3D Hand and Fingers Reconstruction

1. Research Team

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

Graduate Students: Sung Uk Lee

2. Statement of Project Goals

The need for accurate detection and tracking of hand motion has been increasing in the HCI (Human and Computer Interaction) research area. Articulated hand motions can be used as a convenient user interface for various applications in which the state and motion of the hands are inferred from a set of cameras. The goal of this project is to understand and reconstruct human hand motion with natural articulation and occlusion in real time.

3. Project Role in Support of IMSC Strategic Plan

Multimodal communications rely on speech, haptics and visual sensing of the user in order to characterize the user's emotional and/or physical state. The proposed research focuses on augmenting the interaction capabilities of a multimodal system by providing an automatic tracking of hands motion and recognizing the corresponding hands gestures. The objective here is to augment or replace the mouse and keyboard paradigm with functionalities relying on natural hand motions for driving a specific application.

4. Discussion of Methodology Used

The 3D shape of the hands is reconstructed from 2D hands' silhouettes extracted from two calibrated cameras using GC-based approach [1]. Hand pose information can be computed by matching these synthesized shapes to pre-defined hand models or learning hands' postures using 3D shape description [8]. The previous framework we developed for 3D body reconstruction [1] was modified for 3D hands reconstruction. Due to the large amount of self occlusions of the fingers and the incomplete representation of the 3D shape from the silhouettes we segment the detected silhouettes into separate fingers: we compute edge information in the segmented hand region (inside the silhouettes) to reconstruct each finger as a cylindrical form. This local edge analysis allows us to reconstruct the 3D hand shape accurately when the fingers are splayed or when they are not. In Figures 1 and 2 we show the advantage of using such local edge analysis in characterizing accurately the fingers description.

