

NANYANG
TECHNOLOGICAL
UNIVERSITY

Model-based Hand Tracking with the Kinect Sensor

LIANG HUI

– School of Electrical & Electronic Engineering

Supervisors:

Prof. Junsong Yuan

– School of Electrical and Electronic Engineering.

Prof. Daniel Thalmann

– Institute for Media Innovation.

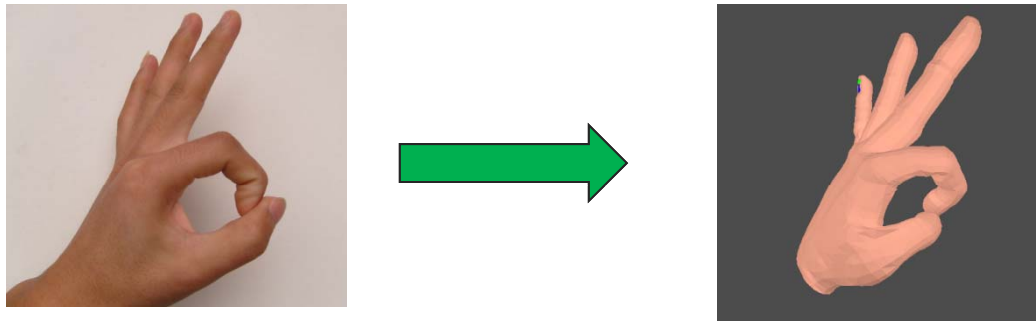


Outline

- Introduction
- Related Work
- Objectives
- System Overview
- Future Work

Introduction

- Restore 27 DOFs of hand motion from the input



- Applications
 - Human-Machine Interaction
 - Hand Gesture Recognition
 - Animation Synthesis

Introduction

- Objectives
 - Speed
 - Accuracy
- Problem
 - High dimension space
 - Self-occlusion

Related Work – By Input

- Color Image Based Tracking [1] [2]
 - Chamfer distance, shape context, color likelihood
 - Limited accuracy
 - Require subsidiary techniques
- Depth Image Based Tracking [3] [4]
 - Robust
 - Restore more degrees of freedom
 - Hand segmentation needs to be refined

[1] Real-time hand tracking with a data glove, in SIGGRAPH 2009 .

[2] Model-Based 3D Hand Pose Estimation from Monocular Video, in TPAMI 2011.

[3] Efficient model-based 3D tracking of hand articulations using Kinect, in BMVC 2011.

[4] A data-driven approach for real-time full body pose reconstruction from a depth camera, in ICCV 2011.

Related Work – By Methodology

- Template-Based Tracking [1] [5] [6]
 - Mapping from feature to pose
 - Quick and robust
 - Limited accuracy
- Model-Based Tracking [2] [3] [7] [8]
 - Enhanced accuracy
 - Slower
 - Sensitive to noisy data and fast motion

[5] Fast Pose Estimation with Parameter-Sensitive Hashing, in ICCV 2003.

[6] Inferring Body Pose without Tracking Body Parts, in CVPR 2000.

