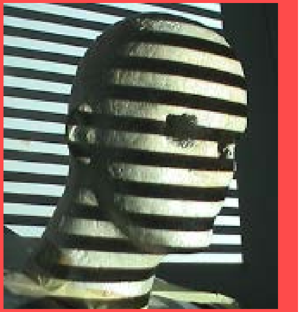


3D Computer Vision Introduction

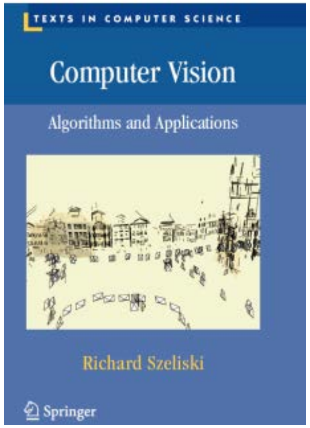
Guido Gerig

CS 6320, Spring 2012

gerig@sci.utah.edu

Administrivia

- Classes: M & W, 1.25-2:45
  Room WEB L126
- Instructor: Guido Gerig
  gerig@sci.utah.edu
  (801) 585 0327
- Prerequisites: CS 6640 ImProc (or equiv.)
- Textbook:
  Computer Vision:
  Algorithms and Applications“
  by Richard Szeliski
- Organization
  web-site
  (slides, documents and assignments)



TEXTS IN COMPUTER SCIENCE

Computer Vision

Algorithms and Applications

Richard Szeliski

Springer

Teaching Assistant

- TA: Kathlea Quebbeman, SoC
  quebbeka@cs.utah.edu
  kathlea.quebbeman@slcc.edu
- HW/SW: Matlab+ev. Imaging Toolbox
  CADE lab WEB 130
  http://www.cade.utah.edu/
- Office Hours: tbd

Prerequisites

- General Prerequisites:
  - Data structures
  - A good working knowledge of MATLAB programming (or willingness and time to pick it up quickly!)
  - Linear algebra
  - Vector calculus
- Assignments include theoretical paper questions and programming tasks (ideally Matlab or C++).
- Image Processing CS 6640 (or equivalent).
- Students who do not have background in signal processing / image processing: Eventually possible to follow class, but requires <u>significant special effort</u> to learn some basic procedures necessary to solve practical computer problems.

Textbook


1.

Computer Vision: Algorithms and Applications (Texts in Computer Science)
by Richard Szeliski (Nov 24, 2010)

Buy new: ~~$89.95~~ **$62.74**

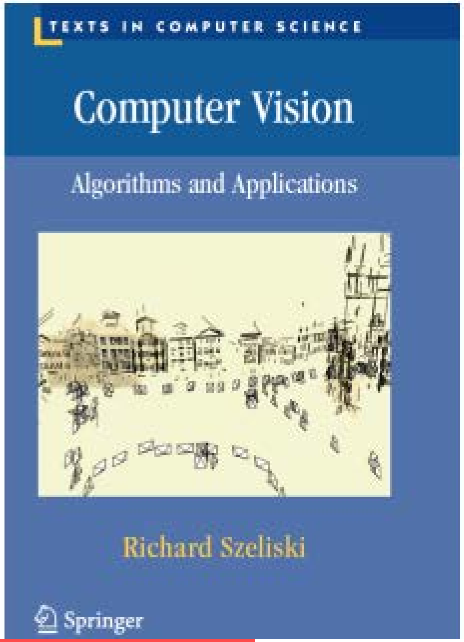
19 new from $62.74

22 used from $59.60

Get it by **Tuesday, Jan 10** if you order in the next **31 hours** and choose one-day shipping.

Computer Vision: Algorithms and Appl

© 2010 Richard Szeliski, Microsoft Research



TEXTS IN COMPUTER SCIENCE

Computer Vision

Algorithms and Applications

Richard Szeliski

Springer

Grading

- Assignments (4-6 theory/prog.): 50%
- Final project (incl. proposal and presentation): 20%
- Class participation: 10%

- Final project replaces final exam

Other Resources

- Cvonline: http://homepages.inf.ed.ac.uk/rbf/CVonline/
- A first point of contact for explanations of different image related concepts and techniques. CVonline currently has about 2000 topics, 1600 of which have content.
- See list of other relevant books in syllabus.

Some Basics

- Instructor and TA do not use WebCT email list (since most students use individual emails)
- It will be <u>your responsibility</u> to regularly read the <u>Announcements</u> on WebCT
- We don't need a laptop for the class, please keep them closed !!!!!
- Please interact, ask questions, clarifications, input to instructor and TA

Syllabus

- See separate syllabus (on blackboard and linked to UofU class list).
- Document

Goal and Objectives

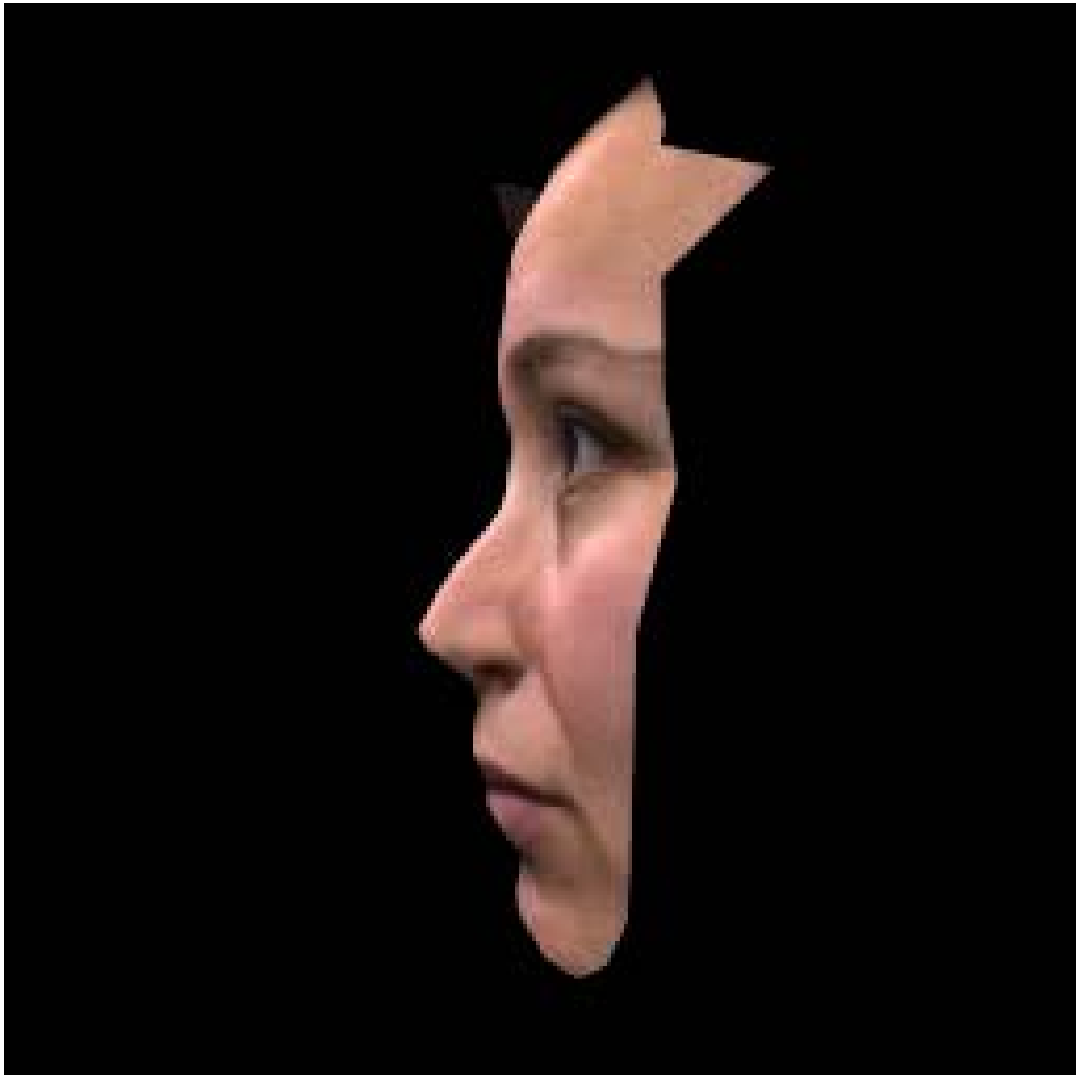
From Snapshots, a 3-D View

NYT, August 21, 2008, Personal Tech

http://www.nytimes.com/2008/08/21/technology/personaltech/21pogue.html

Stuart Goldenberg

Modeling 3D Structure from Pictures or 3D Sensors



Modeling ctd.

Goal and objectives

- To introduce the fundamental problems of computer vision.
- To introduce the main concepts and techniques used to solve those.
- To enable participants to implement solutions for reasonably complex problems.
- To enable the student to make sense of the literature of computer vision.

Why study Computer Vision?

- Images and movies are everywhere
- Fast-growing collection of useful applications
  - building representations of the 3D world from pictures
  - automated surveillance (who’s doing what)
  - movie post-processing
  - CAM (computer-aided manufacturing
  - Robot navigation
  - face finding
- Various deep and attractive scientific mysteries
  - how does object recognition work?
- Greater understanding of human vision

CV: What is the problem?

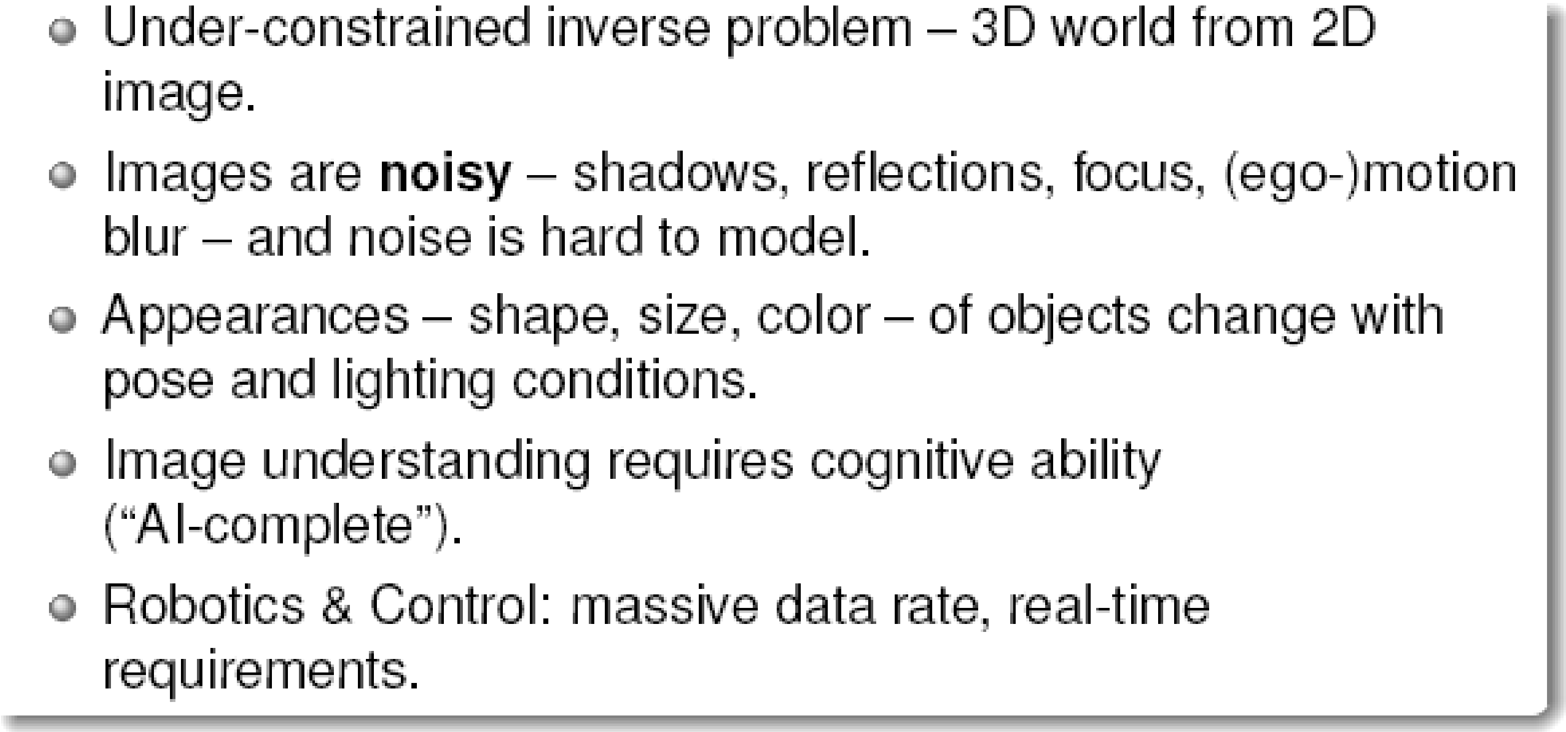
Image Formation: From World to Image

- Camera model (optics & geometry): From points in 3D scene to points on 2D image.
- Photometry: From lights and surfaces in scene to intensity (brightness) and color in image.

Vision: From Image to (Knowledge of the) World

- Reconstruct scene (**world model**) from images.
- Extract sufficient **information** for detection/control **task**.

CV: A Hard Problem



- Under-constrained inverse problem – 3D world from 2D image.
- Images are **noisy** – shadows, reflections, focus, (ego-)motion blur – and noise is hard to model.
- Appearances – shape, size, color – of objects change with pose and lighting conditions.
- Image understanding requires cognitive ability (“AI-complete”).
- Robotics & Control: massive data rate, real-time requirements.

Properties of Vision

- 3D representations are easily constructed
  - There are many different cues.
  - Useful
    - to humans (avoid bumping into things; planning a grasp; etc.)
    - in computer vision (build models for movies).
  - Cues include
    - multiple views (motion, stereopsis)
    - texture
    - shading

Properties of Vision

- People draw distinctions between what is seen
  - "Object recognition"
  - This could mean "is this a fish or a bicycle?"
  - It could mean "is this George Washington?"
  - It could mean "is this poisonous or not?"
  - It could mean "is this slippery or not?"
  - It could mean "will this support my weight?"
  - Great mystery
    - How to build programs that can draw useful distinctions based on image properties.

Main topics

- Shape (and motion) recovery

  “What is the 3D shape of what I see?”
- Segmentation

  “What belongs together?”
- Tracking

  “Where does something go?”
- Recognition

  “What is it that I see?”

