognizing actions and activities for smart kitchen applications

- **Unsupervised Feature Learning for RGB-D Based Object Recognition**. [abstract] [pdf] [C++ code]
  Liefeng Bo, Xiaofeng Ren, and Dieter Fox, at ISER 2012.
  Extending our fast feature learning using hierarchical matching pursuit (NIPS '11) to RGB-D data

- **RGB-(D) Scene Labeling: Features and Algorithms**. [abstract] [pdf] [code]
  Xiaofeng Ren, Liefeng Bo, and Dieter Fox, at CVPR 2012.
  RGB (and D) scene labeling, 76% on NYU Depth (up from 56%) and 83% on Stanford Background (from 79%)

- **Detection-based Object Labeling in 3D Scenes**. [abstract] [pdf] [video]
  Kevin Lai, Liefeng Bo, Xiaofeng Ren and Dieter Fox, at ICRA 2012.
  Detecting 3D objects in RGB-D frames and segmenting them in merged 3D point clouds (using RGB-D dataset)

- **RGB-D Mapping: Using Depth Cameras for Dense 3D Modeling of Indoor Environments** . [abstract] [pdf]
  Peter Henry, Michael Krainin, Evan Herbst, Xiaofeng Ren and Dieter Fox, at International Journal of Robotics Research (IJRR), 2012.
  Journal version of our RGB-D mapping system (ISER), with extended algorithm comparisons and results

- **Hierarchical Matching Pursuit for Image Classification: Architecture and Fast Algorithms**. [abstract] [pdf] [C++ code] [Android code]
  Liefeng Bo, Xiaofeng Ren and Dieter Fox, at NIPS 2011.
  Learning patch features using matching pursuit (sparse coding), 2 orders of magnitude faster than prior work

- **Toward Robust Material Recognition for Everyday Objects**. [abstract] [pdf]
  Diane Hu, Liefeng Bo, Xiaofeng Ren, at BMVC 2011.
  Real-world material recognition, 54% on MIT Flickr Dataset (up from 45%)

- **Combining Self Training and Active Learning for Video Segmentation**. [pdf]
  Alireza Fathi, Maria Florina Balcan, Xiaofeng Ren and Jim Rehg, at BMVC 2011.
  Simple but effective semi-supervised learning for segmenting out objects in video

- **Depth Kernel Descriptors for Object Recognition**. [pdf] [C++ code] [matlab code]
  Liefeng Bo, Xiaofeng Ren and Dieter Fox, at IROS 2011.
  Designing kernel descriptors for recognition in depth images, 53%=>79% on RGB-D benchmark

- **RGB-D Object Discovery via Multi-Scene Analysis**. [pdf]
  Evan Herbst, Xiaofeng Ren and Dieter Fox, at IROS 2011.
  Enable a robot to automatically discover and cluster objects via multiple visits to a scene; works on all objects

- **Interactive 3D Modeling of Indoor Environments with a Consumer Depth Camera**. [abstract] [pdf] [video]
  Hao Du, Peter Henry, Xiaofeng Ren, Marvin Cheng, Dan Goldman, Steve Seitz, Dieter Fox, at Ubicomp 2011.
  First interactive system for RGB-D mapping, runs near real-time and allows user feedback and control on-the-spot

- **HeatWave: Thermal Imaging for Surface User Interaction**. (honorable mention) [abstract] [pdf]
  Eric Larson, Gabe Cohn, Sidhant Gupta, Xiaofeng Ren, Beverly Harrison, Dieter Fox, Shwetak N. Patel at CHI 2011.
  All the cool things you can do with a heat camera! I wish they are mass produced...

- **A Scalable Tree-based Approach for Joint Object and Pose Recognition**. [abstract] [pdf]
  Kevin Lai, Liefeng Bo, Xiaofeng Ren and Dieter Fox, at AAAI 2011.
  Joint tree model for category, instance and pose recognition; sequential decision.