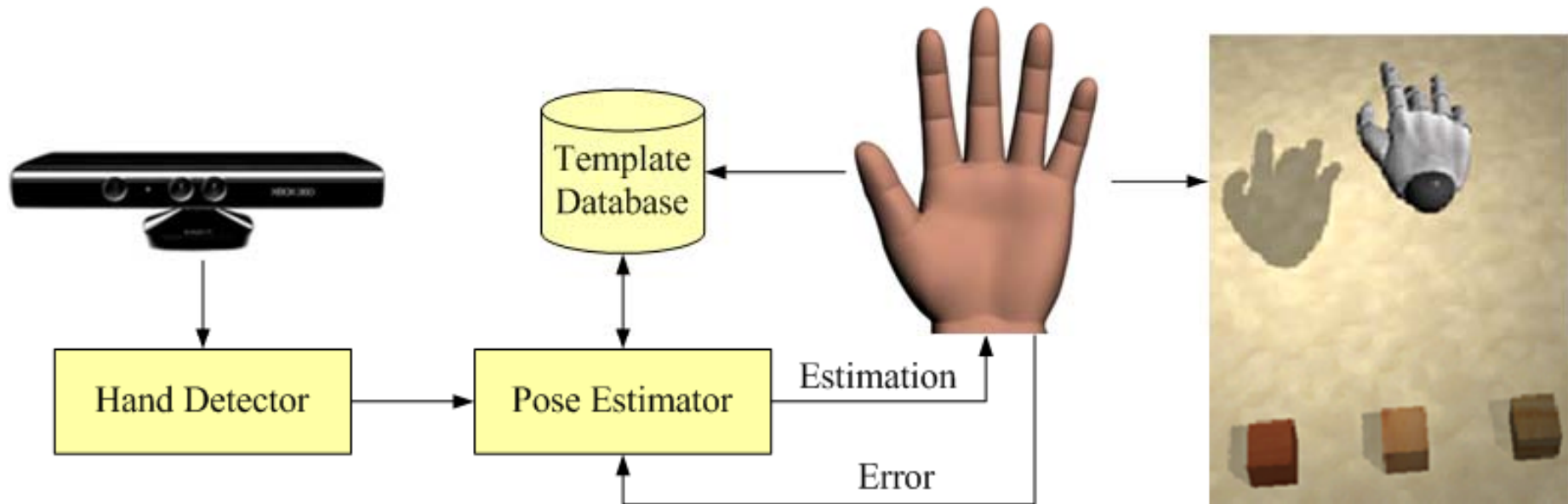
[7] Model-Based Hand Tracking Using a Hierarchical Bayesian Filter, in TPAMI 2006.

[8] Model Based Hand Tracking by Chamfer Distance and Adaptive Color Learning Using Particle