Guest Editors' Introduction to the Special Section on Computational Photography

Ayan Chakrabarti, Kalyan Sunkavalli, and David A. Forsyth

The development of increasingly successful visual inference algorithms has driven progress in a number of different application domains—ranging from photography to autonomous vehicles to graphics and virtual reality systems. As we continue to extend the capabilities of these computational algorithms, a complementary research direction lies in asking what the right visual measurements are for these algorithms to operate on. In computational photography, we seek to investigate both components—computational and sensory—of intelligent visual systems in synergy, to build measurement schemes and inference algorithms that are jointly optimal for a desired task, and thus create functionalities that go beyond what is possible with traditional cameras and computational tools.

This special section on "Computational Photography" presents a number of exciting new works in this direction. The call for papers for this section was co-ordinated with the 2020 IEEE International Conference on Computational Photography (ICCP). As guest-editors for the section and program chairs for the conference, we received a total of seventy-four submissions, out of which nine papers were accepted to this special section. These papers were presented at ICCP held from April 24-26, 2020, along with twenty-three other papers that were accepted to the conference proceedings. The papers in this section can be broadly categorized into two groups: (i) those that introduce novel computational methods that exploit non-traditional sensors; and (ii) those that propose novel sensors and acquisition systems.

1 Novel Inference Methods

The four papers in this category introduce new inference algorithms that leverage visual measurements made by unconventional sensors and camera systems.

The paper "Distance Surface for Event-Based Optical Flow" by M. Almatrafi, R. Baldwin, K. Aizawa, and K. Hirakawa, describes a novel algorithm for estimating optical flow from measurements made by an event camera—which only records instances of intensity changes above a certain threshold. These cameras are able to operate at higher framerates and in high dynamic-range scenes, and reliable optical flow estimation from their measurements has potential uses in diverse applications. The proposed method converts event measurements to a "distance surface" representation, and demonstrates that this representation then allows the use of standard intensity-domain optical flow algorithms. The authors evaluate their approach on real event sensor data and show it to be remarkably successful.

The paper "Deep Slow Motion Video Reconstruction with Hybrid Imaging System" by A. Paliwal and N.K. Kalantari, provides a novel algorithm to reconstruct high-quality videos that are high-resolution in both space and time. The algorithm works on a frames coming in from a hybrid setup with two cameras—one that has a high frame rate, and the other high spatial resolution. The authors introduce a novel deep learning-based approach to harmonize and fuse content from these two sources, while being aware of context and occlusions when reasoning about appearance. The paper includes extensive evaluation on both synthetic videos as well as those captured using a real two-camera system.

The paper "Neural Opacity Point Cloud" by C. Wang, M. Wu, Z. Wang, L. Wang, H. Sheng, and J. Yu, introduces a new neural rendering technique to produce photo-realistic renderings from three-dimensional point-cloud data. A key aspect of this technique is its ability to reason about partial transparency at "fuzzy" borders. The paper's experimental results provide considerable evidence of its success at producing high-quality renderings from a variety of simulated viewpoints, using data captured from standard three-dimensional scanners.

The paper "Differential 3D Facial Recognition: Adding 3D to your State-of-the-Art 2D Method" by J.M. Di Martino, F. Suzacq, M. Delbracio, Q. Qiu, and G. Sapiro, describes a convenient approach to exploit three-dimensional information for face recognition. The paper proposes projecting a high-frequency structured light pattern on the face and capturing a single RGB image. This captured image, while not sufficient to enable a full 3D reconstruction (which would require multiple captures while varying the pattern), encodes geometric information that can be used to compute depth features and aid recognition. The paper describes a neural network-based approach to exploit this additional depth information for face recognition, and demonstrates that it substantially improves accuracy and robustness to spoofing attacks.

- *A. Chakrabarti is with the Department of CSE, Washington University in St. Louis, St. Louis, MO 63130. E-mail: ayan@wustl.edu.*
- *K. Sunkavalli is with Adobe Research, San Jose, CA 95110. E-mail: sunkaval@adobe.com.*
- *D. A. Forsyth is with the Department of CS, University of Illinois, Urbana, IL 61801. E-mail: daf@illinois.edu.*

Digital Object Identifier no. 10.1109/TPAMI.2020.2993888

2 Novel Sensor Designs

The second group of five papers introduce new sensor and camera designs to make measurements that, when coupled with corresponding computational methods, are optimal for inference.

