Human Body Tracking with Shape Estimation and Auxiliary Measurements

1. Research Team

Project Leader: Prof. Isaac Cohen, *Computer Science*

Graduate Students: Mun Wai Lee

Industrial Partner(s): ICT

2. Statement of Project Goals

This goal of this project is to develop techniques for recognizing body motion and gestures for multimodal interactions. We present here our research effort in automatically fitting and tracking an articulated body model to human body silhouettes extracted from synchronous video streams.

3. Project Role in Support of IMSC Strategic Plan

Human body tracking has important potential applications in a number of areas that IMSC is engaged in. Human body tracking is an alternative non-invasive human motion capture technology, which is useful for human animation. Human body tracking is also useful for understanding human body structure and human movement. The technology can also potentially be used as user interface for interactive virtual reality applications (e.g. in games and physiotherapy applications).

4. Discussion of Methodology Used

An articulated human model (as shown in Figure 1) is constructed and used as a reference of the physical structure of the human body and constraints of the human movement. This model consists of 10 joints and 14 segments, representing the head, torso and limbs. Each segment is represented by a tapered 3D cone with an elliptical cross-section. The model has 31 degrees of freedom that include the global translation, rotation and local joint rotations.

The tracking framework is based on particle filter [1,6](a.k.a. Condensation [7]), which is popularly used for human body tracking. [4,5,10,11,12] However particle filter has a number of significant problems. Particle filter requires an impractically large number of particles to sample the high dimensional state space effectively; otherwise, it is easy to lose track and difficult to recover tracking failure because of sample depletion in the state space. In addition, particle filters require accurate model initialization. Often, initialization is done manually, which is undesirable in many applications.

In the project, we are developing techniques to tackle the problems discussed above and to improve the performance of tracking. We developed a novel method for integrating particle filter with analytical inference techniques. We proposed a method for detecting body parts such as the head, the hands and torso and use the detection result to analytically infer a subset of state parameters corresponding to the observed human body pose. This additional inference is used to improve the state estimation within the particle filtering framework. A schematic diagram of this framework is shown in Figure 2.

There are several advantages of combining particle filter with analytical inference. Firstly, the inference helps to reduce the degree of freedom that is dependant on Monte-Carlo simulation during state estimation. This reduces the required number of particles and the computational load. Secondly, the analysis is useful for automatic model initialization and recovery of lost tracks.

