Using Rao-Blackwellised Particle Filter Track 3D Arm Motion based on Hierarchical Limb Model

XueSong Yu JiaFeng Liu XiangLong Tang JianHua Huang

Abstract—For improving the efficiency of human 3D tracking, we present an algorithm to track 3D Arm Motion. First, the Hierarchy Limb Model (HLM) is proposed based on the human 3D skeleton model. Second, via graph decomposition, the arm motion state space, modeled by HLM, can be decomposed into two low dimensional subspaces: root nodes and leaf nodes. Finally, Rao-Blackwellised Particle Filter is used to estimate the 3D arm motion. The result of experiment shows that our algorithm can advance the computation efficiency.

Keywords—Hierarchy Limb Model; Rao-Blackwellised Particle Filter; 3D tracking

I. INTRODUCTION

ALTHOUGH caught the many researchers’ attention because of its widely application, human 3D tracking is still a challenging task because of the exponentially increased computational complexity in terms of the freedom degrees.

Moeslund [1] [2] et al. classified the existing research pose estimation algorithms into learning-based algorithm and model-based algorithm. The model-based category builds the human motion model with human prior knowledge and the human motion constraints, and the use of stochastic sampling techniques in model-based analysis-by-synthesis to obtain the optimal estimation based on the Bayesian network framework. As the nonlinear filter algorithm based on the Bayesian estimation framework, particle filter [3] has been widely application [4] [5] in the area of human 3D motion estimation. Deutcher [6] et al. proposed the annealed particle filter to track the human 3D motion. Markov Chain Monte Carlo [7] [8] is utilized to solve the particle degeneracy problem. Recently, the structure graphical model [9] has been used to facilitate the estimation of human 3D motion. Based on the structure graphical model, Wei [13] proposed a decentralized framework. And facilitated by graph decomposition, they derived a novel Bayesian conditional density propagation rule.

In recent years, Rao-Blackwellised Particle Filter [10] [11] [12] has been widely studied in human tracking. Xinyu [12] et al. learns motion correlation using the Partial Least Square, and proposed the RBPF-PLS algorithm to track the walking pose.

Although these algorithms have achieved the goal, they can’t track any motion in the nature scene but tracking the learned motion. And learning a general probabilistic model in full space is very difficult because of the high dimensionality and the huge amounts of training data to account for motion complexity. Because of the restrictive dependency, RBPF can not be directly used for 3D human tracking.

Focused on these problems, the paper proposes the 3D arm motion tracking algorithm using Rao-Blackwellised Particle Filter based on Hierarchy Limb Model. The arm motion spaces can be divided into two parts: root variables and leaf variables by the Hierarchy Limb Model. Rao-Blackwellised Particle Filter can be used to track arm motion via the decomposition. As a result, our algorithm can advance the computational efficiency because of the lower dimensionality of the search space and the reduced amounts of particle.

The paper is organized as follows. Section 2 describes the Hierarchy Limb Model. The 3D arm tracking algorithm based on RBPF is proposed in Section 3. Experimental results and analysis are shown in Section 4, and finally concludes the paper.

II. HIERARCHY LIMB MODEL

The arm can be represented by an arm graphical model such as shown Fig. 1 (a). The circle nodes corresponds to a part of right arm, such as the right upper arm and the right lower arm. The square nodes are the observation values associated with each circle nodes. The undirected links represent physical constraints among different parts of the right arm. The directed link from a part’s state to its associated observation represents the local observation likelihood. In order to describe the motion of an articulated object, we accommodate the state dynamics by a dynamical graphical model such as shown in Fig. 1 (b). It contains two consecutive time frames. The directed links between consecutive states represents the dynamics translation from time $t-1$ to time $t$.

We model the arm as 2 cylinders connected at revolute joints, and denote the state of each part of arm at time $t$ by $x_i$, where $i=1,2$ is the index of two parts of arm. The arm motion state space is 12-dimensional, including 6D for the global (shoulder) position and orientation, 3D for the elbow as $x_i$, 3D for the wrist as $x_i$. 

XueSong Yu is with the Harbin Institute of Technology, Harbin, Heilongjiang China (corresponding author to provide phone: 086-013304819068.; e-mail: yyz001@hit.edu.cn).

