Range Data Processing

Guido Gerig

CS 6320, Spring 2015

Book F&P: New book Ch14, pp 422 ff

Old book: Ch21, pp 467ff

(credit: some slides from F&P book Computer Vision)
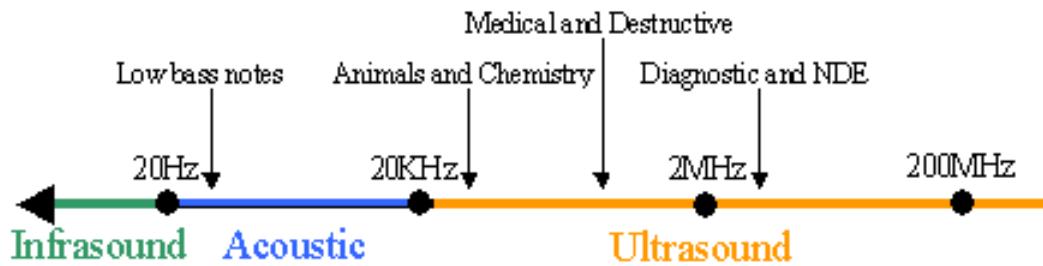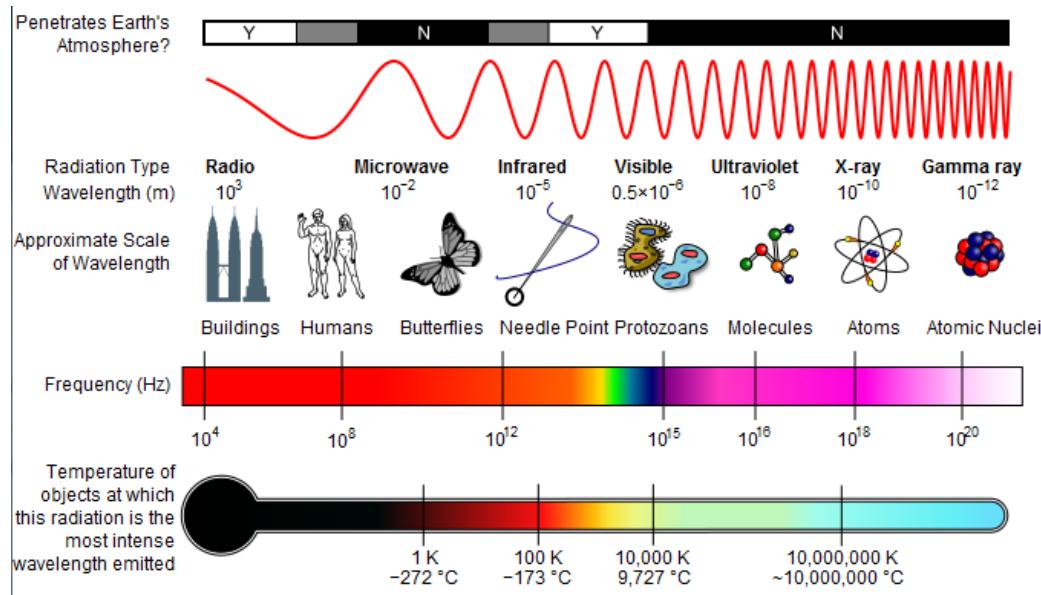
Contents

- Range cameras and physical principle
- Processing of range image data
  - Patches of homogeneous properties
  - Extended Gaussian Image (EGI)
  - Discontinuities: Local curvatures
  - Registration of Range Data: ICP
- Applications

Material: Szelisky coursebook Computer Vision 12.2/12/3
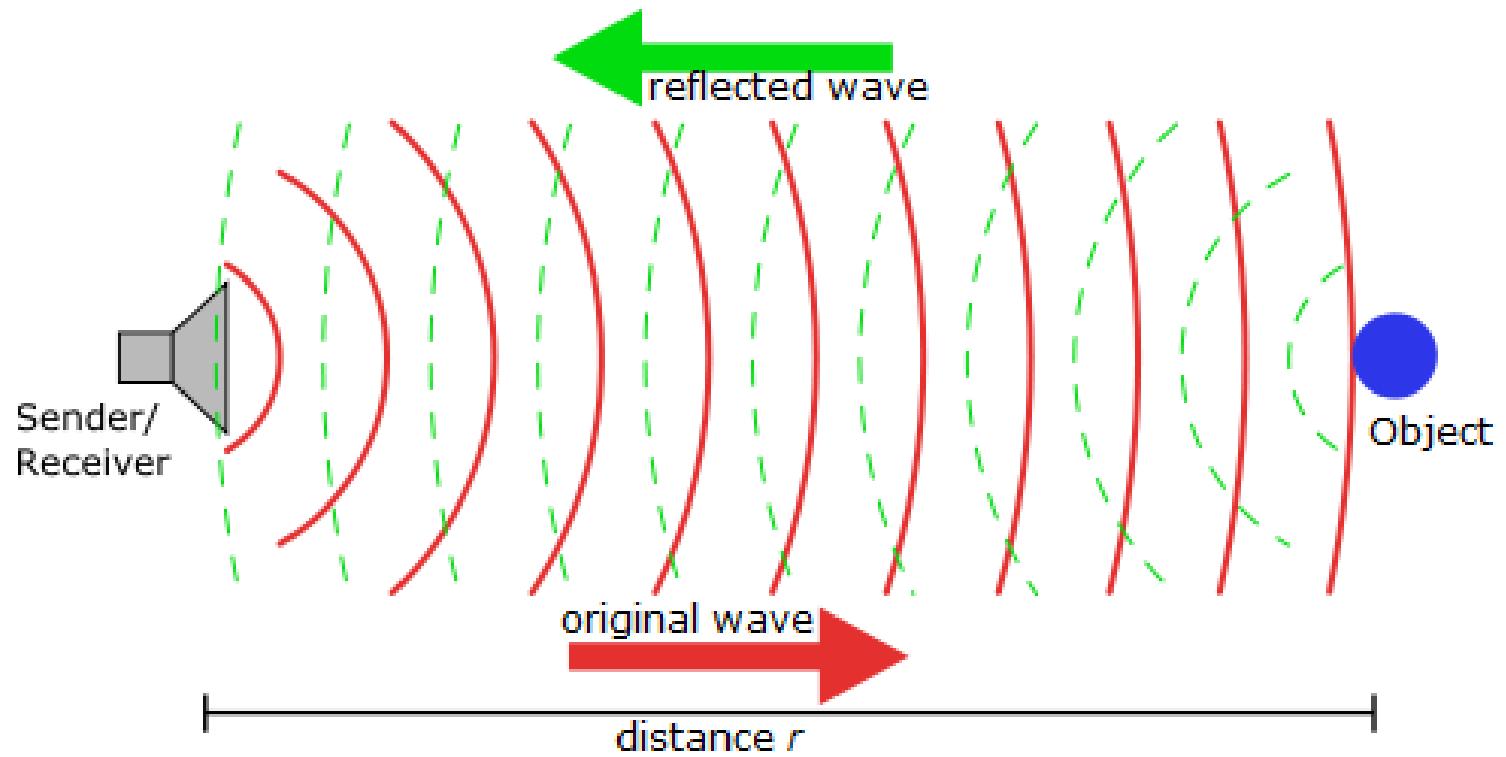
Physical Principles
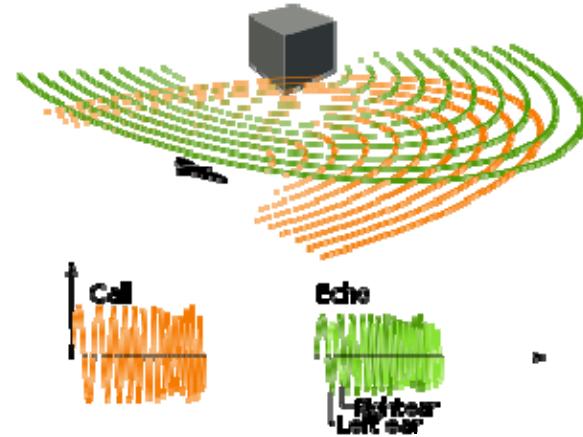
SONAR (Sound navigation and ranging)

Principle: Wave with known velocity v traveling distance $2*r$ → takes time t_f

Bats
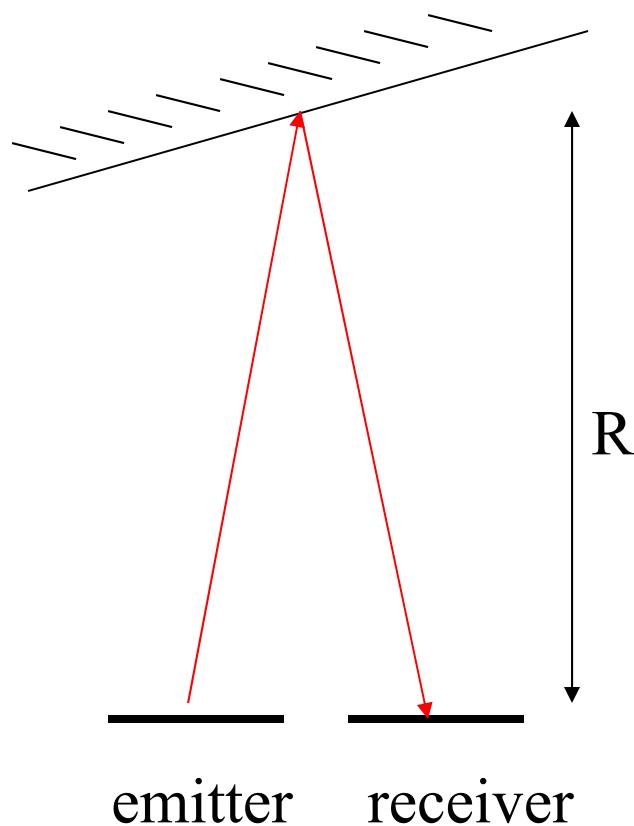
source: wikipedia

Bats use a variety of ultrasonic ranging (echolocation) techniques to detect their prey. They can detect frequencies as high as 100 kHz, although there is some disagreement on the upper limit.[22]
(see also dolphins, shrews, whales).

Time of Flight (TOF)

Basic principle: Time of Flight (TOF):   Emit signal, wait for echo, measure time difference
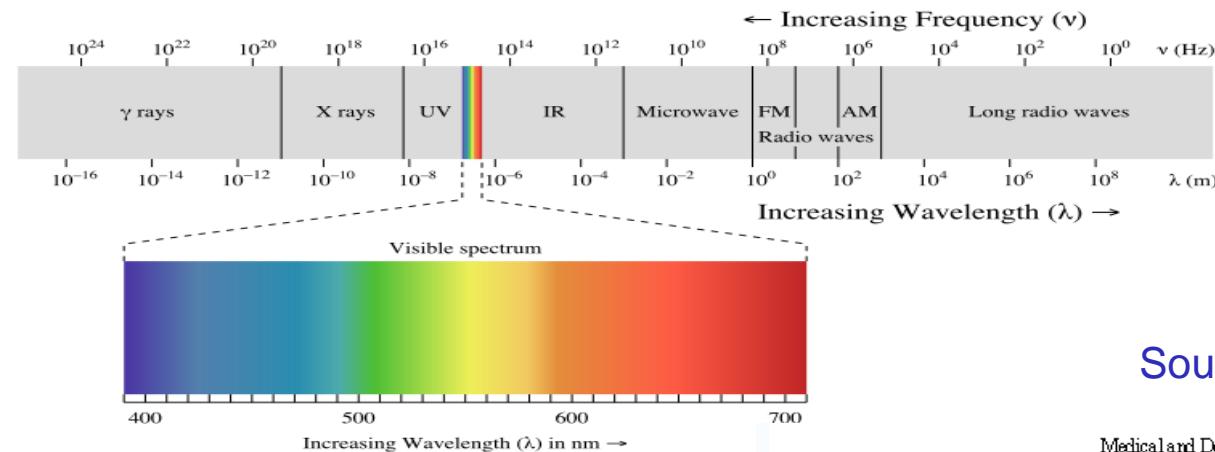
Range measurement:
- Velocity v is known
- t_f to be measured

R

emitter   receiver

$$2R = v * t_f$$
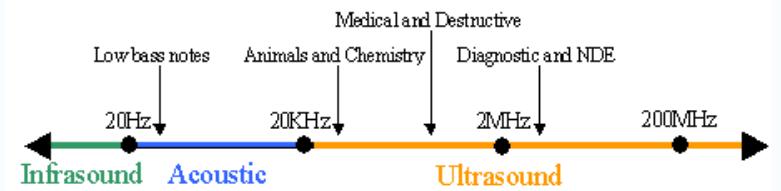
$$R = \frac{v * t_f}{2}$$

Wavelengths for TOF

Source: Wikipedia

- Radar (microwaves: $c=c_{light}$, λ = 0.02m, f = 15GHz)
- Light/Laser (light: c_{light} = $3*10^{-8}$ m/sec, λ = 400nm to 700nm, f= 7 to $4*10^6$ GHz)
- Sound (sound: c = 331 m/sec, λ = 0.02m, f = 20Hz to 20kHz)
- Ultrasound (sound: c = 331 m/sec, λ = 0.017mm, f = 2MHz)

TOF ctd.

