Computer Vision and Image Understanding 153 (2016) 1–2

Contents lists available at ScienceDirect

Computer Vision and Image Understanding

journal homepage: www.elsevier.com/locate/cviu

Editorial

Special Issue on Visual Tracking

This special issue of the Computer Vision and Image Understanding (CVIU) journal provides a venue to present recent advances and trends in visual tracking that are driven by real-world applications. In particular, it aims to bridge the gaps among researchers in computer vision, robotics, machine learning, intelligent systems and related fields by providing insight on how to design visual tracking modules that can robustly perform in complex environments.

Visual tracking is an important and essential component of perception that has been an active research topic in the computer vision community for decades. The developments of the visual tracking algorithms have enjoyed rapid progress thanks to the explosive growth of video data which in turn creates high demand for speed and accuracy of tracking algorithms. Researchers are motivated to design faster and better methods in spite of the challenges that exist in visual tracking, especially robustness to heavy occlusions, drastic scale change, accurate localization, multi-object tracking, and recovery from failure. Despite the success in addressing numerous challenges under a wide range of circumstances, the core problems remain complex and challenging. Developing robust visual tracking algorithms for a wide range of practical applications will remain an active research topic in a foreseeable future.

The special issue includes 15 high-quality papers (30% of total submissions). We received a total of 50 submissions from all over the world, of which all were reviewed. The reviewers were selected to cover a broad range of expertise, to balance junior and senior members, and to represent a variety of geographical locations. We invited experienced reviewers from the computer vision and pattern recognition communities, of those most are specialized in visual tracking. Each paper was reviewed by at least two experts and considered by at least two guest editors before a decision was made. In the following, we provide a brief overview of the accepted papers.

Robust representation schemes: We have three papers on robust represenation shemes for visual tracking. In the paper “Convolutional Neural Net Bagging for Online Visual Tracking”, Li et al. propose a new CNN bagging strategy to simultaneously address label noise and model uncertainty of target appearances. Yang et al. propose an effective representation scheme to handle large appearance variations in visual tracking in “Robust Object Tracking by Online Fisher Discrimination Boosting Feature Selection”. The key contribution is an online Fisher discrimination boosting feature selection mechanism. In “Clustering based Ensemble Correlation Tracking”, Zhu et al. propose a clustering based ensemble correlation tracker, which jointly captures target appearance by multi-scale kernelized correlation filter and exploits long-term object properties with object templates.

Multi-object tracking: There are three papers for multi-object tracking. In “Multi-Object Tracking via Discriminative Appearance Modeling”, Huang et al. propose a metric learning and hierarchical multiple hypotheses tracking method that fuses multiple visual cues. The paper “Temporal Dynamic Appearance Modeling for Online Multi-Person Tracking” by Yang and Jia describes an online tracking method to use temporal consistency of target appearances for ensuring more accurate data association. In “An Online Variational Bayesian Model for Multi-Person Tracking from Cluttered Scenes”, a variational EM algorithm with a closed-form expression is explored to estimate posterior distribution and model parameters for multi-person tracking.

Dataset and performance evaluation: A benchmark data set with diverse scenes for human running detection is proposed in “Human Running Detection: Benchmark and Baseline” by Lao et al. A data set with ground truth annotations, two evaluation metrics and a baseline method are presented in this paper.

Real-time tracking: In “PICASO: PIxel Correspondences and Soft Match Selection for Real-time Tracking”, Timofte et al. note the trade-off between performance and speed in visual tracking, and propose an efficient algorithm based on pixel correspondence and soft match selection. Occlusion is also addressed to a certain extent by an efficient mask and median filters. In the paper “Real-Time 3D Reconstruction of Non-Rigid Shapes with a Single Moving Camera”, Agudo et al. consider the 3D reconstruction of deformable objects from a calibrated monocular video. A sequential algorithm based on the Navier-Cauchy equations is devised to simultaneously estimate camera motion and time-varying shape.

Multi-tracker fusion: The proposed method in “Online Adaptive Hidden Markov Model for Multi-Tracker Fusion” by Vojir et al. fueses observations from complementary trackers and a detector through a hidden Markov model (HMM). An unsupervised training method for HMM is exploited to alleviate the problems of inaccurate updates caused by false detections.

Motion estimation: In “Object-Guided Motion Estimation” by Perez-Rua et al., the combination of object tracking and dense motion estimation is investigated. The proposed method is applied to point tracking and yields improved performance over existing methods.

Generic visual tracking algorithms: The paper “A Local-Global Coupled-Layer Puppet Model for Robust Online Human Pose Tracking” by Ma el al. considers the problem on online tracking of articulated human body poses. A cross-coupled global-local model is

http://dx.doi.org/10.1016/j.cviu.2016.11.001

designed to handle human pose variations for tracking. The tracking method proposed in "Tracking by Switching State Space Models" by Kwon et al. is able to adaptively switch between state-space models over time and take advantage of both the image and ground planes. It is shown that the camera motion can be factored out by referring motions to the ground plane, making target motion more smooth. The paper "High-order Framewise Smoothness-constrained Globally-optimal Tracking" by Ukita and Okada proposes a smoothness-constrained globally optimal tracking method. The tracking task is formulated based on a graph where high-order constraints on motion smoothness is enforced. In "Visual Object Trapping" by Crivelli et al., a new criterion for object localization in videos is presented where conventional tracking constraints and formulations based on position errors are relaxed.

We would like to thank authors for their efforts and cooperation as this issue came together. We would also like to express our appreciation to all reviewers who responsibly reviewed the special issue submissions. We appreciate the suggestions from the Editor-in-Chief, and thank the support from the Elsevier publication staff members. Without their timely assistance, it would not have been possible to handle the record number of submissions and bring this high-quality special issue on visual tracking to fruition as scheduled.

Guest Editors:

Xue Mei
Toyota Research Institute North America, USA

Tianzhu Zhang
Chinese Academy of Sciences, China

Huchuan Lu
Dalian University of Technology, China

Ming-Hsuan Yang
University of California at Merced, USA

Kyoung Mu Lee
Seoul National University, Korea

Horst Bischof
Graz University of Techonlogy, Austria

E-mail addresses: xue.mei@toyota.com (X. Mei), tzzhang@nlpr.ia.ac.cn (T. Zhang), lhchuan@dlut.edu.cn (H. Lu), mhyang@ucmerced.edu (M.-H. Yang), kyoungmu@snu.ac.kr (K.M. Lee), bischof@icg.tugraz.at (H. Bischof)