Image and Vision Computing 31 (2013) 419–420

Contents lists available at SciVerse ScienceDirect

Image and Vision Computing

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

Preface

# Machine learning in motion analysis: New advances

With the ubiquitous presence of video data and its increasing importance in a wide range of real-world applications such as visual surveillance, human–machine interfaces and sport event interpretation, there is a growing demand for automated analysis and understanding of object motions from large amounts of video footage.

Vision-based motion analysis aims to detect, track, and identify objects, and, more generally, to understand their behaviors, from video sequences. Over the past decade, we have witnessed the exciting development and successful application of machine learning algorithms in a number of challenging motion analysis problems that demonstrate the great potential of this area. A well-known commercial example is the Kinect, the vision-based pose estimation system used in Microsoft X-box. Despite significant progress over the years, many challenging problems remain unsolved, such as robust object detection and tracking, unconstrained object activity recognition, and communicative behavior analysis. We believe that novel learning technologies have a strong potential to further contribute to the development of robust yet flexible vision systems. Interestingly, advances in computer vision algorithms have also brought numerous new challenges and positive feedback to the field of machine learning, leading to approaches such as learning from partial or limited annotations (e.g., one shot learning), or with a large number of categories (e.g., ImageNet), or with the scenarios of deploying a trained model to different target domains (e.g., domain adaptation).

The purpose of this Special Issue is to present and highlight the latest developments in dedicated motion analysis and machine learning algorithms to address a range of problems that can benefit from general vision-based object motion analysis and understanding from a machine learning perspective, including biological and medical applications. The aims are to highlight relevant work by researchers from related disciplines (machine learning, computer vision, motion analysis, pattern recognition, biomedicine, and beyond), to provide a selection of papers that represents different perspectives and that will promote further research in the area. To this end, we solicited original research contributions that address vision-based object motion using machine learning approaches, or that develop new machine learning and motion analysis approaches.

We received 24 submissions and one invited survey paper from around the world, with the majority coming from the USA, Japan, and India. Each submission was rigorously reviewed by at least two experts in the related fields based on the criteria of originality, significance, quality, and clarity. Eventually, eight papers were accepted for the special issue, spanning a variety of topics in terms of machine learning and its utilization for analyzing motion including biological trajectory matching. The first contribution, an invited survey by Metaxas, introduces the concept and applications of Nonverbal Communication Computing and studies its corresponding motion analysis methods, including face tracking, expression recognition, body reconstruction, and group activity analysis. This paper also discusses some open problems and the future directions of Nonverbal Communication Computing research.

A new machine learning based strategy to build the observation model of tracking systems is presented by Penne, Tilmant, Chateau, and Barra. The global observation function results from a linear combination of several modules (one per visual cue). Each module is built using an Adaboost-like algorithm, derived from the Ensemble Tracking Algorithm. The importance of each module is estimated using an original probabilistic sequential filtering framework with a joint state model composed by both the spatial object parameters and the importance parameters of the observation modules.

In the third paper, Ukita proposes human motion models of multiple actions for 3D pose tracking. A training pose sequence of each action, such as walking and jogging, is separately recorded by a motion capture system and modeled independently. Unlike existing approaches with similar motion models (e.g., switching dynamical models), this pose tracking method uses multiple models simultaneously to cope with ambiguous motions. For robust tracking with the multiple models, particle filtering is employed so that particles are distributed simultaneously in the models. Efficient use of the particles is achieved by locating many particles in the model corresponding to an action that is currently observed. For transferring the particles among the models in quick response to changes in the action, transition paths are synthesized between the different models in order to virtually prepare inter-action motions.

Nayak, Zhu, and Roy-Chowdhury propose an end-to-end system to recognize multi-person behaviors in video, unifying different tasks like segmentation, modeling and recognition within a single optical flow based motion analysis framework. Activities are modeled using underlying motion patterns formed in the optical flow. Critical points of the Helmholtz decomposition of motion field are used to identify interesting motion regions. Streaklines formed via integration of the flow field are clustered to identify motion patterns. Shape comparison and subspace analysis are used to compare motion patterns for activity recognition. They conducted experiments on realistic multi-object outdoor scenes.

Fan and Zhang present a two-layer gait representation framework for video-based human motion estimation. This work studies the idea of gait manifold learning and its related metrics. One whole-based and two part-level gait manifolds are involved by integrating the visual gait generative model (VGGM) and kinematic gait generative model (KGGM) for part–whole gait modeling. Moreover, this paper develops a two-stage Monte Carlo Markov Chain (MCMC) inference algorithm for part–whole gait estimation. The proposed method achieves the state-of-the-art results on the Human Eva dataset and CMU mobcap library databases.

An on-line learning of the transition model via Support Vector Regression is proposed by Salti and Stefano. The specialization of this general framework for linear/Gaussian filters, Support Vector Kalman (SVK), is

0262-8856/$ – see front matter 
http://dx.doi.org/10.1016/j.imavis.2013.04.003

then introduced and shown to outperform a standard, non-adaptive Kalman filter as well as a widespread solution to cope with unknown transition models such as the Interacting Multiple Models (IMM) filter.

Many existing feature transforms for trajectory matching are often based on shape matching, tending to perform poorly for biological trajectories, such as cell motion, because similar biological behavior often results in dissimilar trajectory shape. Additionally, the criteria used for similarity may differ depending on the user's particular interest or the specific query behavior. Fasciano, Souvenir and Shin present a rank-based distance metric learning method that combines user input and a new set of biologically-motivated features for biological trajectory matching. With a small amount of user effort, this method is shown to outperform existing trajectory methods.

Last, but not least, is the contribution from Ohnishi and Imiya who develop an algorithm for navigating a mobile robot using the visual potential. The visual potential is computed from an image sequence and optical flow computed from successive images captured by a camera mounted on the robot. The direction to the destination is provided at the initial position of the robot. The robot dynamically selects a local pathway to the destination without collision with obstacles and without any knowledge of the robot workspace. Furthermore, the guidance algorithm to destination allows the mobile robot to return from the destination to the initial position.

Papers in this Special Issue provide a glimpse of the current progress in the area of machine learning and motion analysis. We hope they can be useful towards providing insights into existing and new directions in video-based motion analysis. We also hope that this Special Issue will motivate more researchers to enter this exciting research area.

**Acknowledgments**

This special issue would have been impossible without the efforts and interests from all the authors who contributed their submissions. We would like to take this opportunity to thank them. We are grateful to the reviewers for their careful and valuable comments on the submitted manuscripts and for their detailed and helpful suggestions for improvement. We would also like to express our sincere acknowledgement to the Editors-in-Chief of the Image and Vision Computing, Profs. Frahm and Pantic, and to the Elsevier publication staff members, who all have contributed much of their time helping to bring this Special Edition to fruition.

Matti Pietikäinen*
*Center for Machine Vision Research, University of Oulu, Finland*
*Corresponding author.
*E-mail address:* matti.pietikainen@ee.oulu.fi.

Matthew Turk
*Department of Computer Science, University of California, Santa Barbara, USA*

Liang Wang
*National Laboratory of Pattern Recognition (NLPR), Institute of Automation, Chinese Academy of Sciences, PR China*

Guoying Zhao
*Center for Machine Vision Research, University of Oulu, Finland*

Li Cheng
*Bioinformatics Institute, A*STAR, Singapore School of Computing, National University of Singapore, Singapore*