Resolution:
Challenge for electronics:

$$t_f = \frac{2*R}{v}$$

$$\Delta t_f = \frac{2*\Delta R}{v}$$

Example:

- Sound:
  v=330m/sec

  R=1cm
  t=60µs

- Light:
  c=3*10⁸m/sec

  R=1cm
  t=67ps
  (picoseconds)

Ultrasound

- Example: **Polaroid**
- Material or topology may absorb arbitrary frequencies: Transmits several frequences (Polaroid: 60,57,53,50kHz)
- Engineering principle: Use pulsed frequency (f) and digital counter (n)
- Range of counter: 2^k-1 (e.g. 16bit)
- Range of unique depth measurement: R*
- Example: f=50kHz, v=330m/sec, k=16: R*=216m, 1count: 6.6mm)
- Problem: wide bundle (30°)

$$t_f = \frac{n}{f}$$

$$R = \frac{t_f * v}{2} = \frac{n * v}{2f}$$

$$R* = \frac{(2^k - 1) * v}{2f}$$
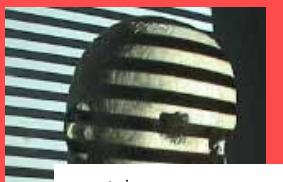
Pulsed Time of Flight

- Advantages:
  - Large working volume (up to 100 m.)
- Disadvantages:
  - Not-so-great accuracy (at best ~5 mm.)
    - Requires getting timing to ~30 picoseconds
    - Does not scale with working volume
- Often used for scanning buildings, rooms, archeological sites, etc.
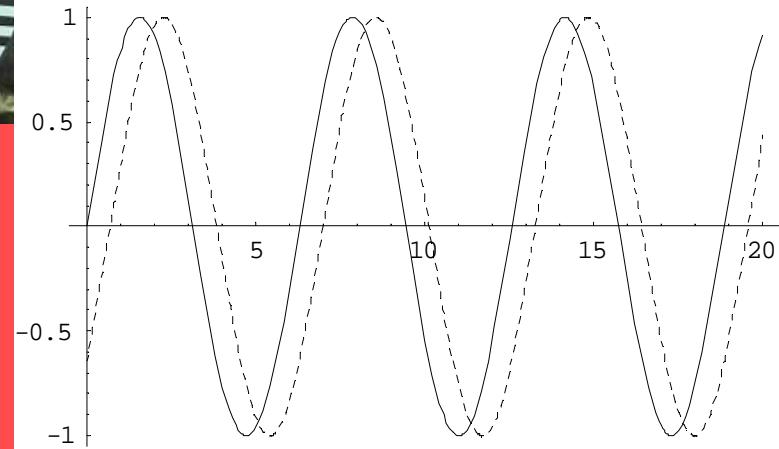
Laser

- Very narrow bundle: high spatial resolution
- But: High temporal resolution of measurement electronics (pico-seconds)
- Example: 1cm depth resolution: 70 pico sec
- Reliable measurements: Large #pulses
- Alternative to TOF:
  - Phase Shift encoding
  - Modulation of laser with sin-wave of frequency f_{AM}
  - Phase shift due to time of flight

Laser ctd.

High depth resolution:

- Either: high fAM
- Or: high resolution Φ of electronics
- Example: Φ for 1cm: 0.22°

$$t_f = \frac{2R}{c}$$

$$\Delta\phi = \frac{t_f}{f_{AM}^{-1}} * 2\pi = \frac{4\pi * f_{AM} R}{c}$$

$$R* = \frac{c}{2 * f_{AM}} = \frac{\lambda_{AM}}{2}$$

Pulsed Time of Flight

- Basic idea: send out pulse of light (usually laser), time how long it takes to return

$$d = \frac{1}{2}c\Delta t$$

DeltaSphere by http://www.3rdtech.com/
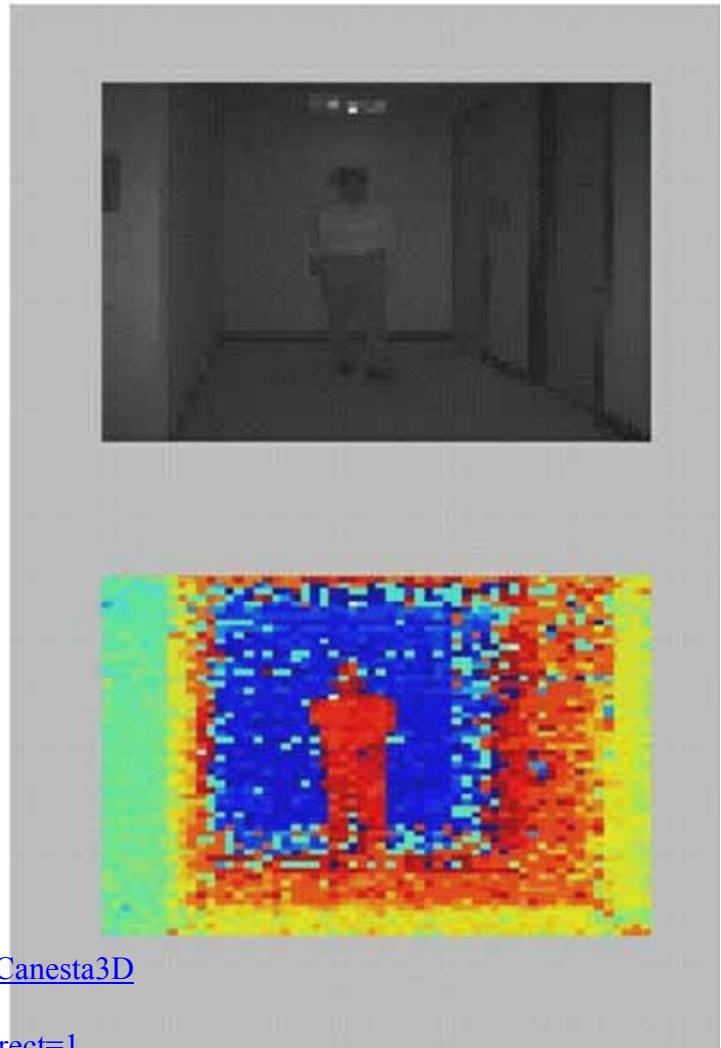
Depth cameras

2D array of
time-of-flight sensors

e.g. Canesta's CMOS 3D sensor

jitter too big on single measurement,

but averages out on many
(10,000 measurements⇒100x improvement)

Canesta: Principle: http://en.wikipedia.org/wiki/Canesta
Casnesta Inc Demo Videos: https://www.youtube.com/user/Canesta3D
Demos:
http://www.youtube.com/watch?v=5_PVx1NbUZQ&noredirect=1
http://www.youtube.com/watch?v=TmKShSHOSYU

Depth cameras

3DV's Z-cam
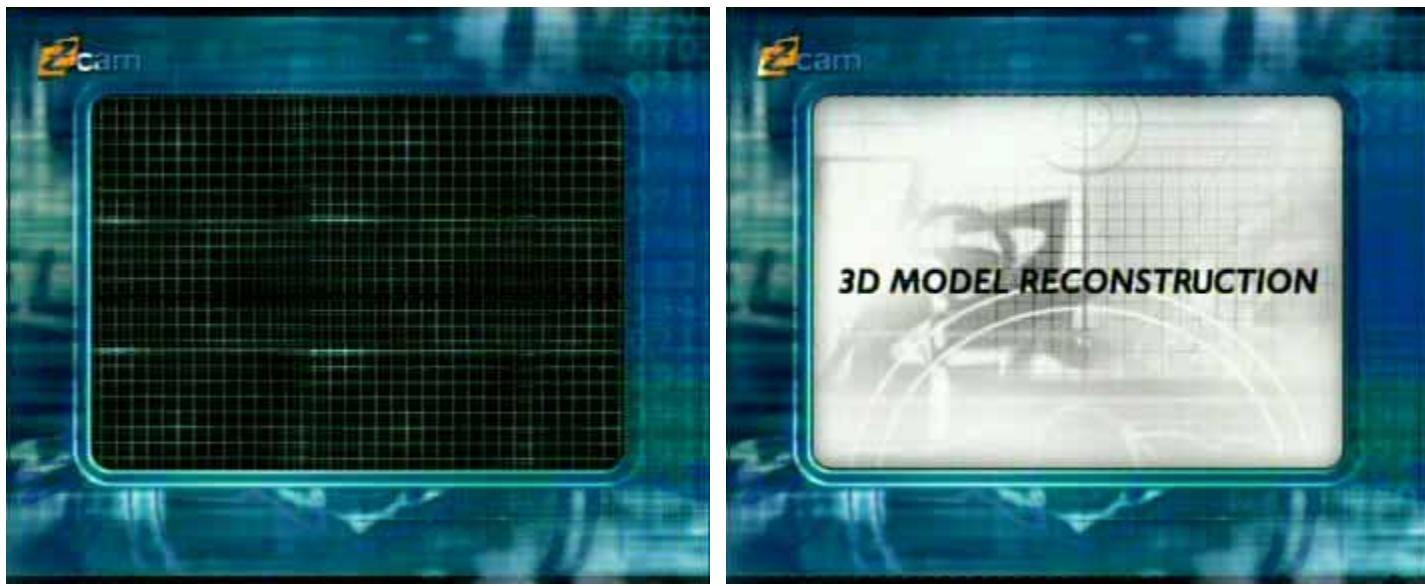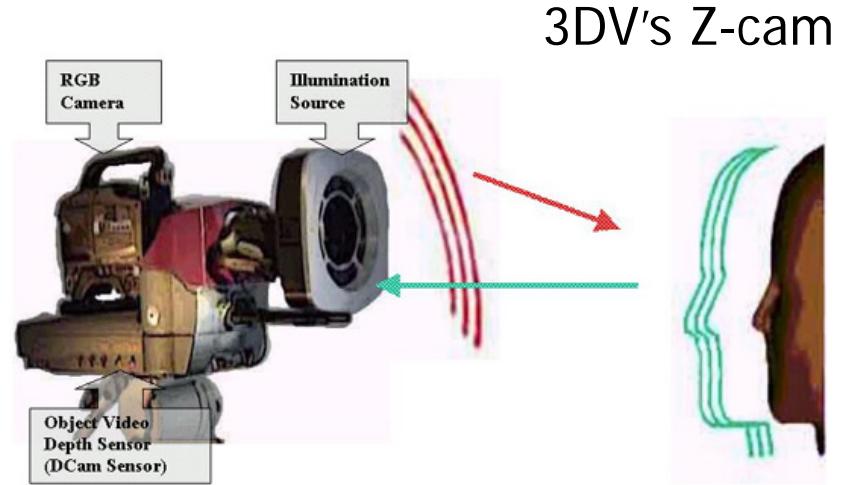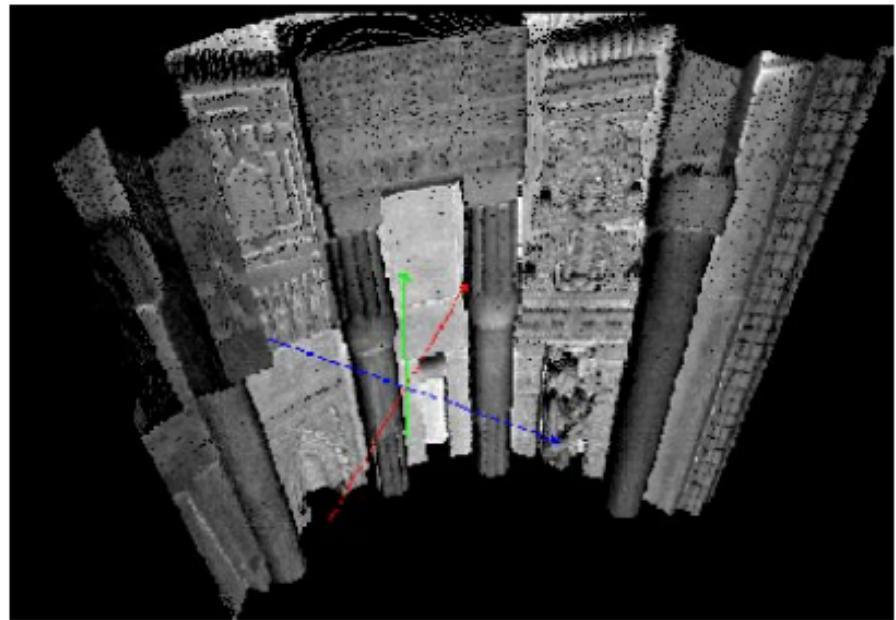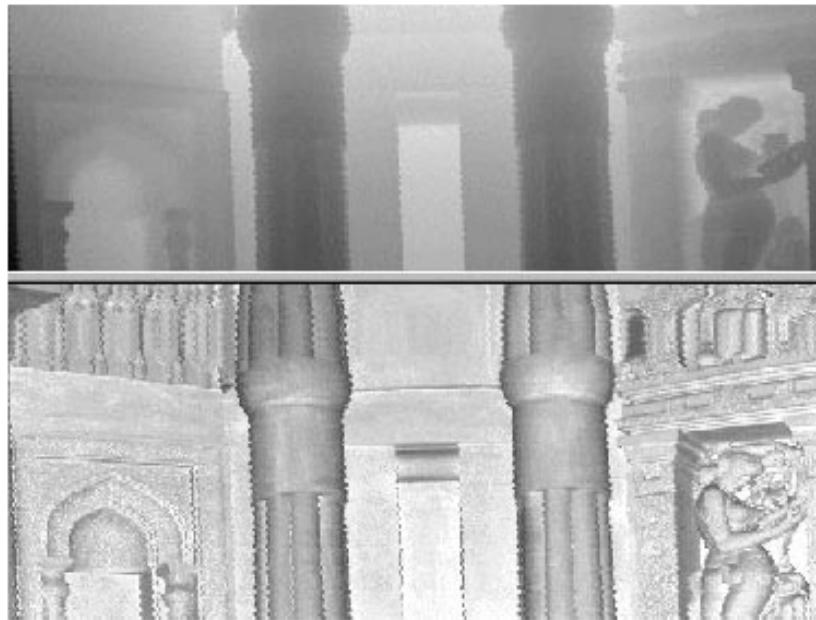
Superfast shutter + standard CCD

- cut light off while pulse is coming back, then $I \sim Z$
- but $I \sim$ albedo (use unshuttered reference view)

RGB Camera

Illumination Source

Object Video Depth Sensor (DCam Sensor)

3D MODEL RECONSTRUCTION
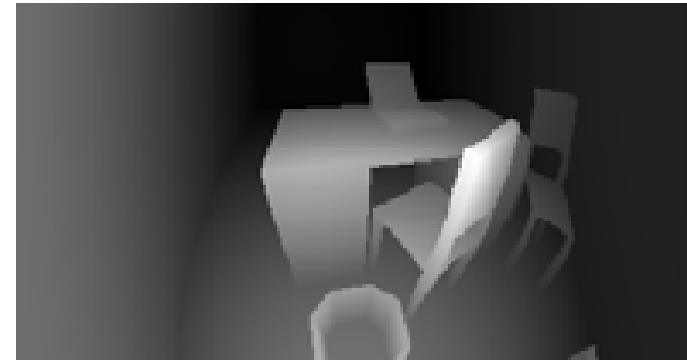
Range Image Data

Figure 24.3. Range data captured by the AM phase shift range finder described in [Hancock *et al.*, 1998]: (left) range and intensity images; (right) perspective plot of the range data. Reprinted from [Hebert, 2000], Figure 5.

