

3D image and video technology

Akihiro SUGIMOTO¹, Yoichi SATO², and Reinhard KLETTE³

¹National Institute of Informatics

²The University of Tokyo

³The University of Auckland

3D imaging and video technologies are of growing interest in recent times because of having potential applications in many fields such as robotics, visualization, 3DTV, autonomous vehicles, driver assistance, “flying eyes”, intelligent human-machine interfaces, and so forth. To advance those technologies, computer vision and pattern recognition, which originated separately, are today increasingly interacting. In fact, papers published in this special issue address areas of both computer vision and pattern recognition. After going through a rigorous anonymous peer reviewing process and several revisions, seven papers are finally presented in this special issue.

The first paper by Kobayashi, Sakaue and Sato, entitled “Multiple view geometry of projector-camera systems from virtual mutual projection” presents a calibration method for projector-camera systems. The proposed method generates virtual mutual projections between projectors and cameras by considering the shadow of cameras and the shadow of projectors and then discusses multiple view geometry for calibration. This work advances existing multiple view geometry so that it can also deal with projector-camera systems.

The second paper by Jarusirisawad, Hosokawa and Saito, entitled “Diminished reality using plane-sweep algorithm with weakly-calibrated cameras” addresses an on-line method for generating free viewpoint views using captured images at different viewpoints. In this paper, the plane-sweep algorithm already proposed by the authors is extended from in the Euclidean space to in the projective space. The advantage of this method exists in excluding occluding objects for view synthesis.

In “Object segmentation under varying illumination: stochastic background model considering spatial local-

ity”, Tanaka, Shimada, Arita and Taniguchi propose a background modeling method for object detection under varying illumination conditions. Detecting object in outside scene is one of important issues in visual surveillance. The challenge here is to deal with illumination changes to robustly detect objects in the scene. The contribution of this paper is fusing two complementary approaches: modifying background model and considering invariance of features in a local region.

In “People detection based on co-occurrence of appearance and spatio-temporal features”, Yamauchi, Fujiyoshi, Iwahori and Kanade present a people detection method. The basic idea of the proposed method exists in focusing on the co-occurrence of appearance features and spatio-temporal features. Detecting or tracking pedestrians is now a hot topic in the computer vision community, and the proposed method incorporates recent results such as HOG or the cascade AdaBoost classifier for pedestrian detection from videos.

The last three papers are more pattern recognition oriented than computer vision oriented. Okabe, Kondo, Kitani and Sato present a method for object recognition that is able to deal with objects from multiple categories. In difference to existing methods, they take advantage of a scene’s context that is represented by the co-occurrence between object categories. Experiments using a standard dataset called PASCAL show the effectiveness of the proposed method. Object recognition has advanced fast in recent years, thanks to publically available datasets, and this work also benefited by such datasets.

In videos, in particular TV programs, human faces are the most important object to focus on. Yamamoto, Yamaguchi and Aoki present a face clustering method to catalogue television titles, supposing an application for browsing recorded TV programs by faces. In order to catalogue titles in tolerable waiting time, high speed

Received February 16, 2010.

¹⁾sugimoto@nii.ac.jp, ²⁾ysato@iis.u-tokyo.ac.jp, ³⁾r.klette@auckland.ac.nz

DOI: 10.2201/NiiPi.2010.7.1

processing becomes mandatory. The main contribution of this paper is to use similarities of shots but not to rely on face recognition. Due to this strategy, the computational time is drastically reduced.

The final paper by Kitahara and Yanai, entitled “Associating faces and names in Japanese photo news articles” describes a system that extracts faces and names of pictured people from news articles on the net. The proposed system is based on an automatic association of faces that appear in photos in news articles on the net, and names of persons extracted from accompanying text. Not only detecting faces but also giving them names is an important issue. The contribution by this paper will draw attention from both the image analysis community and the natural language processing community.

We would like to thank the authors for submitting excellent papers to this special issue. We also express our thanks to all the reviewers that dedicated a significant amount of time and helped to improve the quality of the papers. Finally, we thank the staff of the journal office, in particular Ayumi Shimizu, for their support throughout the editorial process.



Akihiro SUGIMOTO

Akihiro SUGIMOTO received his B.Sc., M.Sc., and D.Eng. degrees in mathematical engineering from the University of Tokyo in 1987, 1989, and 1996, respectively. After working at Hitachi Advanced Research Laboratory, ATR, and Kyoto University, he joined the National Institute of Informatics, Japan, where he is currently a professor. From 2006 to 2007, he was a visiting professor at ESIEE, France. He received a Paper Award from the Information Processing Society in 2001. He is a member of IEEE. He is interested in mathematical methods in engineering. In particular, his current main research interests include discrete mathematics, approximation algorithm, vision geometry, and modeling of human vision.



Yoichi SATO

Yoichi SATO is an associate professor jointly affiliated with the Graduate School of Interdisciplinary Information Studies, and the Institute of Industrial Science, at the University of Tokyo, Japan. He received the BSE degree from the University of Tokyo in 1990, and the M.S. and Ph.D. degrees in robotics from the School of Computer Science, Carnegie Mellon University, Pittsburgh, Pennsylvania, in 1993 and 1997 respectively. His research interests include physics-based vision, reflectance analysis, image-based modeling and rendering, tracking and gesture analysis, and computer vision for HCI.



Reinhard KLETTE

Reinhard KLETTE is professor at the Computer Science Department at The University of Auckland, New Zealand. He received the MSc, Dr. rer. nat., and Dr. sc. nat. degrees from Jena University, Germany. He was professor for Computer Vision at the Technical University Berlin before moving to New Zealand. His research interests are in geometric methods and algorithms for computer vision, or image and pattern analysis. He is co-author of books or monographs on image processing operators (with the late Piero Zamperoni), on computer vision (with Karsten Schlüns and Andreas Koschan), on digital geometry (with the late Azriel Rosenfeld), and on panoramic vision (with Fay Huang and Karsten Scheibe).