

Shape from X

Readings:

Szeliski: Chapter 11 (11.1,11.2,11.3,11.4)
Chapter 12 (12.1,12.2)

Some Slides adapted from Univ. of Washington
<http://www.cs.washington.edu/education/courses/cse576/08sp/>

Horst Bischof Computer Vision I SS11 2

Shape from X


or

How to get 3D Information

Horst Bischof Computer Vision I SS11 2

Ill-posed


- In trying to extract 3d structure from 2d images, vision is an *ill-posed* problem.



Horst Bischof Computer Vision I SS11 3

Ill-posed


- In trying to extract 3d structure from 2d images, vision is an *ill-posed* problem.



Horst Bischof Computer Vision I SS11 4

Ill-posed

- In trying to extract 3d structure from 2d images, vision is an *ill-posed* problem.



- An image isn't enough to disambiguate the many possible 3d worlds that could have produced it.

Horst Bischof Computer Vision I SS11 5

Depth Cues

- Size of known objects
- Occlusion of Objects
- Color change...far away is blueish
- Motion...Slow motion if far away
- Stereo
- Focus
- Texture
- Vergence

Horst Bischof Computer Vision I SS11 6

Shape from x

- Profile (3D Scanning)
- Shape from Shading
- Shape from Shadow
- Structured Light
- Shape from Focus
- Photometric Stereo
- Interferometry
- Stereopsis
- Shape from Motion
- Tomography (CAT)

Horst Bischof Computer Vision I SS11 7

Shape from Stereo

Horst Bischof Computer Vision I SS11 8

Why Stereo Vision?

2D images project 3D points into 2D:

- 3D Points on the same viewing line have the same 2D image:
 - 2D imaging results in depth information loss

Horst Bischof Computer Vision I SS11 9

Pinhole Camera

From Th. Svoboda

Horst Bischof Computer Vision I SS11 10

1D Pinhole (2D → 1D)

image plane

optical axis

$$\frac{x_1}{-f} = \frac{X_1}{Z_1}$$

$$x_1 = -f \frac{X_1}{Z_1}$$

Horst Bischof From Th. Svoboda Computer Vision I SS11 11

Perspective projection

camera world

Π'

$P \begin{bmatrix} x \\ y \\ z \end{bmatrix}$

$C' \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$

Cartesian coordinates:

We have, by similar triangles, that $(x, y, z) \rightarrow (f \cdot x/z, f \cdot y/z, -f)$

Ignore the third coordinate, and get

$$(x, y, z) \rightarrow \left(f \frac{x}{z}, f \frac{y}{z} \right)$$

Horst Bischof Computer Vision I SS11 12

Geometric properties of projection

- Points go to points
- Lines go to lines
- Planes go to whole image or half-planes.
- Polygons go to polygons
- Degenerate cases
 - line through focal point to point
 - plane through focal point to line

Horst Bischof Computer Vision I SS11 13

Scene Projection

- 3D lines project to 2D lines
- but the angles change, parallel lines are no more parallel.
- area ratios change, note the front and backside of the house

Horst Bischof From Th. Svoboda Computer Vision I SS11 14

Other projection models: Orthographic projection

$(x, y, z) \rightarrow (x, y)$

Horst Bischof Computer Vision I SS11 15

Other projection models: Weak perspective

- Issue
 - perspective effects, but not over the scale of individual objects
 - collect points into a group at about the same depth, then divide each point by the depth of its group
 - Adv: easy
 - Disadv: only approximate

$(x, y, z) \rightarrow \left(\frac{fx}{z_0}, \frac{fy}{z_0} \right)$

Horst Bischof Computer Vision I SS11 16

Recap: Camera parameters

A camera is described by several parameters

- Translation T of the optical center from the origin of world coords
- Rotation R of the image plane
- focal length f , principle point (x'_c, y'_c) , pixel size (s_x, s_y)
- blue parameters are called "extrinsics," red are "intrinsic"

Projection equation

$$\mathbf{x} = \begin{bmatrix} sx \\ sy \\ s \end{bmatrix} = \begin{bmatrix} * & * & * & * \\ * & * & * & * \\ * & * & * & * \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \mathbf{P}\mathbf{X}$$