Filter, in EURASIP Journal on Image and Video Processing 2009.



Objectives

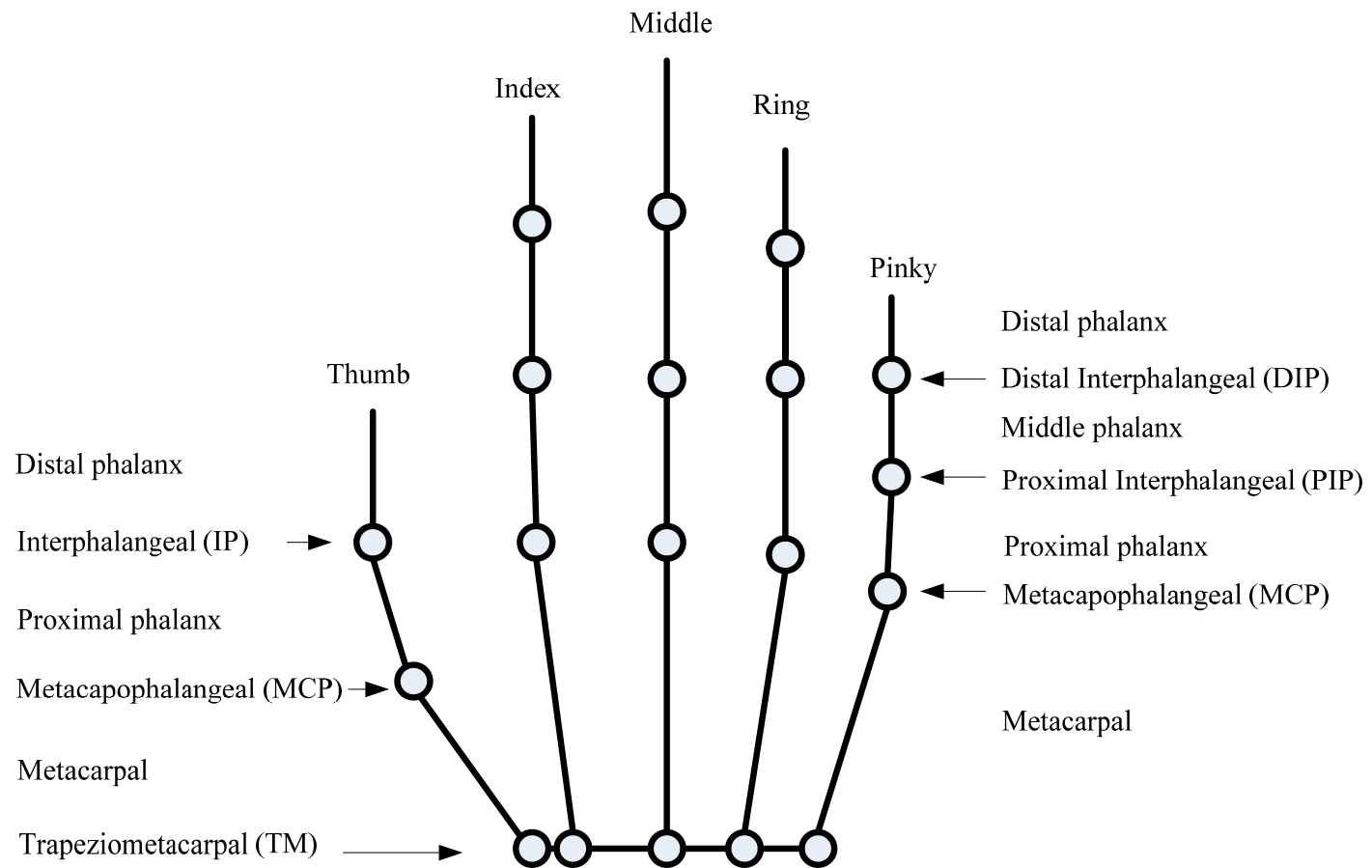
- Object manipulation task recognition
- Modeling of object manipulation in a Virtual Environment