Input Data

(a)

(b)
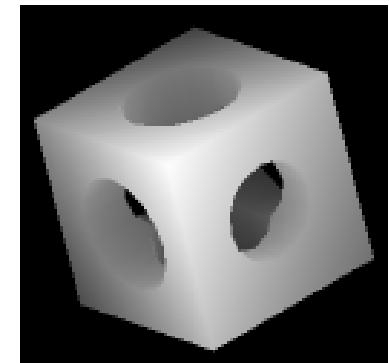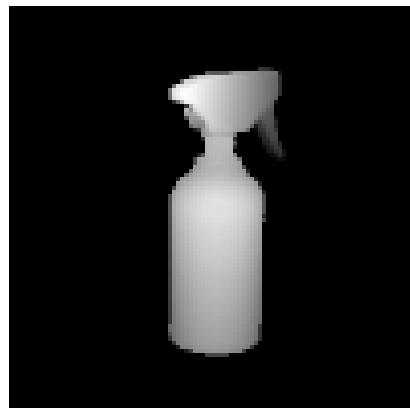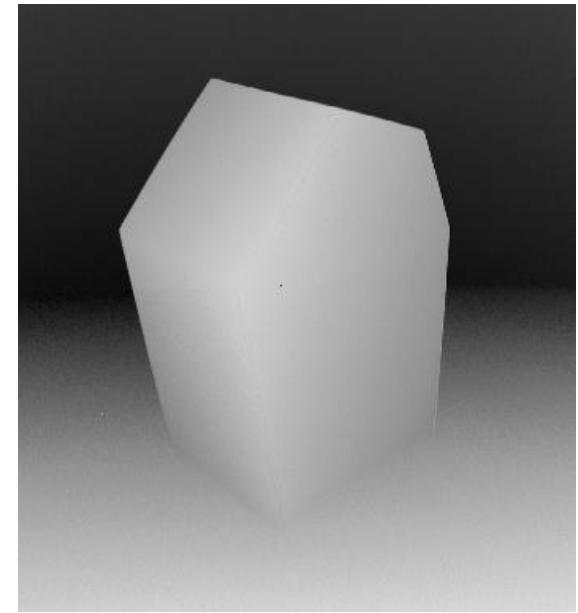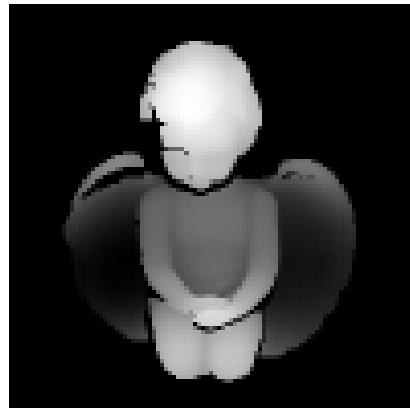
(c)

Simulated and real range images
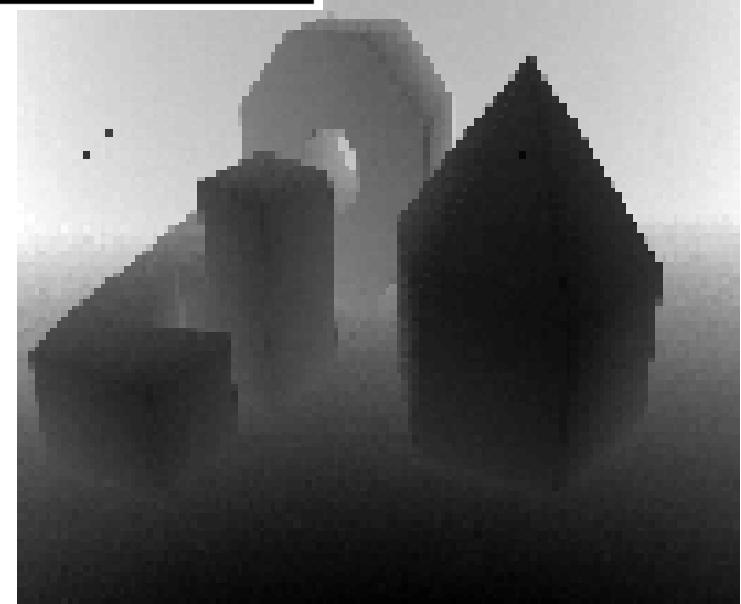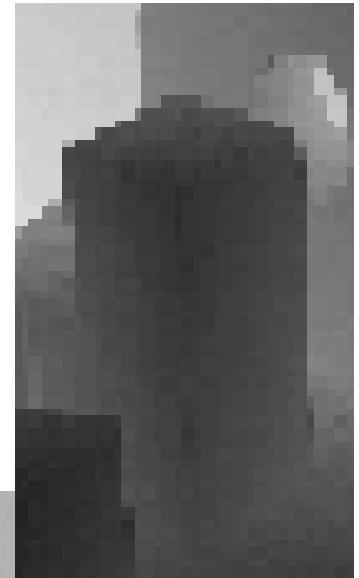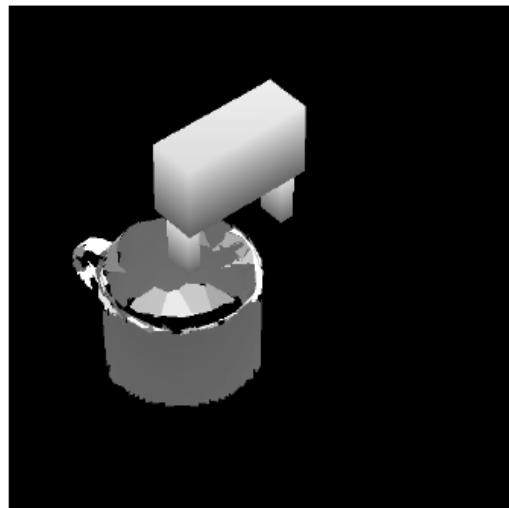
What is special about range images?

Object faces?
Object boundaries?
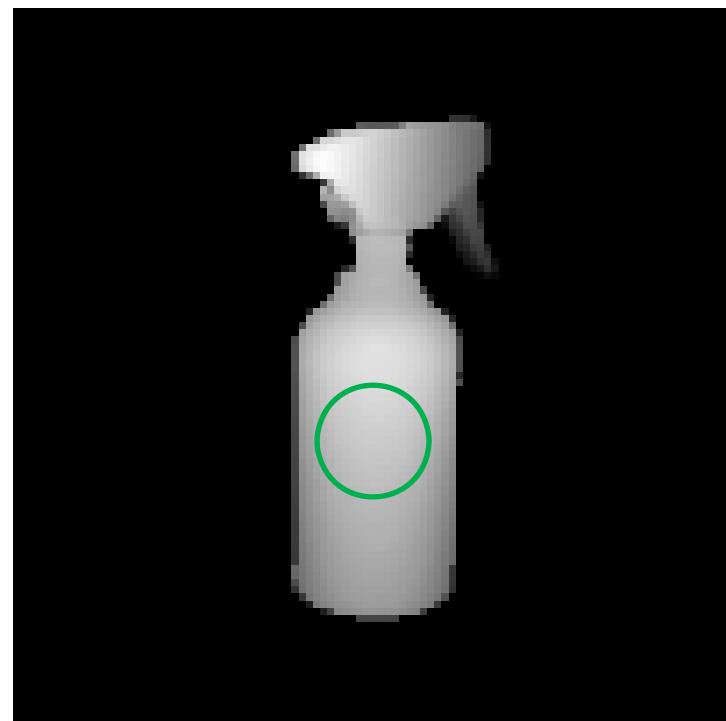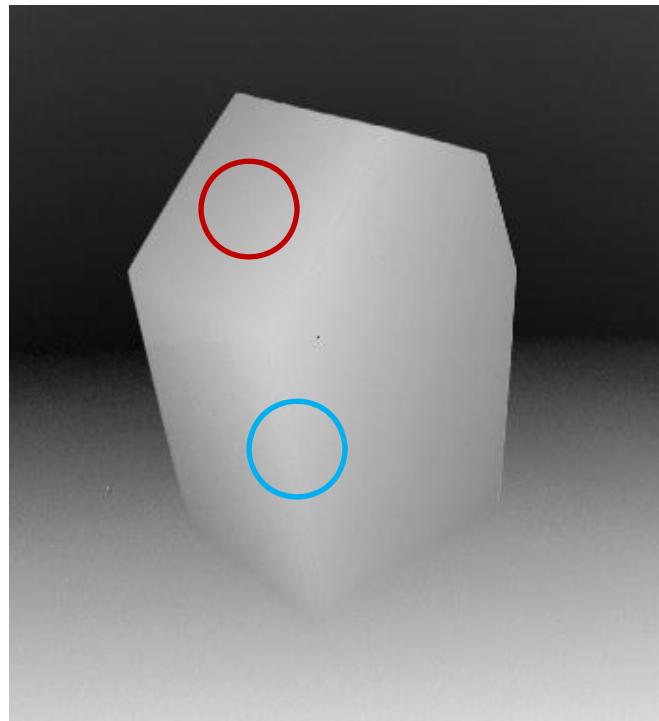
What is different in range images?

Object faces?
Object boundaries?
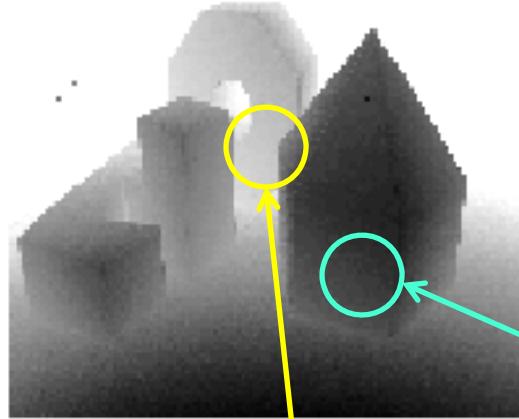
What is special about range images?
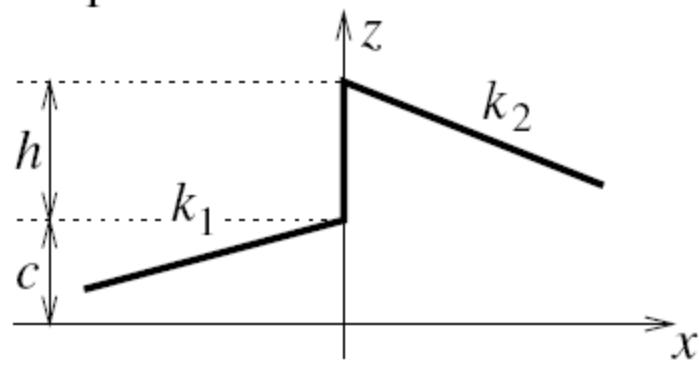
- Homogeneous in surface normals
- Crest line: Abrupt change of surface normals
- Continuous change of normals, homogeneous in curvature
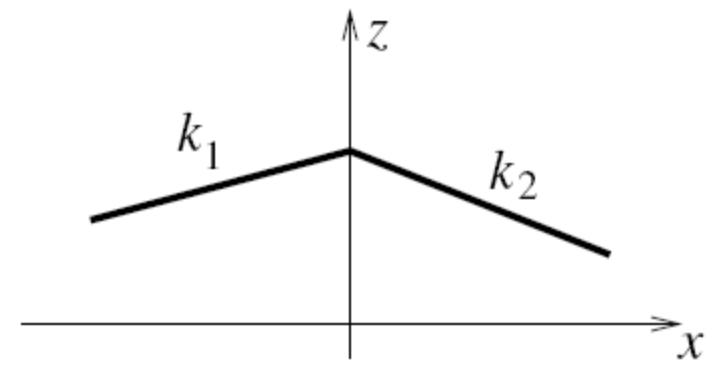
Types of Discontinuities in Range Images

