Computer Vision: Reconstruction, Recognition, and Interaction

Linda G. Shapiro
Department of Computer Science and Engineering
University of Washington

OUTLINE

- Reconstruction
  - Kari Pulli's Reconstruction from Range Data
  - Zhenrong Qian's Reconstruction from Visible Human Data
- Recognition
  - Sal Ruiz's 3D Object Recognition and Localization from Range Data
  - Pam Neal's Approach to Class Recognition
- Interaction
  - Habib Abi Rached's Work on Gesture Recognition
  - Mark Billinghurst's Augmented Reality Work

Surface Modeling and Display from Range and Color Data

Kari	Pulli	UW
Michael	Cohen	MSR
Tom	Duchamp	UW
Hugues	Hoppe	MSR
John	McDonald	UW
Linda	Shapiro	UW
Werner	Stuetzle	UW

UW = University of Washington, Seattle, WA USA
MSR = Microsoft Research, Redmond, WA USA

Introduction

Goal

- develop robust algorithms for constructing 3D models from range & color data
- use those models to produce realistic renderings of the scanned objects

Surface Reconstuction

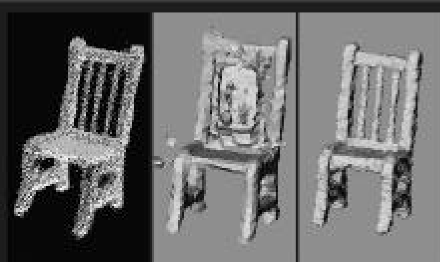
Step 1: Data acquisition
Obtain range data that covers the object. Filter, remove background.

Step 2: Registration
Register the range maps into a common coordinate system.

Step 3: Integration
Integrate the registered range data into a single surface representation.

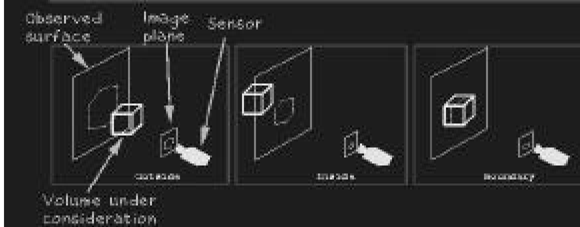
Step 4: Optimization
Fit the surface more accurately to the data, simplify the representation.

Problem



Noisy registered data | Signed distance fn & marching cubes | Hierarchical & directional space carving

Carve space in cubes

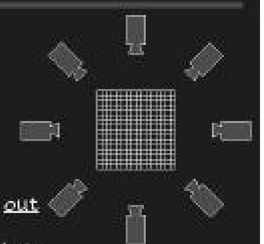


Label cubes

- Project cube to image plane (hexagon)
- Test against data in the hexagon

Several views

Processing order:
FOR EACH cube
FOR EACH view

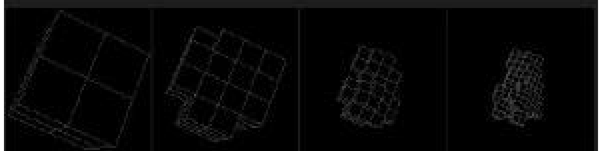


Rules:
<u>any</u> view thinks cube's <u>out</u>
=> it's <u>out</u>
<u>every</u> view thinks cube's <u>in</u>
=> it's <u>in</u>
else
=> it's at boundary

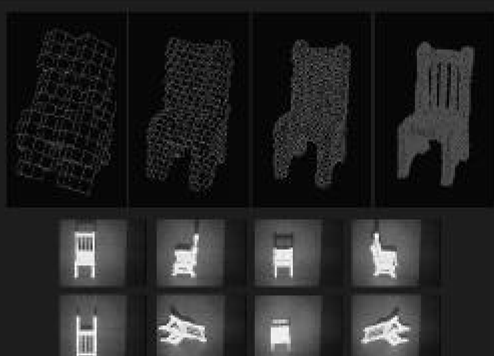
Hierarchical space carving

- Big cubes => fast, poor results
- Small cubes => slow, more accurate results
- Combination = octrees