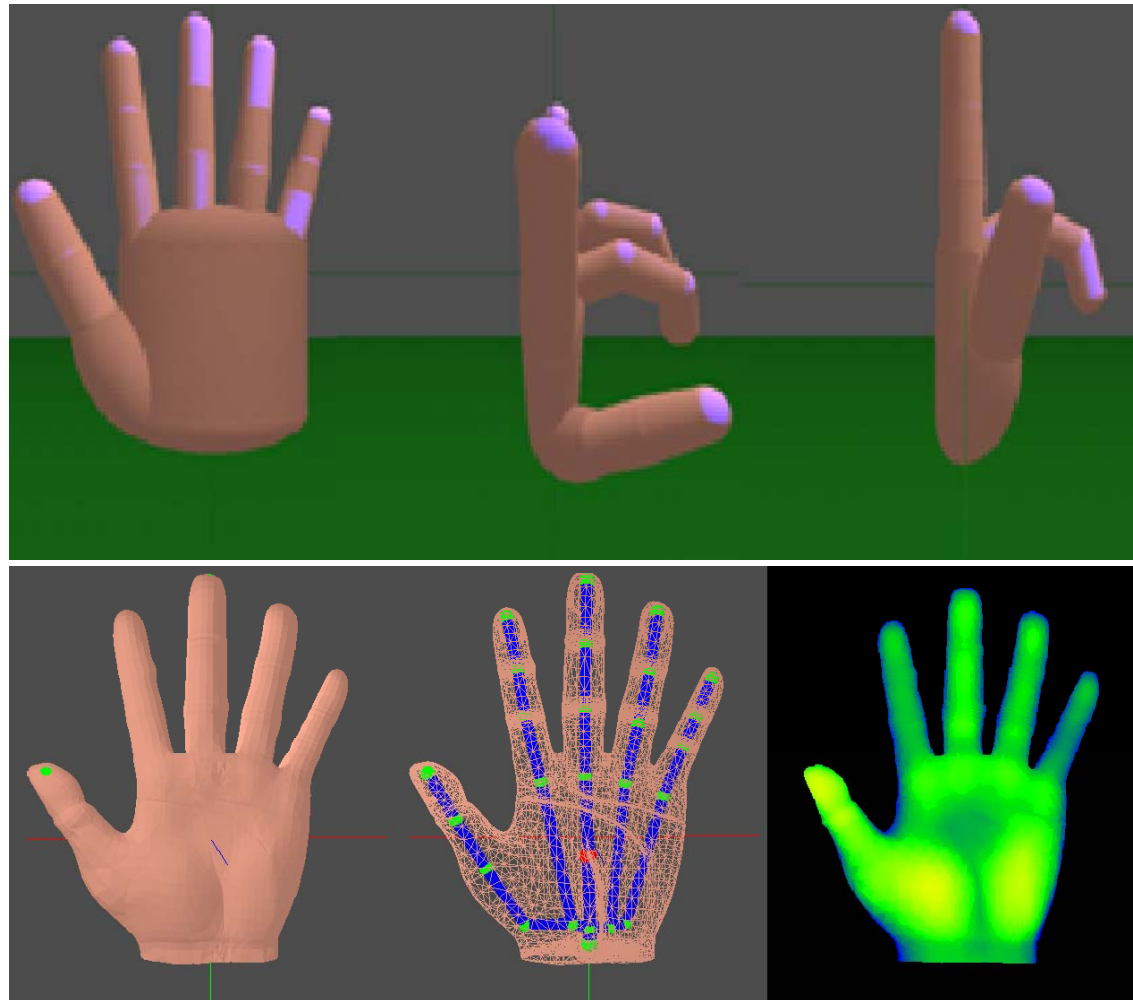
System Overview

- 3D hand model
- Hand detector
- Pose estimator

3D Hand Model



3D Hand Model



Hand Detector – Color based



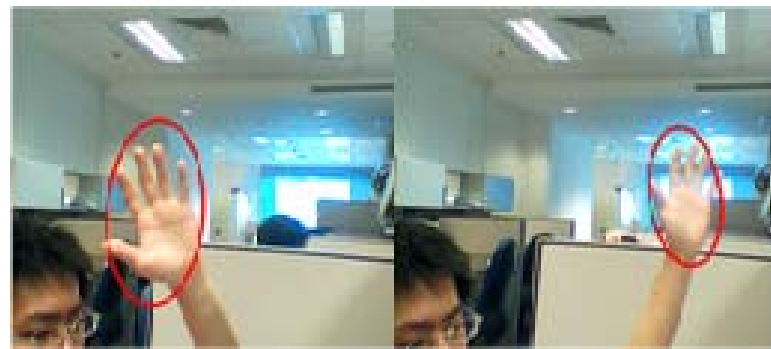
RGB



HSV



YCbCr



MeanShift

Hand Detector – Depth based



(1)



(2)



(3)



(4)

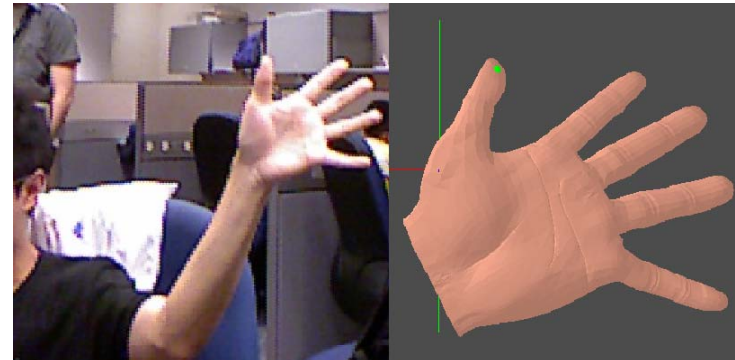
Pose Estimation

- Global Motion
 - 3D Global Translation and Rotation
 - Basis for local hand pose estimation
- Local Hand Pose
 - 15 Degrees of Freedom
 - Objective Function Minimization
 - Initial Estimation
 - Local Optimization