Step Model

Roof Model
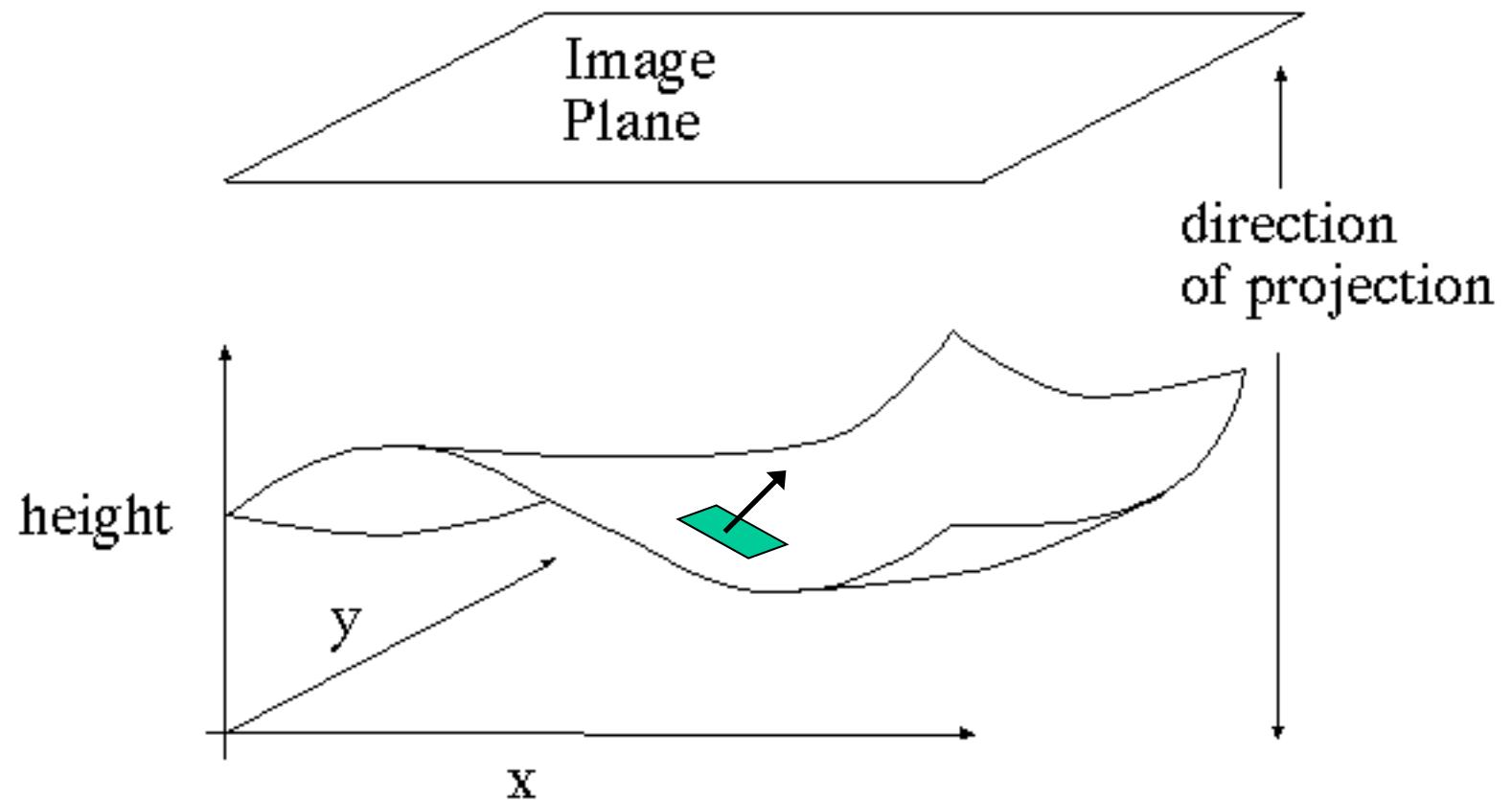
Properties of object surfaces in range images

- Homogeneity of surface properties in:
  - Surface normals
  - Curvature
- Discontinuities between surfaces:
  - "roof edges": locations with change of normals
  - "step edges": discontinuous depth (e.g. hidden objects)
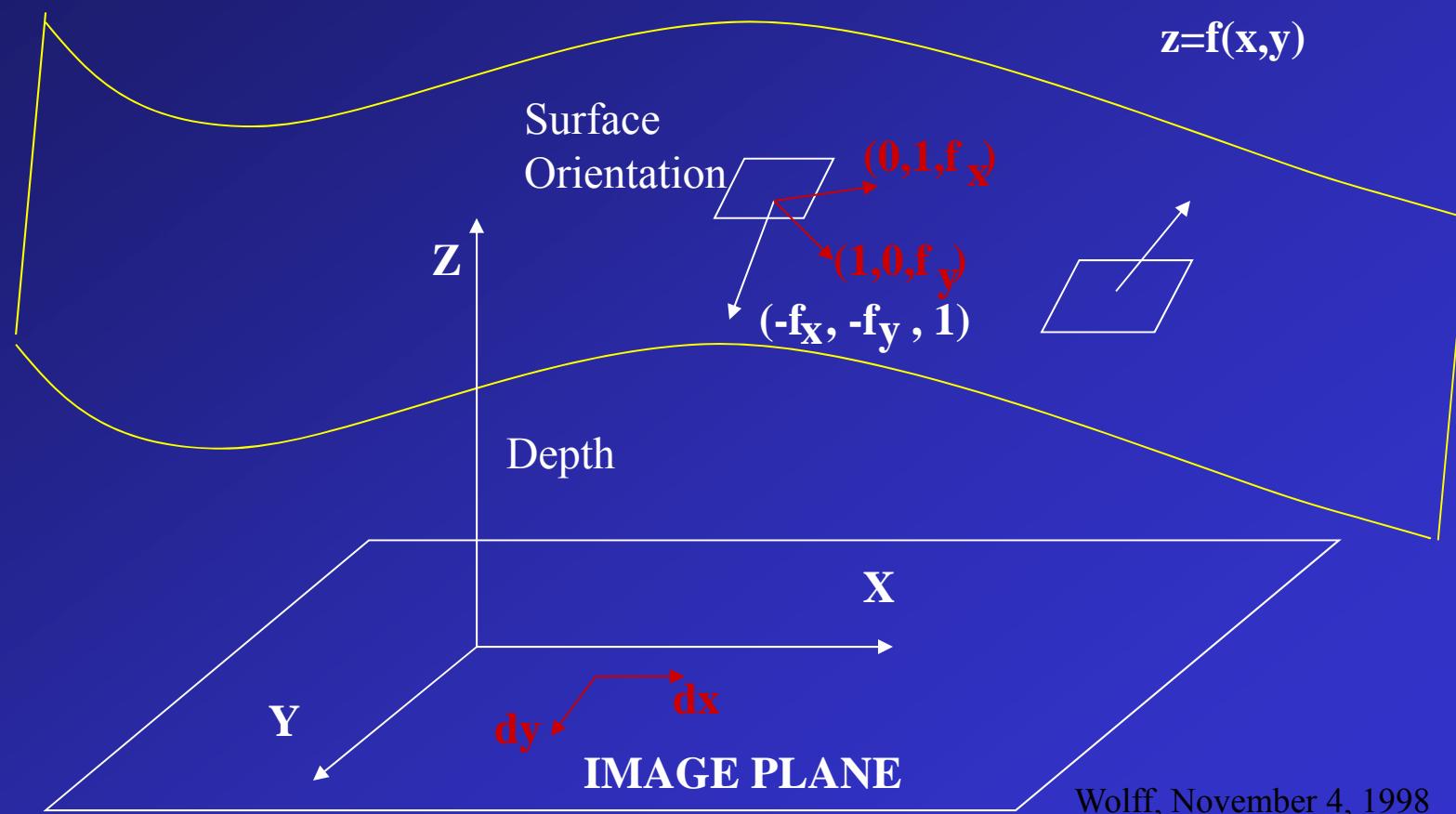
Remember:
Shape from Shading: "Monge" Patch

Image Plane

direction of projection

height

y

x

Surface Orientation and Surface Normal

$(f_x, f_y, -1) = (0,1,f_x) \times (1,0,f_y)$

$z = f(x,y)$

Surface Orientation

$(0,1,f_x)$

$(1,0,f_y)$

Z

$(-f_x, -f_y, 1)$

Depth

X

dx

Y

dy

IMAGE PLANE

Surface Orientation and Surface Normal
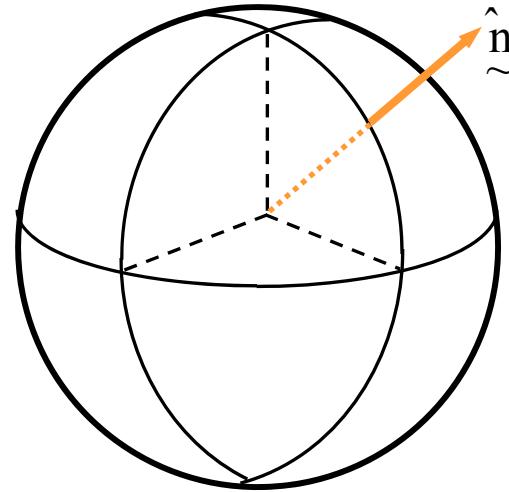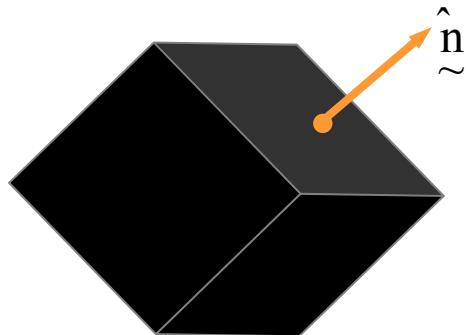
$$(-f_x, -f_y, 1) = (-p, -q, 1)$$

p, q comprise a gradient or gradient space representation for local surface orientation.
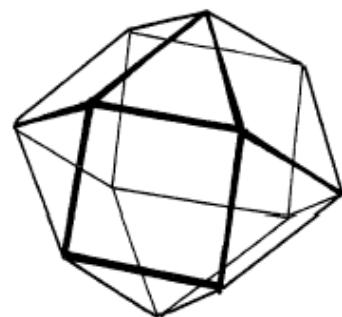
Object Representation: The Gaussian Image (EGI)

- Surface normal information for any object is mapped onto a unit (Gaussian) sphere by finding the point on the sphere with the same surface normal:
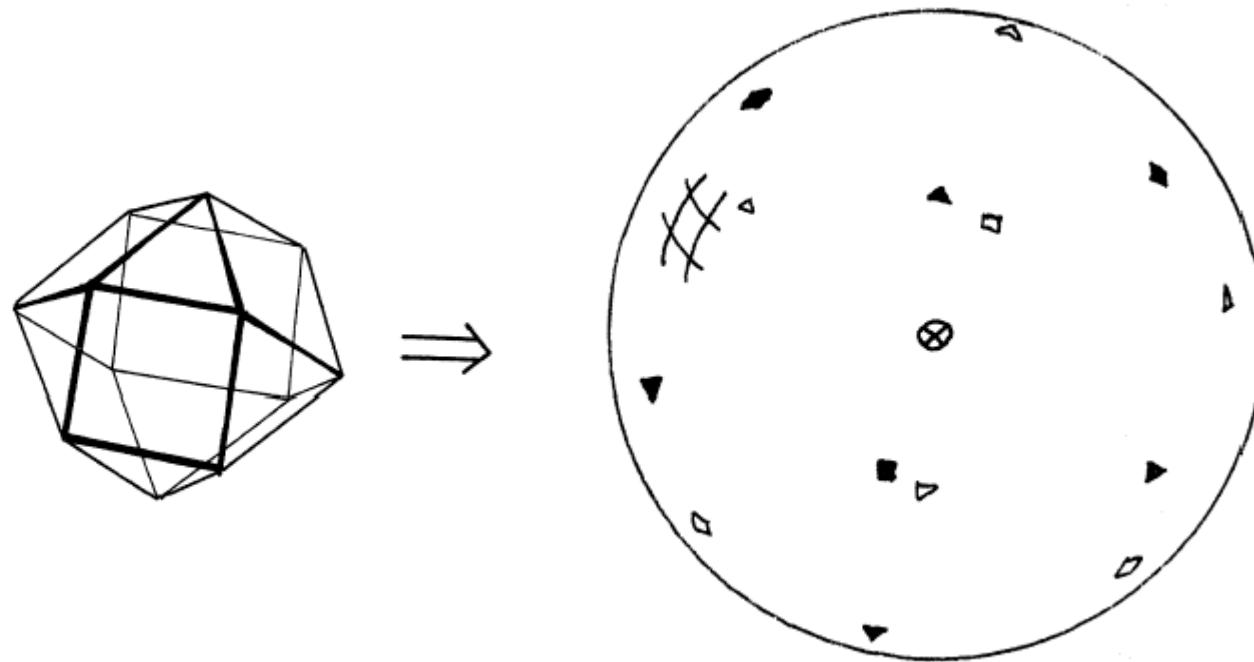
Example (K. Horn)

Example (K. Horn)

