

JINGMING DONG

Research Scientist

Facebook
1101 Dexter Ave N
Seattle, WA 98109

Email

djingming@gmail.com

Website

<http://vision.ucla.edu/~jingming>

EDUCATION

Ph.D., Computer Science

University of California, Los Angeles (UCLA), Los Angeles, CA

September 2011 – May 2017

Advisor: Professor Stefano Soatto

Dissertation: Optimal Visual Representation Engineering and Learning for Computer Vision

B.Sc., Computer Science

Fudan University, Shanghai, China

September 2007 – June 2011

Advisor: Professor Hong Lu

Thesis: A Visual System for Movie Genre Categorization (English abstract)

GPA: 3.85 / 4.00, Rank: 1 / 114

Visiting Student, Computer Science and Engineering

The Hong Kong University of Science and Technology (HKUST), Hong Kong

February – June 2009

GPA: 11.75 / 12, Dean's List

RESEARCH INTERESTS

Computer Vision, Machine Perception and Deep Learning. Local representation learning for general object recognition, wide-baseline image matching and augmented reality. Semantic scene representation in visual-inertial navigation systems.

PROFESSIONAL EXPERIENCE

Facebook

Research Scientist

June 2017 – present

UCLA Vision Lab

Graduate Student Researcher

September 2011 – June 2017

Summer Research Intern

July – September 2010

NVIDIA Research

Research Scientist Intern

June – September 2015

◦ Research project: A general adaptive image processing framework.

Media Computing and Web Intelligence Lab, Fudan University

Undergraduate Research Assistant

September 2009 – June 2011

Microsoft

Program Manager Intern, *Developer Division, Server and Tools Business, China*

February – April 2010

◦ Visual Studio 2010 Out of Branch release (OOB) product design for Team Architecture features; Software development management.

SELECTED RESEARCH PROJECTS

Visual-Inertial Semantic Scene Representation

We describe a representation of a scene that captures geometric and semantic attributes of objects within, along with their uncertainty. Objects are assumed persistent in the scene, and their likelihood computed from intermittent visual data using a convolutional architecture, integrated within a Bayesian filtering framework with inertials and a simple scene model. Our method yields a posterior estimate of geometry and semantics for a variable number of objects within the scene, causally and in real-time on commodity hardware.

Current Convolutional Architectures' Ability to Manage Nuisance Location and Scale Variability

We conduct an empirical study to test the ability of convolutional neural networks (CNNs) to reduce the effects of nuisance transformations of the input data, such as location, scale and aspect ratio. Empirical tests lead us to conclude that – at the current level of complexity of convolutional architectures and scale of the data sets used to train them, CNNs are not very effective at marginalizing nuisance variability. We propose improved sampling techniques

for heuristic proposal schemes that improve end-to-end performance to state-of-the-art levels on large-scale image classification dataset.

Domain-Size Pooling in Local Descriptors and its Application to Network Architectures

We introduce a simple modification of local image descriptors, such as SIFT, based on pooling gradient orientations across different domain sizes, in addition to spatial locations. The resulting descriptor outperforms other methods in wide-baseline matching benchmarks, including those based on convolutional neural networks, despite having the same dimension of SIFT and requiring no training. Extensions of the same idea have been shown to also improve the performance of other low- to middle-level representations, scattering transform and deep neural networks.

Past Projects

Multiple View Descriptors; Real-time Object Recognition on Mobile Devices (Android and iOS); Content-based Web Video Genre Categorization; Distance Metric Learning in Heterogeneous Domains; NDB – Not a Database.

PUBLICATIONS

Book Chapters

- S. Soatto, **J. Dong**. Visual Correspondence, the Lambert-Ambient Shape Space and the Systematic Design of Feature Descriptors. *Registration and Recognition in Images and Videos*, Pages 63-93, 2014.

Conference Papers & Technical Reports

- **J. Dong** and S. Soatto. Domain-size Pooling in Local Descriptors: DSP-SIFT. (Oral Presentation) In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- **J. Dong**, N. Karianakis, D. Davis, J. Hernandez, J. Balzer and S. Soatto. Multi-view Feature Engineering and Learning. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015.
- S. Soatto, **J. Dong** and N. Karianakis. Visual Scene Representations: Scaling and Occlusion in Convolutional Architectures. *The International Conference on Learning Representations (ICLR) Workshop*, 2015.
- N. Karianakis, **J. Dong** and S. Soatto. An Empirical Evaluation of Current Convolutional Architectures' Ability to Manage Nuisance Location and Scale Variability. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.
- **J. Dong**, I. Frosio and J. Kautz. Learning Adaptive Parameter Tuning for Image Processing. *arXiv preprint arXiv:1610.09414*, 2016.
- **J. Dong**, X. Fei and S. Soatto. Visual Inertial Semantic Scene Representation for 3D Object Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017.

PATENTS

DSP-SIFT: DOMAIN-SIZE POOLING FOR IMAGE DESCRIPTORS FOR IMAGE MATCHING AND OTHER APPLICATIONS. U.S. Patent, filed November, 2016.

PROFESSIONAL SERVICES

Reviewer: IEEE Transactions on Pattern Analysis and Machine Intelligence.

Reviewer: IEEE International Conference on Computer Vision, 2017.

External Reviewer: SIAM International Conference on Data Mining, 2016.

TEACHING EXPERIENCE

Teaching Assistant. *Comp. Sci. 268: Machine Perception*, Master Online Program (MSOL), UCLA.

Fall 2014.

HONORS AND AWARDS

National Scholarship, 1st Prize (Nov. 2008). Dean's List, The Hong Kong University of Science and Technology (Jun. 2009). Tung OCCL Scholarship, 1st Prize (Nov. 2009). Google Excellence Scholarship (Jun. 2010). Best Poster Design Award, UCLA 2015 Tech Forum (Feb. 2015).

TECHNICAL SKILLS

Proficient in C/C++, Matlab, VLFeat, Linux.

Familiar with Python, Java, Lua, Caffe, MatConvNet, Torch, OpenCV.