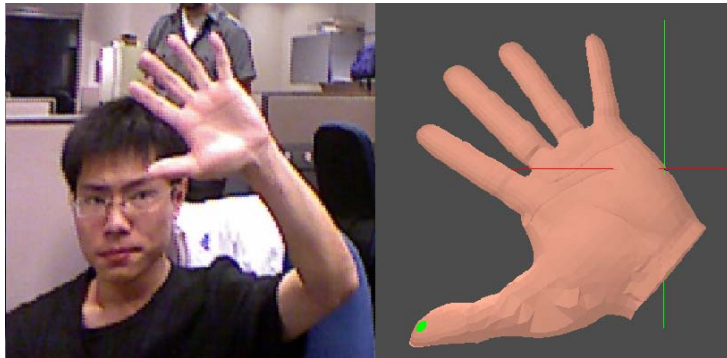
Global Rotation



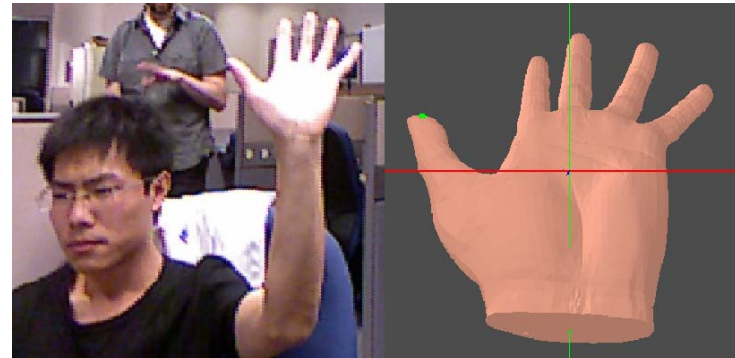
(1)



(2)



(3)

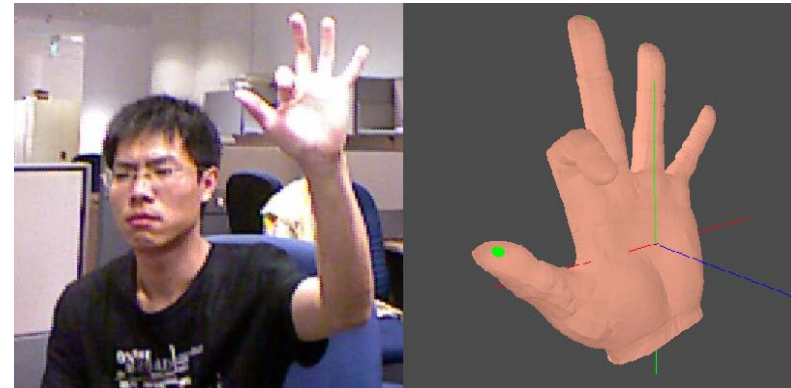


(4)

Local Motion – Initial Estimation



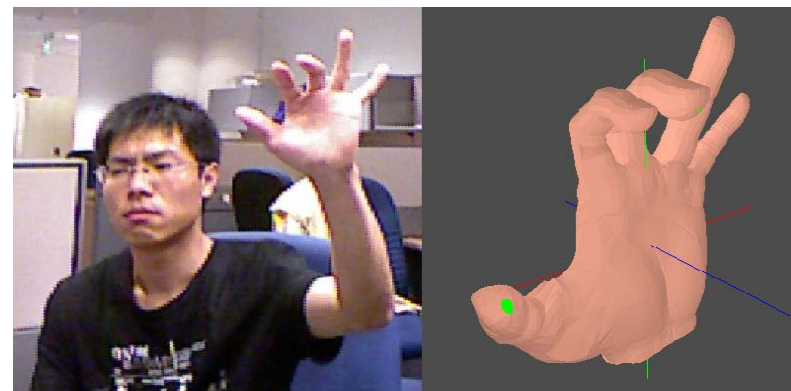
(1)



(2)