*www.cs.jhu.edu/~misha/Fall04/**EGI**1.ppt*
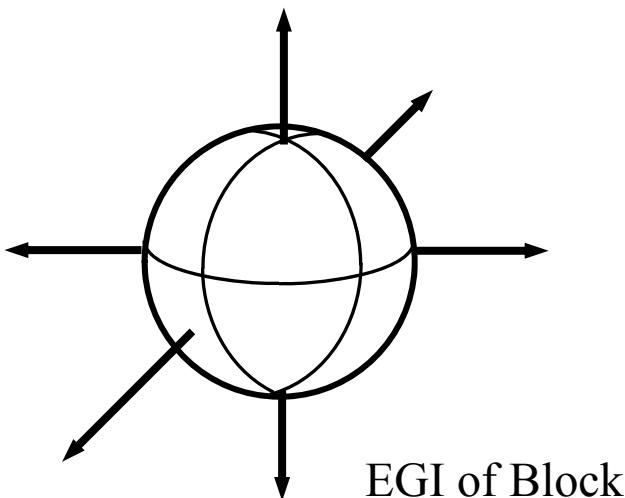
The Extended Gaussian Image

- We can extend the Gaussian image by
  - placing a mass at each point on the sphere equal to the area of the surface having the given normal
  - masses are represented by vectors parallel to the normals, with length equal to the mass (VOTING)
- An example:
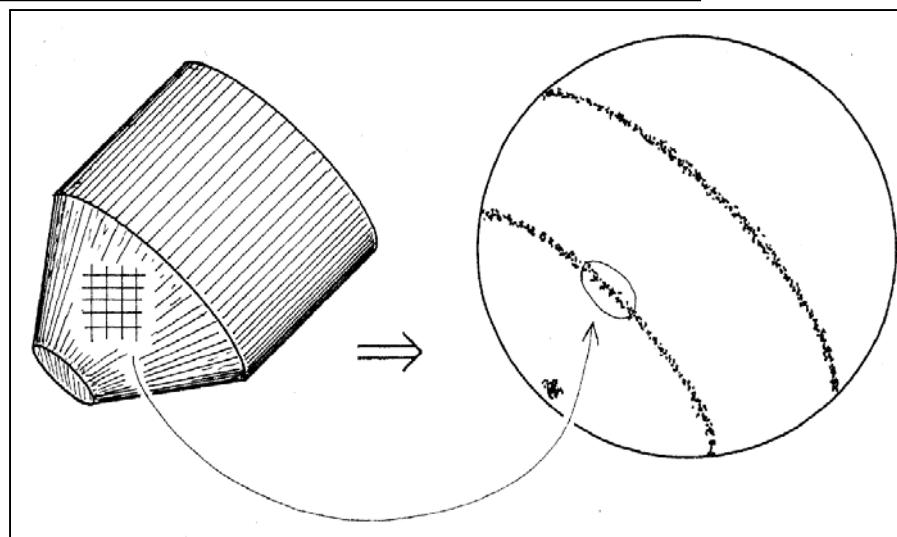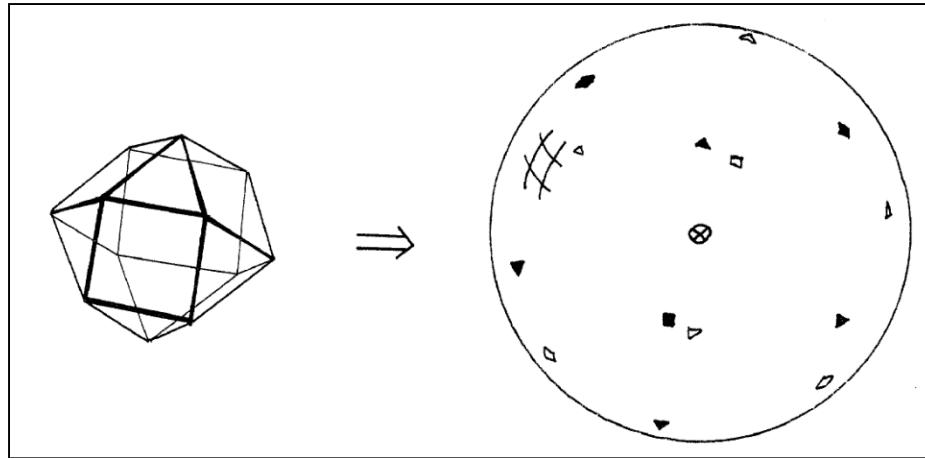
Block

EGI of Block

K. Horn, MIT, 1983
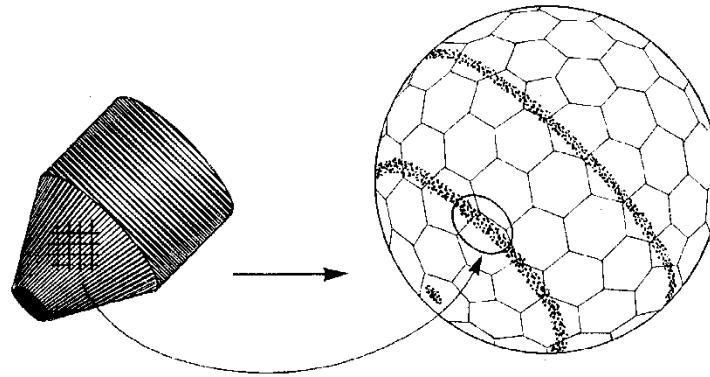
The Discrete Case EGI

- To represent the information of the Gaussian sphere in a computer, the sphere is divided into cells:

- For each image cell on the left, a surface orientation is found and accumulated in the corresponding cell of the sphere.

Properties of the Gaussian Image

- This mapping is called the *Gaussian image* of the object when the surface normals for each point on the object are placed such that:
  - tails lie at the center of the Gaussian sphere
  - heads lie on the sphere at the matching normal point
- In areas of convex objects with positive curvature, no two points will have the same normal.
- Patches on the surface with zero curvature (lines or areas) correspond to a single point on the sphere.
- Rotations of the object correspond to rotations of the sphere.

Using the EGI

- EGIs for different objects or object types may be computed and stored in a model database as a surface normal vector histogram.

- Given a depth image, surface normals may be extracted by plane fitting.

- By comparing EGI histogram of the extracted normals and those in the database, the identity and orientation of the object may be found.

Properties of object surfaces in range images

- Homogeneity of surface properties in:
  - Surface normals
  - Curvature
- Discontinuities between surfaces:
  - "roof edges": continuous depth but change of normals
  - "step edges": discontinuous depth (e.g. hidden objects)

Segmentation into planar patches

- F&P page 476/477
- Idea: Break object surface into sets of flat pieces
  - Clustering of surface normals via EGI
  - Region growing: Iterative merging of planar patches via graph/arc-costs

Segmentation into planar patches
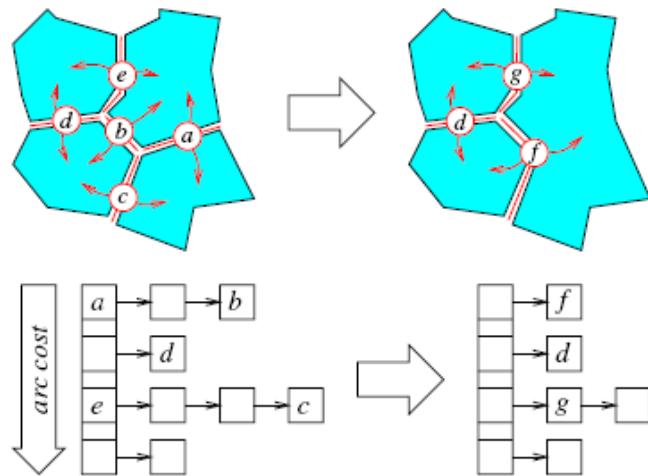
Figure 24.11. This diagram illustrates one iteration of the region growing process during which the two patches incident to the minimum-cost arc labelled a are merged. The heap shown in the bottom part of the figure is updated as well: the arcs a, b, c and e are deleted, and two new arcs f and g are created and inserted in the heap.

Iterative merging of planar patches:
- Graph nodes: Patches with best fitting plane
- Graph arcs: costs corresponding to average error between combined set of points and plane that best fits these points
- Iteration: Find best arc, merge, next …

Segmentation into planar patches
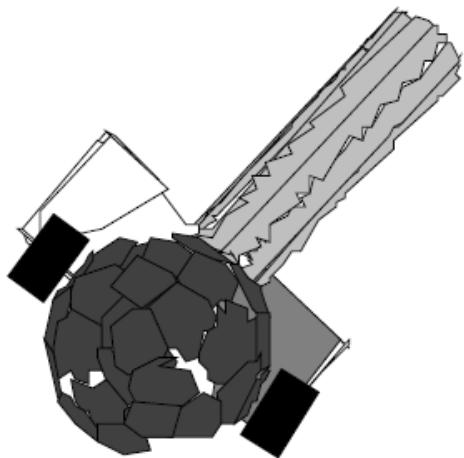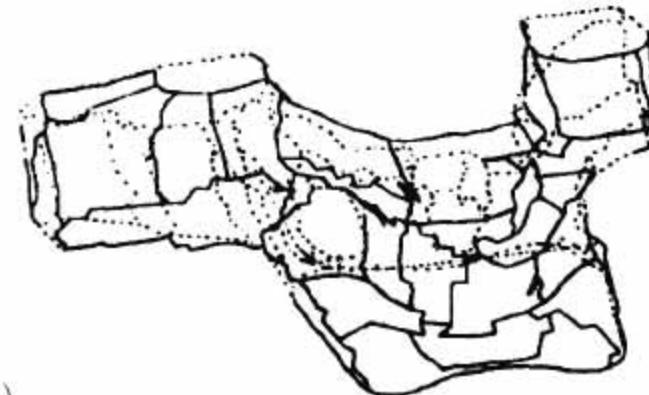
Figure 7: One of several grasping possibilities.
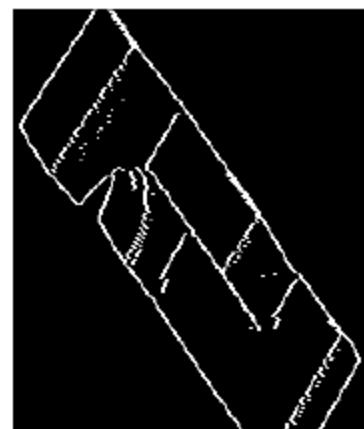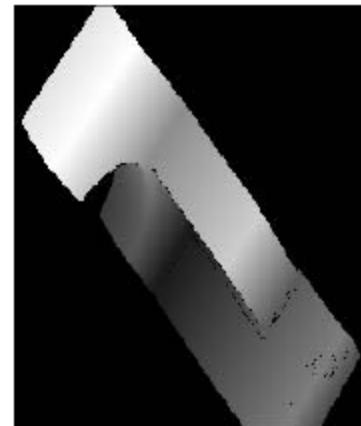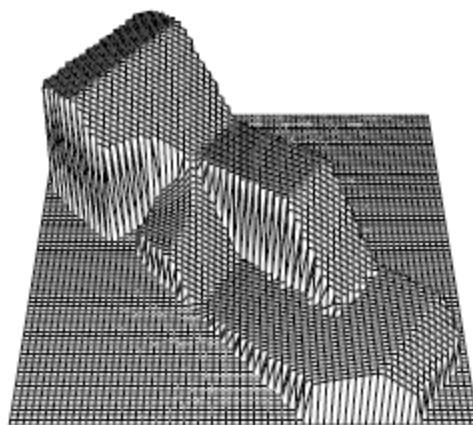

(a)   (b)
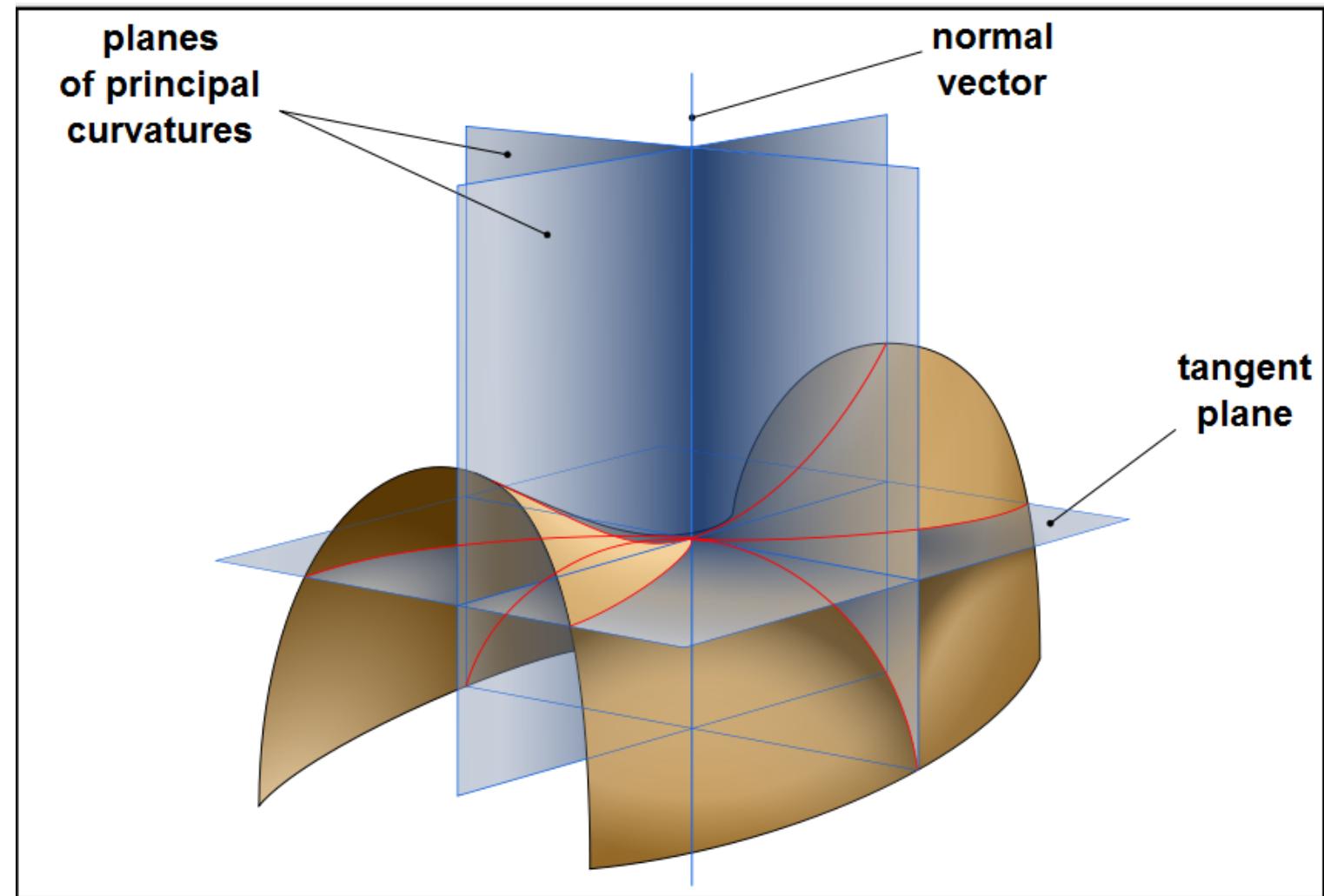
Figure 24.12. The Renault part: (a) photo of the part and (b) its model. Reprinted from [Faugeras and Hebert, 1986], Figures 1 and 6.

