

A Self-Training Approach for Visual Tracking and Recognition of Complex Human Activity Patterns

Jan Bandouch · Odest Chadwicke Jenkins ·
Michael Beetz

Received: 18 January 2011 / Accepted: 6 March 2012 / Published online: 23 March 2012
© Springer Science+Business Media, LLC 2012

Abstract Automatically observing and understanding human activities is one of the big challenges in computer vision research. Among the potential fields of application are areas such as robotics, human computer interaction or medical research. In this article we present our work on unintrusive observation and interpretation of human activities for the precise recognition of human fullbody motions. The presented system requires no more than three cameras and is capable of tracking a large spectrum of motions in a wide variety of scenarios. This includes scenarios where the subject is partially occluded, where it manipulates objects as part of its activities, or where it interacts with the environment or other humans. Our system is self-training, *i.e.* it is capable of learning models of human motion over time. These are used both to improve the prediction of human dynamics and to provide the basis for the recognition and interpretation of observed activities. The accuracy and robustness obtained by our system is the combined result of several contributions. By taking an anthropometric human model and optimizing it towards use in a probabilistic tracking framework we obtain a detailed biomechanical representation of human shape, posture and motion. Furthermore, we introduce a sophisticated hierarchical sampling strategy for tracking

that is embedded in a probabilistic framework and outperforms state-of-the-art Bayesian methods. We then show how to track complex manipulation activities in everyday environments using a combination of learned human appearance models and implicit environment models. Finally, we discuss a locally consistent representation of human motion that we use as a basis for learning environment- and task-specific motion models. All methods presented in this article have been subject to extensive experimental evaluation on today's benchmarks and several challenging sequences ranging from athletic exercises to ergonomic case studies to everyday manipulation tasks in a kitchen environment.

Keywords Markerless human motion capture · Probabilistic state estimation · Self-trained models of human motion · Activity recognition

1 Introduction

Observing and interpreting human activities is a constant topic of interest in *artificial intelligence* (AI) and *computer vision* (CV) research. The ability to understand human behavior and to act with respect to human actions or intentions is an ambitious goal. Once achieved, it will prove beneficial in application areas ranging from *robotics* to *human computer interaction* (HCI) or *medical research* such as *gait analysis*. However, despite the broad progress made in the field in the last decades, many challenges remain.

In this article we will present our contributions to this highly active field of research. We present a markerless system for human fullbody motion tracking from three or more cameras that utilizes a realistic human model to estimate the observed motions at high accuracy. Our system is capable to extract this information for arbitrary types of motions

J. Bandouch (✉) · M. Beetz
Intelligent Autonomous Systems Group, Technische Universität
München, Boltzmannstr. 3, 85748 Garching bei München,
Germany
e-mail: bandouch@cs.tum.edu

M. Beetz
e-mail: beetz@cs.tum.edu

O.C. Jenkins
Department of Computer Science, Brown University,
115 Waterman St., Providence, RI, 02912-1910, USA
e-mail: cjenkins@cs.brown.edu