Chapter 7

A Generic Framework for 2D and 3D Upper Body Tracking

Lei Zhang
Rensselaer Polytechnic Institute, USA

Jixu Chen
Rensselaer Polytechnic Institute, USA

Zhi Zeng
Rensselaer Polytechnic Institute, USA

Qiang Ji
Rensselaer Polytechnic Institute, USA

ABSTRACT

Upper body tracking is a problem to track the pose of human body from video sequences. It is difficult due to such problems as the high dimensionality of the state space, the self-occlusion, the appearance changes, etc. In this paper, we propose a generic framework that can be used for both 2D and 3D upper body tracking and can be easily parameterized without heavily depending on supervised training. We first construct a Bayesian Network (BN) to represent the human upper body structure and then incorporate into the BN various generic physical and anatomical constraints on the parts of the upper body. Unlike the existing upper body models, we aim at handling physically feasible body motions rather than only some typical motions. We also explicitly model the body part occlusion in the model, which allows to automatically detect the occurrence of self-occlusion and to minimize the effect of measurement errors on the tracking accuracy due to occlusion. Using the proposed model, upper body tracking can be performed through probabilistic inference over time. A series of experiments were performed on both monocular and stereo video sequences to demonstrate the effectiveness and capability of the model in improving upper body tracking accuracy and robustness.

DOI: 10.4018/978-1-60566-900-7.ch007
Human body tracking from 2D images is a challenging problem in computer vision. Assuming human body is composed of $N$ rigid body parts, the whole body pose can be represented as a long vector $\Theta = (\phi_1, \ldots, \phi_N)$, where $\phi_i = \{T, R\}$ represents the pose (i.e. translation and rotation) of each body part. The whole body pose is usually in a high dimensional continuous space (25-50 dimensions is not uncommon (Deutscher, Blake, & Reid, 2000)). If one simply estimates the whole pose, the high dimensionality of the state space will lead to the problem of intractable computational complexity.

In order to efficiently and robustly search in the high-dimensional body pose space, people use either sampling-based methods or learning-based methods. Sidenbladh et al. (Sidenbladh, Black, & Fleet, 2000), Deutscher et al (Deutscher, et al., 2000) and MacCormick et al. (MacCormick & Isard, 2000) attempt to handle this problem by importance sampling, annealed sampling and partitioned sampling, respectively. In the sampling based method, the posteriori probability distribution of the human pose is represented by a certain number of particles (pose hypothesis). During tracking, these particles are propagated using the dynamical model and weighted by the image likelihood. However, in this kind of basic sequential importance sampling (Sidenbladh, et al., 2000), the required number of particles grows exponentially with the dimension of the pose space (MacCormick & Isard, 2000), which makes it inefficient. To reduce the required samples, Deutscher et al (Deutscher, et al., 2000) use the annealed sampling, which generate the samples through several “annealing” steps, and show that the required number of particles can be reduced by over a factor of 10. MacCormick et al. (MacCormick & Isard, 2000) use partitioned sampling to “partition” the pose space into sub-spaces sequentially, so the number of required samples will not significantly grow with the dimensionality. Although these methods can reduce the samples to around 200-700, they add much more computation load to the sample generation step, which makes them still computationally inefficient.

The learning-based methods attempt to learn a direct mapping from the image feature space to the pose space, and this mapping is learned from the training data. Currently, the most popular learning-based methods in body tracking are the regression learning techniques, such as the regression method in (Agarwal & Triggs, 2006) and the Gaussian process latent variable model (GPLVM) in (Tian, Li, & Sclaroff, 2005). However, the learning-based methods usually can only give good results for specific persons and specific motions that are similar to the training data.

So far, many robust and efficient head and limb detectors have been proposed (Ramanan & Forsyth, 2003). Since the whole body pose $\Theta$ is difficult to be recovered directly, more and more body tracking techniques are proposed to track the pose of each body part independently. This independent tracking reduces the problem of high dimensionality, but it can still be difficult due to part occlusion as well as significant changes in part appearances.

On the other hand, two adjacent human body parts are anatomically connected with joints and muscles, and the feasible poses of an upper body must satisfy some anatomical constraints. Researchers have exploited the relationships among human body parts for constraining the body tracking problem. The problem is how to efficiently and effectively capture these relationships in a systematic way.

Probabilistic graphical model (PGM) is a natural way to represent the relationships among the body parts. If the only relationships captured by the PGM are those between the adjacent body parts, the model can be represented as a tree-structured PGM. The simple tree structure allows efficient
Related Content

An Experimental Investigation of Pulse Measurement by Means of a Plethysmographic Sensor Integrated in a ZigBee Medical Network

Estimation of the Energy Potential of the Euripus’ Gulf Tidal Stream Using Channel Sea-surface Slope

Matching Score Level Fusion
David Zhang, Fengxi Song, Yong Xu and Zhizhen Liang (2009). *Advanced Pattern Recognition Technologies with Applications to Biometrics* (pp. 305-327).
[www.igi-global.com/chapter/matching-score-level-fusion/4286?camid=4v1a](www.igi-global.com/chapter/matching-score-level-fusion/4286?camid=4v1a)

Profile-Based Text Classification for Children with Dyslexia
[www.igi-global.com/article/profile-based-text-classification-for-children-with-dyslexia/145351?camid=4v1a](www.igi-global.com/article/profile-based-text-classification-for-children-with-dyslexia/145351?camid=4v1a)