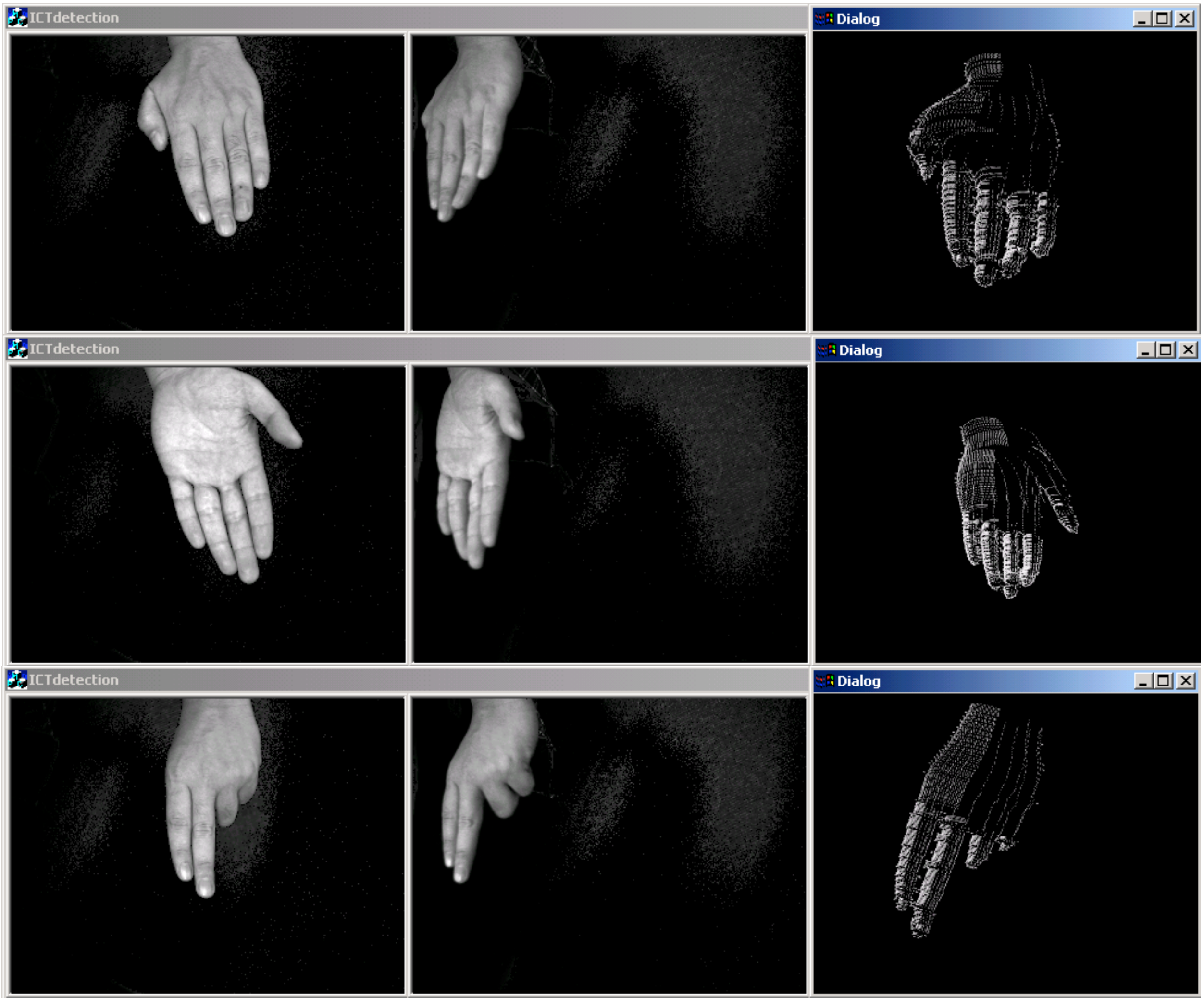

Figure 1. 3D hand reconstruction from two cameras. Each finger is reconstructed as a generalized cylinder by local edge analysis.

Figure 2. 3D shape of the inferred from the detected silhouettes without local edge analysis.

Using 3D articulated hand model and the corresponding hand motion constraints; we can fit and track an articulated model to the detected silhouettes. Figure 3 shows our 3D hand model, it consists of 15 joints and 20 degrees of freedom. Each finger consists of 3 joints and a fingertip.

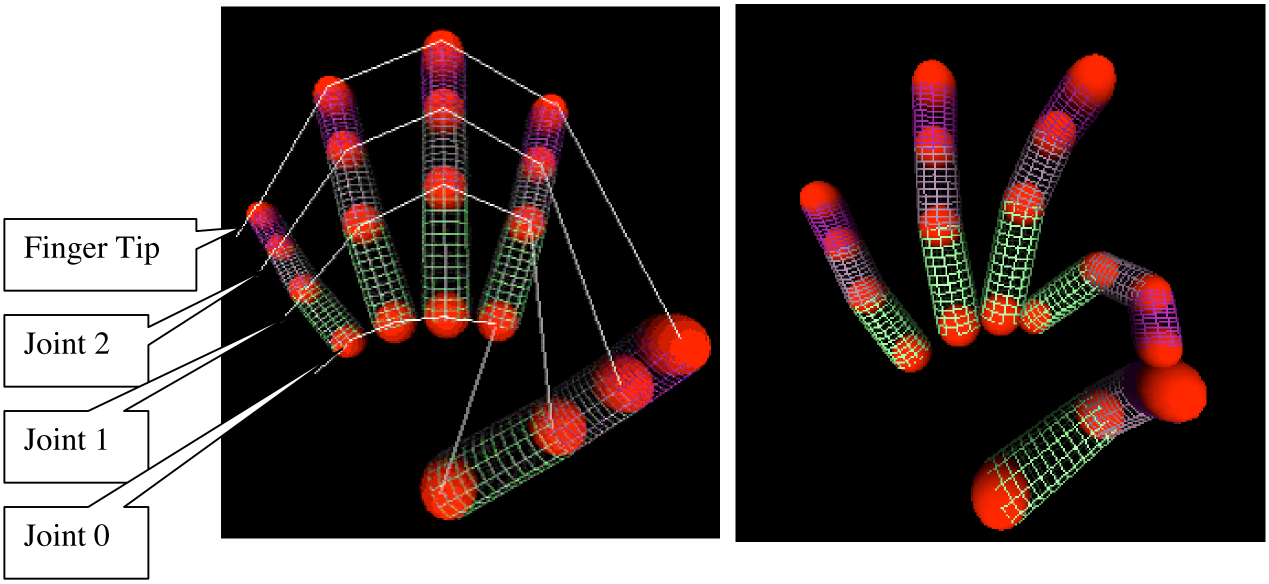

Figure 3. The considered 3D hand model.