The paper "Shape and Reflectance Reconstruction Using Concentric Multi-Spectral Light Field" by M. Zhou, Y. Ding, Y. Ji, S.S. Young, J. Yu, and J. Ye, proposes a novel acquisition setup for single-shot shape and appearance capture for non-Lambertian surfaces. This setup features a concentric arrangement of cameras and light-sources, with co-ordinated narrow-band spectral filters, that allows the capture of different viewpoints and lighting directions in a single shot through spectral multiplexing. The paper then describes an algorithm to estimate shape and reflectance from these measurements, which is found to yield high-quality reconstructions in synthetic and real experiments.

The paper "SweepCam - Depth-Aware Lensless Imaging using Programmable Masks" by Y. Hua, S. Nakamura, M.S. Asif, and A.C. Sankaranarayanan, extends the capabilities of lensless imaging to handle scenes with large depth variations. It proposes a novel scheme to capture a series of images by varying patterns on a programmable mask in the aperture. The acquisition scheme is paired with fast and accurate reconstruction algorithm that produces high-quality images. The authors demonstrate the practical utility of this approach by building and showing results on a prototype of their setup.

The paper "PhlatCam: Designed Phase-Mask Based Thin Lensless Camera" by V. Boominathan, J.K. Adams, J.T. Robinson, and A. Veeraraghavan, also targets lensless imaging, but with the use of "phase" rather than amplitude masks. The paper introduces an approach to design phase masks that are able to induce "contour"-based point spread functions (PSFs) proposed by the authors. While these PSFs have large spatial supports like other lensless systems, they are sparse within that support. This allows high-fidelity image and depth reconstruction, and enables computational refocusing. The paper provides extensive experimental evaluation on a prototype system.

The paper "One-Bit Time-Resolved Imaging" by A. Bhandari, M.H. Conde, and O. Loffeld, describes a novel acquisition scheme for time-resolved imaging, to capture the time profile at each sensor pixel with a very high temporal resolution—for use in Time-of-Flight and LiDAR sensors. This scheme trades off bit depth for temporal resolution, by making one-bit measurements at a sampling rate that is much higher than the Nyquist rate. The paper then proposes a low-complexity non-iterative algorithm that is able to reconstruct the time profile with high-fidelity. Interestingly, the one-bit scheme provides better reconstructions than those that allocate more bits per sample, but sample at a lower rate for the same total number of bits. The success of this method opens up the possibility of realizing mega-pixel resolution time-resolved imaging sensors in the future.

The paper "Neural Sensors: Learning Pixel Exposures for HDR Imaging and Video Compressive Sensing with Programmable Sensors" by J.N.P. Martel, L.K. Müller, S.J. Carey, P. Dudek, and G. Wetzstein, introduces a framework to design optimal coded image-acquisition schemes in a setup that allows programmable per-pixel exposure times. A framework, to learn optimal and physically realizable exposure codes, is applied to two application settings: snapshot HDR imaging and high-speed compressive imaging. Experiments on simulated and real data show that the framework is successful in boosting performance over past approaches and sensor codes.

3 Research Outlook

The papers in this section describe recent strides in designing and leveraging new kinds of sensors and measurements towards reasoning about the visual world. We believe these methods and frameworks will benefit both researchers and practitioners working on various applications in the domains of computer vision, graphics, imaging, and robotics and autonomous systems.

Ayan Chakrabarti
Kalyan Sunkavalli
David A. Forsyth
Guest Editors

Ayan Chakrabarti received the SM and PhD and degrees from Harvard University, in 2008 and 2011, respectively. He is an assistant professor with the CSE department at Washington University in St. Louis, where he directs the vision and learning group. He works on applying tools from machine learning to problems in computer vision and computational photography—dealing with the design of accurate and efficient algorithms for visual inference, and of new kinds of high-capability sensors and cameras. He was a research assistant professor at the Toyota Technological Institute at Chicago from 2014-17.

Kalyan Sunkavalli received the masters degree from Columbia University, in 2006, and the PhD degree from Harvard University, in 2012. He is a senior research scientist at Adobe Research. His research interests lie at the intersection of computer vision, graphics, and machine learning and focus on understanding, reconstructing, and editing visual appearance from images and videos.

David A. Forsyth is currently the Fulton-Watson-Copp chair in computer science at University Illinois at Urbana-Champaign. He has published more than 170 papers on computer vision, computer graphics and machine learning. He has served as program or general co-chair for multiple international computer vision conferences. He received an IEEE technical achievement award for 2005 for his research, and became an IEEE fellow in 2009, and an ACM fellow in 2014. He served two terms as editor in chief, *IEEE Transactions on Pattern Analysis and Machine Intelligence*.