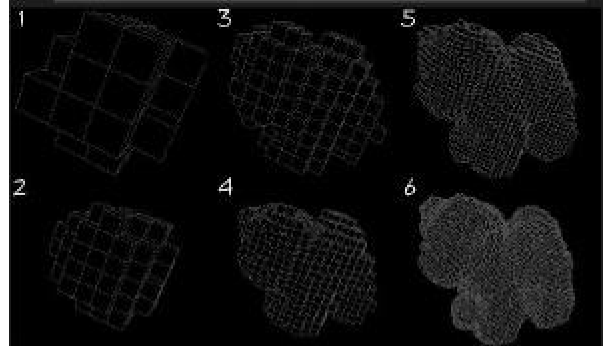
RULES:
- cube's out => done
- cube's in => done
- else => recurse



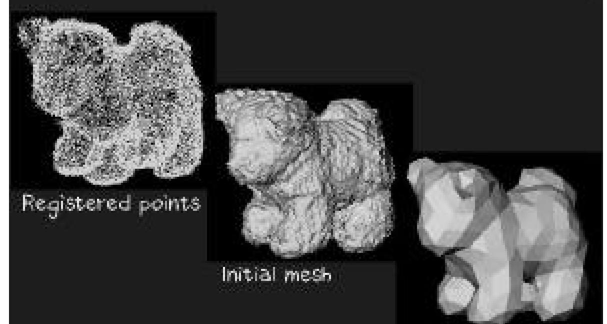
The rest of the chair

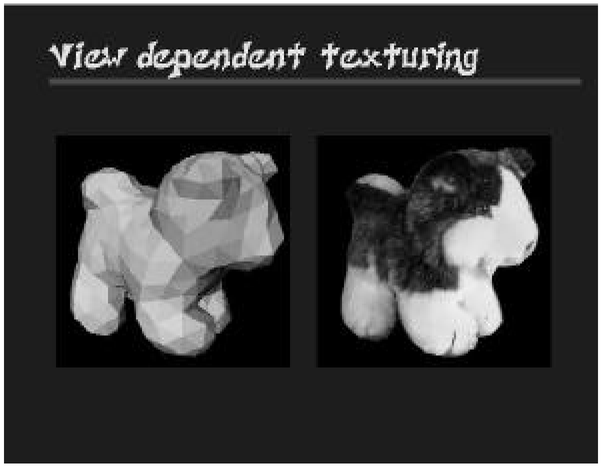


Same for a husky pup

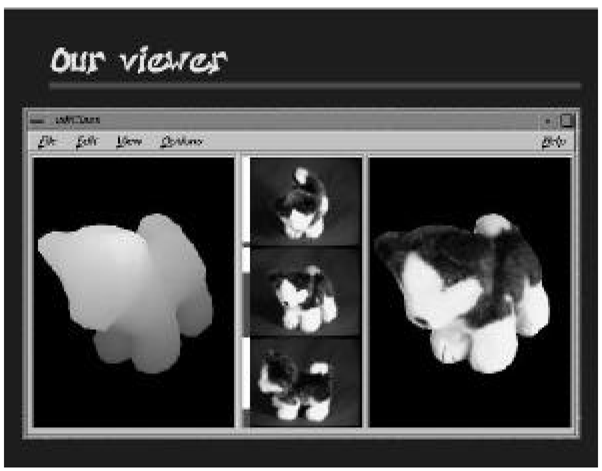


Optimizing the dog mesh

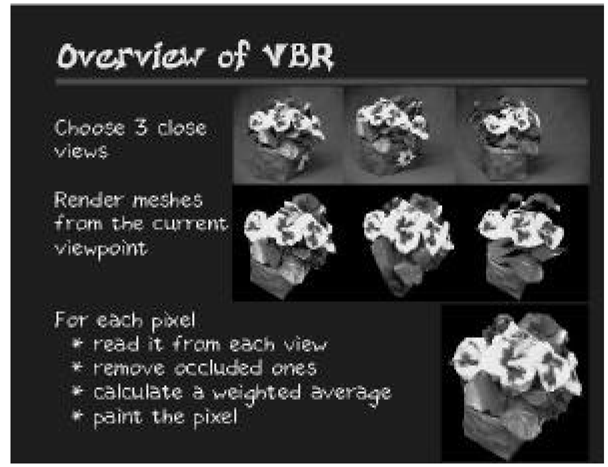



View dependent texturing

13


Our viewer

14


Overview of VBR
Choose 3 close views
Render meshes from the current viewpoint
For each pixel
* read it from each view
* remove occluded ones
* calculate a weighted average
* paint the pixel

15

16

Reconstruction of Blood Vessel Trees from Visible Human Data

Zhenrong Qian and Linda Shapiro

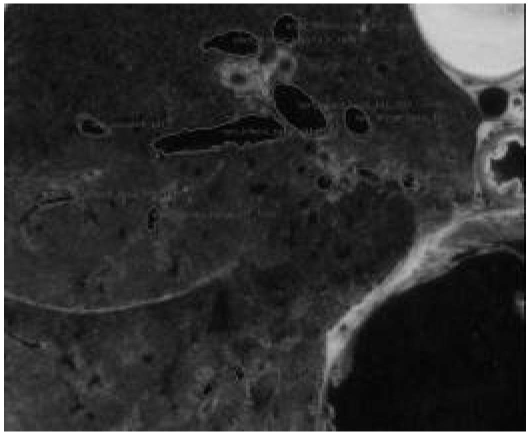
Computer Science & Engineering Department

University of Washington

Introduction

- **Goal**
  - to reconstruct the blood vessels of the lungs from Visible Human Data
- **Computer vision**
  - semi-automation
  - low-level image processing
  - model construction

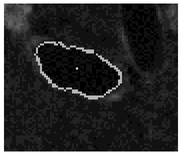
Visible Human Data: Slice through the Lung



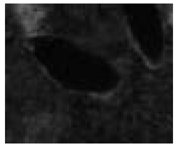
Problems Encountered

- **Data source**
  - black spots that are not blood vessels