Main topics

- Camera & Light
  - Geometry, Radiometry, Color
- Digital images
  - Filters, edges, texture, optical flow
- Shape (and motion) recovery
  - Multi-view geometry
  - Stereo, motion, photometric stereo, …
- Segmentation
  - Clustering, model fitting, probalistic
- Tracking
  - Linear dynamics, non-linear dynamics
- Recognition
  - templates, relations between templates

Camera and lights

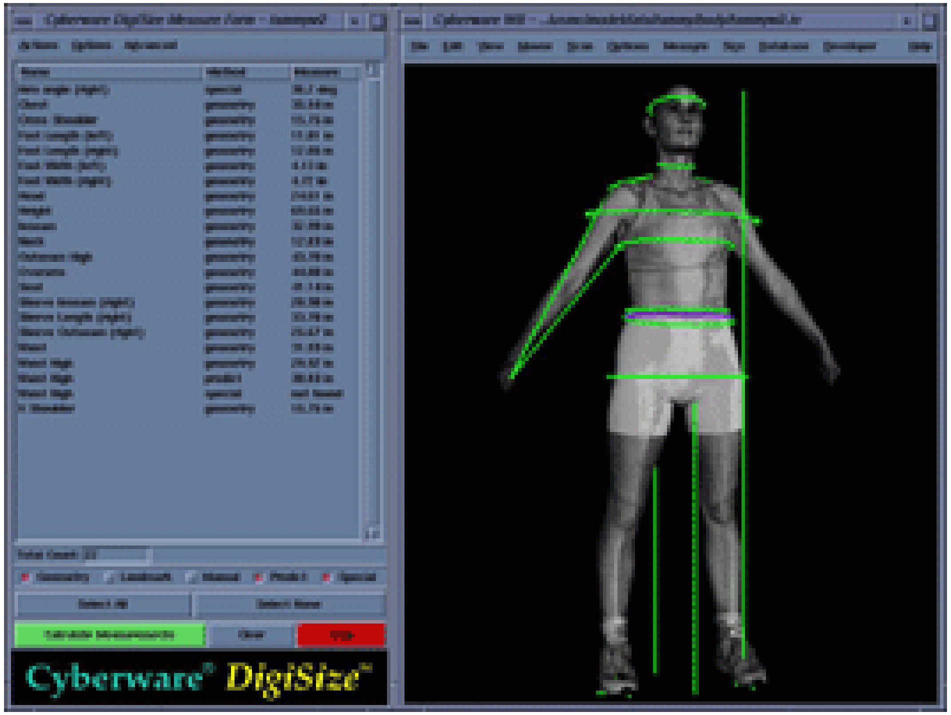
- How images are formed
  - Cameras
    - What a camera does
    - How to tell where the camera was
  - Light
    - How to measure light
    - What light does at surfaces
    - How the brightness values we see in cameras are determined
  - Color
    - The underlying mechanisms of color
    - How to describe it and measure it

Motivation

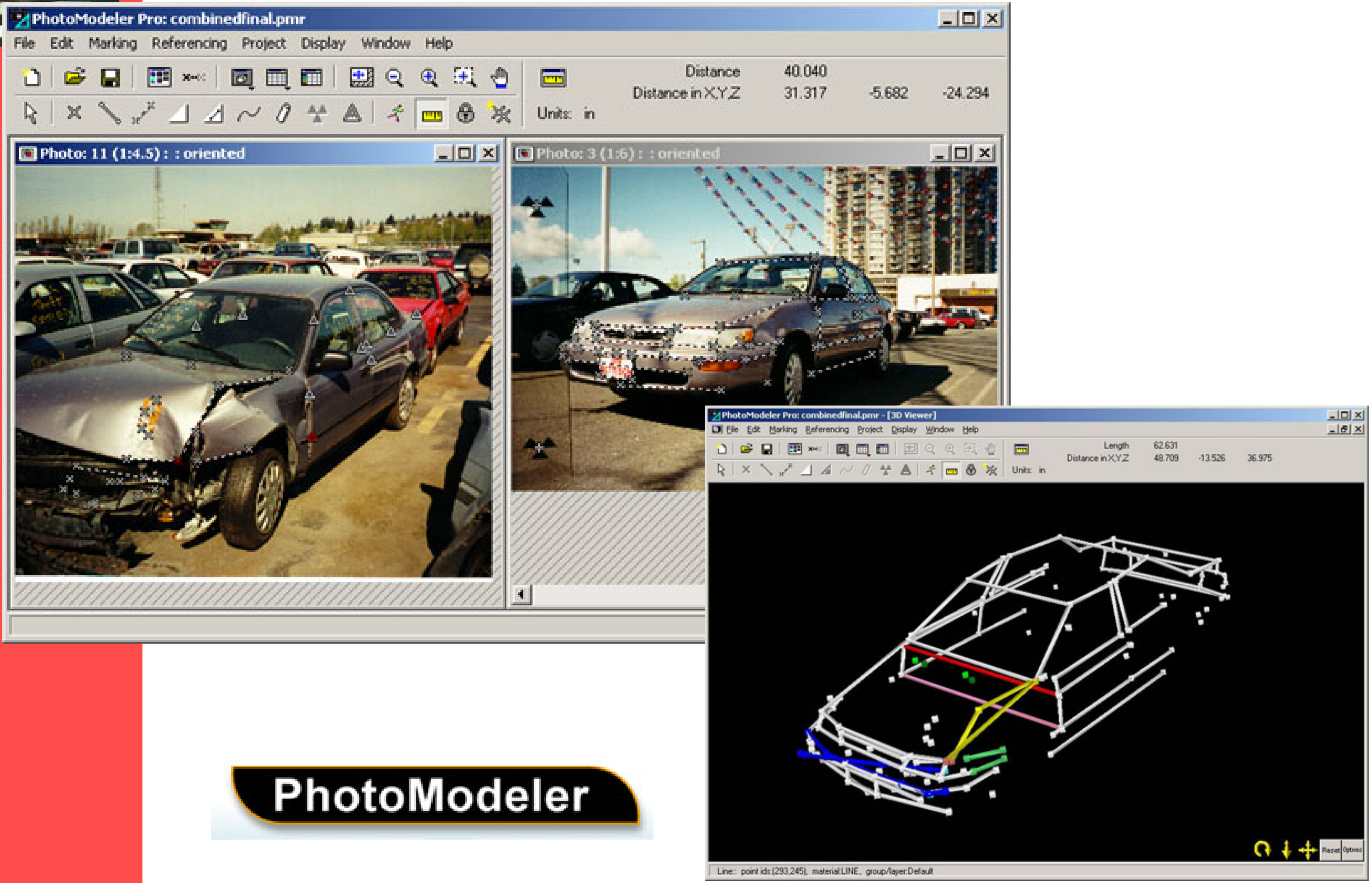
- (some slides modified from Marc Pollefeys, UNC Chapel Hill & ETH Zurich)

Clothing

- Scan a person, custom-fit clothing



Forensics



Distance 40.040
Distance in X,Y,Z 31.317 -5.682 -24.294
Units: in

Length 62.631
Distance in X,Y,Z 48.709 -13.526 36.975
Units: in

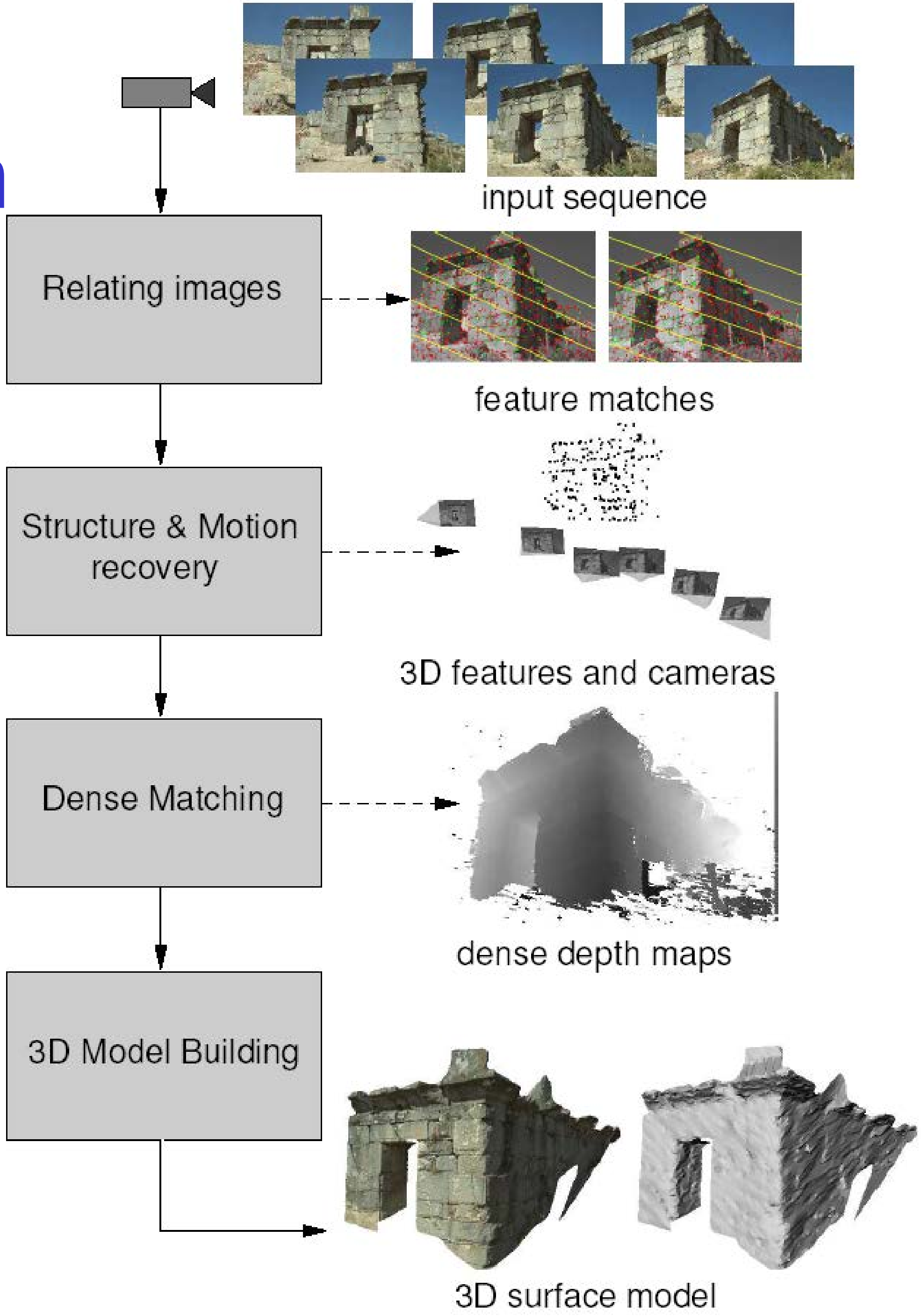
Line:: point ids:(293,245), material:LINE, group/layer:Default