JiaFeng Liu, XiangLong Tang and JianHua Huang were with the Harbin Institute of Technology, Harbin, Heilongjiang China.
According to the characteristics of arm motion, the motion of any node of the arm only interacts with its children nodes. For example, the motion of lower arm is not constrained by any limbs but only the motion of corresponding upper arm. Fig. 2 (a) show the decomposition result for the right arm in Fig. 1 (b), and Fig. 2 (b) is the associated moral graph via the separation theorem and the characteristics of the dynamic Markov network [13].

Using the arm hierarchy model, the problem of tracking right arm motion can be formulated as the prediction of $x$ at time $t$.

III. 3D ARM TRACKING WITH RBPF

By the HLM, arm motion can be decomposed into two parts: $x_u$ (upper arm) and $x_l$ (lower arm). We denote $x_v$ as root variable $R_v$ and $x_l$ as leaf variable $L_v$. In the algorithm, we marginalize the lower arm pose variable by using the motion speed correlation matrix $V$. The strategy of Rao-Blackwellised Particle Filter is to partition the full state space into two parts, $R_v$ (root variable) and $L_v$ (leaf variable), so the $P(L_v \mid R_v, Z_v)$ is a distribution that can be computed exactly conditional on $R_v$, and the distribution $P(X \mid Z_v)$ can be estimated using sample methods, such as Markov Chain Monte Carlo.

The posterior probability distribution for the right arm motion can be factorized by Equation (1).

$$P(R_v, L_v \mid Z_v) = P(L_v \mid R_v, Z_v)P(R_v \mid Z_v)$$  

Given the speed correlation matrix $V$, and the image measurements $Z_v$, the Equation (1) can be approximated by the following expression:

$$P(R_v, L_v \mid Z_v) = P(L_v \mid R_v, Z_v)P(R_v \mid Z_v)$$  

In Equation (2), the integral parts is propagated the root variables $R_v$ utilizing the state transition model according to the standard Particle Filter. The leaf variables $L_v$ can be predicted analytically using Kalman Filtering prediction. The image likelihood model, as $P(Z \mid R, L_v)$, is computed in the state space, including root and leaf variable.

A. State Transition Model

In particle filter theoretical framework, the state transition model, by which particle is generated, is described as shown Equation 3.

$$x_t = x_{t-1} + v_t, \quad v_t \sim N(\mu, \Sigma)$$  

Where $v_t$ is drawn from the Gaussian noise that the expectation $\mu$ is a 3×1 scalar, which is defined as the motion speed of root variables, and the variance $\Sigma$ is the 3×3 diagonal matrix. Utilizing hard prior, we can eliminate the particles corresponding to implausible arm pose to reduce the search space.

B. Prediction for the Leaf Variables

The leaf variables can be predicted by the use of the motion speed correlation matrix $V$ using the Kalman Filter while the root variables are propagated. Condition on the root variables and the correlation matrix $V$, the leaf variable $L_v$ are described as the following expression:

$$L_v = AL_v + V(R_v - \hat{R}_v) + \xi_v, \quad p(\xi) \sim N(0, \Phi)$$  

$$Z_v = HL_v + \xi_v, \quad p(\xi) \sim N(0, \Psi)$$

Where $\xi$ represent the process noise and $\xi$ represent the measurement noise. $V$ is defined as the motion speed correlation matrix. Then the leaf variables are predicted by Equation (5).

$$L_v' = AL_v + V(R_v' - R_v)$$

$$P_r' = AP_r A^t + \Phi$$

In Equation (5), $i$ is defined as the particle index. In Equation (4) and Equation (5), $A$ and $H$ are assumed to be the diagonal matrix.

According to the Equation (4) and Equation (5), Kalman Updates can be computed as the following expression:
In Equation (8) and Equation (9), the frames of test video are described as $T$, and $T=796$. $x$ is the prediction value, $X_i$ is the ground truth at frame $t$.