  - variations of lighting
- **Characteristics of blood vessels**
  - similar color surrounds
  - lack of knowledge
  - close location
  - shape variety
  - continuous change not expected
  - dense data

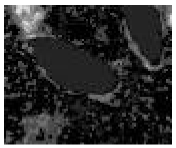
Finding the contours of a vessel being tracked (1)



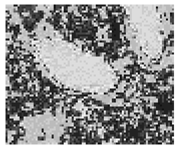
Previous contour



Current slice



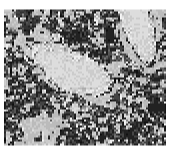
EM Segmentation



False color for the segmentation

Finding the contours of a vessel being tracked (2)

- The results after selecting regions of similar color to the tracked region



Segmentation result

Selected regions

Finding the contours of a vessel being tracked (3)

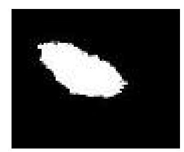
- The results after selecting the region that overlaps most with the previous contour

Selected regions

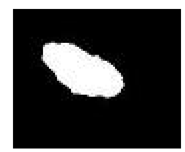
Region that overlaps most

Find the contours of a vessel being tracked (4)

- The results after morphology to close holes and remove noise



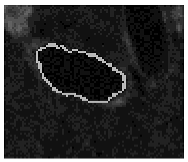
Selected region



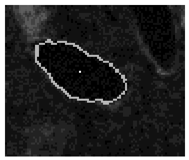
After noise removal

Find the contours of a vessel being tracked (5)