PhotoModeler

3D urban modeling

drive by modeling in Baltimore

Structure from Motion



input sequence

Relating images

feature matches

Structure & Motion recovery

3D features and cameras

Dense Matching

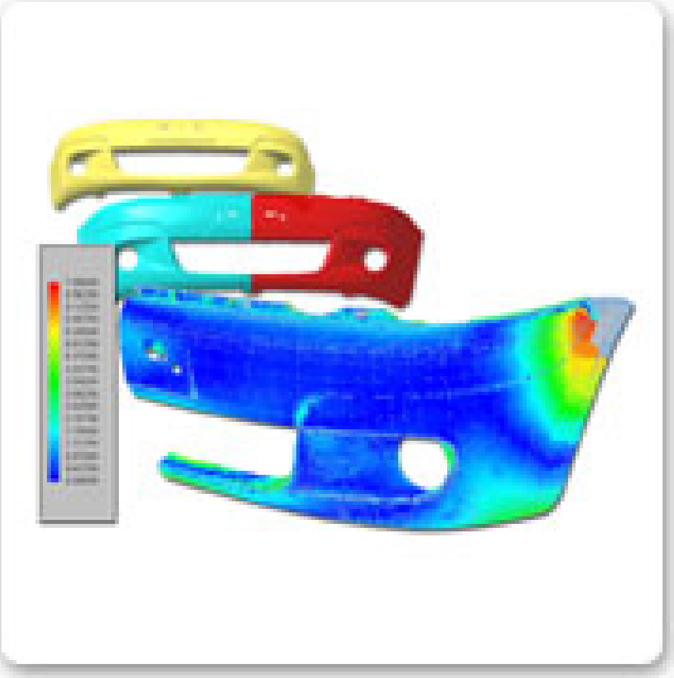
dense depth maps

3D Model Building

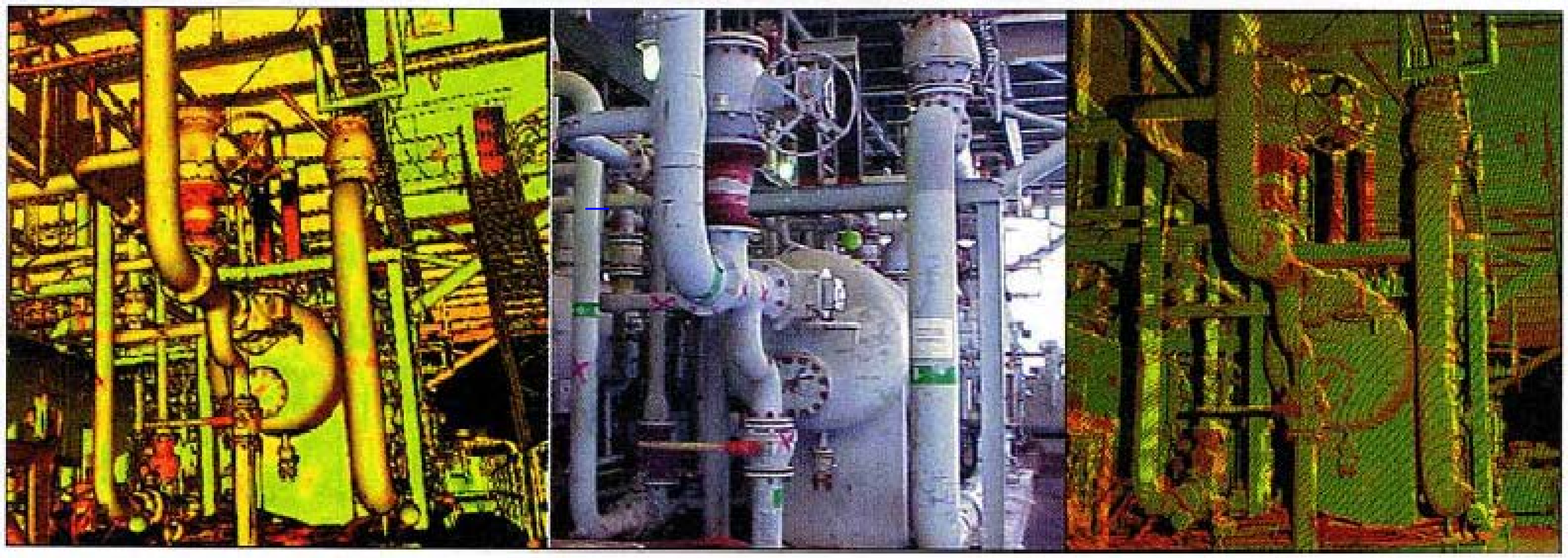
3D surface model

Industrial inspection

- Verify specifications
- Compare measured model with CAD



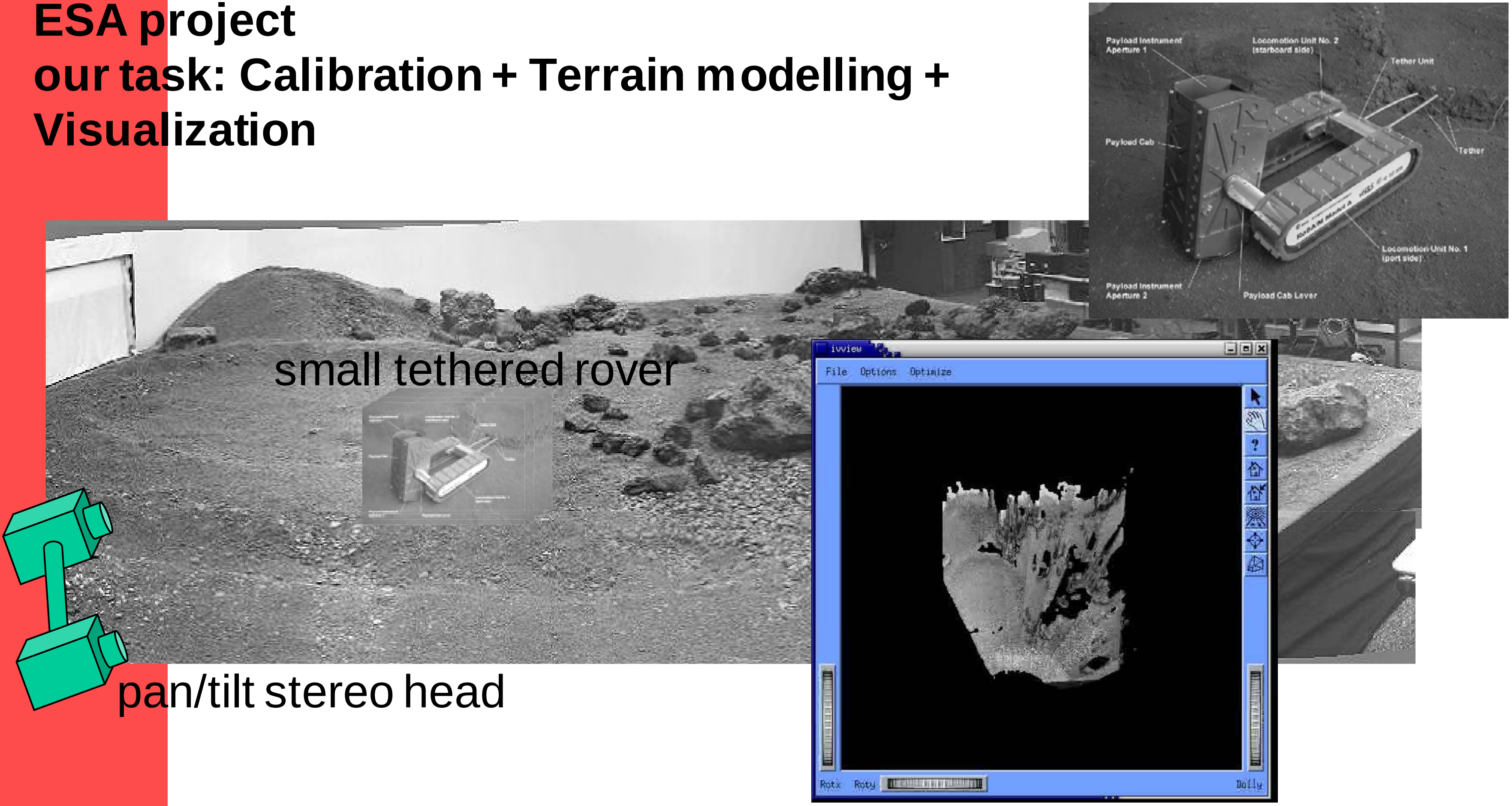
Scanning industrial sites



as-build 3D model of off-shore oil platform

Robot navigation

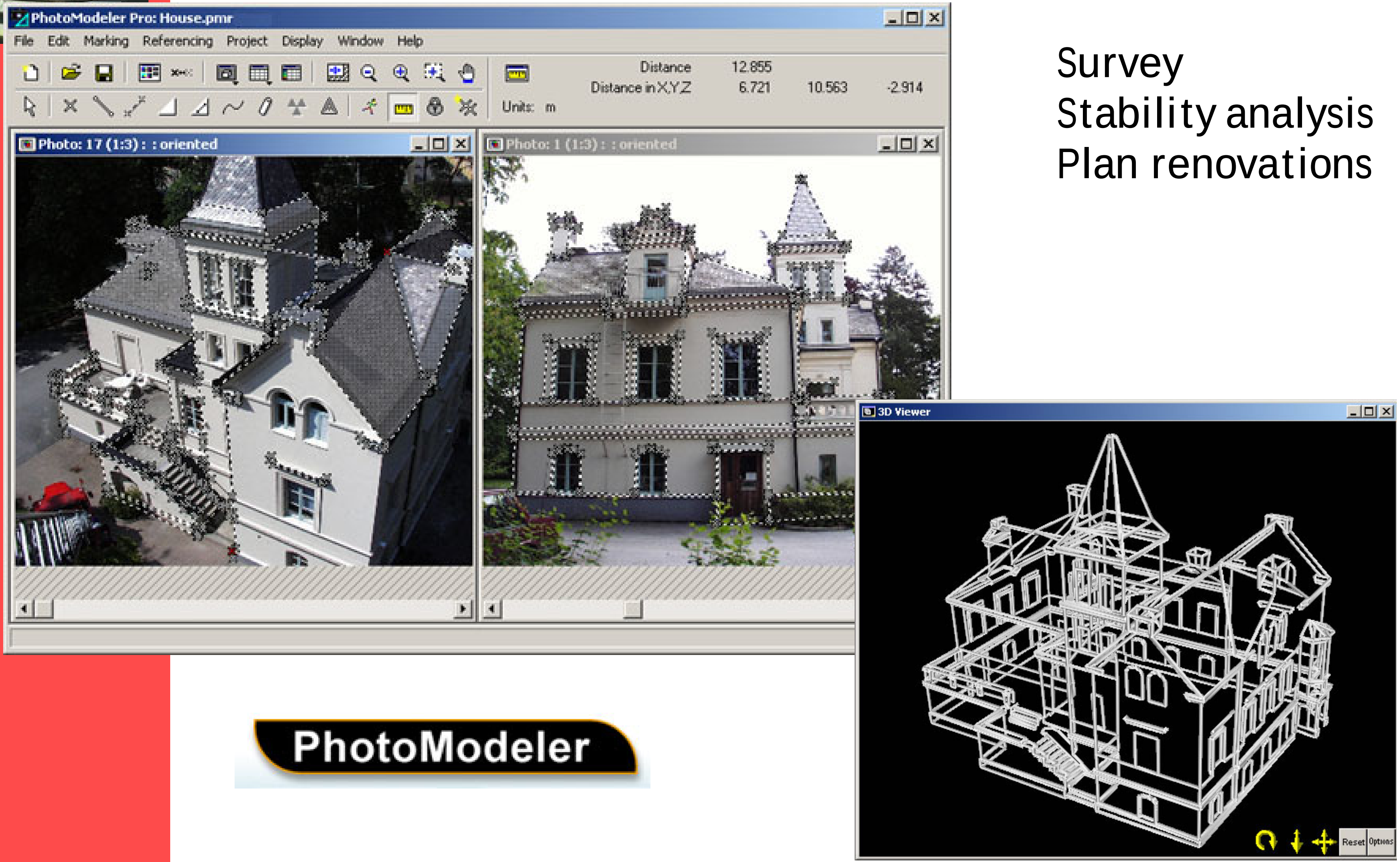
ESA project
our task: Calibration + Terrain modelling + Visualization