Segmentation into planar patches

From flat pieces to curvature: Differential Geometry

Elements of Analytical Differential Geometry (see F&P)

- Parametric surface: $\quad x : U \times \mathbb{R}^2 \rightarrow \mathbb{R}^3$

- Unit surface normal: $\quad N = \dfrac{1}{|x_u \times x_v|} (x_u \times x_v)$

- First fundamental form:

$$I(t, t) = Eu'^2 + 2Fu'v' + Gv'^2 \qquad \begin{cases} E = x_u.x_u \\ F = x_u.x_v \\ G = x_v.x_v \end{cases}$$

- Second fundamental form:

$$II(t, t) = eu'^2 + 2fu'v' + gv'^2 \qquad \begin{cases} e = -N.x_{uu} \\ f = -N.x_{uv} \\ g = -N.x_{vv} \end{cases}$$

- Normal (direction t) and Gaussian curvatures:
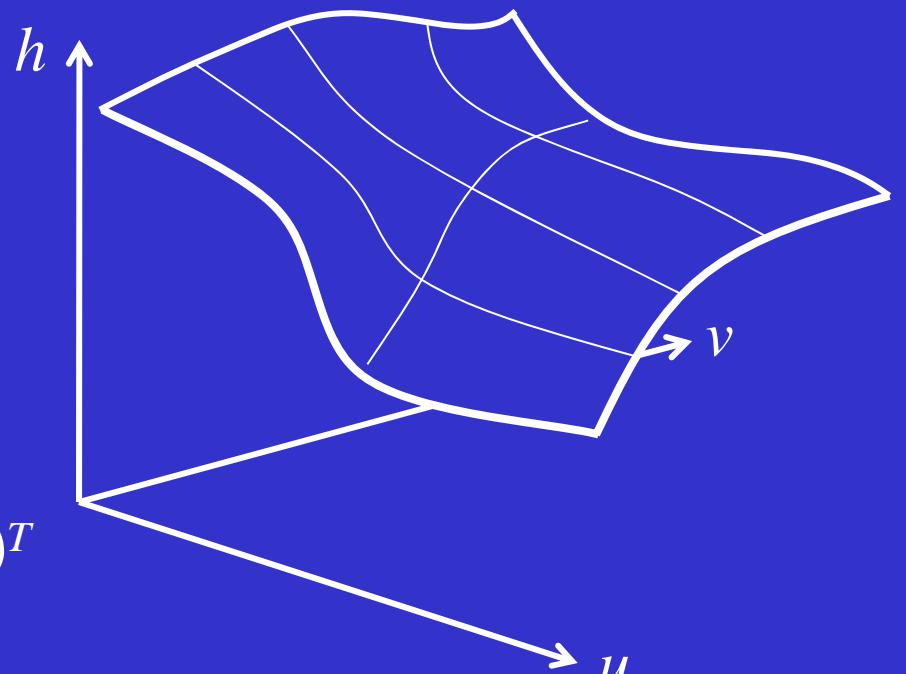
$$\kappa_t = \frac{II(t, t)}{I(t, t)} \qquad\qquad K = \frac{eg - f^2}{EG - F^2}$$

Example: Monge Patches

$$\boxed{x\,(\,u,\,v\,)=\,(u,\,v,\,h(\,u,\,v\,))}$$

In this case

- $N= \dfrac{1}{(1+h_u^2+h_v^2)^{1/2}}\,(\,-h_u,\,-h_v,\,1)^T$

- $E = 1+h_u^2;\ \ F = h_uh_v;\ G = 1+h_v^2$

- $e = \dfrac{-h_{uu}}{(1+h_u^2+h_v^2)^{1/2}}\,;\ \ f = \dfrac{-h_{uv}}{(1+h_u^2+h_v^2)^{1/2}}\,;\ \ g = \dfrac{-h_{vv}}{(1+h_u^2+h_v^2)^{1/2}}$

And the Gaussian curvature is: $\quad K = \dfrac{h_{uu}h_{vv}-h_{uv}^2}{(1+h_u^2+h_v^2)^2}$.
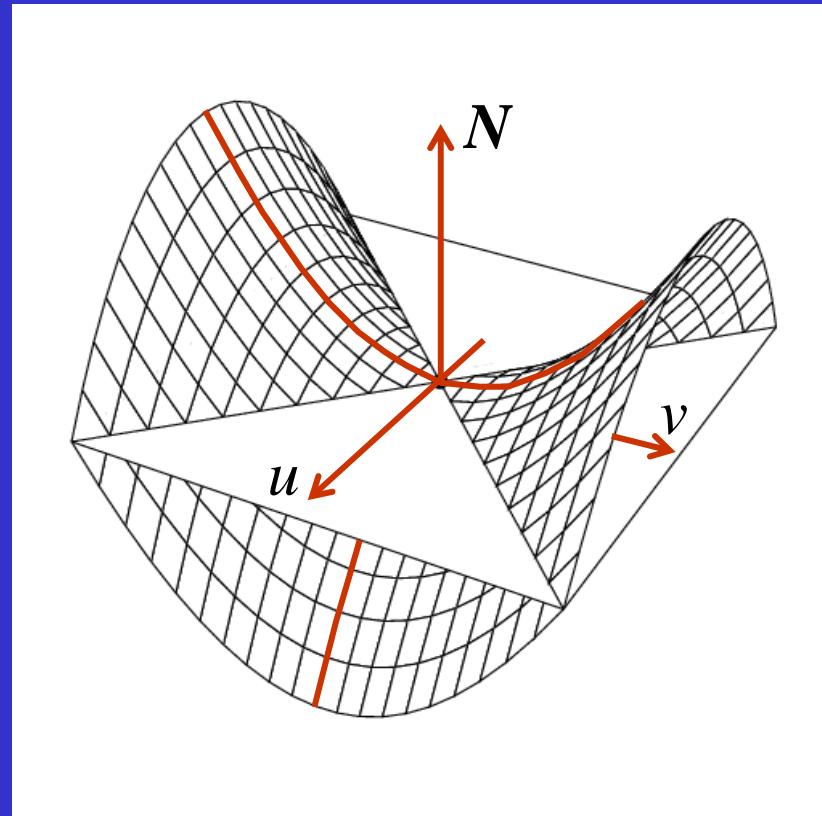
Example: Local Surface Parameterization

- u,v axes = principal directions

- h axis = surface normal

In this case:
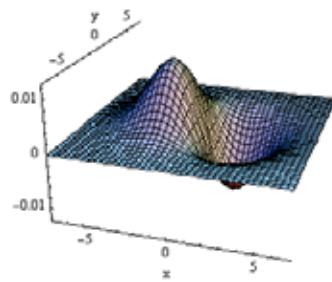
- $h(0,0)=h_u(0,0)=h_v(0,0)=0$

- $N=(0,0,1)^T$

- $h_{uv}(0,0)=0$, $\kappa_1=-h_{uu}(0,0)$, $\kappa_2=-h_{vv}(0,0)$

Taylor expansion of order 2 $\Longrightarrow$ $h(u,v) = -\tfrac{1}{2}\,(\kappa_1\,u^2 + \kappa_2\,v^2)$

Calculation of Partial Derivatives
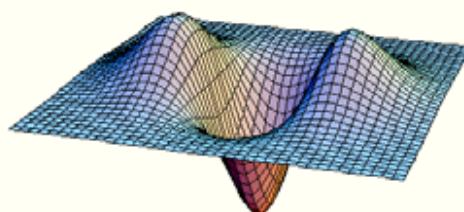
by convolving the smoothed image with the masks:

$$\frac{\partial}{\partial x} = \frac{1}{6} \begin{array}{|c|c|c|} \hline -1 & 0 & 1 \\ \hline -1 & 0 & 1 \\ \hline -1 & 0 & 1 \\ \hline \end{array} \quad \text{and} \quad \frac{\partial}{\partial y} = \frac{1}{6} \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline 0 & 0 & 0 \\ \hline -1 & -1 & -1 \\ \hline \end{array},$$

and the Hessian is computed by convolving the smoothed image with the masks

$$\frac{\partial^2}{\partial x^2} = \frac{1}{3} \begin{array}{|c|c|c|} \hline 1 & -2 & 1 \\ \hline 1 & -2 & 1 \\ \hline 1 & -2 & 1 \\ \hline \end{array}, \quad \frac{\partial^2}{\partial x \partial y} = \frac{1}{4} \begin{array}{|c|c|c|} \hline -1 & 0 & 1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 0 & -1 \\ \hline \end{array} \quad \text{and} \quad \frac{\partial^2}{\partial y^2} = \frac{1}{3} \begin{array}{|c|c|c|} \hline 1 & 1 & 1 \\ \hline -2 & -2 & -2 \\ \hline 1 & 1 & 1 \\ \hline \end{array}.$$

Principal Directions
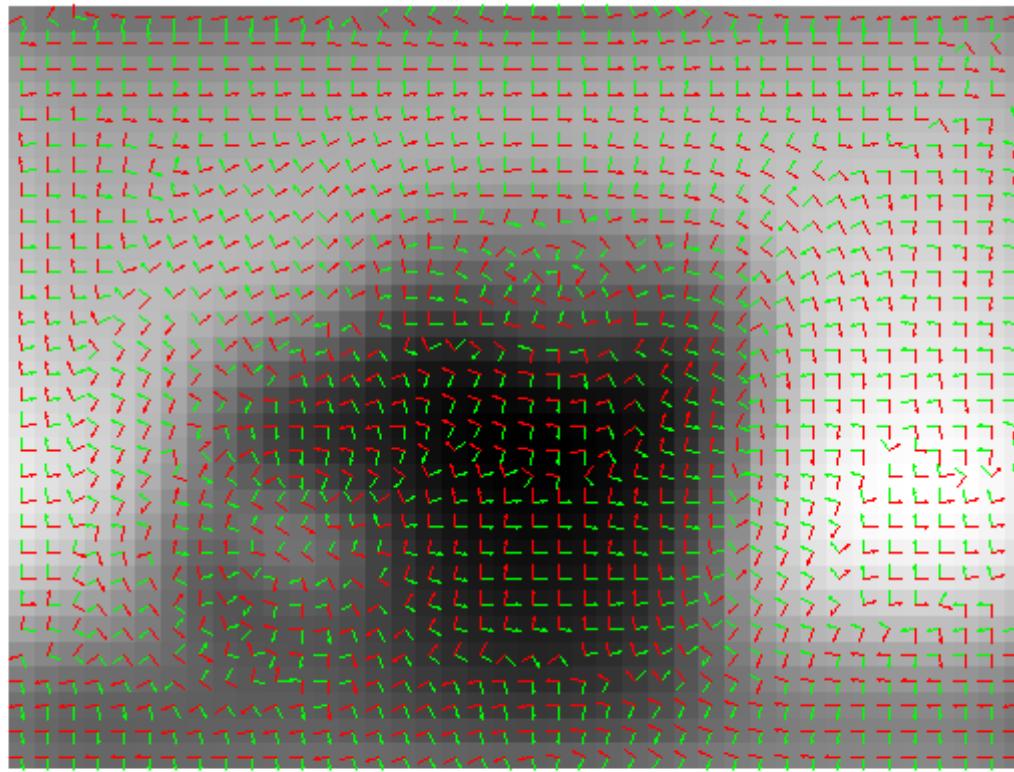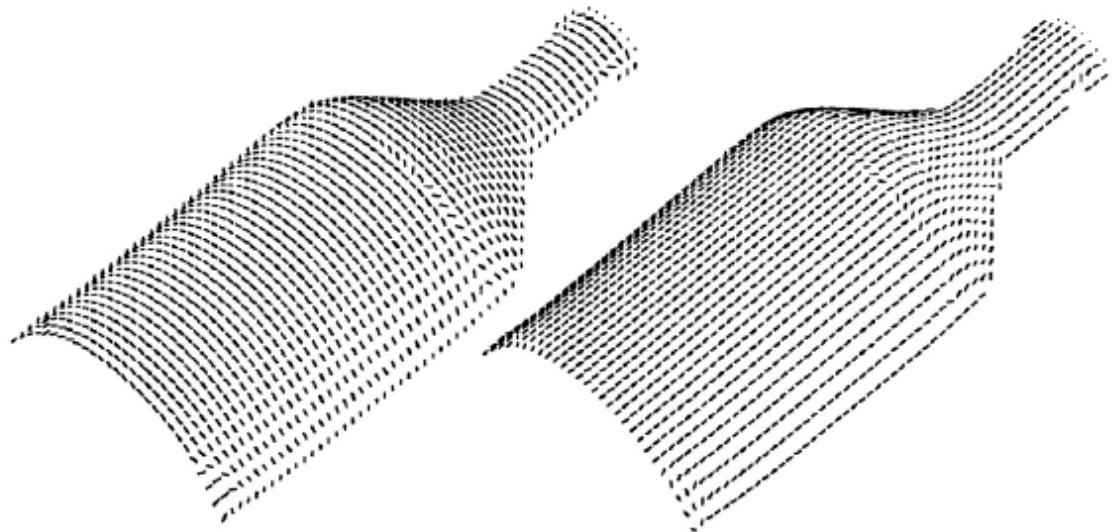
Figure 6.1 Frames of the normalized principal curvature directions at a scale of 1 pixel. Image resolution 32^2 pixels. Green: maximal principal curvature direction; red: minimal principal curvature direction.