- The contour is determined through a fast-marching level-set approach



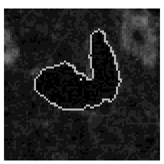
Previous contour

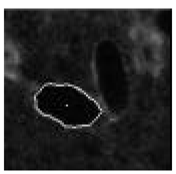


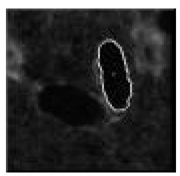
Current contour

How branching is handled

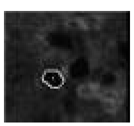
- One contour divides into two

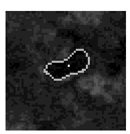


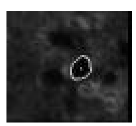




- Two contours merge into one





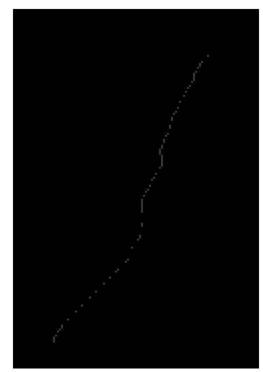


The use of resampling when the axis is not vertical

- **Track** the axis through the center points of found contours
- **Fit** a spline curve
- **Resample** the data perpendicular to the spline curve
- Use the resampled contours for model creation

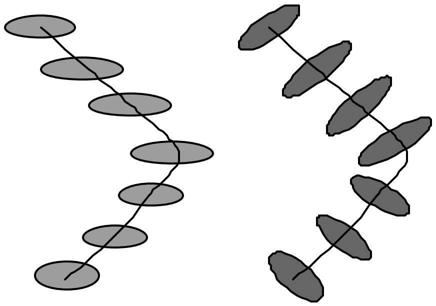
Detect the axis

Center points of found contours



Spline-fitted axis

Resample the data perpendicular to the spline curve



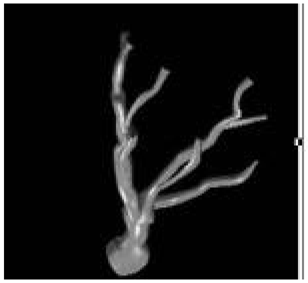
29

Overall Procedure for finding Vessel Trees

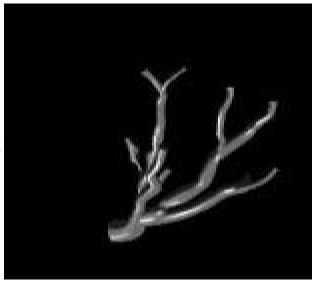
- The user **selects** a starting point
- The program automatically **tracks** the selected vessel and any branches it finds
- The program creates a **generalized cylinder** representation of the vessel tree
- The user may select more starting points

30

Some Initial Results



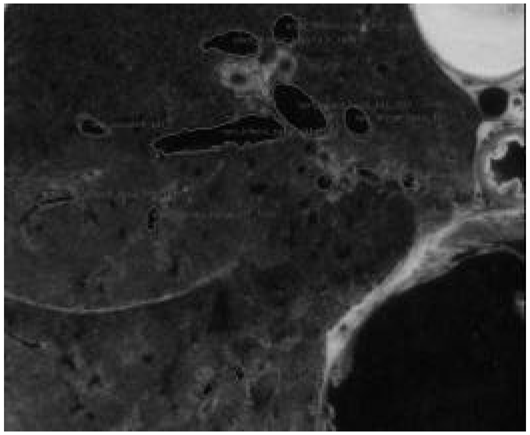
Artery tree from single seed



Vein tree from single seed

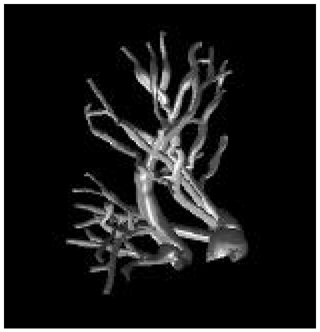
31

Typical Cross Section



32

Results : blood vessels in right lung from previous section



33

34

A Signature-Based Method for Efficient 3-D Recognition

Salvador Ruiz Correa

Linda G. Shapiro

Department of Electrical Engineering

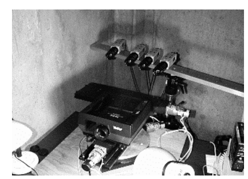
Department of Computer Science & Eng.