small tethered rover

pan/tilt stereo head

Architecture

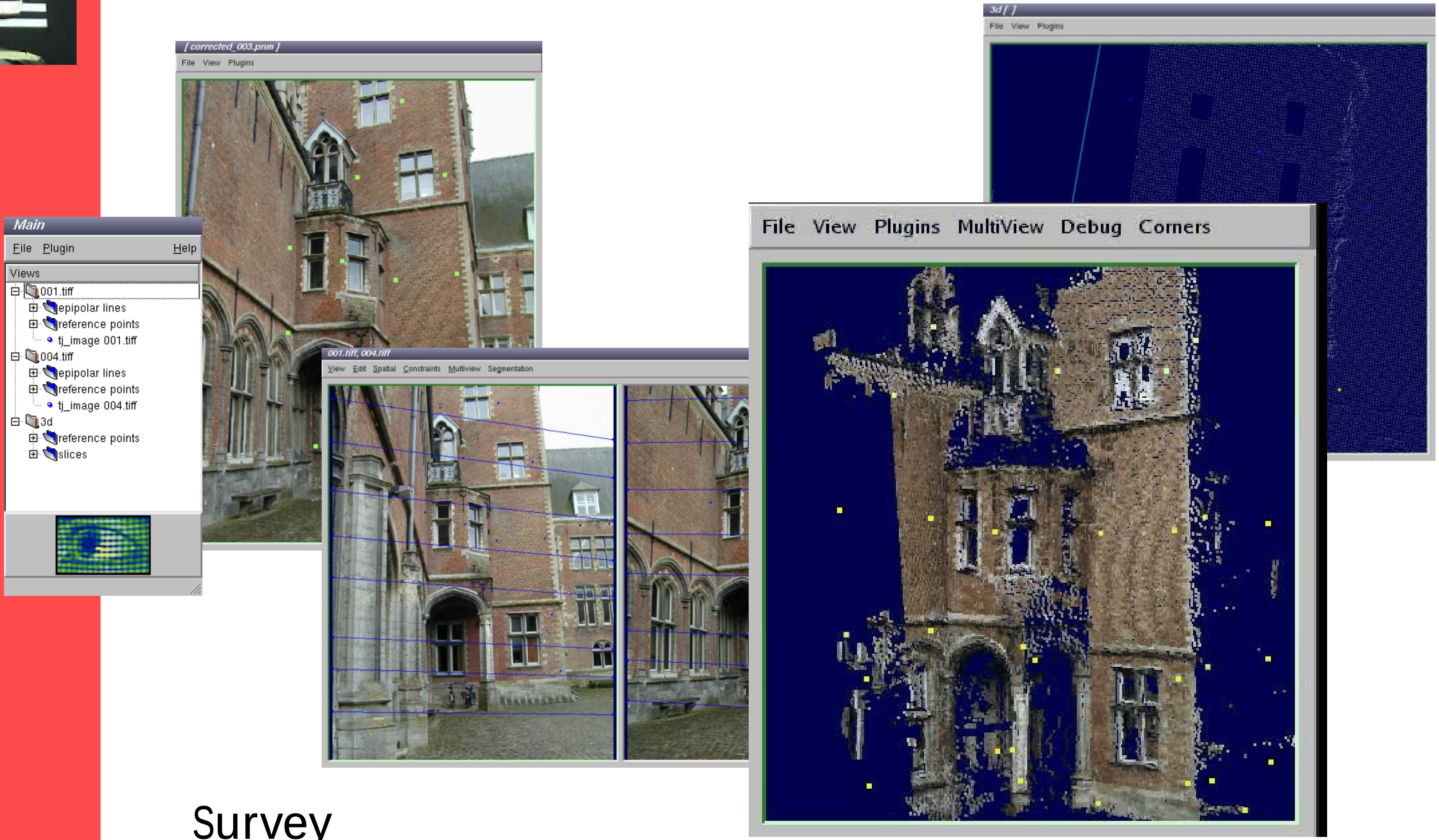


Distance 12.855
Distance in X,Y,Z 6.721 10.563 -2.914
Units: m

Survey
Stability analysis
Plan renovations

PhotoModeler

Architecture



Survey
Stability analysis
Plan renovations

Cultural heritage

Virtual Monticello

Allow virtual visits

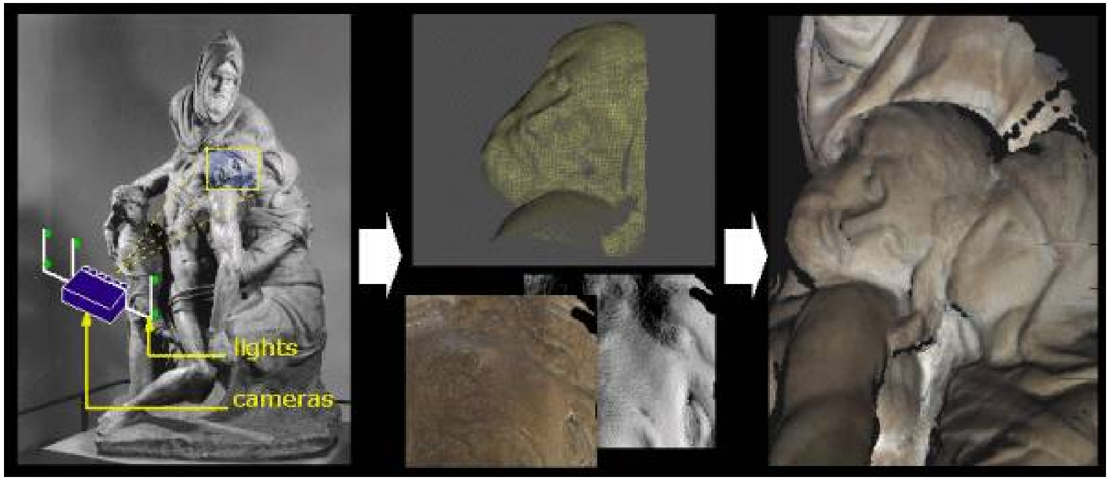
Cultural heritage

Stanford’s Digital Michelangelo

Digital archive
Art historic studies

IBM’s pieta project

Photometric stereo + structured light



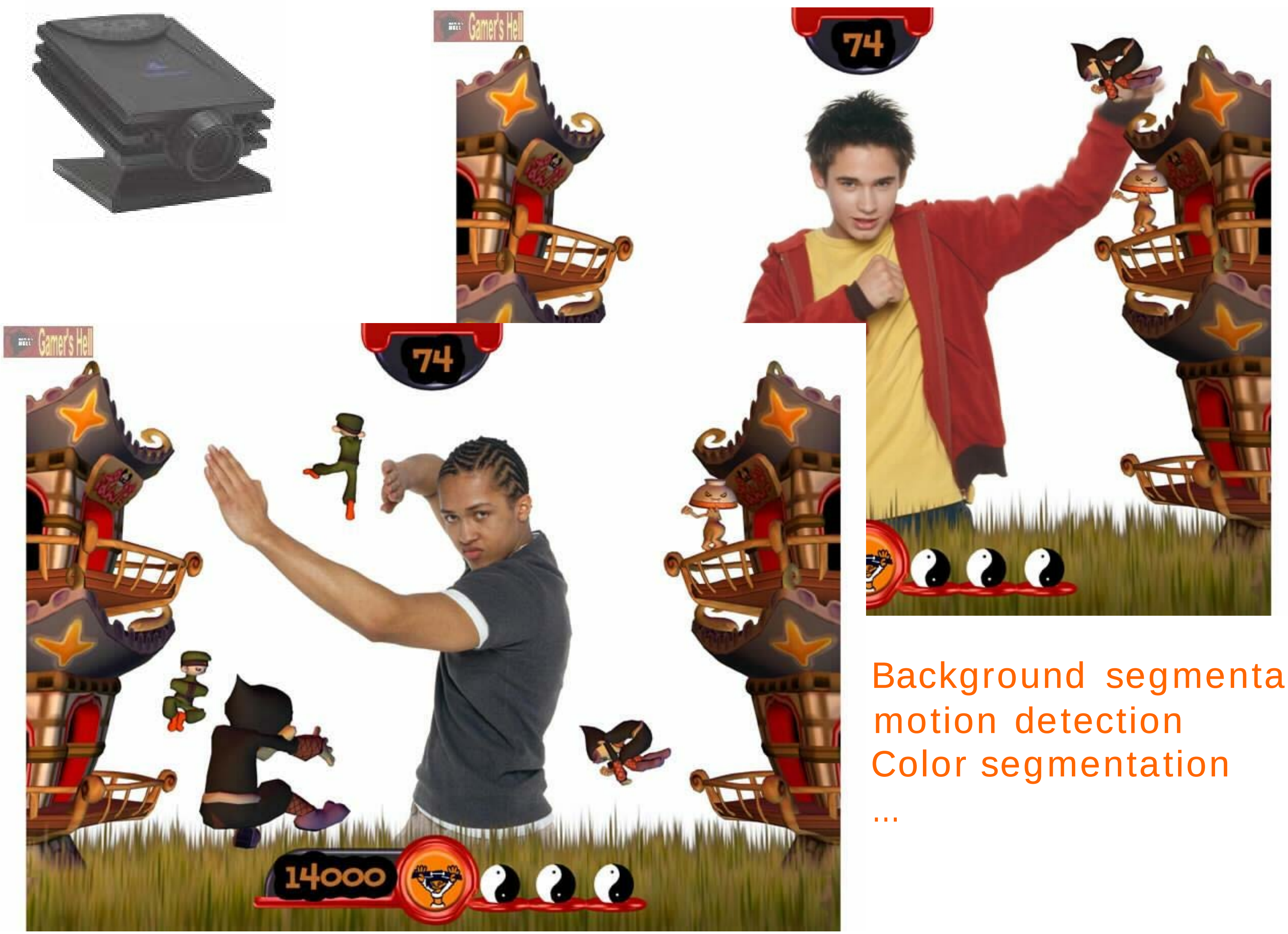
more info:

http://researchweb.watson.ibm.com/pieta/pieta_details.htm

Archaeology

accuracy ~1/500 from DV video
(i.e. 140kb jpegs 576x720)

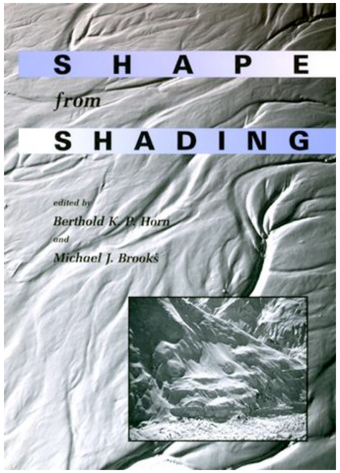
Sony's Eye Toy:
Computer Vision for the masses



Background segmentation/
motion detection
Color segmentation
...