From Table I, the mean error and error variance between the prediction and ground truth have not evidently changed as the particle count of limbs increases. Then we can draw the conclusion that the count of particle for limbs can not affect the tracking result of our algorithm. Figure 3 shows the tracking results of 3D arm motion by our algorithm. It is not evidently different between the tracking results of our algorithm and the real pose of arm motion.

| Table I | THE ERROR MEAN AND STD. OF TRACKING RESULT USING OUR ALGORITHM |
|---------|-------------------------------------------------------------|
| Direction | X | Y | Z               |
| Mean | 13.4550 | 11.4206 | 10.0029 |
| Std. | 10.4712 | 9.1102 | 10.3215 |

| Table II | THE COMPARISON OF ERROR MEAN AND STD. OF TRACKING RESULT USING HLMRBF AND SPF |
|----------|-------------------------------------------------------------|
|          | X Direction | Y Direction | Z Direction |
| Mean     | Std.        | Mean | Std. |
| HLM-RBPF | 12.87 | 10.25 | 13.10 | 11.88 | 12.30 | 11.99 |
| SPF      | 15.55 | 13.60 | 12.18 | 10.81 | 13.09 | 11.77 |

V. CONCLUSIONS

The paper proposes 3D arm motion fast tracking algorithm. Based on the HLM, the algorithm can transfer the global optimal search of the whole state space to the top-bottom search based on the joints under the case that the dimension of state space is unchanged. In the process of tracking, the particle count is reduced by the prediction of each joint of HLMRBPF. The experiment shows that the tracking result using our algorithm is not evidently different compared with the standard...
particle filter under the same dimension of state space. The algorithm can effectively apply to track 3D arm motion based on Particle Filter.

**REFERENCES**

[1] Thomas B. Moeslund and E. Granum. “A survey of computer vision-based human motion capture”. *Computer Vision and Image Understanding*, 2001, vol. 81, pp. 231–268.

[2] Thomas B. Moeslund, Adrian Hilton and Volker Kru¨ger. “A survey of advances in vision-based human motion capture and analysis”. *Computer Vision and Image Understanding*, Vol.104, No.2, 2006, pp.90-126.

[3] A. Blake and M. Isard. “Condensation—Conditional Density Propagation for Visual Tracking”. *Int’l J. Computer Vision*, vol. 29, No. 1, pp. 5-28, 1998.

[4] P. Azad, A. Ude, R. Dillmann, and G. Cheng. “A full body human motion capture system using particle filtering and on-the-fly edge detection”. *4th IEEE/RAS International Conference on Humanoid Robots*, 2004, pp. 941–959.

[5] Jannal. Saboune and Francois. Charpillet. “Using Interval Particle Filtering for Marker less 3D Human Motion Capture”. *Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence*, page(s):7 pp, 2005.

[6] J. Deutscher and I. Reid. “Articulated body motion capture by stochastic search”. *Proc. International Journal of Computer Vision*, pp 185- 205, Vol 2, 2005.

[7] C. Sminchisescu and B. Triggs. “Kinematic Jump Processes for Monocular 3D Human Tracking”. *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp. 69-76, 2003.

[8] MunWai. Lee and Isaac. Cohen. “A Model-Based Approach for Estimating Human 3D Poses in Static Images”. *IEEE Transactions on Pattern Analysis And Machine Intelligence*, Vol. 28, NO. 6, June 2006.

[9] L. Sigal, S. Bhatia, S. Roth, M. J. Black, and M. Isard. “Tracking loose-limbed people”. *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, pp.421-428, Vol 1, 2004.

[10] X. Xu, B. Li. “Adaptive Rao-Blackwellised Particle Filter and its evaluation for tracking in surveillance”. *IEEE Trans. Image Processing*, 16(3):838-849, 2007.

[11] K. Murphy and S. Russell. *Rao-Blackwellised particle filtering for dynamic Bayesian networks*. In *Sequential Monte Carlo Methods in Practice*, A. Doucet, et al Eds. New York: Springer-Verlag, 2001, ch. 24.

[12] Xinyu Xu, Baoxin Li. “Learning Motion Correlation for Tracking Articulated Human Body with a Rao-Blackwellised Particle Filter”. *Proc. International Conference on Computer Vision*, Page(s):1-8, Vol 1, 2007.

[13] Wei Qu and Dan Schonfeld. “Real-time decentralized articulated motion analysis and object tracking from videos”. *IEEE Trans Image Processing*, VOL. 16, NO. 8, Aug. 2007.

---

**Fig. 3** 3D animation Comparison between the tracking result by our algorithm and ground truth; (a) 3D animation for the tracking value of our algorithm; (b) 3D animation for the ground truth.