University of Washington

35

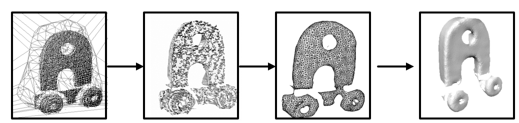
Goal

To develop a compact representation of shape for 3-D object recognition in complex 3-D scenes.



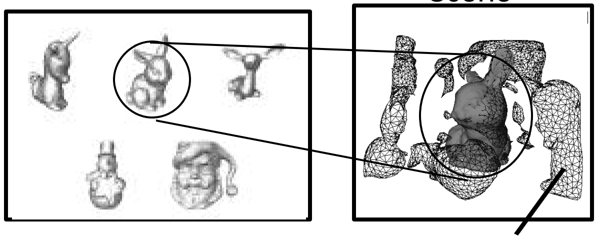
Range Data

Polygonal Mesh



36

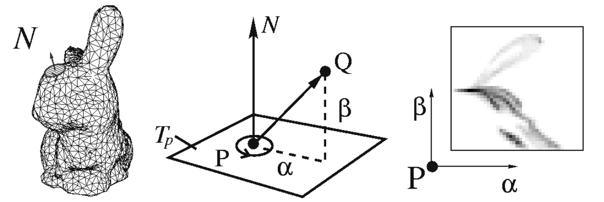
Task



Previous Work
Nonparametric Representations

- Splashes (Stein and Medioni, 1992)
- SAI (Hebert et. al., 1995).
- Point Signatures (Chua and Jarvis, 1997).
- Shape Spectrum (Dorai and Jain, 1997).
- Harmonic Images (Zang, 1999).
- Spin Images (Johnson, 1999).

A Technique for Generating Signatures (Johnson, 1999)



- Accumulate coordinates of contributing points like Q.

Spherical Spin Images

Spin Images: Γ^n

Equivalence Classes Spin Images: Γ^n/E

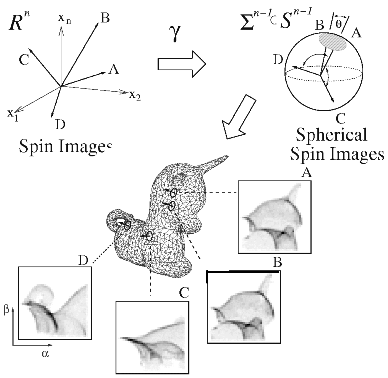
$\Gamma^n \xrightarrow{e} \Gamma^n/E$, $\Gamma^n/E \xrightarrow{g} \Sigma^{n-1}$, $\Gamma^n \xrightarrow{\gamma} \Sigma^{n-1}$

Spherical Spin Images: Σ^{n-1}

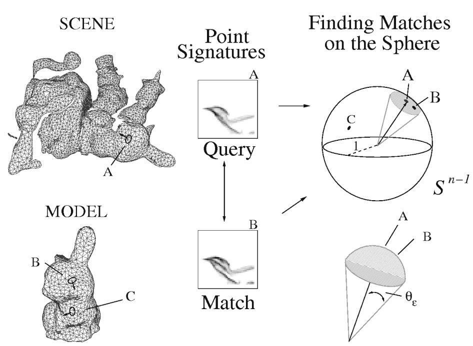
Generalizes to arbitrary surface signatures

Metric Space

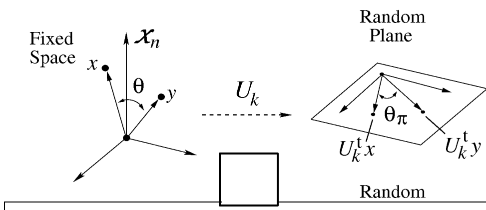
Spherical Signatures



Surface Matching



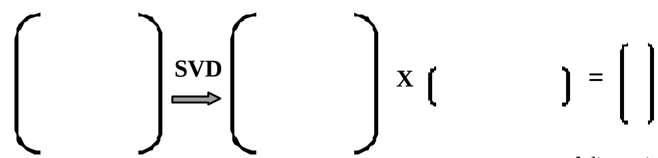
Random Projection



A random projection is a transformation matrix U_k whose columns span a random hyperplane. The plane is distributed according to a unique probability measure that is invariant under rotations.