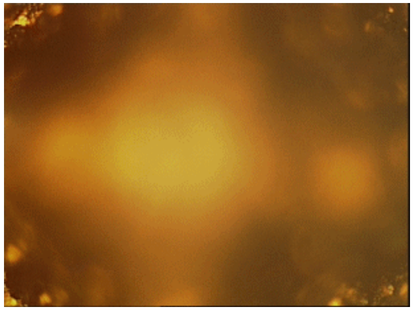
Shape from …

many different approaches/cues



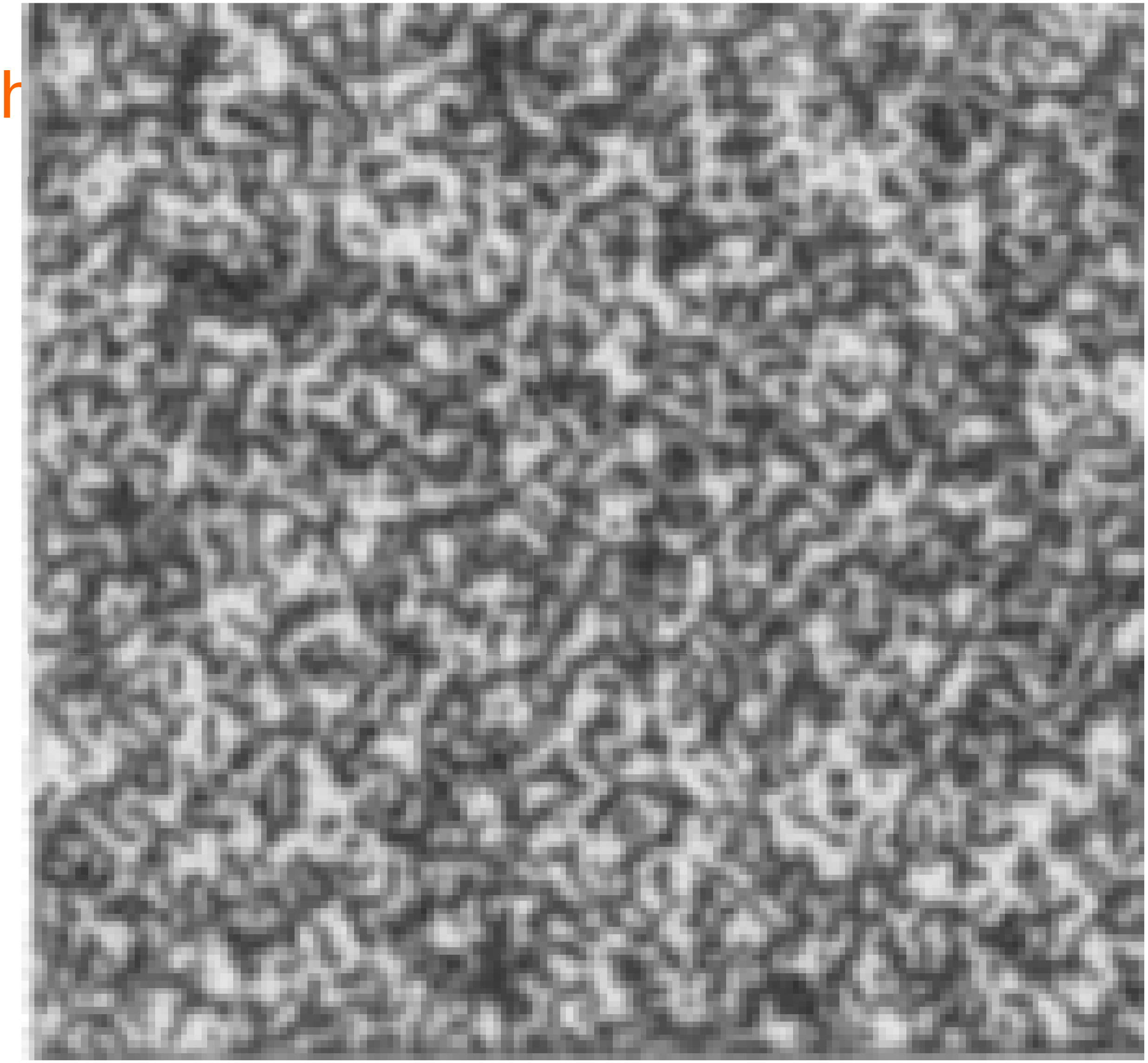
SHAPE
from
SHADING

edited by
Berthold K. P. Horn
and
Michael J. Brooks



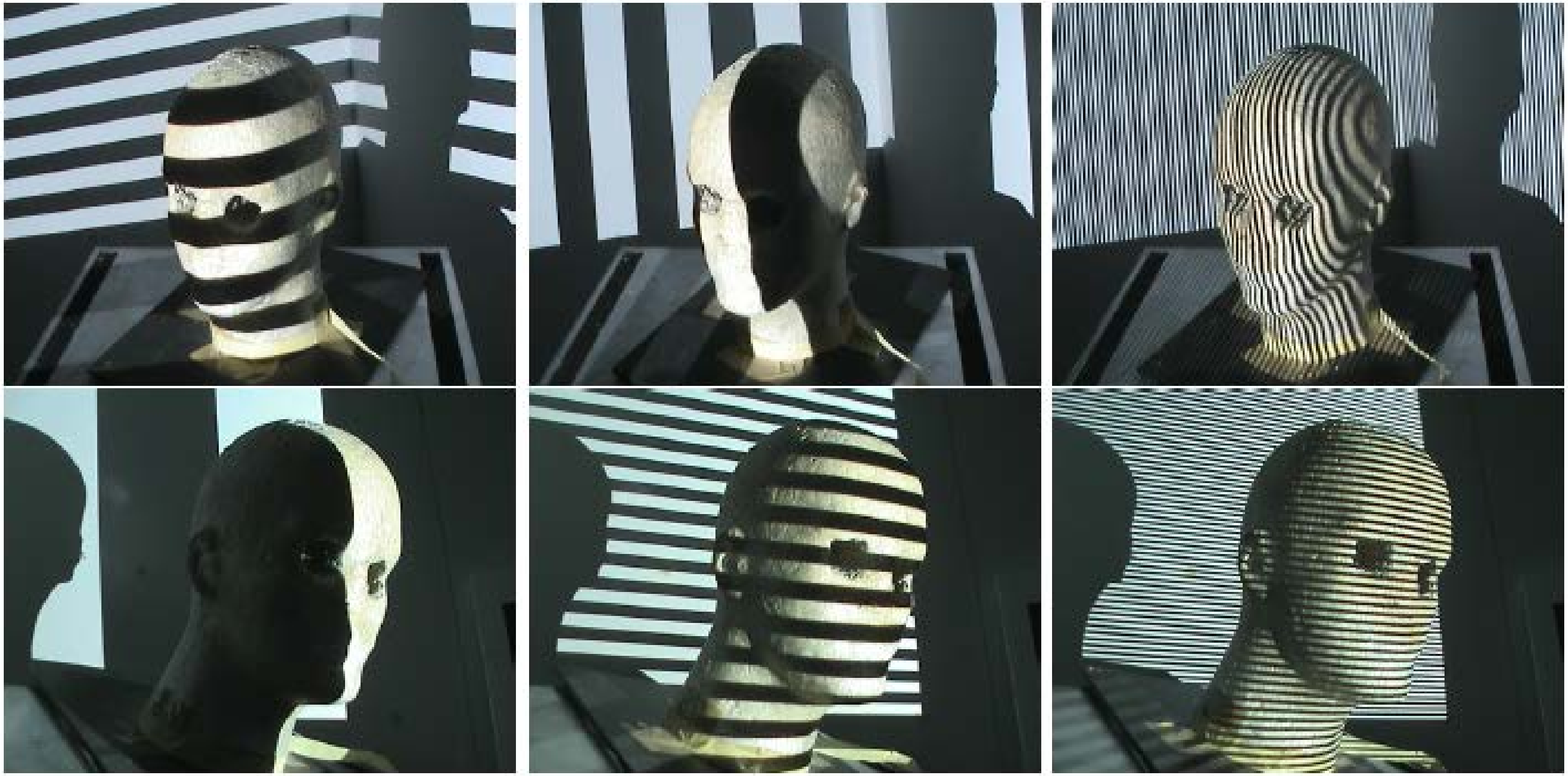
Optical flow

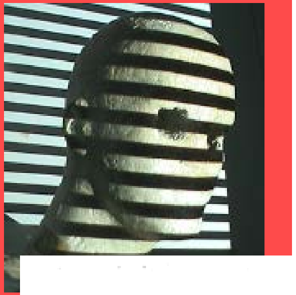
Wh



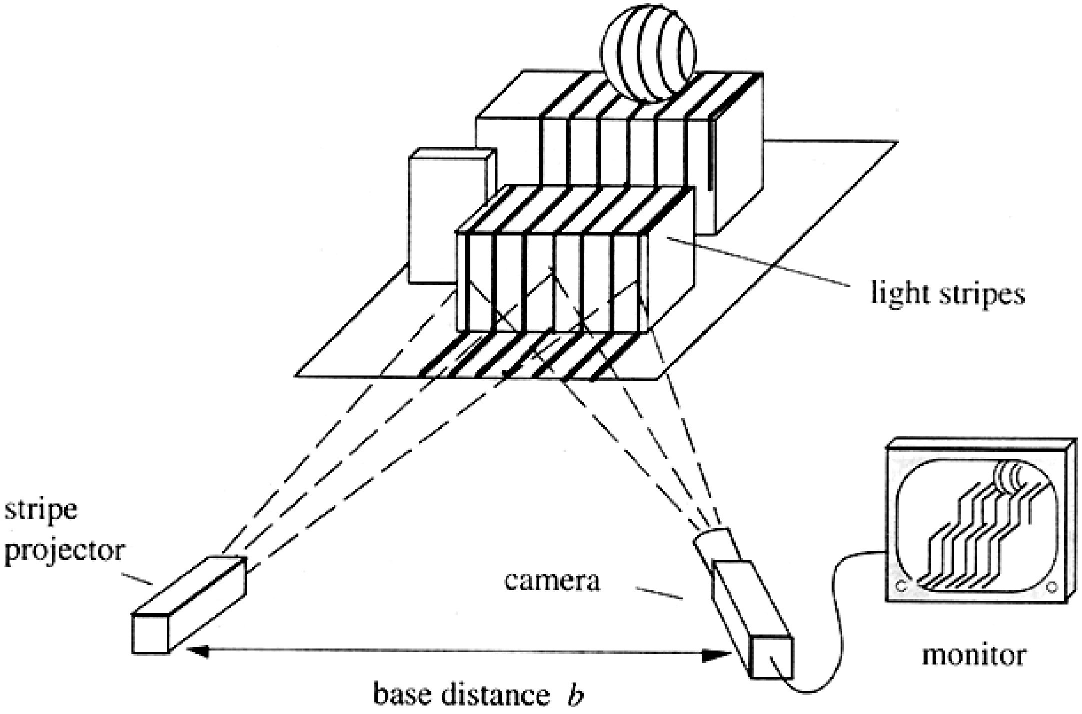
Optical flow

Active Vision: Structured Light

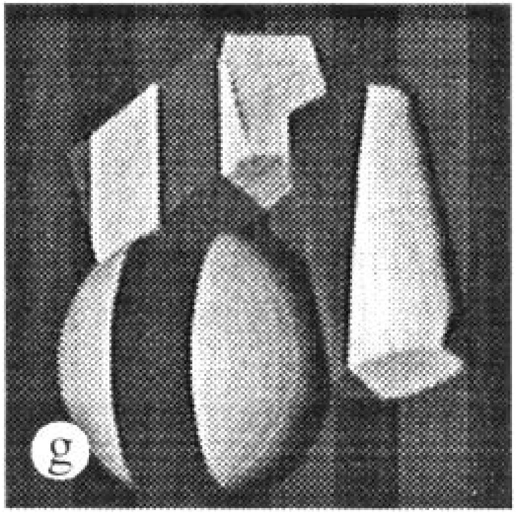




Static Light Pattern Projection



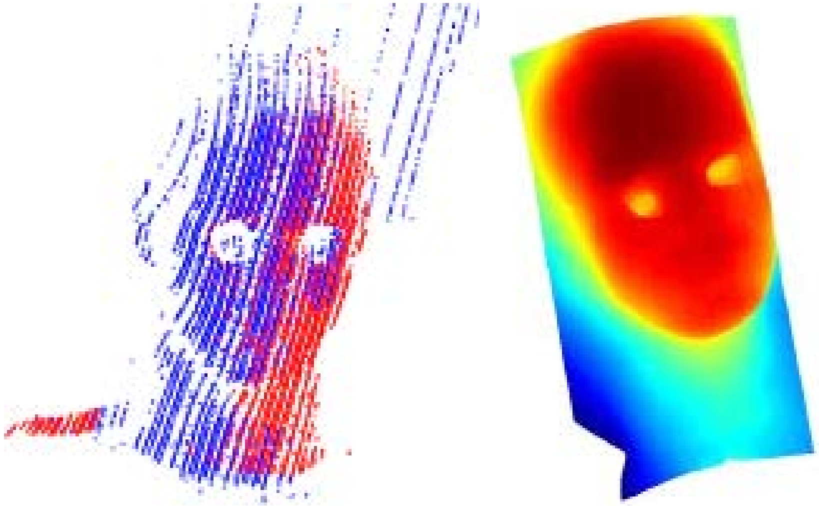
Geometry between projector, sensor and objects



Segmentation of stripes from set of images

Active Vision: Structured Light

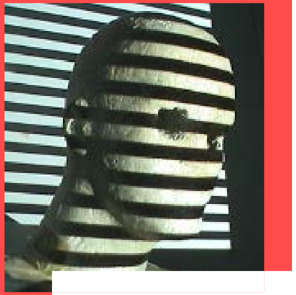
Segmentation: Binarization and coding of stripes



3D model extracted from stripe pattern

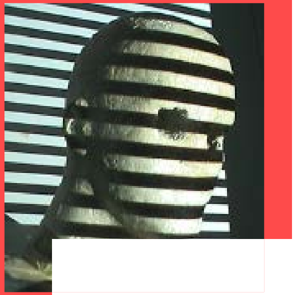
Spatiotemporal Volumes

- Eric Bennet, UNC, PhD thesis



Spatiotemporal Volumes

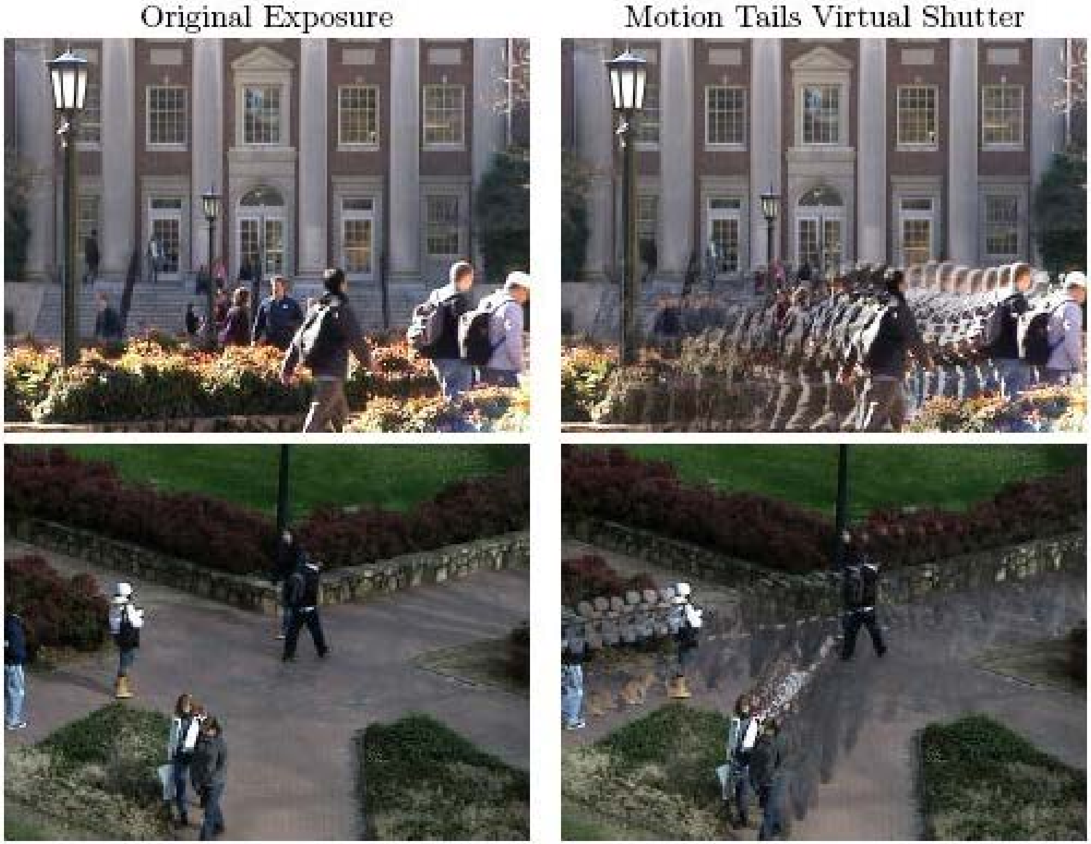
Figure 3.3: Visualization of a spatio-temporal volume and a spatio-temporal cut plane. On the left, a 10 second video is presented as a spatio-temporal volume. The front of the volume shows the first frame, the right side shows the right-most vertical line through time, and the top shows the top-most scanline through time. On the right, the volume has been rotated and been cut using two planar cuts. The first, parallel to the front face, has shortened the video. The second has revealed a different scanline which shows the motion of people walking during the duration of the video.



Video-sequence editing II

Figure 3.12: Sheared volume after the Frisbee has been stabilized. Three visualizations of the volume are shown above. The top-left image shows the sheared volume at a given time. The right image shows a fixed column through time and the bottom image shows a fixed scan line after the Frisbees has been stabilized.

Motion Tails

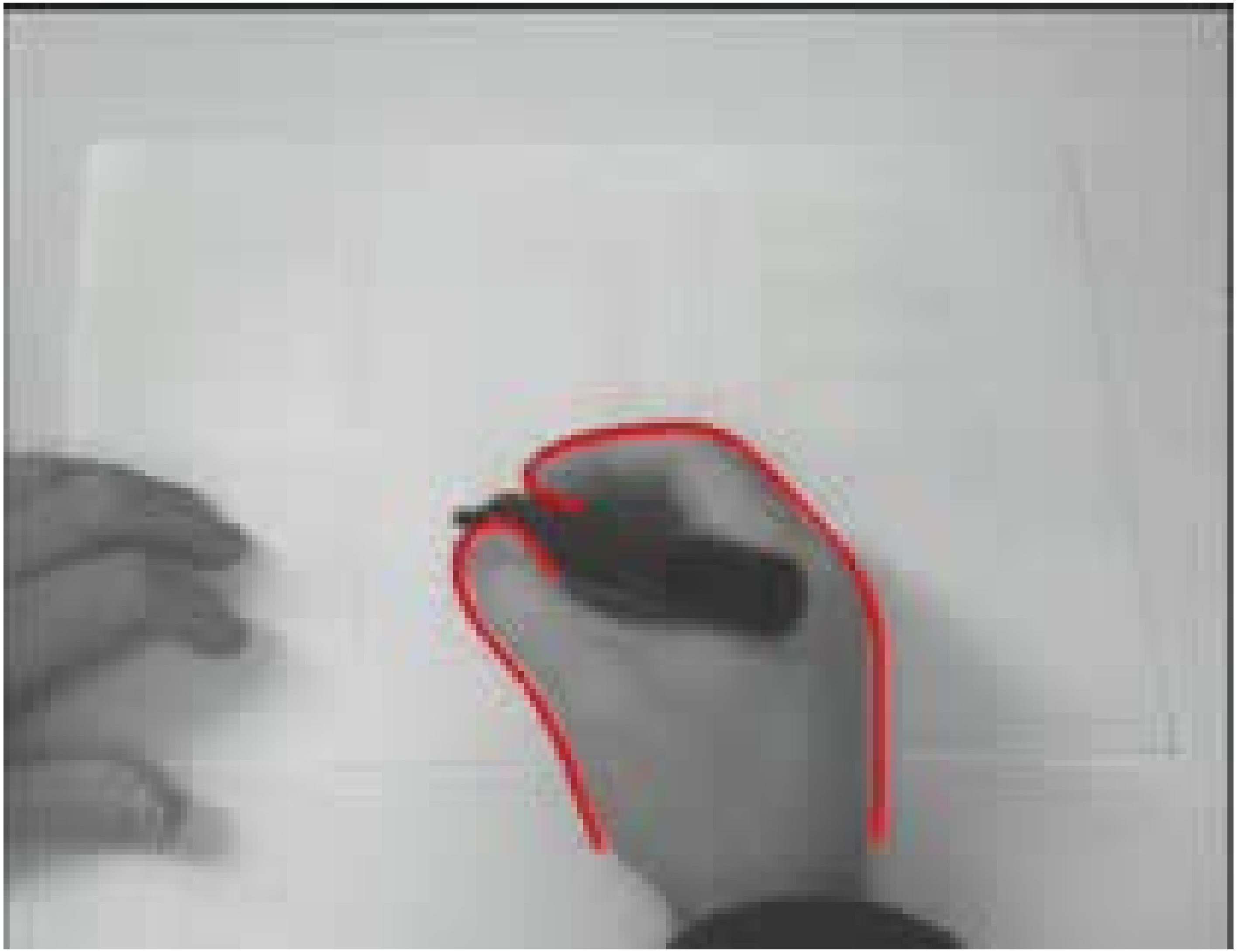


Original Exposure

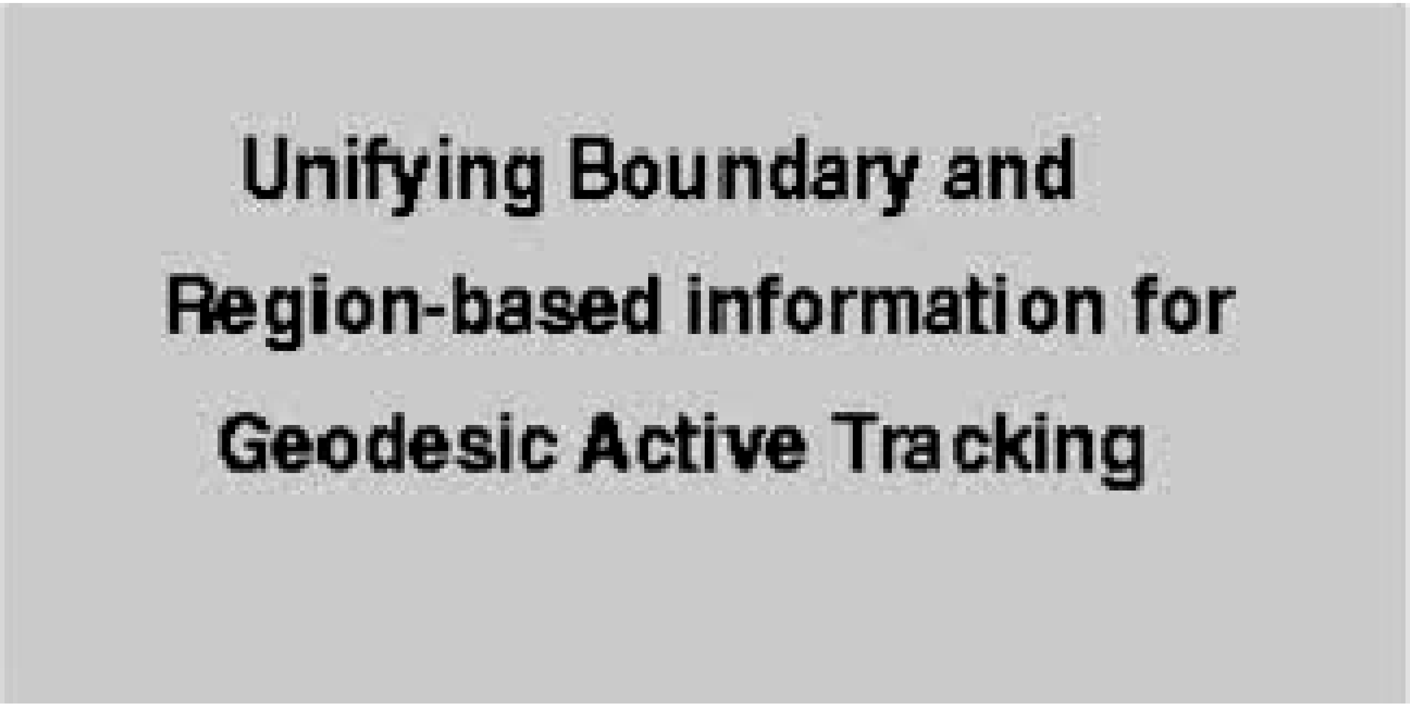
Motion Tails Virtual Shutter

Figure 5.9: Two examples of using motion tails to depict dense motion paths between sampled time-lapse frames. The building front result (above) uses uniform sampling, while the crowded sidewalk (below) is non-uniformly sampled.

