Accepted Manuscript

Marker based human pose tracking using adaptive annealed particle swarm optimization with search space partitioning

Ashraf Sharifi, Ahad Harati, Abedin Vahedian

PII: S0262-8856(17)30065-3
DOI: doi:10.1016/j.imavis.2017.03.003
Reference: IMAVIS 3609

To appear in: *Image and Vision Computing*

Received date: 21 April 2016
Revised date: 1 January 2017
Accepted date: 22 March 2017

Please cite this article as: Ashraf Sharifi, Ahad Harati, Abedin Vahedian, Marker based human pose tracking using adaptive annealed particle swarm optimization with search space partitioning, *Image and Vision Computing* (2017), doi:10.1016/j.imavis.2017.03.003

Marker based human pose tracking using adaptive annealed particle swarm optimization with search space partitioning

Ashraf Sharifi[a], Ahad Harati[a], Abedin Vahedian[a,*]

[a]*Department of Computer Engineering,Ferdowsi University of Mashhad,Mashhad, Iran*

Abstract

Pose estimation and tracking of an articulated structure based on data from multiple cameras has seen numerous applications in recent years. In this paper, a marker based human pose tracking algorithm from multi view video sequences is proposed. The purpose of the proposed algorithm is to present a low cost motion capture system that can be used as an alternative to high cost available commercial human motion capture systems. The problem is defined as the optimization of 45 parameters which define body pose model and is solved using a modified version of particle swarm optimization (PSO) algorithm. The objective of this optimization is to maximize a fitness function which formulates how much the body model matches with $2D$ markers coordinates in video frames. A sampling covariance matrix is used in the first part of the velocity equation of PSO and is annealed with iterations. The sampling covariance matrix is computed adaptively, based on variance of parameters in the swarm. One of the concerns in this algorithm is the high number of parameters to define the model of body pose. To tackle this problem, we partition the optimization state space into six stages that exploit the hierarchical structure of the skeletal model. The first stage optimizes the six parameters that define the global orientation and position of the body. Other stages relate to optimization of right and left hand, right and left leg and head orientation. In the proposed partitioning method previously optimized parameters are allowed some variation in each step that is called soft partitioning. Experimental results on Pose Estimation and Action Recognition (PEAR) database indicate that the proposed algorithm achieves lower estimation error in tracking human motion compared with Annealed Particle Filter (APF) and Parametric Annealing (PA) methods.

Keywords: Pose tracking, Marker based human pose estimation, Particle swarm optimization, Search space partitioning

1. Introduction

Capturing and tracking the human motion have found numerous applications in recent years. Motion capture is the process of recording human motion as a sequence of 3D Cartesian coordinates called motion data [1]. Human pose tracking is the process of determining the configuration (orientation and location) of body parts at consecutive time instants using motion data. There are three major goals for human pose tracking [2]: smart surveillance, object control and research purposes. The purpose of surveillance applications is human body pose tracking while monitoring for specific actions such as shop lifting. Animating virtual characters in games and movies can be considered as control applications. The aim of these applications is avatar control within virtual worlds based on human motion in the real world. Motion data in research applications are used for diagnostics orthopedic patients in clinical studies or train athletes to improve their performance.

General structure of motion analyzing systems consists of four steps, namely initialization, tracking, body pose estimation and action recognition. Determining an appropriate model of subject in model based systems and camera calibration in image based systems are examples of initialization. Aim of the tracking step is to determine the position of corresponding segments of the body parts in successive frames. In the pose estimation step relative orientation and location of body parts related to each other is determined. In the last step of this process the estimated pose in consecutive frames is analyzed to recognize the action performed by the subject. There are different sensor types to capture human motion, which are categorized as active and passive sensors [3]. Active sensors transmit or receive signals from the other sensors while passive sensors have no effects on the other sensors. Accelerometers, mechanical, electromagnetic [4] and acoustic sensors are examples of active sensors already used for human motion capture. The methods based on these sensors usually require devices to be attached to the body parts such as skeletal-like structures in mechanical approaches and magnetic or acoustic sensors in other approaches [5]. The major problem about the above methods is that the

*Corresponding author

Email addresses: `sharifi.a7@alumni.um.ac.ir` (Ashraf Sharifi), `a.harati@um.ac.ir` (Ahad Harati), `vahedian@um.ac.ir` (Abedin Vahedian)