Implementation

Let **n** be the original dimension and **d** the desired dimension.



n x d matrix of N(0,1) random values

n x d matrix whose d columns are orthonormal vectors

n-dimensional signature

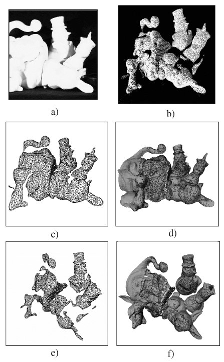
d-dimensional signature

Recognition Examples

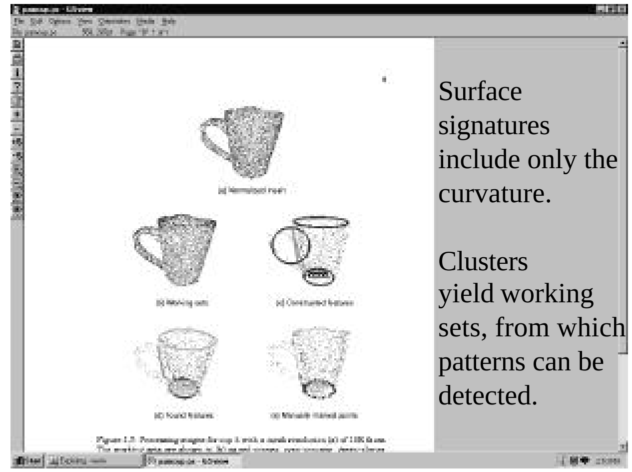
color image | range data

surface mesh front view | recognized models front view

surface mesh side view | recognized models side view



Pam Neal's Prior Mesh Analysis Work



Surface signatures include only the curvature.

Clusters yield working sets, from which patterns can be detected.

Improvements(%)

R=recognition; O=occlusion; C=clutter; L=localization; T=time

Algorithm	RO	RC	LO	LC	T
SS vs SI	3.36	4.24	4.74	6.68	76.11
SS+RP vs SI+RP	13.55	25.82	4.12	-	-
SS+RP vs SI+PCA	26.18	27.73	21.13	13.81	16.12
SI+RP vs SI+PCA	12.67	-	17.01	29.30	15.71

Algorithms

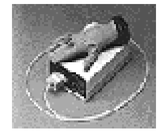
- SI = Spin Images
- SI + RP = Spin Images + Random Projections
- SI+ PCA = Spin Images + Principal Component Analysis
- SS = Spherical Signatures
- SS + RP = Spherical Signatures + Random Projections
- Signature dimension: n=400
- Reduced dimension: k=400,80,60,40,20.

Stereo-based Hand Gesture Tracking and Recognition in Immersive Stereoscopic Displays

Habib Abi-Rached
HITLab (Human Interface Technology Lab)
Electrical Engineering Department
University of Washington

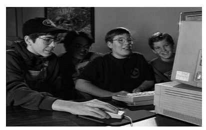
Tuesday December 18th 2001

Limitation of Current Technology.




- Limited efficiency.
  - Mouse Keyboard…
- No 3D. (Monitors).
- Small FOV. (Monitors).
- Few Degrees of Freedom. (Joysticks, Mice).
- Limited intuitiveness.
- Physical connection.
  - (Gloves, Mice, HMD, phantom, polhemus).
- Precision depends on distance.