Object Tracking: Using Deformable Models in Vision



Object Tracking: Using Deformable Models in Vision: II



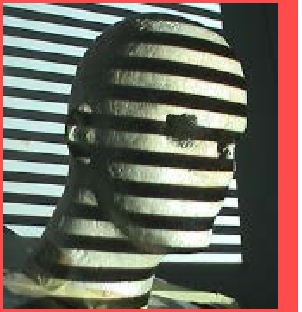
Unifying Boundary and
Region-based information for
Geodesic Active Tracking

Object Tracking III

Face detection

http://www.umiacs.umd.edu/~aswch/test.html

http://vasc.ri.cmu.edu/cgi-bin/demos/findface.cgi

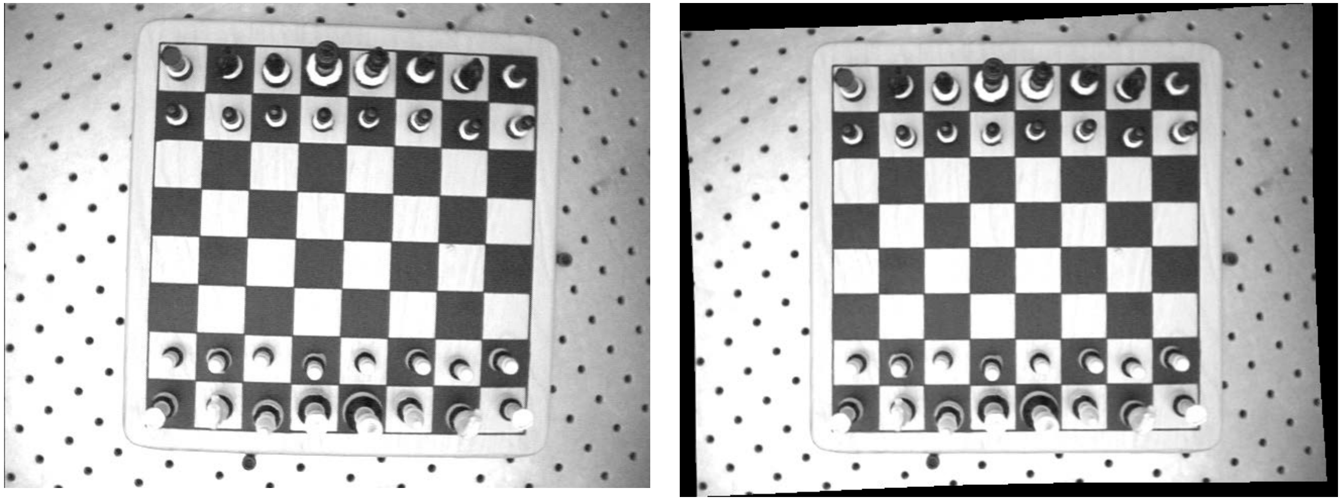


Examples of Student Projects

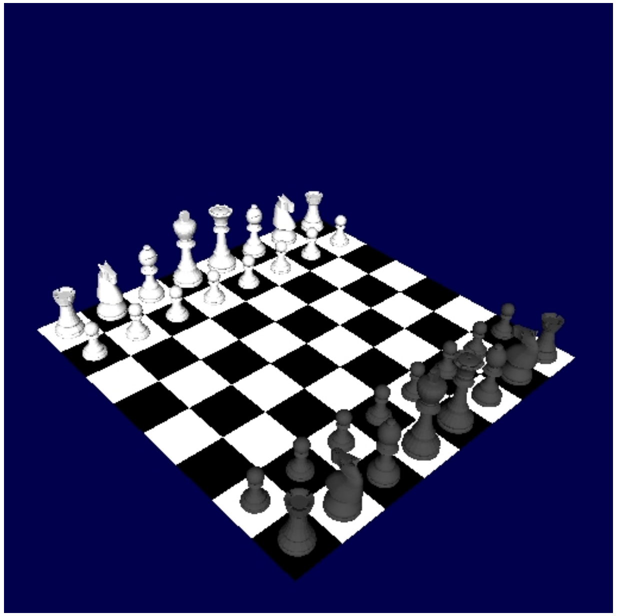
Student Project: Playing Chess, Recognition and Simulation

- Track individual chess pieces
- Maintain state of board
- Graphically represent state changes and state
- D. Allen, D. McLaurin UNC
- Major ideas:
  - 3D from stereo
  - detect and describe changes
  - Use world knowledge (chess)

Calibration, Rendering & Replay

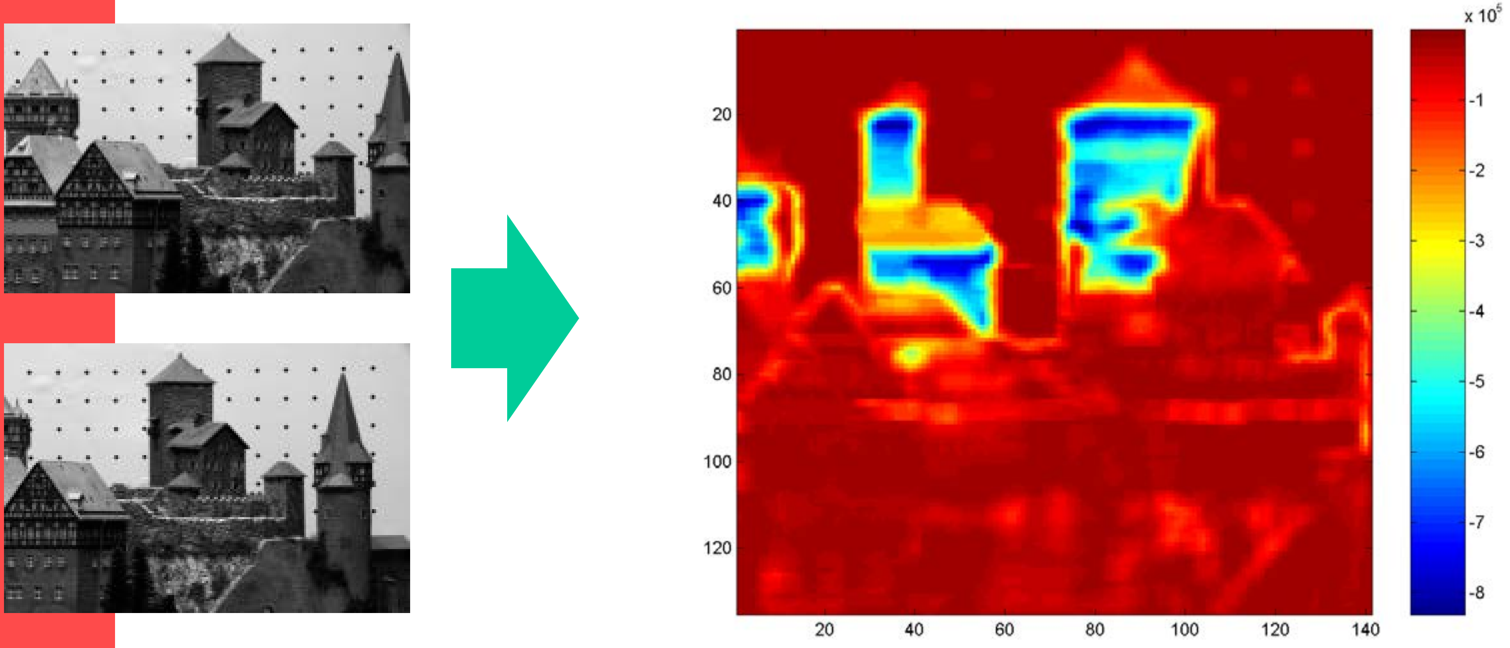


Movie



Enhanced Correlation for Stereo Vision

- Andrew Nashel
- Correlation map produced by precomputation method:

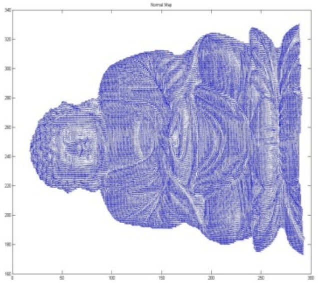


Results – Lord Buddha Images – Pre-Processed Images

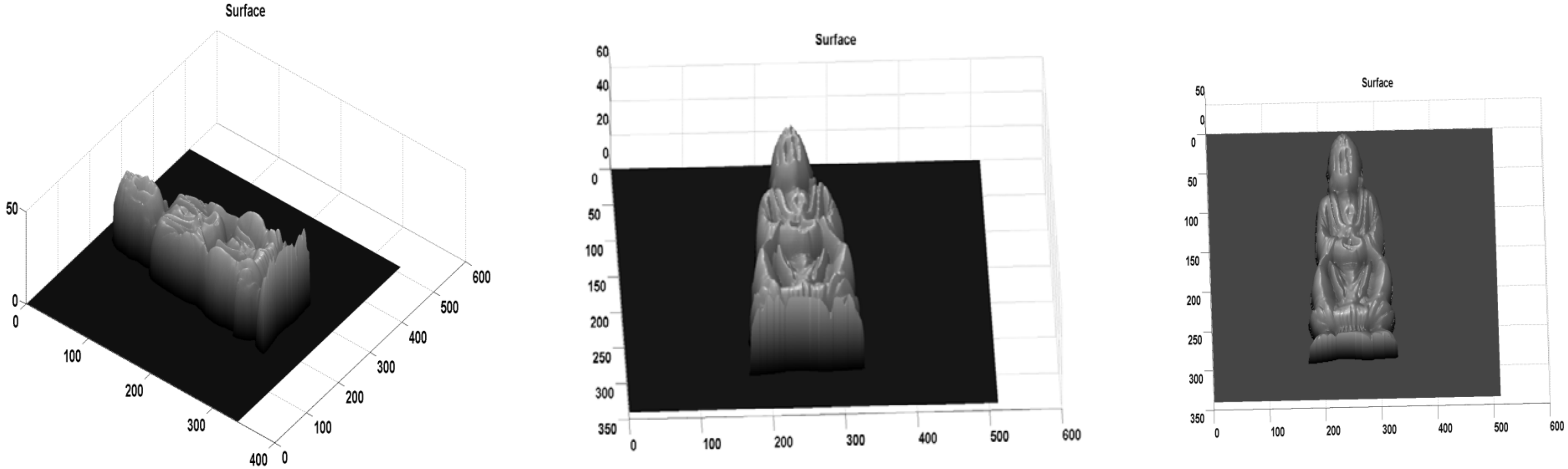
Guozhen Fan and Aman Shah

Original Image

Albedo



Surface Normals



Obtained Surfaces from different angles

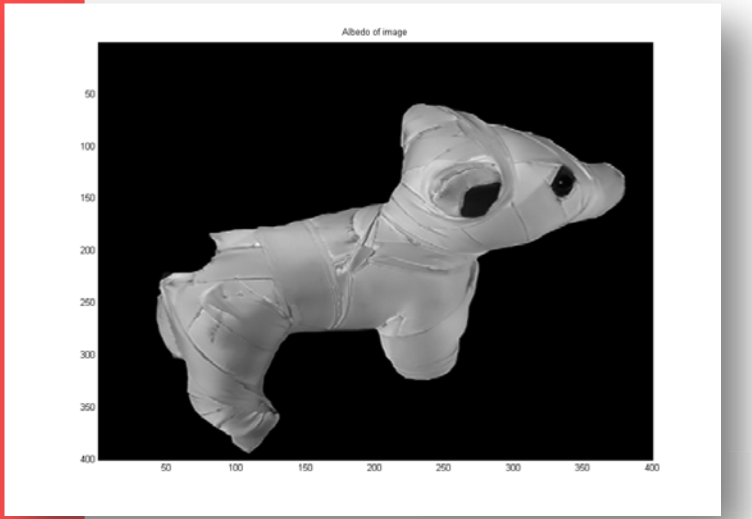
Photometric Stereo

Christopher Bireley

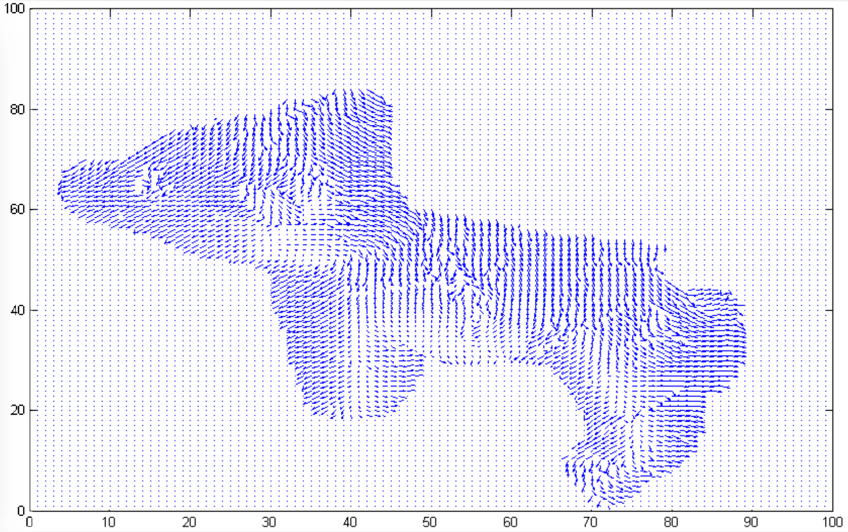
Bandage Dog

Imaging Setup

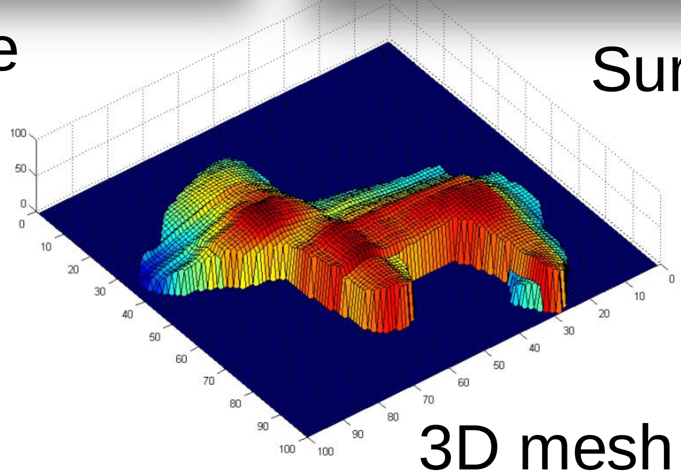
Photometric Stereo
Christopher Bireley



Albedo image

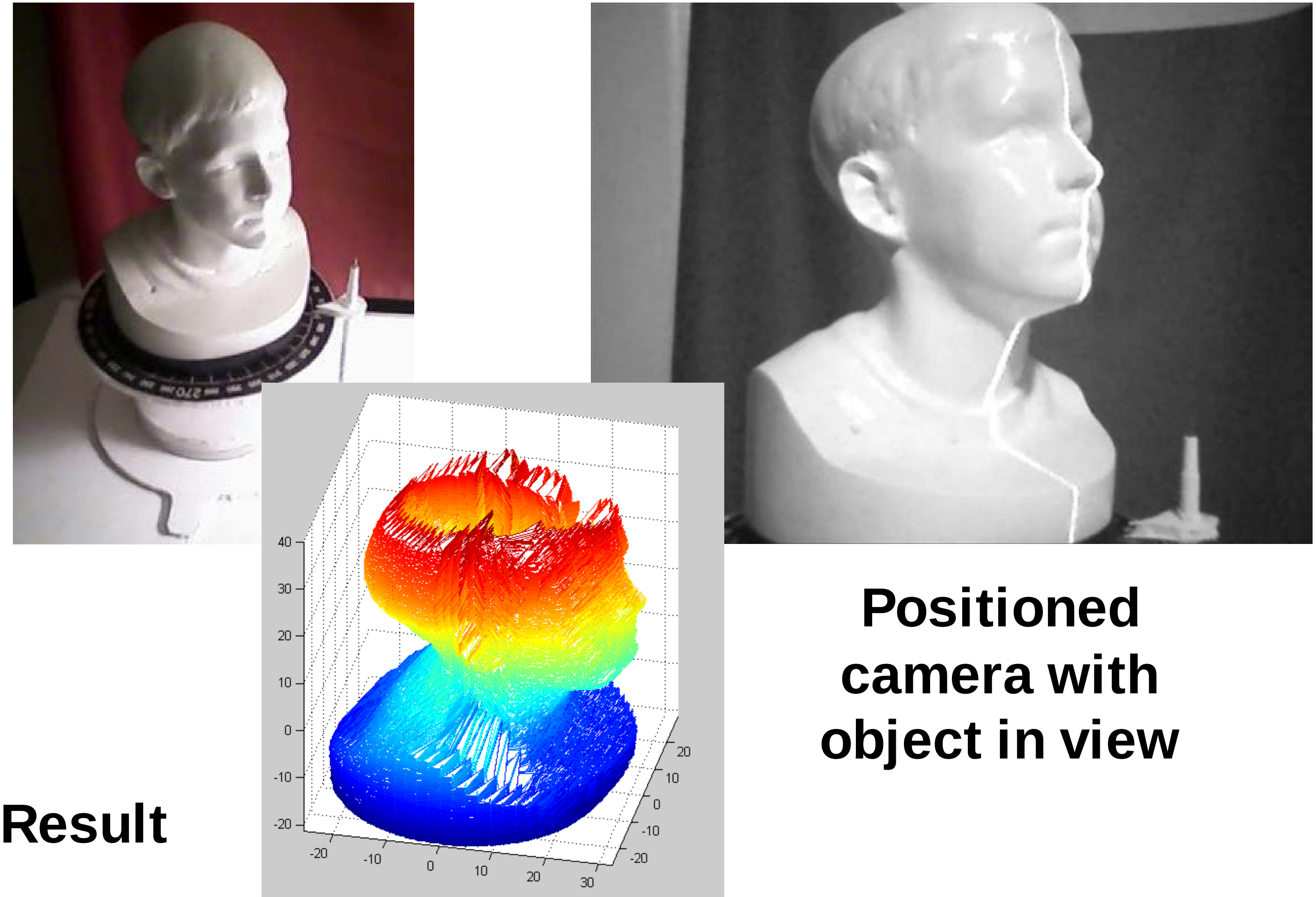


Surface Normals



3D mesh

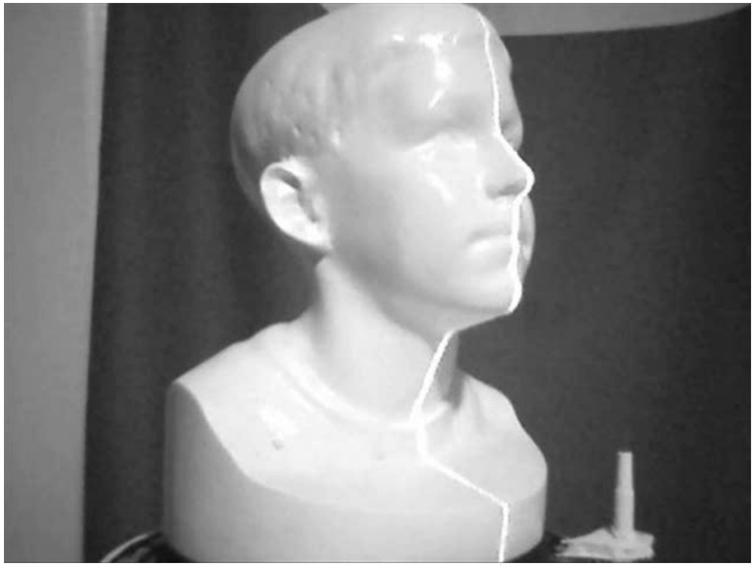
Structured Light
James Clark

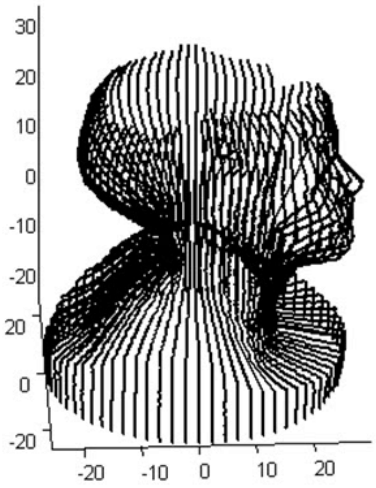


Positioned camera with object in view

Result

Structured Light ctd.
James Clark





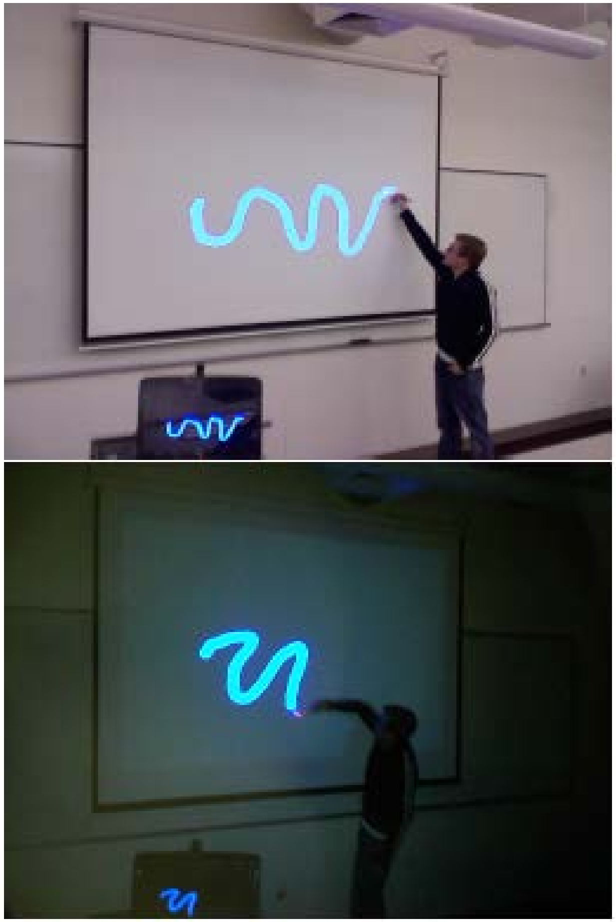
Webcam Based Virtual Whiteboard

Jon Bronson James Fishbaugh

- Blackboards came first
- Whiteboards eventually followed
- Virtual Whiteboards are coming
- Basic Idea:
    - Write on any surface
    - Use no ink/chalk
    - Store all information to disk

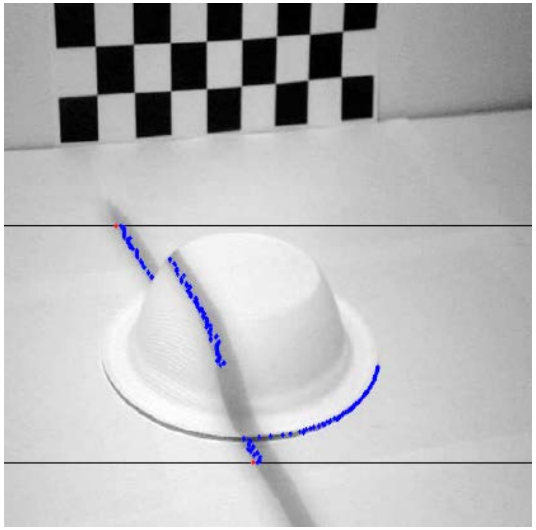
Webcam Based Virtual Whiteboard

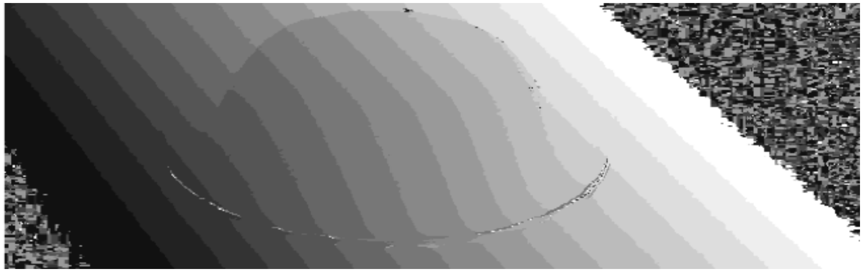
Jon Bronson James Fishbaugh



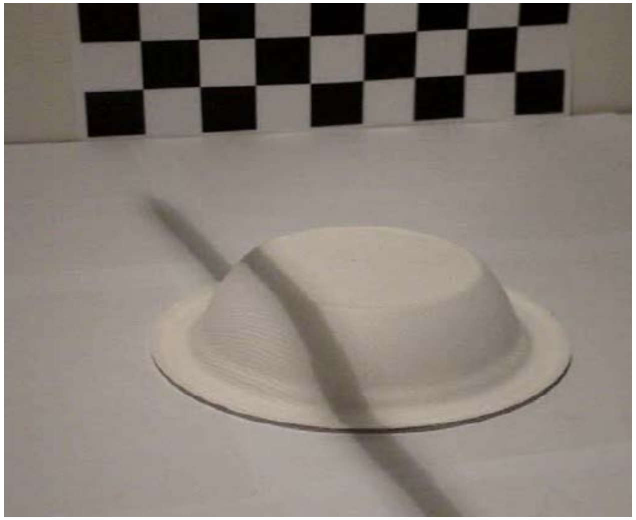
Structured Light

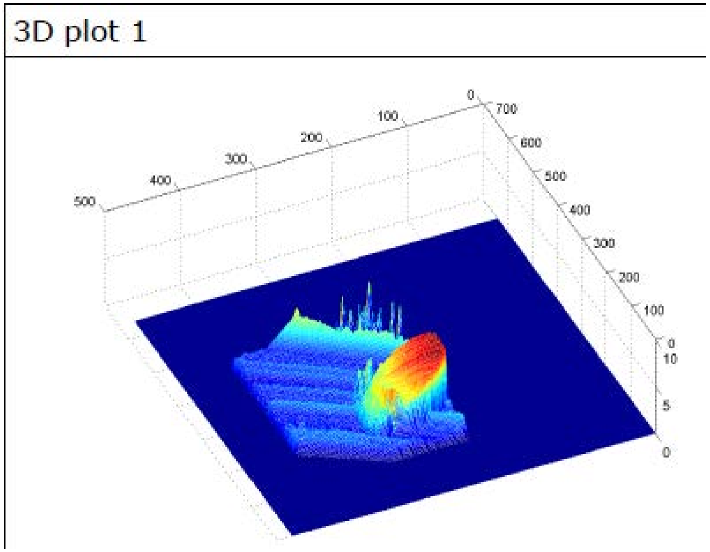
Anuja Sharma, Abishek Kumar



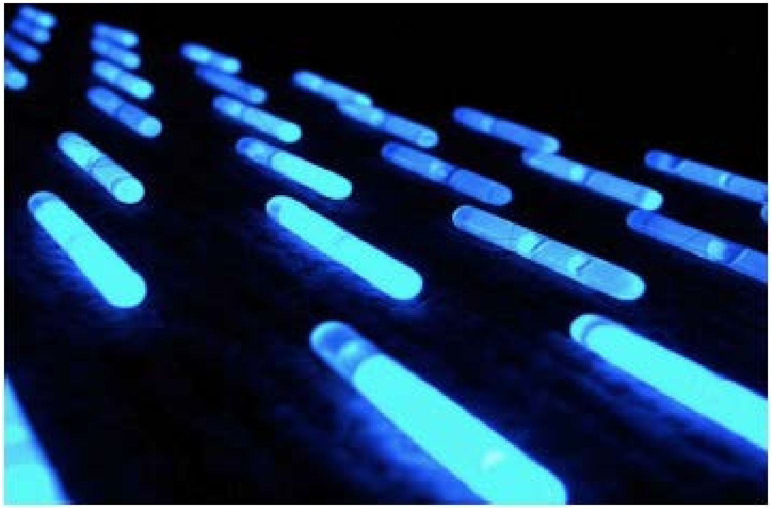


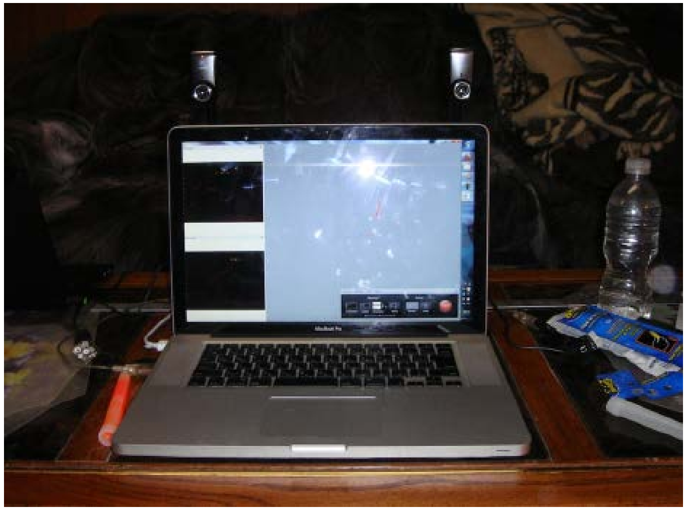
Structured Light
Anuja Sharma, Abishek Kumar