Cylinders of variable radius connect each joint. The thickness and length of the cylinders can also be defined. 0^{th} joints of all fingers have 2 degrees of freedom and other joints have a single rotation axis. Therefore, the suggested model has 20 degrees of freedom. All joints have margin of joint rotation and the constraints of hand motion have been modeled in two ways: the intra-finger motion constraint, including the joint relations in a finger and the inter-finger motion constraint for relationship between adjacent fingers.

The 3D hand model is initialized automatically by detecting the initial pose. When concave and convex positions of hand are detected, they are examined by the hand model definition and then the size and position of each joint for 3D hand model are estimated proportionally (see Figure 4).

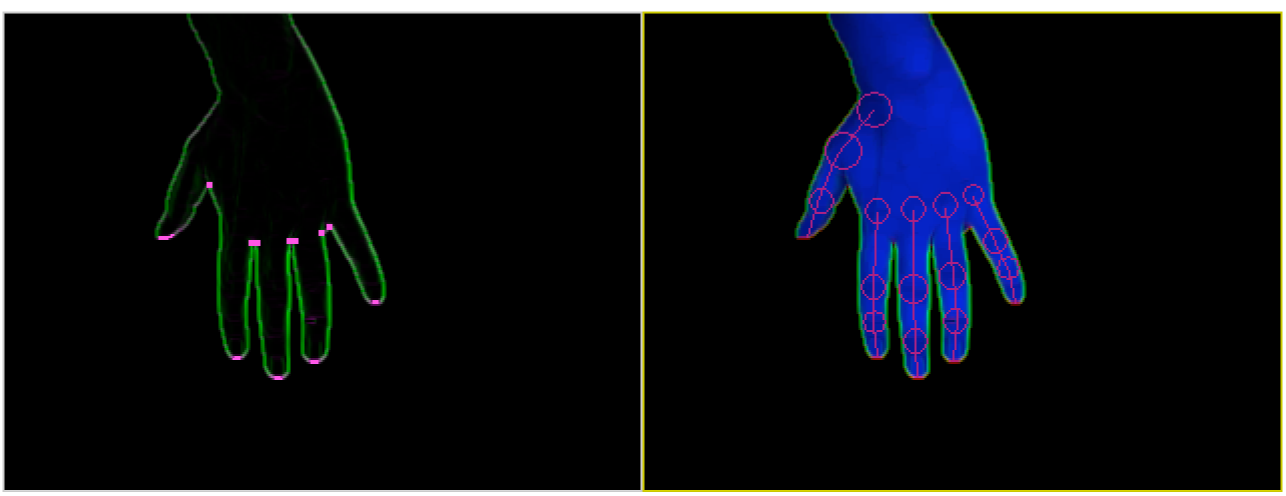
Figure 4. Automatic initial pose detection and model initialization.

5. Short Description of Achievements in Previous Years

This is a new research activity that started in Fall 2002.

5a. Detail of Accomplishments During the Past Year

The goal of this project is to understand and reconstruct human hand motion with natural articulation and occlusion in real time. The objective of this research effort is to augment or replace the mouse and keyboard paradigm with functionalities relying on natural hand motions for driving a specific application. During this year we have focused on defining the problem and modeling the 3D hand and its motion and shape constraints. The local shape analysis depicted in Figures 1 and 2 allowed us to generate accurate 3D shapes of the hand and fingers using a GC-based reconstruction. We have started to implement a particle filter technique for fitting and tracking hands and fingers. The particle filter technique relies on a good estimation of the initial hand and finger shape in order to track them efficiently. We have devised a method for automatic initialization of the 3D hand pose that will allow us to use efficiently a particle filtering technique for fitting and tracking the user's hand motion.

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

A large number of studies focused on the 3D hand modeling. We can group them into three classes: model-based approaches, appearance-based approaches and hybrid approaches. Model based approaches [2,3] rely on the physical properties of the hand and finger motion in order to infer it from a single or multiple views. Appearance based approaches [4] rely on the appearances of the various hand configurations in order to characterize the actual hand and finger pose. They rely on learning-based approaches trained on a large data set that depicts all possible configurations. Mixed (Hybrid) approaches [5-7] employ variety visual cues and hand motion constraints for efficiently searching the 20 dimension space. In our approach, we employ the visual cues and hand motion constraints to enhance the accuracy and efficiency of the fitting. However, we are using two views to accurately track hand position and motion in the 3D space.

In Figure 5 we show an example of the proposed approach.