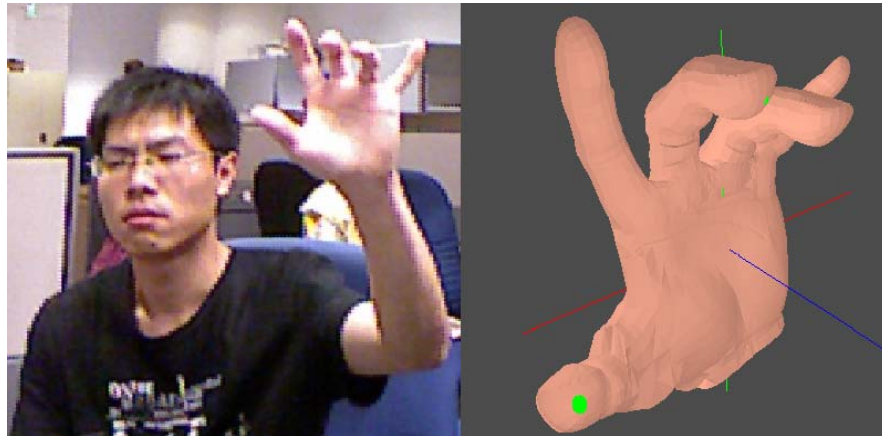


(3)

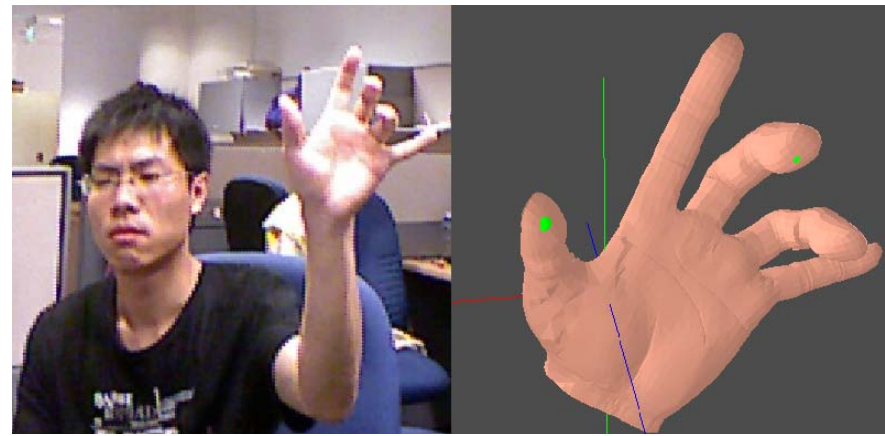


(4)

Local Motion – Initial Estimation



(5)



(6)

Local Motion – Local Optimization

- Objective Function

- Silhouette

- Depth

$$f_C(\theta) = \lambda \frac{|A_I - A_H|}{\max(A_H, A_I)} + (1 - \lambda) \frac{1}{z_M A_O} \sum_{z_n, z_h \in A_O} \min(|z_n - z_h|, z_M)$$

- Optimization Methods

- Nelder-Mead Simplex

- Genetic Algorithm

Local Motion – Local Optimization

- Some Previous Results
 - Optimization using Silhouette difference
 - Simplex Search and Genetic Algorithm



