International Scholarly Research Network
ISRN Machine Vision
Volume 2012, Article ID 235396, 11 pages
doi:10.5402/2012/235396

Research Article

3D Human Motion Tracking and Reconstruction Using DCT Matrix Descriptor

Alireza Behrad and Nadia Roodsarabi

Electrical Engineering Department, Faculty of Engineering, Shahed University, Tehran 33191-18651, Iran

Correspondence should be addressed to Alireza Behrad, behrad@shahed.ac.ir

Received 30 January 2012; Accepted 19 February 2012

Academic Editors: C.-C. Han and J. Heikkilä

One of the most important issues in human motion analysis is the tracking and 3D reconstruction of human motion, which utilizes the anatomic points' positions. These points can uniquely define the position and orientation of all anatomical segments. In this work, a new method is proposed for tracking and 3D reconstruction of human motion from the image sequence of a monocular static camera. In this method, 2D tracking is used for 3D reconstruction, which a database of selected frames is used for the correction of tracking process. The method utilizes a new image descriptor based on discrete cosine transform (DCT), which is employed in different stages of the algorithm. The advantage of using this descriptor is the capabilities of selecting proper frequency regions in various tasks, which results in an efficient tracking and pose matching algorithms. The tracking and matching algorithms are based on reference descriptor matrixes (RDMs), which are updated after each stage based on the frequency regions in DCT blocks. Finally, 3D reconstruction is performed using Taylor's method. Experimental results show the promise of the algorithm.

1. Introduction

One of the challenging issues in machine vision and computer graphic applications is the modeling and animation of human characters. Especially body modeling using video sequences is a difficult task that has been investigated a lot in the last decade. Nowadays, 3D human models are employed in various applications like movies, video games, ergonomic, e-commerce, virtual environments, and medicine.

3D scanners [1, 2] and video cameras are two sample tools that have been presented for 3D human model reconstruction. 3D scanners have limited flexibility and freedom constraints. In addition, the higher cost of these devices put them out of reach for general use.

Video cameras are nonintrusive and flexible devices for extraction of human motion. However, due to the high number of degrees of freedom for the human body, human motion tracking is a difficult task. In addition, self-occlusion of human segments and their unknown kinematics make the human tracking algorithm more challenging.

Existing vision-based approaches for human motion analysis may be divided in two groups, including model-based and model-free methods [3]. In model-based methods [4–8], *a priori* known human model is employed to represent human joints and segments as well as their kinematics. Model-free approaches do not employ a predefined human model for motion analysis; instead, the motion information is derived directly from video sequences. Model-free approaches mostly use a database of exemplars [9] or a learning machine [10, 11] for motion reconstruction. They are mostly restricted to known environments or images taken from a known viewpoint. Model-based approaches are more general and typically support the viewpoint independent processing or multiple viewpoints. However, they need initialization.

Various algorithms may also be divided into different categories based on the acquisition system. Some approaches are based on monocular cameras [4–14], while others employ multicamera video streams [15–20]. Also, some approaches benefit from calibrated views or cameras [15–20], while others utilize uncalibrated images [5–14].

Nowadays, monocular uncalibrated video sequences such as sports video footage are the most common source