- The projection matrix models the cumulative effect of all parameters
- Useful to decompose into a series of operations

$$\mathbf{P} = \begin{bmatrix} -fx & 0 & x'_c & 1 & 0 & 0 \\ 0 & -fy & y'_c & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mathbf{R}_{2x3} & \mathbf{0}_{2x3} \\ \mathbf{0}_{3x3} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{T}_{2x3} & \mathbf{T}_{3x1} \\ \mathbf{0}_{3x3} & 1 \end{bmatrix}$$

intrinsic projection rotation translation identity matrix

Horst Bischof Computer Vision I SS11 17

Camera calibration

- Given n points with known 3D coordinates X_i and known image projections x_i , estimate the camera parameters

Horst Bischof Computer Vision I SS11 18

Camera calibration

$$\lambda x_i = PX_i \quad x_i \times PX_i = 0 \quad \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \times \begin{bmatrix} P_1^T X_i \\ P_2^T X_i \\ P_3^T X_i \end{bmatrix} = 0$$

$$\begin{bmatrix} 0 & -X_i^T & y_i X_i^T \\ X_i^T & 0 & -x_i X_i^T \\ -y_i X_i^T & x_i X_i^T & 0 \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} = 0$$

Two linearly independent equations

Horst Bischof Computer Vision I SS11

Camera calibration

$$\begin{bmatrix} 0^T & X_1^T & -y_1 X_1^T \\ X_1^T & 0^T & -x_1 X_1^T \\ \dots & \dots & \dots \\ 0^T & X_n^T & -y_n X_n^T \\ X_n^T & 0^T & -x_n X_n^T \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} = 0 \quad Ap = 0$$

- P has 11 degrees of freedom (12 parameters, but scale is arbitrary)
- One 2D/3D correspondence gives us two linearly independent equations
- Homogeneous least squares
- 6 correspondences needed for a minimal solution

Horst Bischof Computer Vision I SS11

Camera calibration

$$\begin{bmatrix} 0^T & X_1^T & -y_1 X_1^T \\ X_1^T & 0^T & -x_1 X_1^T \\ \dots & \dots & \dots \\ 0^T & X_n^T & -y_n X_n^T \\ X_n^T & 0^T & -x_n X_n^T \end{bmatrix} \begin{bmatrix} P_1 \\ P_2 \\ P_3 \end{bmatrix} = 0 \quad Ap = 0$$

- Note: for coplanar points that satisfy $\Pi^T X = 0$, we will get degenerate solutions $(\Pi, 0, 0)$, $(0, \Pi, 0)$, or $(0, 0, \Pi)$


Horst Bischof Computer Vision I SS11

Camera calibration

- Once we've recovered the numerical form of the camera matrix, we still have to figure out the intrinsic and extrinsic parameters
- This is a matrix decomposition problem, not an estimation problem

Horst Bischof Computer Vision I SS11

Alternative: multi-plane calibration



Images courtesy Jean-Yves Bouquet, Intel Corp.

Advantage

- Only requires a plane
- Don't have to know positions/orientations
- Good code available online!
 - Intel's OpenCV library: <http://www.intel.com/research/mrl/research/opencv/>
 - Matlab version by Jean-Yves Bouquet: http://www.vision.caltech.edu/bouquet/calib_doc/index.html
 - Zhengyou Zhang's web site: <http://research.microsoft.com/~zhang/Calib/>

Horst Computer Vision I SS11

Stereo Vision

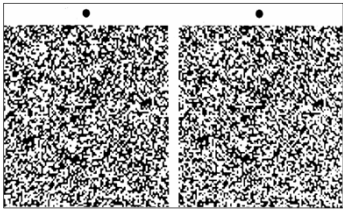
Refers to the ability of:

The ability to infer information on the 3D structure and distance of a scene from two or more images taken from different viewpoints.