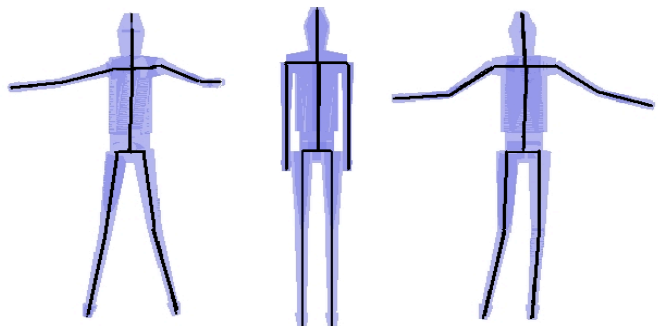


Figure 1. Articulated Human Body

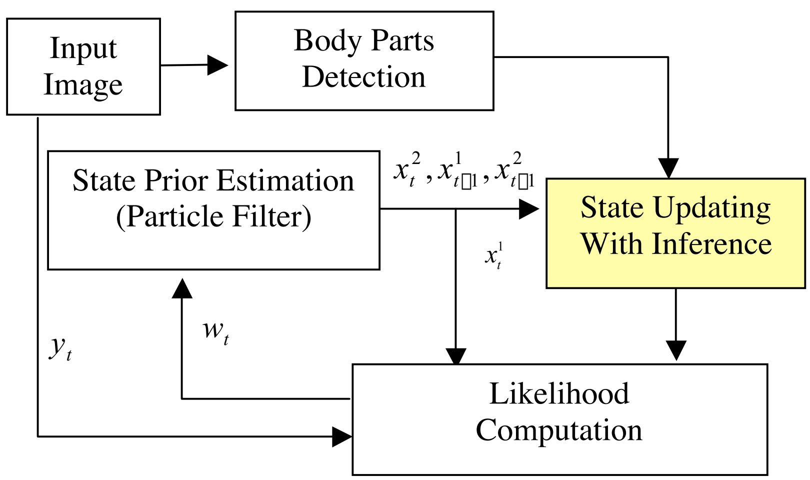
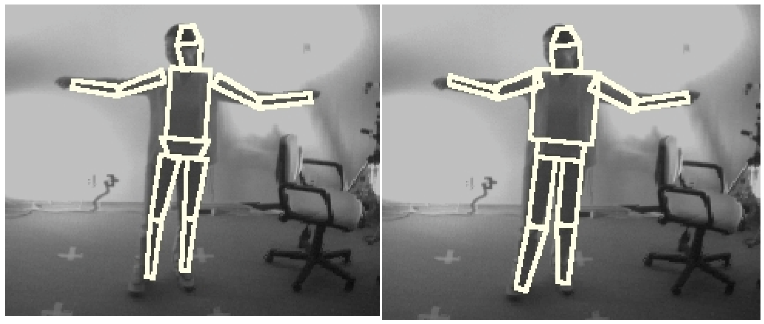


Figure 2. Particle Filtering with Inference

This project also addresses the problem of estimating unknown static parameters (i.e. height and shape) during tracking. Shape information of the human body is useful in the inference of body visual appearance for accurate tracking. In many previous works, the shape of the human model is manually initialized. In our proposed technique, the density distribution of static parameters is represented by mixture Gaussians where the locations and widths of the Gaussians are represented in the particles. At each time step, the Gaussian locations and width are updated using current observations. A result of this shape estimation method is shown in Figure 3.

With the above, discussed proposed technique, we conducted a number of experiments to track a person in video sequences captured in an indoor environment. A tracking result is shown in Figure 4.



(a) (b)

Figure 3. Shape Parameters Estimation. (a) Inaccurate fit of the human model without adjusting shape parameters, (b) well-fitted human model with estimated shape.

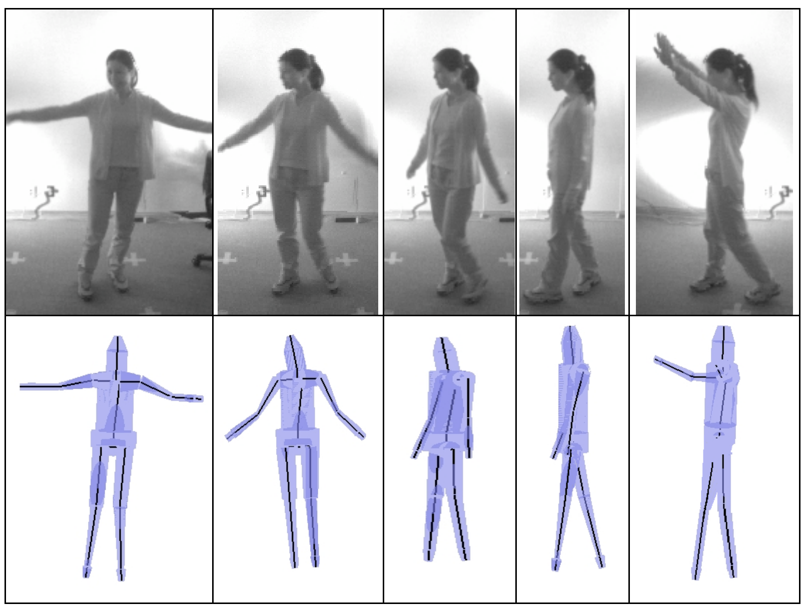


Figure 4. Tracking Result of Turning Person.

5. Short Description of Achievements in Previous Years

This is a new activity within the Communication Vision. We propose to use the visual cues for achieving multimodal interaction with an information system.

5a. Detail of Accomplishments During the Past Year

We have implemented and proposed a number of techniques for articulated body model fitting to silhouettes. These include the integration of inference with particle filter and estimation of human body shape. A portion of this work is published in [3,9].