- **Object Recognition with Hierarchical Kernel Descriptors**. [abstract] [pdf]
  Liefeng Bo, Kevin Lai, Xiaofeng Ren and Dieter Fox, at CVPR 2011.
  Use the kernel descriptor framework (NIPS10) to both pixel->patch and patch->image

- **Learning to Recognize Objects in Egocentric Activities**. [abstract] [pdf]
  Alireza Fathi, Xiaofeng Ren and Jim Rehg, at CVPR 2011.
  Focus on object-in-hand in egocentric video; clustering and discovery of objects in activities

- **A Large-Scale Hierarchical Multi-View RGB-D Object Dataset**. [abstract] [pdf] [dataset]
  Kevin Lai, Liefeng Bo, Xiaofeng Ren and Dieter Fox, at ICRA 2011.
  RGB-D object recognition benchmark: 51 categories, 300 objects, 250K total views

- **Sparse Distance Learning for Object Recognition Combining RGB and Depth Information**. (best vision paper) [abstract] [pdf]
  Kevin Lai, Liefeng Bo, Xiaofeng Ren and Dieter Fox, at ICRA 2011.
  Local distance learning and feature selection using instance-to-class distance

- **Toward Object Discovery and Modeling via 3-D Scene Comparison**. [abstract] [pdf]
  Evan Herbst, Xiaofeng Ren and Dieter Fox, at ICRA 2011.
  How can a robot discover objects robustly? By visiting a scene and finding out the changes

- **Kernel Descriptors for Visual Recognition**. [abstract] [pdf] [C++ code] [matlab code and more info]
  Liefeng Bo, Xiaofeng Ren and Dieter Fox, at NIPS 2010.
  Flexible way to derive local descriptors; significantly outperform SIFT on many benchmarks

- **RGB-D Mapping: Using Depth Cameras for Dense 3D Modeling of Indoor Environments** . [abstract] [pdf] [news]
  Peter Henry, Michael Krainin, Evan Herbst, Xiaofeng Ren and Dieter Fox, at ISER 2010.
  Our first RGB-D paper, 3D modeling of large indoor environments with RGB-D cameras

- **Discriminative Mixture-of-Templates for Viewpoint Classification**. [abstract] [pdf]
  Chunhui Gu and Xiaofeng Ren, at ECCV 2010, Crete, Greece, 2010.
  First paper on discriminative models for viewpoint/pose recognition, large improvement: 57%=>74% on 3DObject

- **Manipulator and Object Tracking for In-Hand 3D Object Modeling**. [pdf]
  Michael Krainin, Peter Henry, Xiaofeng Ren and Dieter Fox, in International Journal of Robotics Research (IJRR), 2011.
  Enable a robot to automatically study objects and build 3D models through manipulation

- **Manipulator and Object Tracking for In Hand Model Acquisition**. [pdf]
  Michael Krainin, Peter Henry, Xiaofeng Ren and Dieter Fox, at the Mobile Manipulation and Best Practices in Robotics Workshops at ICRA 2010.

- **Figure-Ground Segmentation Improves Handled Object Recognition in Egocentric Video**. [abstract] [pdf] [video] [dataset]
  Xiaofeng Ren and Chunhui Gu, at CVPR 2010, San Francisco, 2010.
  Egocentric recognition can work!! ~90% accuracy on a very challenging dataset for objects-in-hand. Check out the videos.

- **Egocentric Recognition of Handled Objects: Benchmark and Analysis**. [abstract] [pdf] [dataset]
  Xiaofeng Ren and Matthai Philipose, in *Egovision Workshop '09*, Miami, 2009.
  Can we recognize objects in a user's hand? A large benchmark (43 objects, 2 hours of video) using a wearable camera

- **Multi-Scale Improves Boundary Detection in Natural Images**. [abstract] [pdf]
  Xiaofeng Ren, in *ECCV '08*, Marseille, 2008.
  Multi-scale does help contour detection on natural images - extensive benchmarking and analysis

- **Finding People in Archive Films through Tracking**. [abstract] [pdf] [video]
  Xiaofeng Ren, in *CVPR '08*, Anchorage, 2008.

- **Local Grouping for Optical Flow**. [abstract] [pdf]
  Xiaofeng Ren, in *CVPR '08*, Anchorage, 2008.

- **Tracking as Repeated Figure/Ground Segmentation**. [abstract] [pdf] [ps] [bibtex] [video]
  Xiaofeng Ren and Jitendra Malik, in *CVPR '07*, Minneapolis 2007.

- **Learning and Matching Line Aspects for Articulated Objects**. [abstract] [pdf] [ps] [bibtex]
  Xiaofeng Ren, in *CVPR '07*, Minneapolis 2007.