# Calculation of principal curvatures
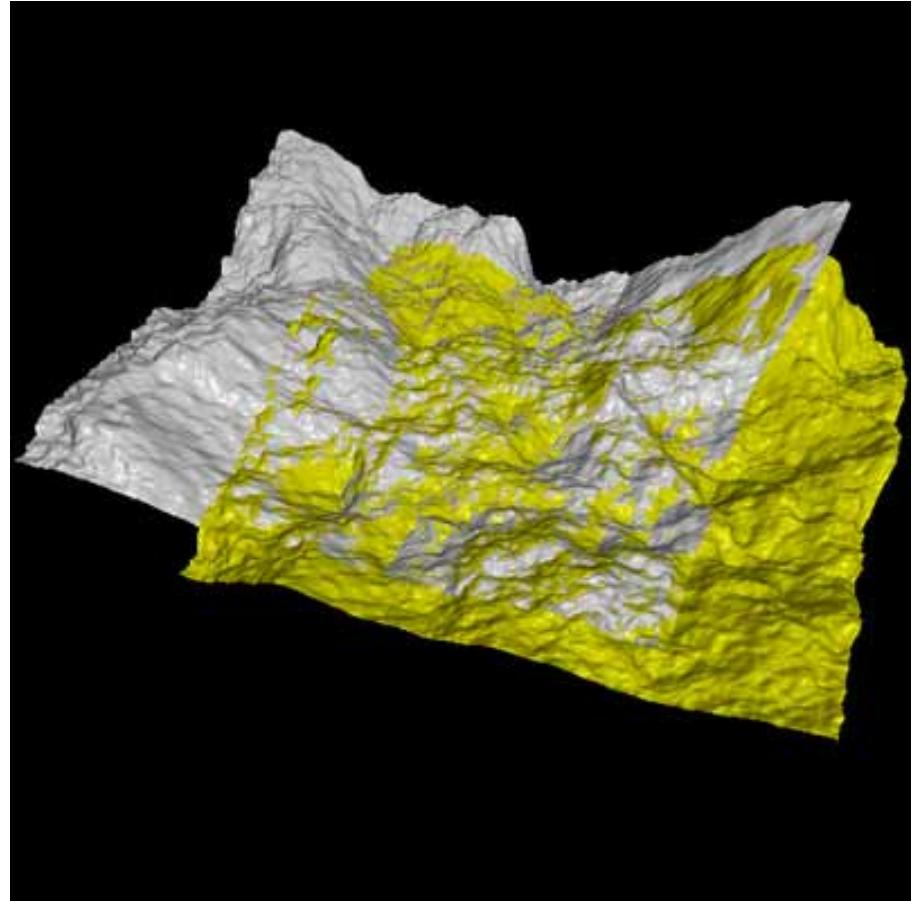
Figure 24.8. The two principal direction fields for the oil bottle. Reprinted from [Brady *et al.*, 1985], Figure 18.

Note that the principal curvatures are homogeneous across the large lower part of the bottle → can serve as <u>homogeneous</u> features for clustering

The Problem

Align two
partially-
overlapping
meshes
given initial
guess
for relative
transform

Range Image Registration ctd.

- Concept:
  - Determine rigid transformation between pairs of range surfaces
  - Minimize average distance between point sets
  - **ICP: I**terative **C**losest **P**oint algorithm (Besl & McKay 1992)
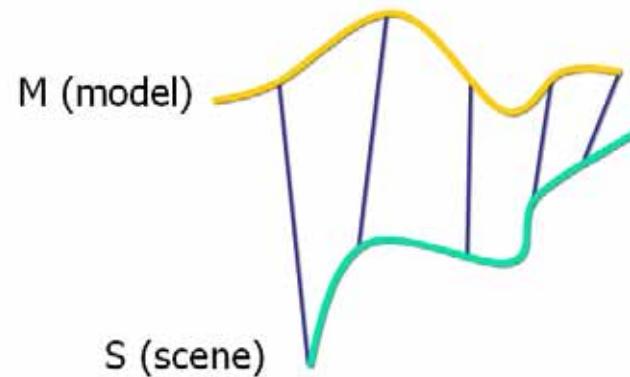
Corresponding Point Set Alignment

- Let M be a model point set.
- Let S be a scene point set.

We assume :

1. $N_M = N_S$.
2. Each point S_i correspond to M_i .

M (model)

S (scene)

Corresponding Point Set Alignment

The MSE objective function :

$$f(R,T) = \frac{1}{N_S} \sum_{i=1}^{N_S} \left\| m_i - Rot(s_i) - Trans(si) \right\|^2$$
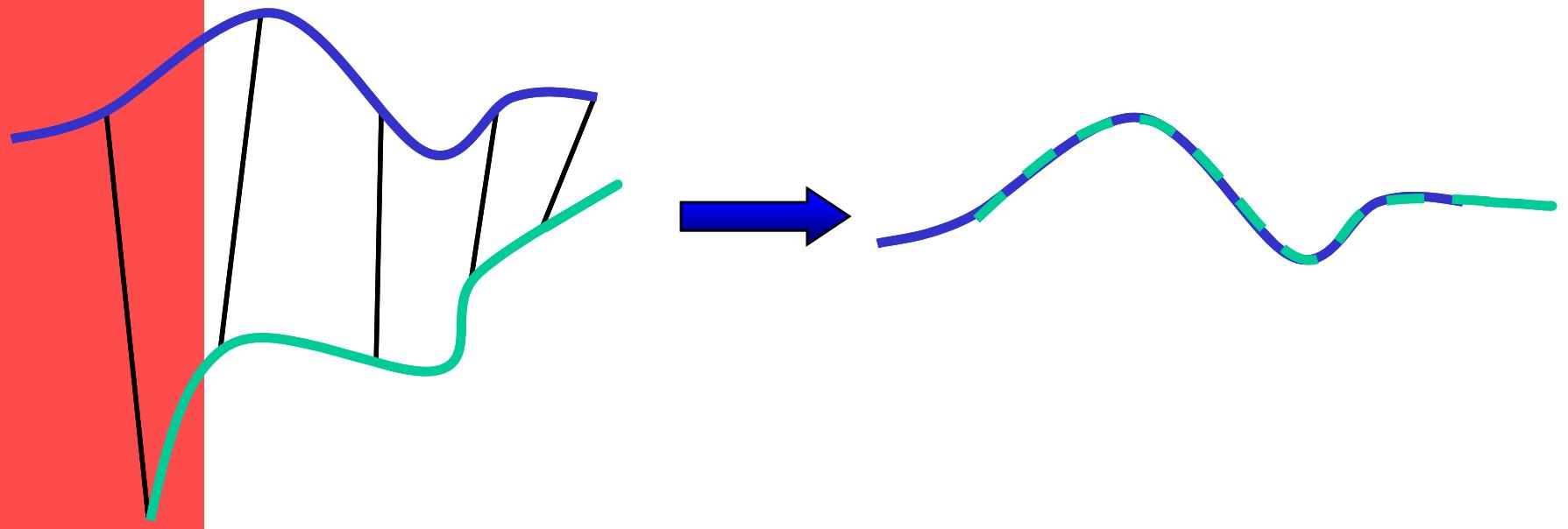
The alignment is :

$$(rot, trans, d_{mse}) = \Phi(M, S)$$

Aligning 3D Data
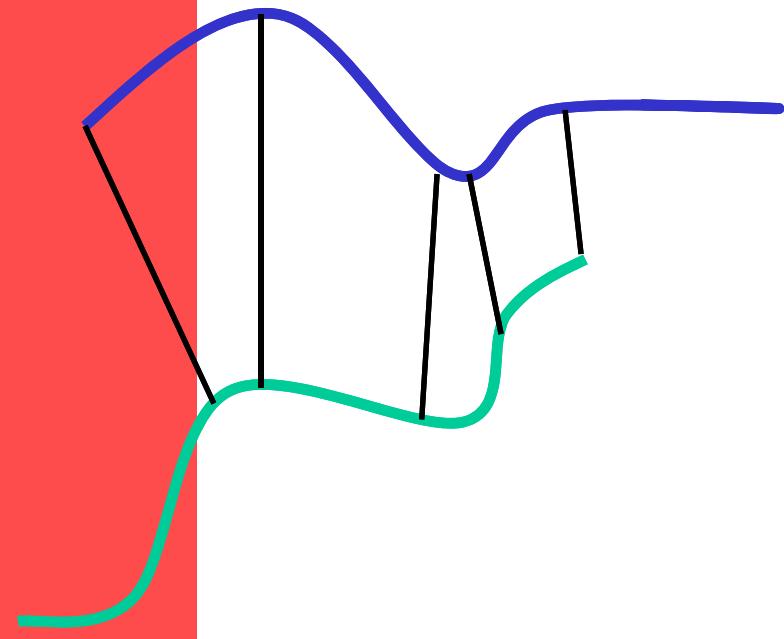
- If correct correspondences are known, can find correct relative rotation/translation

Aligning 3D Data
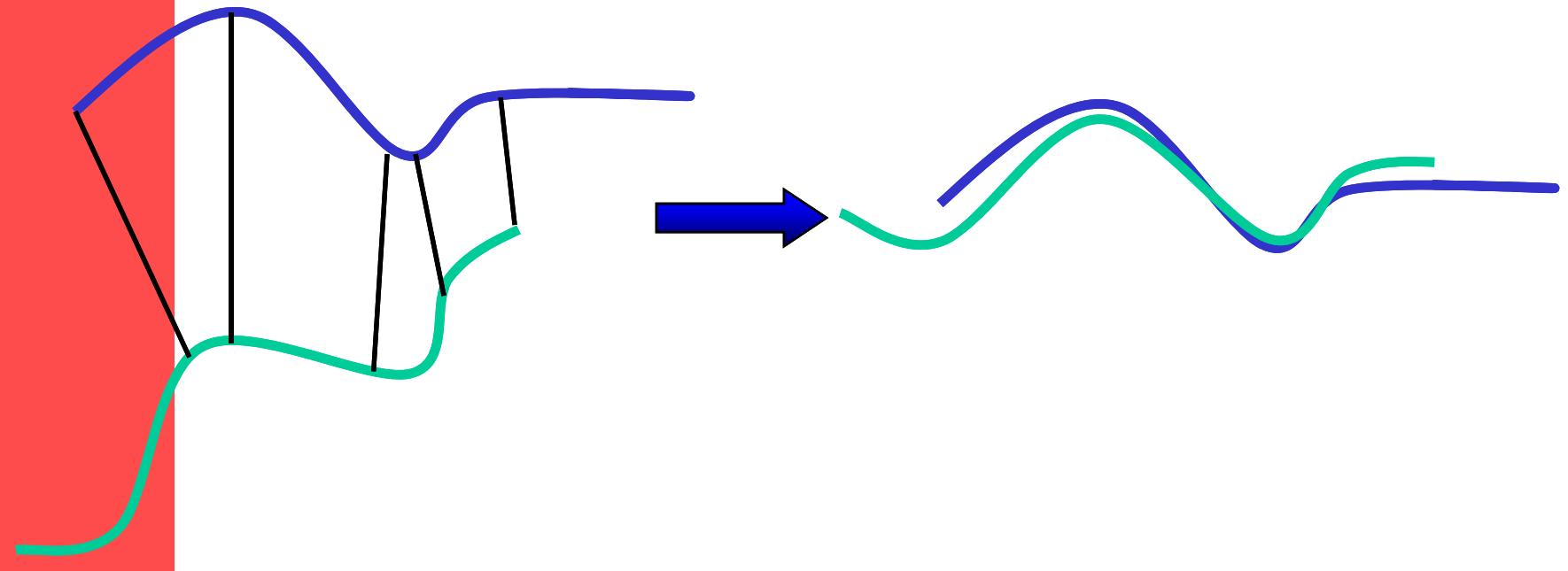
- How to find correspondences: User input? Feature detection? Signatures?
- Alternative: assume closest points correspond
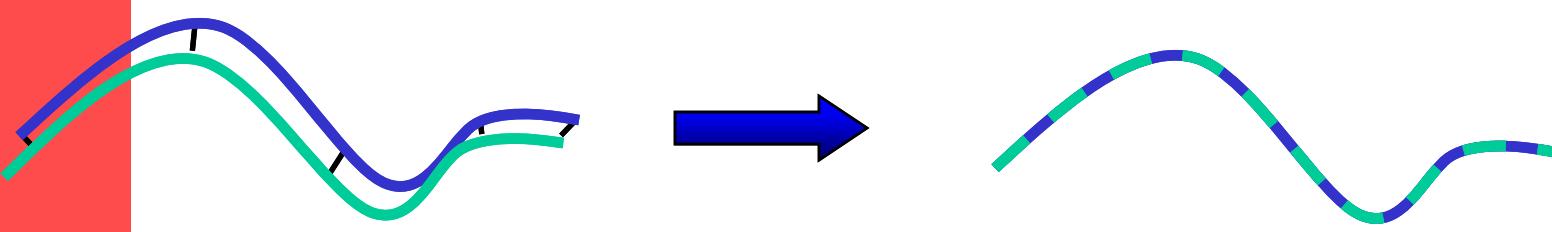
Aligning 3D Data

- How to find correspondences: User input? Feature detection? Signatures?