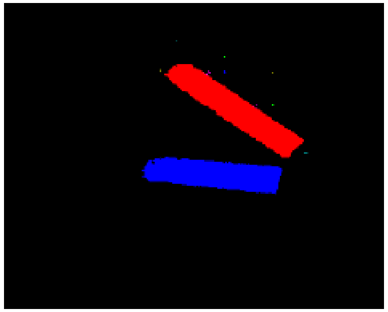


Realtime Glowstick Detection
Andrei Ostanin





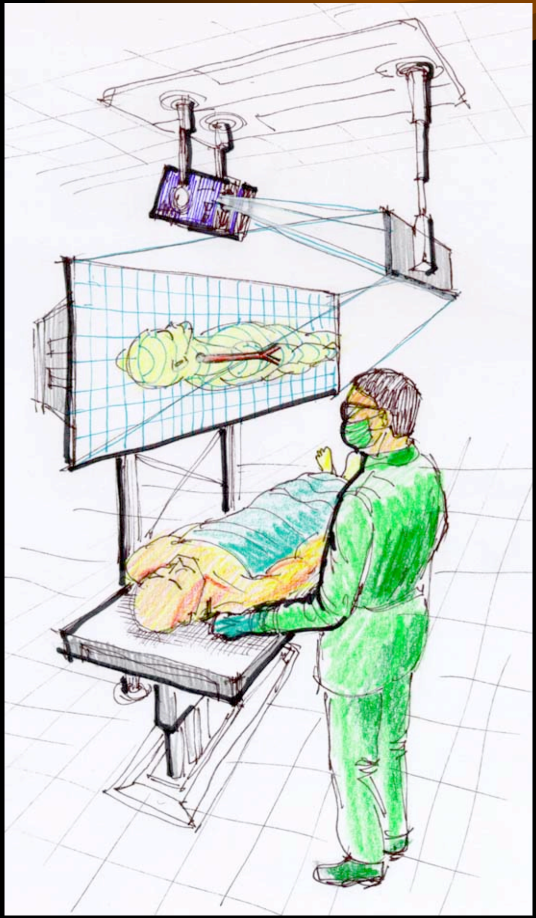
- Capture the 3D position of glowsticks in real-time using two webcams
- Environment dark enough that glowsticks are easily segmented out
- Prefer speed over correctness



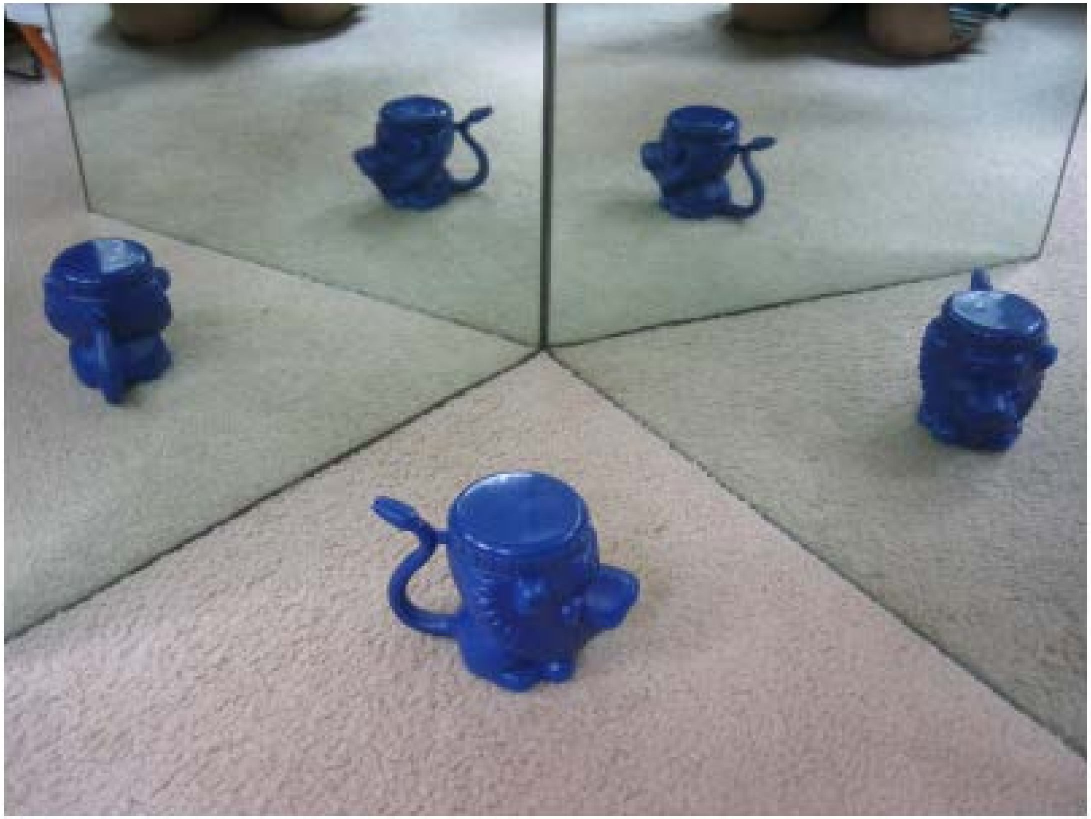
movie

Passive Object Tracking from Stereo Vision

- 3D visualization - blood vessels, organs, teapots
- Stereo displays require head tracking
- Cumbersome trackers are undesirable
- Michael Rosenthal



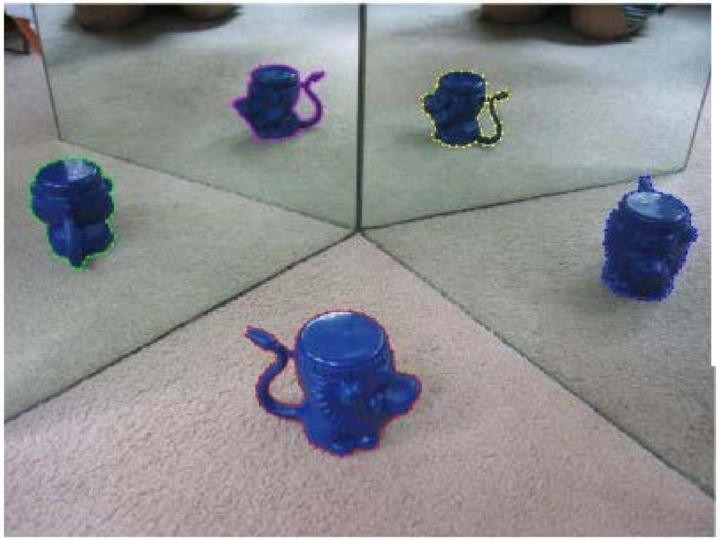
3D shape from silhouettes: Two Mirrors and uncalibrated camera

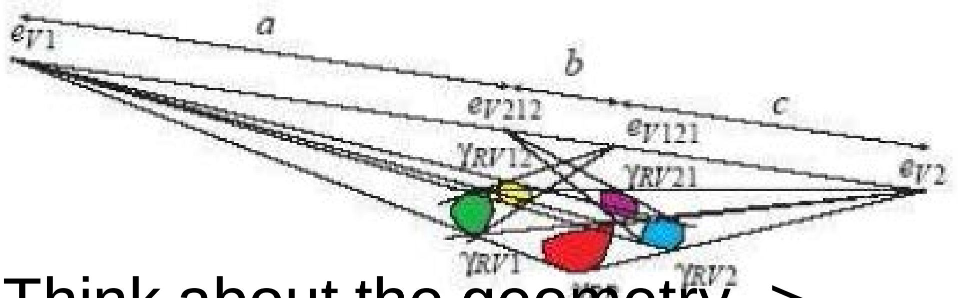


Forbes et al., ICCV2005

Christine Xu, Computer Vision Student Project

3D shape from silhouettes

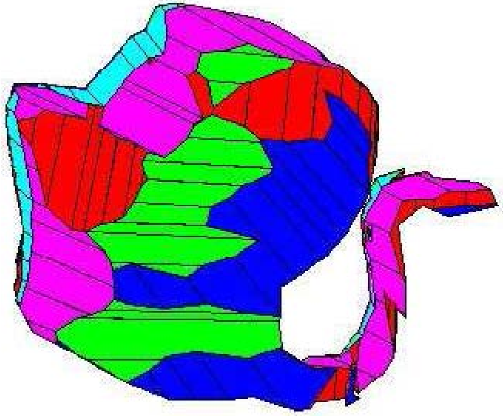




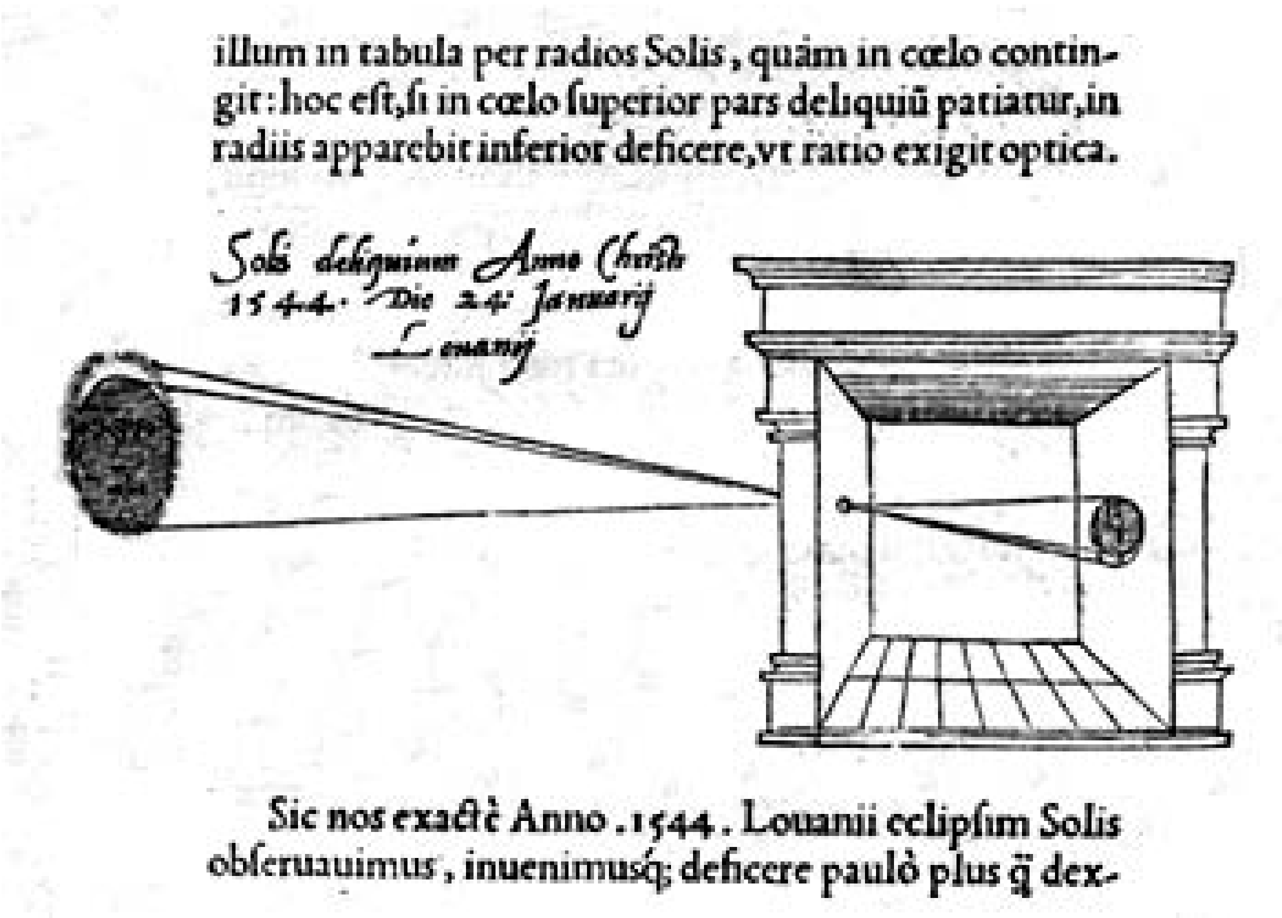
Think about the geometry -> calculate relationship between silhouettes

Segmentation of contours

Result: 3D Object



Next class: Cameras
Chapter 2: Image Formation



illum in tabula per radios Solis, quàm in cœlo contingit: hoc eſt, ſi in cœlo ſuperior pars deliquiũ patiatur, in radiis apparebit inferior deficere, vt ratio exigit optica.

Solis deliquium Anno Christi
1544. Die 24. Januarij
Louanij

Sic nos exactè Anno .1544 . Louanii eclipſim Solis obſeruauimus, inuenimusq; deficere paulò plus q̃ dex-

Next class: Image Formation

Chapter 2: Textbook

- Please find pdf copies of Chapters 1&2 on the website.

Assignment:

- Read Chapter 1 for additional materials
- Read Chapter 2 for preparation of 2nd lecture