Horst Bischof Computer Vision I SS11

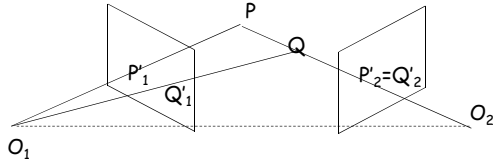
Random Dot Stereograms

1960 to prove that only Stereo information is necessary [Julez59]



Horst Bischof Computer Vision I SS11 25

Recovering Depth Information:



Depth can be recovered with two images and triangulation.

Horst Bischof Computer Vision I SS11 26

Stereo Vision Problems:

Correspondence Problem:

- Determining which pixel on the left corresponds to which pixel on the right.

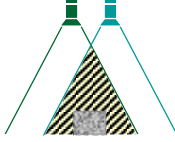
Reconstruction Problem:

- Given a number of correspondence pairs and camera geometry information, find location and 3D structure of the observed objects.

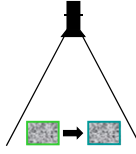
Horst Bischof Computer Vision I SS11 27

Setups

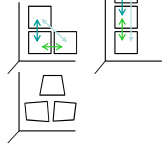
Static Stereo:



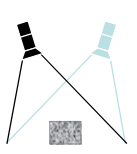
Binocular Stereo:



Moving Object



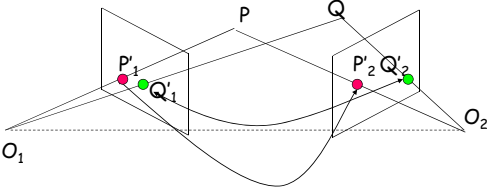
Trinocular Stereo:



Moving Camera

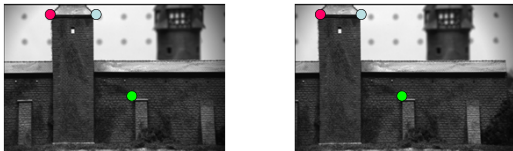
Horst Bischof Computer Vision I SS11 28

Finding Correspondences:



Horst Bischof Computer Vision I SS11 29

Finding Correspondences:



Horst Bischof Computer Vision I SS11 30

Epipolar geometry

Given an image point in one view, where is the corresponding point in the other view?

epipolar line
epipole
baseline

- A point in one view “generates” an epipolar line in the other view
- The corresponding point lies on this line

Horst Bischof Computer Vision I SS11

Epipolar line

Epipolar constraint

- Reduces correspondence problem to 1D search along an epipolar line

Horst Bischof Computer Vision I SS11

Epipolar geometry continued

Epipolar geometry is a consequence of the coplanarity of the camera centres and scene point

The camera centres, corresponding points and scene point lie in a single plane, known as the epipolar plane

Horst Bischof Computer Vision I SS11

Rectification

- Project each image onto same plane, which is parallel to the epipole
- Resample lines (and shear/stretch) to place lines in correspondence, and minimize distortion

- [Loop and Zhang, CVPR'99]

Horst Bischof Computer Vision I SS11

Rectification

(a) Original image pair overlaid with several epipolar lines.

(b) Image pair transformed by the specialized projective mapping H_L and H'_L . Note that the epipolar lines are now parallel to each other in each image.

BAD!

Horst Bischof Computer Vision I SS11

Rectification

(c) Image pair transformed by the similarity H_L and H'_L . Note that the image pair is now rectified (the epipolar lines are horizontally aligned).

(d) Final image rectification after shearing transform H_L and H'_L . Note that the image pair remains rectified, but the horizontal distortion is reduced.

GOOD!

Horst Bischof Computer Vision I SS11

3D Reconstruction

We must solve the correspondence problem first!

Horst Bischof Computer Vision I SS11 37

A simple stereo system

$$\frac{T + x_l - x_r}{Z - f} = \frac{T}{Z}$$

$$Z = f \frac{T}{x_r - x_l}$$

Disparity: $d = x_r - x_l$

$$Z = f \frac{T}{d}$$

T is the stereo baseline
d measures the difference in retinal position between corresponding points

Horst Bischof Computer Vision I SS11 38

A simple stereo system

$$Z = f \frac{T}{d}$$

•Depth is inversely proportional to disparity

Horst Bischof Computer Vision I SS11 39

Disparity

	Images		Disparity
Position 1			d_1
Position 2			d_2
Position 3			d_3

Horst Bischof Computer Vision I SS11 40

Stereo Reconstruction

Steps

- Calibrate cameras
- Rectify images
- Compute disparity
- Estimate depth

$$disparity = u - u' = \frac{baseline * f}{z}$$

Horst Bischof Computer Vision I SS11 41

Choosing the Baseline

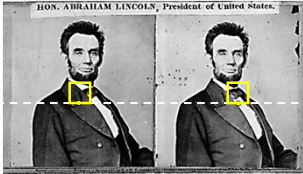
Large Baseline Small Baseline

•What's the optimal baseline?