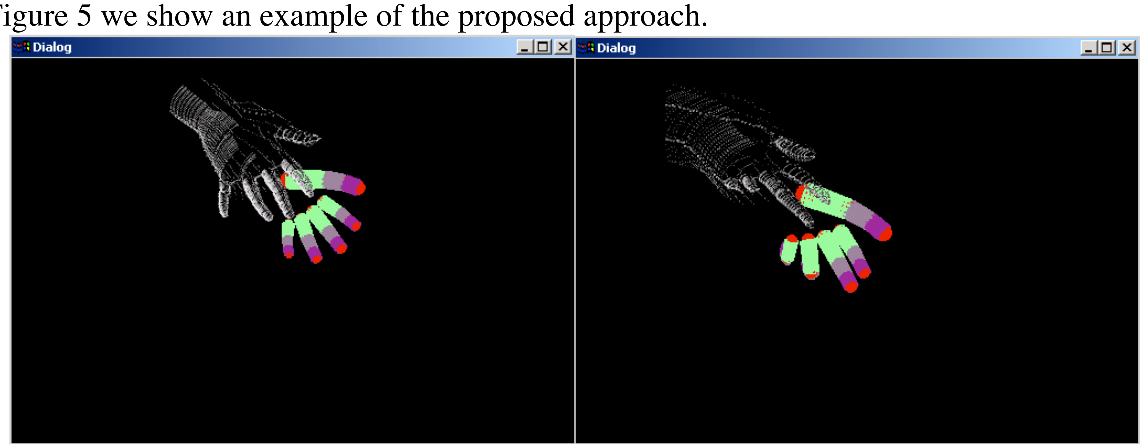

Figure 5. Example of fitting an articulated hand model (in color) to the GC-based hand reconstruction (in white).

7. Plan for the Next Year

We plan to extend the capabilities of the current system to automatically characterize hand position and motion by fitting and tracking an articulated model of the hand to the detected silhouettes.

8. Expected Milestones and Deliverables

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

- Articulated model fitting and tracking
- Efficient use of particle filtering technique for real time performances

9. Member Company Benefits

We are investigating two potential applications: manipulation of 3D objects displayed on a 3D LCD and evaluation of handwriting deficiencies. We initiated a collaboration with occupational therapy specialists that identifies this technology to be very important for characterizing hand-writing deficiencies among persons afflicted with ADD syndrome. A potential use of the technology for virtual prototyping is under evaluation by the Japanese Aerospace Corporation

10. References

[1] Isaac Cohen and Mun Wai Lee, “3D Body Reconstruction for Immersive Interaction”, *Second International Workshop on Articulated Motion and Deformable Objects Palma de Mallorca,* Spain, 21-23 November, 2002.

[2] James M. Rehg and Takeo Kanade, "Model-Based Tracking of Self-Occluding Articulated Objects", *ICCV1995,* pp.612-617, 1995.

[3] B. Stenger, P. R. S. Mendonça, and R. Cipolla, “Model-Based 3D Tracking of an Articulated Hand”, *In Proc. Conference on Computer Vision and Pattern Recognition (CVPR 2001),* Volume II, pages 310-315, Kauai, USA, December 2001.

[4] Romer Rosales, Vassilis Athitsos, Leonid Sigal, and Stan Sclaroff, "3D Hand Pose Reconstruction Using Specialized Mappings" *IEEE International Conf. on Computer Vision (ICCV).* Canada. Jul. 2001.

[5] John Lin, Ying Wu and Thomas S. Huang, "Modeling Human Hand Constraints", *In Proc. Workshop on Human Motion (Humo2000)*, Austin, TX, Dec., 2000.

[6] Ying Wu, John Lin and Thomas S. Huang, "Capturing Natural Hand Articulation", *in Proc. IEEE Int'l Conf. on Computer Vision (ICCV’01)*, Vol.II, pp.426-432, Vancouver, Canada, July, 2001.

[7] Shan Lu, Gang Huang, Dimitris Samaras and Dimitris Metaxas, “Model-based Integration of Visual Cues for Hand Tracking” *In Proc. WMVC 2002*

[8] I. Cohen, H. Li, M. W. Lee. “Human Body Posture Inference for Immersive Interaction”, ACM, International Workshop on Immersive Telepresence, ACM Multimedia 2002, Juan Les Pins, France. 6 December 2002.