- Alternative: assume closest points correspond

Aligning 3D Data

- Converges if starting position "close enough"

Closest Point

- Given 2 points r_1 and r_2 , the Euclidean distance is:

$$d(r_1, r_2) = \|r_1 - r_2\| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2}$$
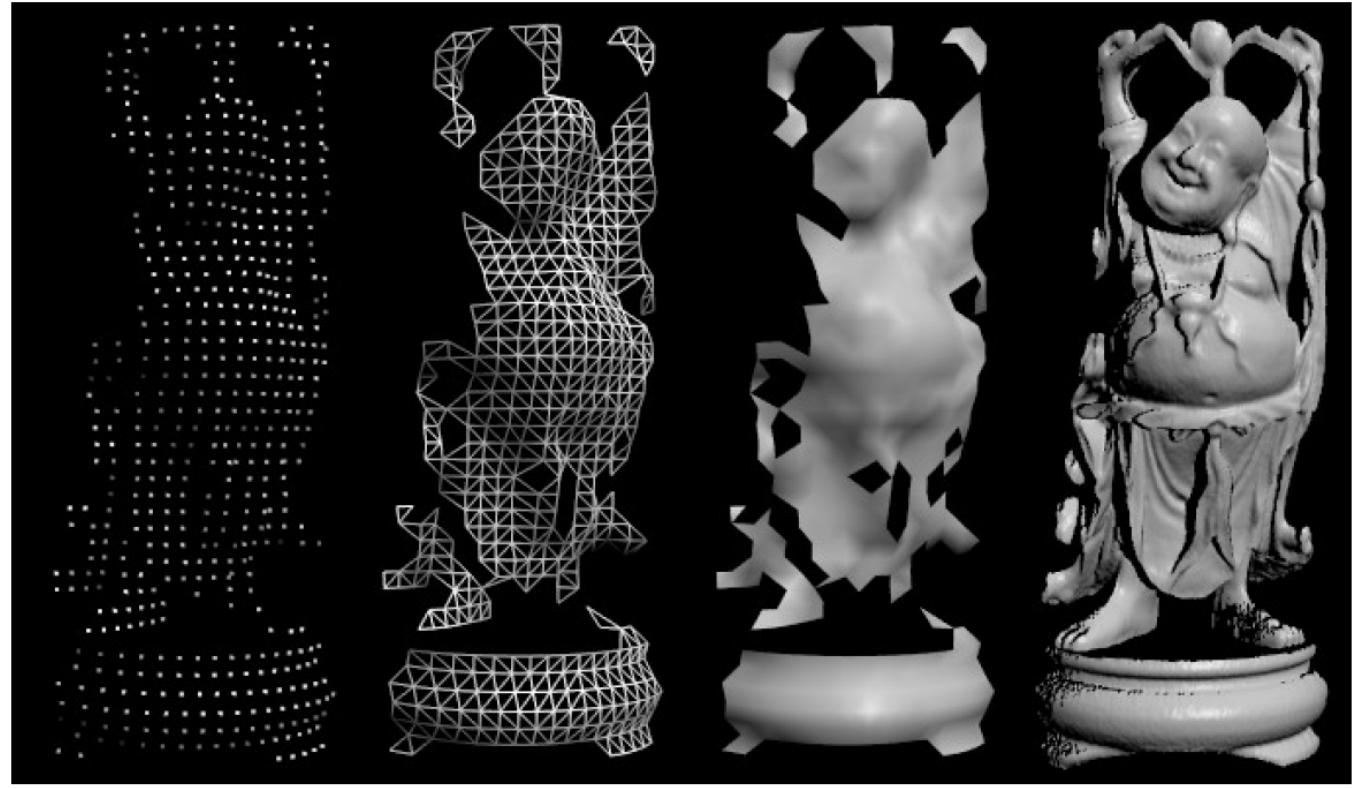
- Given a point r_1 and set of points A , the Euclidean distance is:

$$d(r_1, A) = \min_{i \in 1..n} d(r_1, a_i)$$

Example: 3D Data Integration

- Range image registration
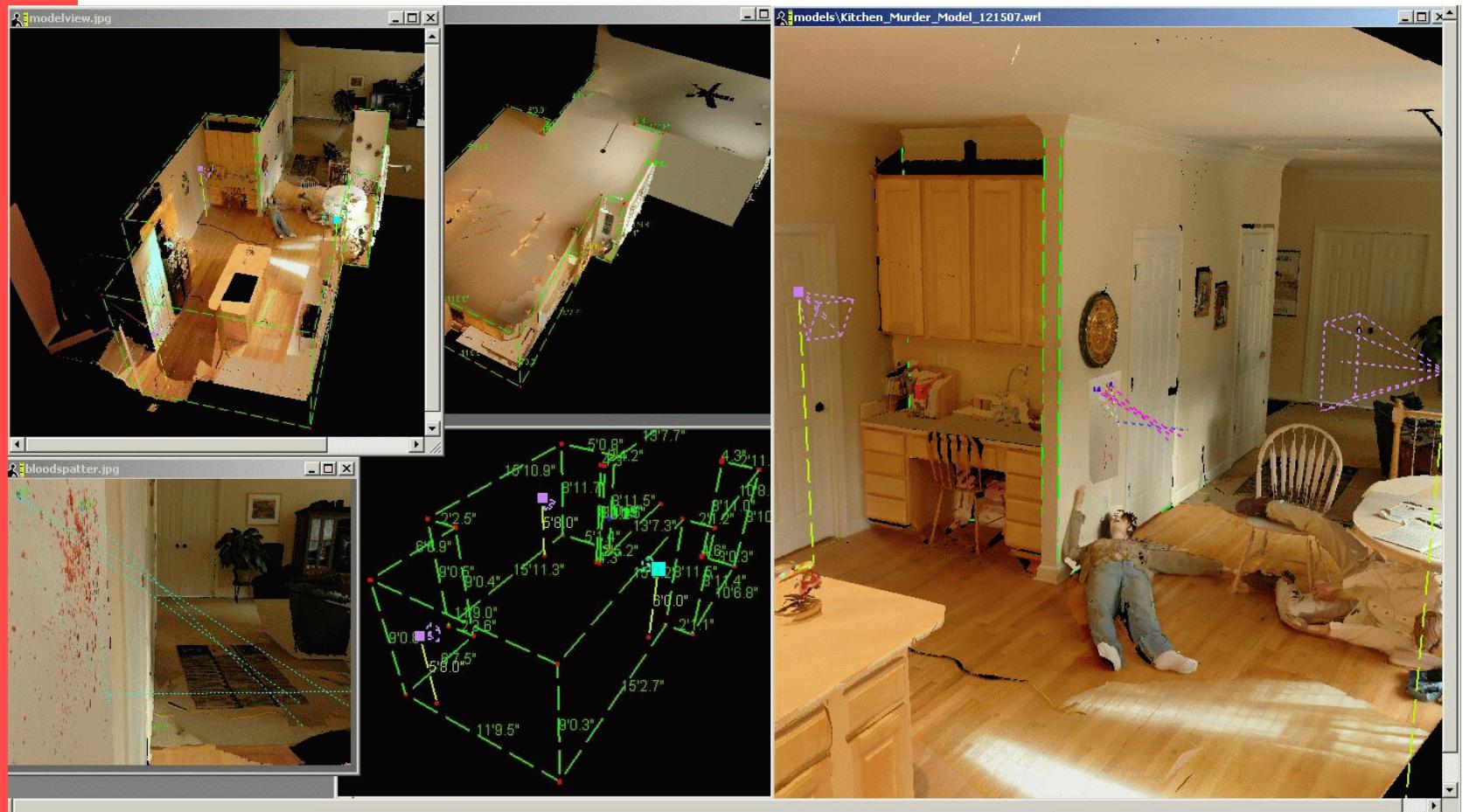
Example: 3D Data Integration

Figure 24.17. 3D Fax of a statuette of a Buddha. From left to right: photograph of the statuette; range image; integrated 3D model; model after hole filling; physical model obtained via stereolithography. Reprinted from [Curless and Levoy, 1996], Figure 10.
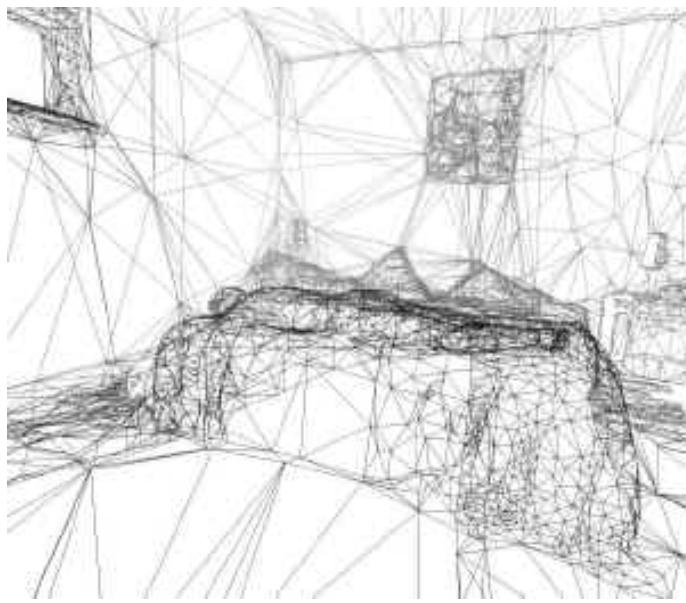
Applications:
Crime Scene, Forensic Analysis

http://www.deltasphere.com/

Applications:
Crime Scene, Forensic Analysis

http://www.deltasphere.com/
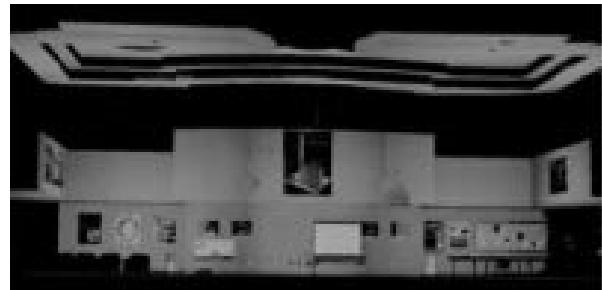
Applications

Museums, Cultural Exhibits

Archeology

Military Simulation and Training

Architecture and Construction

Range Finders: Some References

- P.J. Besl. Active, optical range imaging sensors. *Machine Vision and Applications*,1:127-152, 1988.
- R.A. Jarvis Range sensing for computer vision. In A.K. Jain and P.J. Flynn, editors, *Three-Dimensional Object Recognition Systems*, pages 17-56. Elsevier Science Publishers, 1993.
- T.G. Stahs and F.M. Wahl, "Fast and Robust Range Data Acquisition in a Low-Cost Environment", in *SPIE #1395: Close- Range Photogrammetry Meets Mach. Vis.*, Zurich, 1990, 496-503.

Conclusions

Wide range of application areas including:

- Action recognition and tracking
- Object pose recognition for robotic control
- Obstacle detection for automotive control
- Human-computer interaction
- Video surveillance
- Scene segmentation and obstacle detection
- Computer assisted surgical intervention
- Industrial applications of TOF cameras
- Automotive applications of TOF cameras
- Virtual reality applications
- Integration of range and intensity imaging sensor outputs

References

- Horn, B.K.P. 1984. Extended Gaussian images. *In Proceedings of the IEEE 72*, 12 (Dec.), pp. 1656-1678.

- Horn, B.K.P. 1986. *Robot Vision*. MIT Press, Cambridge, MA, pp. 365-399.

- Kamvysselis, M. 1997. *2D Polygon Morphing using the Extended Gaussian Image.* http://web.mit.edu/manoli/ecimorph/www/ecimorph.html

- Kang, S.B. and K. Ikeuchi. 1990. *3-D Object Pose Determination Using Complex EGI*. tech. report CMU-RI-TR-90-18, Robotics Institute, Carnegie Mellon University.