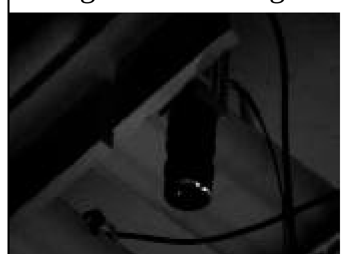
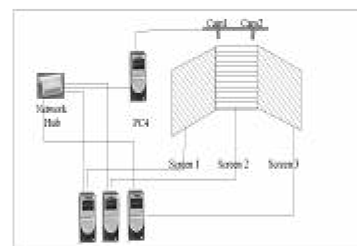
Objective



- Mission: Facilitate communication:
  - Bandwidth.
  - Intuitiveness.
  - Efficiency.
- Means:
  - Visual (Displays, HMD …).
  - Gestural.

Our Approach.

- Inexpensive immersive PC-based gesture tracking / recognition System.

Gesture-based Interaction With 3D Displays.

- Intuitive interaction, easy to learn.

Proposal: Stereo-based Hand Gesture Tracking and Recognition.

- Camera Calibration.
- Stereo matching & reconstruction of the hand.
- Hand modeling.
- Initial pose of the hand model.
- Tracking of the hand.
- Building a gesture library.
- Gesture recognition.
- Selecting a task to measure the goodness of the system.

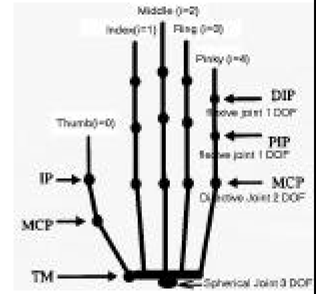
Hand Modeling.

- Dynamic Constraints for all four fingers.

$$\vartheta_{DIP,fe}(i)=\frac{2}{3}\vartheta_{PIP,fe}(i)$$

$$\vartheta_{MCP,aa}=\frac{\vartheta_{MPC,fe}}{90}(\vartheta_{MPC,converge}-\vartheta_{MCP,aa,s})+\vartheta_{MCP,aa,s}$$

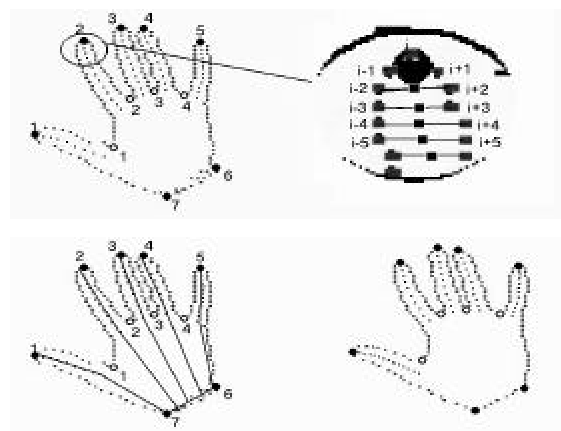
- Static Constraints for all four fingers.



$$0\le\vartheta_{DIP,fe}(i)\le s_{max}(\vartheta_{DIP,fe}(i)) \quad with \quad s_{max}(\vartheta_{DIP,fe}(i))=90$$
$$0\le\vartheta_{PIP,fe}(i)\le s_{max}(\vartheta_{PIP,fe}(i)) \quad with \quad s_{max}(\vartheta_{PIP,fe}(i))=110$$
$$0\le\vartheta_{MCP,fe}(i)\le s_{max}(\vartheta_{MCP,fe}(i)) \quad with \quad s_{max}(\vartheta_{MCP,fe}(i))=90$$
$$-1\le\vartheta_{MCP,aa,s}(2)\le 1$$
$$-15\le\vartheta_{MCP,aa,s}(1,3,4)\le 15$$

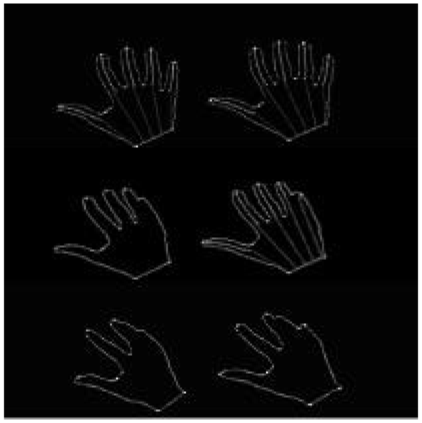
- Kush, Wu.
- Agee 1982.