6. Other Relevant Work Being Conducted and How this Project is Different

Recent works using particle filter for human body tracking have focused on improving efficiency using various methods such as variance analysis [13], simulated annealing approach [5], partitioned sampling [10] and hybrid Monte Carlo approach [2]. Other works include combining particle filter with deterministic gradient descent search [8], and using dynamic motion model [11,12]. However, the high computational complexity remains an important issue for body tracking using particle filter. In addition, these works have not addressed the issues of automatic initialization and robustness of tracking.

This project contributes to this field by developing techniques that integrate the use of multiple cues and inference methods to enable automatic initialization of pose and improving the efficiency and robustness of tracking. We also proposed a method of estimating the human body shape automatically.

7. Plan for the Next Year

We are currently developing techniques to incorporate color-based features into the tracking framework. Color features are richer and generally more stable compared to edge and silhouette features. Incorporating color information can potentially improve the tracking performance. We are exploiting the use of histogram-based tracking technique (Mean-shift algorithm) for tracking local body parts such as hands and legs. We are also addressing the problem of learning the appearance model of the human.

These developments aim to improve the accuracy of the body tracking. In addition, we hope to expand the applications of developed techniques to more challenging tasks such as tracking multiple persons and gesture recognition.

8. Expected Milestones and Deliverables

For the next five-years, we will focus on expanding current method for recognizing a larger set of postures and body gestures for multimodal interactions.

- Improving the tracking using color cues
- Improving the numerical complexity for achieving near real-time performances
- Gesture recognition from the fitted articulated body model
- Computer interaction using basic gestures.
- Multimodal interaction using body motion

9. Member Company Benefits

This work was partially funded by the Institute for Creative Technologies. This work was partially funded by the Institute for Creative Technologies. We initiated collaboration with occupational therapy specialists that identify this technology to be very important for the locomotor rehabilitation process. A potential use of the technology for virtual prototyping is under evaluation by the Japanese Aerospace Corporation.

10. References

[1] M. S. Arulampalam, S. Maskell, N. Gordon, T. Clapp, "A Tutorial on Particle Filters for Online Nonlinear/Non-Gaussian Bayesian Tracking", IEEE Transactions Signal Processing, vol. 50, No. 2, Feb 2002.

[2] K. Choo, and D.J. Fleet, "People tracking with hybrid Monte Carlo." ICCV 2001, vol II, pp. 321-328.

[3] I. Cohen, M. Lee, "3D Body Reconstruction for Immersive Interaction", Second International Workshop on Articulated Motion and Deformable Objects Palma de Mallorca, Spain, 21-23 November, 2002.

[4] J. Deutscher, A. Blake, B. North, and B. Bascle, "Tracking through singularities and discontinuities by random sampling." ICCV 1999, vol. 2, 1144-1149.

[5] J. Deutscher, A. Blake, I. Reid, "Articulated Body Motion Capture by Annealed Particle Filtering," CVPR 2000. vol 2, 126-133.

[6] N. J. Gordon, D. J. Salmond, A. F. M. Smith, "Novel approach to nonlinear/non-Gaussian Bayesian state estimation," Proc. Inst. Elect. Eng. F, v140, n2, 1993.

[7] M. Isard, A. Blake, "CONDENSATION - conditional density propagation for visual tracking," IJCV, 1998.

[8] M. Isard and A. Blake, "ICondensation: Unifying low-level and high-level tracking in a stochastic framework", ECCV 1998. vol. 1, pp 893-908.

[9] Mun Wai Lee, Isaac Cohen, Soon Ki Jung, "Particle Filter with Analytical Inference for Human Body Tracking" Workshop on Motion and Video Computing, 2002, pp. 159-165.

[10] J. MacCormick and M. Isard, "Partitioned sampling, articulated objects, and interface-quality hand tracking," ECCV 2000, vol 2, pp. 3-19.

[11] H. Sidenbladh, Michael J. Black, D. J. Fleet, "Stochastic Tracking of 3D Human Figures Using 2D Image Motion" ECCV 2000, pp. 702-718.

[12] H. Sidenbladh, M. J. Black, and L. Sigal, "Implicit probabilistic models of human motion for synthesis and tracking," ECCV 2002. vol. 1, pp. 784-800.

[13] C. Sminchisecu, B. Triggs, "Covariance Scaled Sampling for Monocular 3D Body Tracking" CVPR 2001 v1, 447-454.