- **Figure/Ground Assignment in Natural Images**. [abstract] [pdf] [ps] [bibtex]
  Xiaofeng Ren, Charless Fowlkes and Jitendra Malik, in *ECCV '06*, volume 2, pages 614-627, Graz 2006.

- **Learning Probabilistic Models for Contour Completion in Natural Images**. [abstract] [pdf]
  Xiaofeng Ren, Charless Fowlkes and Jitendra Malik, in *IJCV* Special Issue on Machine Learning for Vision, May 2008.

- **Cue Integration in Figure/Ground Labeling**. [abstract] [pdf] [bibtex]
  Xiaofeng Ren, Charless Fowlkes and Jitendra Malik, in *NIPS '05*, Vancouver 2005.

- **Recovering Human Body Configurations using Pairwise Constraints between Parts**. [abstract] [pdf] [bibtex]
  Xiaofeng Ren, Alex Berg and Jitendra Malik, in *ICCV '05*, volume 1, pages 824-831, Beijing 2005.

- **Scale-Invariant Contour Completion using Conditional Random Fields**. [abstract] [pdf] [ps] [bibtex]
  Xiaofeng Ren, Charless Fowlkes and Jitendra Malik, in *ICCV '05*, volume 2, pages 1214-1221, Beijing 2005.

- **Familiar Configuration Enables Figure/Ground Assignment in Natural Scenes**. [abstract] [poster] [bibtex]
  Xiaofeng Ren, Charless Fowlkes and Jitendra Malik, in *VSS 05*, Sarasota, FL 2005.

- **Mid-level Cues Improve Boundary Detection**. [abstract] [pdf] [ps] [bibtex]
  Xiaofeng Ren, Charless Fowlkes and Jitendra Malik, Berkeley Technical Report 05-1382, CSD 2005.

- **Recovering Human Body Configurations: Combining Segmentation and Recognition**. [abstract] [pdf] [ps] [bibtex]
  Greg Mori, Xiaofeng Ren, Alyosha Efros and Jitendra Malik, in *CVPR '04*, volume 2, pages 326-333, Washington, DC 2004.

- **Learning a Classification Model for Segmentation**. [abstract] [pdf] [ps] [bibtex]
  Xiaofeng Ren and Jitendra Malik, in *ICCV '03*, volume 1, pages 10-17, Nice 2003.

- **The Ecological Statistics of Good Continuation: Multi-scale Markov Models for Contours**. [abstract] [talk] [bibtex]
  Xiaofeng Ren and Jitendra Malik, in *VSS 02*, Sarasota, FL 2002.

- **A Probabilistic Multi-scale Model for Contour Completion Based on Image Statistics**. [abstract] [pdf] [ps] [bibtex]
  Xiaofeng Ren and Jitendra Malik, in *ECCV '02*, volume 1, pages 312-327, Copenhagen 2002.

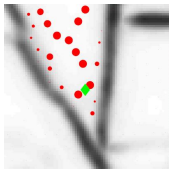
(Old) Research Projects

Discriminative Viewpoint Classification

RGB-D Mapping

Egocentric Object Recognition

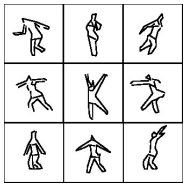
Multi-Scale Improves Boundary Detection



Local Grouping for Optical Flow

Finding and Tracking People in Archive Films

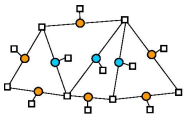
Tracking as Repeated Figure/Ground Segmentation



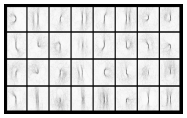
Line-based Aspect Learning and Matching

Figure-ground organization in natural images

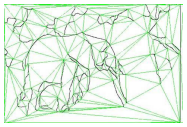
Cue Integration in Figure/Ground Labeling



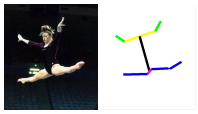
Scale-Invariant Contour Completion using Conditional Random Fields



Using Shapemes for Mid-level Vision



A Scale-Invariant Image Representation: the CDT Graph



Pairwise Constraints between Human Body Parts

Learning Discriminative Models for Image Segmentation

Human Body Configuration from Bottom-Up: a Segmentation-based Approach

Contours in Natural Images and Scale Invariance

Superpixel: Empirical Studies and Applications