Initial Pose of the Hand Model.



Hand Tracking.

- Real time model tracking and 3D reconstruction.
- Occlusion and order constraint problem.



Uniqueness of Our Approach:

- Stereo + detailed hand model will give:
  - Precision.
  - Real time performance.
  - 27 degrees of freedom.
- Wire-free system.
- Accuracy independent of distance.

Shared Space: Explorations in Collaborative Augmented Reality

Mark Billinghurst

grof@hitl.washington.edu

HIT Lab, University of Washington

Collaborative Augmented Reality

- Attributes:
  - Virtuality
  - Augmentation
  - Cooperation
  - Independence
  - Individuality
- Seamless Interaction
- Natural Communication

Collaborative AR Interfaces

- Face to Face Collaboration

  WebSpace, Shared Space, Table Top Demo, Interface Comparison, AR Interface Comparison
- Remote Collaboration

  SharedView, RTAS, Wearable Info Space, WearCom Conferencing, BlockParty
- Transitional Interfaces

  MagicBook
- Hybrid Interfaces

  AR PRISM, GI2VIS

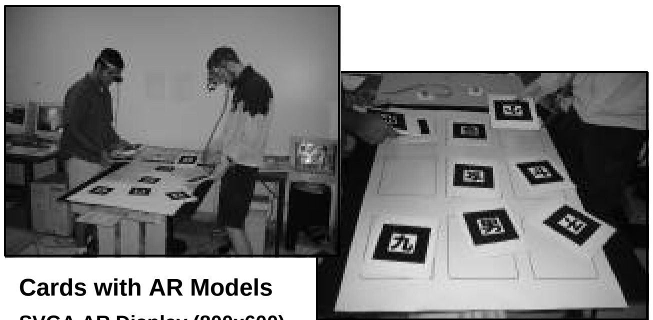
Interface Comparison

- Compare two person collaboration in:
  - Face to Face, AR, Projection Display
- Task
  - Urban design logic puzzle
    - Arrange 9 building to satisfy 10 rules in 7 minutes
- Subjects
  - Within subjects study (counter-balanced)
  - Pilot – 8 pairs grade school children
  - Full – 12+2 pairs of college students

Face to Face Condition

Moving Model Buildings

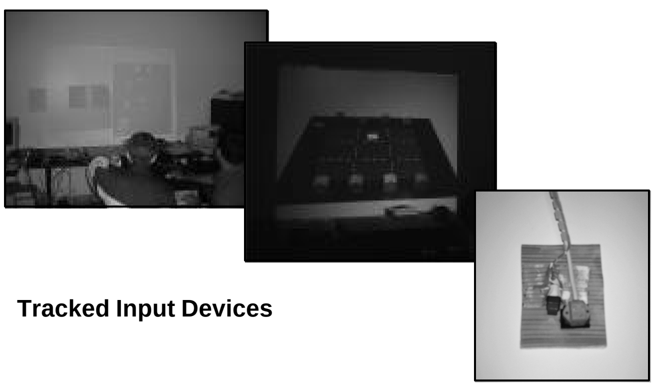
AR Condition



Cards with AR Models

SVGA AR Display (800x600)
Video see-through AR

Projection Condition



Tracked Input Devices

Transitions

- Interfaces of the future will need to support transitions along the RV continuum
- Augmented Reality is preferred for:
  - co-located collaboration
- Immersive Virtual Reality is preferred for:
  - experiencing world immersively (egocentric)
  - sharing views
  - remote collaboration

MagicBook Metaphor