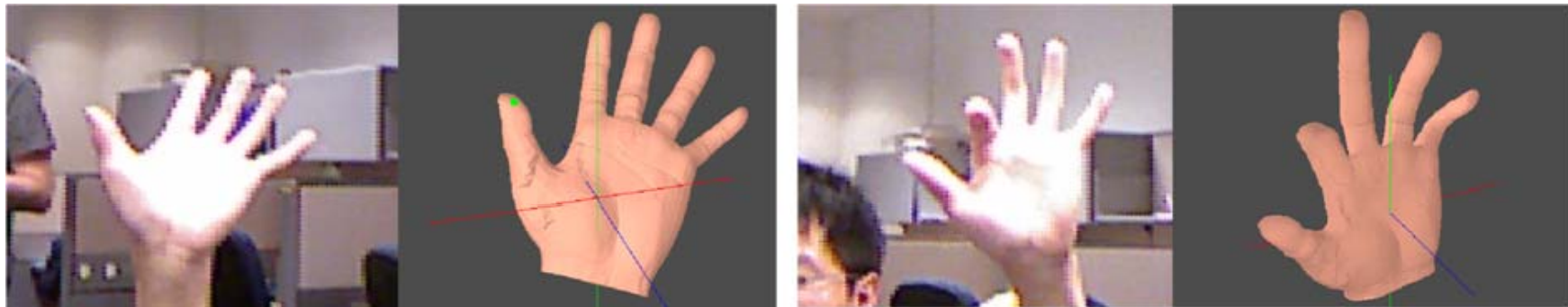
Nelder-Mead Simplex



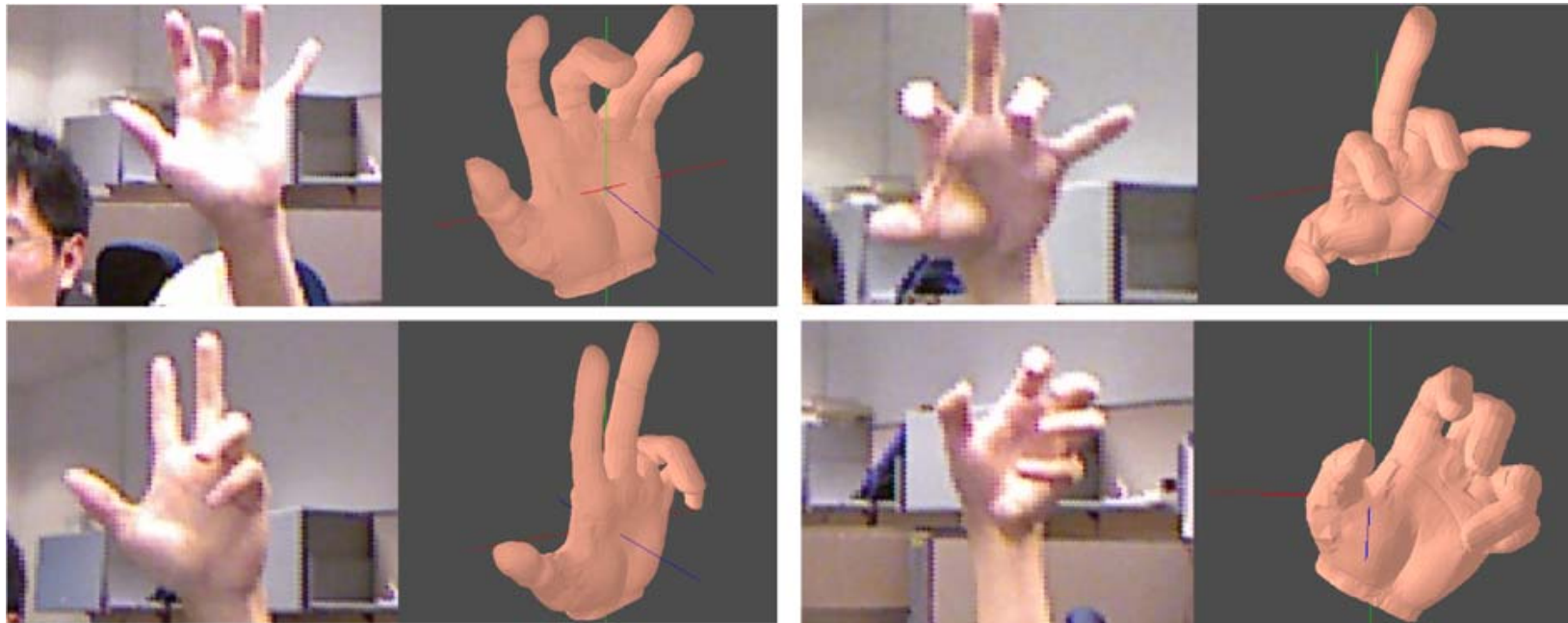
Genetic Algorithm

Local Motion – Local Optimization

- Some Preliminary Results
 - Simplex Search
 - Silhouette and Depth



Local Motion – Local Optimization



Future Work

- Dimension reduction
- Particle filter
- Extended to multi-cameras

Thank You!