- Too small: large depth error
- Too large: difficult search problem

Horst Bischof Computer Vision I SS11 42

Your basic stereo algorithm



HON. ABRAHAM LINCOLN, President of United States.

For each epipolar line

For each pixel in the left image

- compare with every pixel on same epipolar line in right image
- pick pixel with minimum match cost

Improvement: match **windows**

- This should look familiar...

Horst Bischof Computer Vision I SS11 43

The Correspondence Problem

Basic assumptions:

- Most scene points are visible in both images
- Corresponding image regions are similar

These assumptions hold if:

- The distance of the fixation point from the cameras is much larger than the stereo baseline: $Z \gg T$

Horst Bischof Computer Vision I SS11 44

The Correspondence Problem

Is a "search" problem:

- Given an element in the left image, search for the corresponding element in the right image.

We must choose:

- Elements to match
- A similarity measure to compare elements

Horst Bischof Computer Vision I SS11 45

Correspondence Problem

Two classes of algorithms:

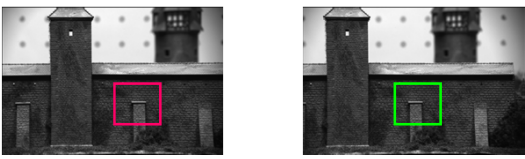
- Correlation-based algorithms
 - Produce a DENSE set of correspondences
- Feature-based algorithms
 - Produce a SPARSE set of correspondences

Horst Bischof Computer Vision I SS11 46

Correlation-based Algorithms

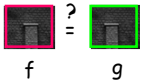
Elements to be matched:

- image WINDOWS of fixed size.



Horst Bischof Computer Vision I SS11 47

Comparing Windows:



Some possible measures:

$$\max_{[i,j] \in R} |f(i,j) - g(i,j)|$$

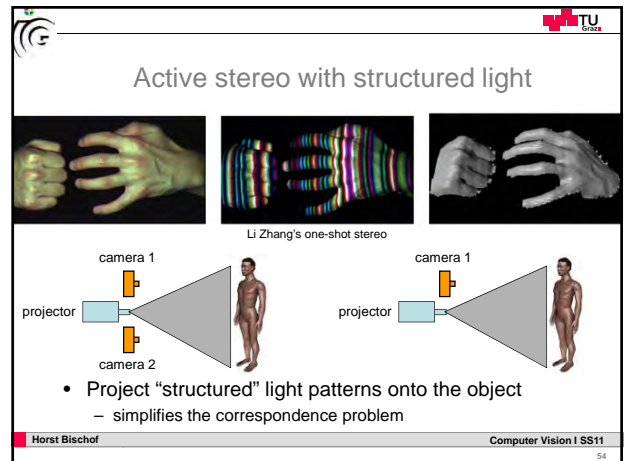
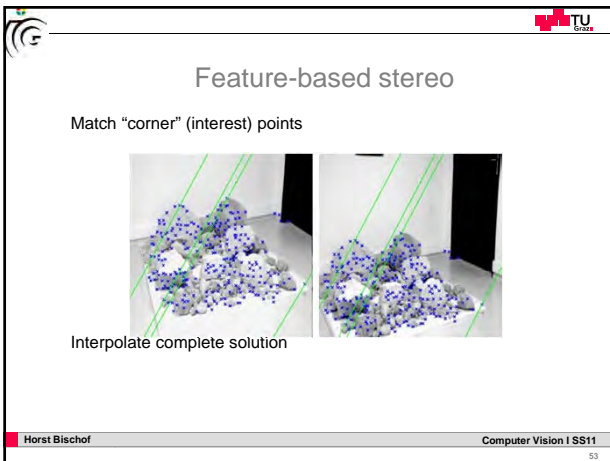
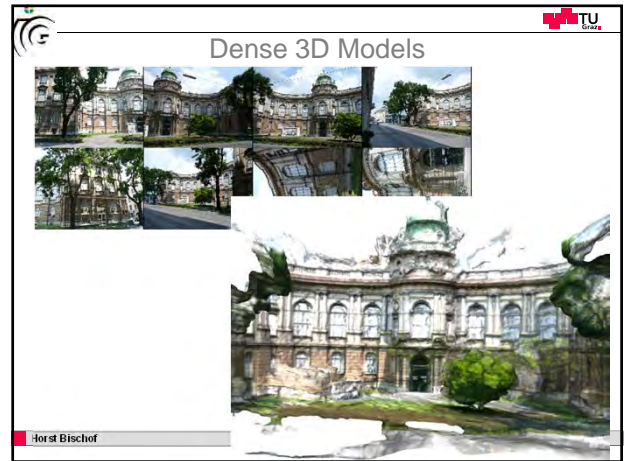
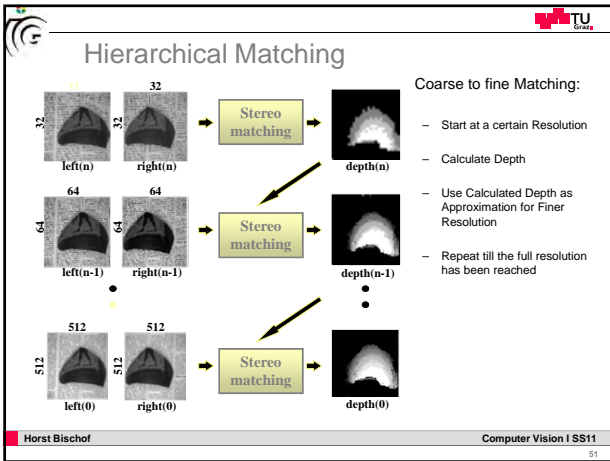
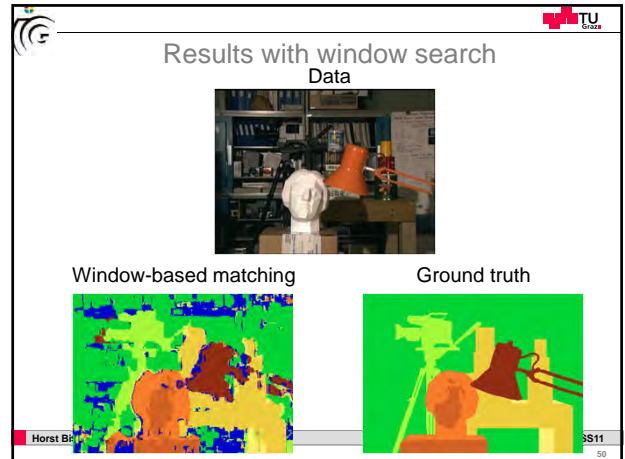
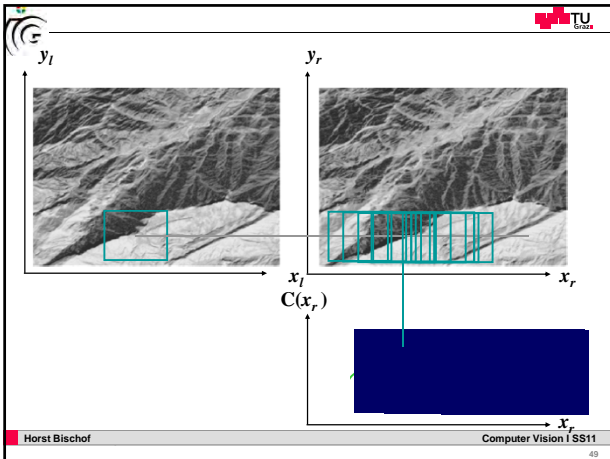
$$\sum_{[i,j] \in R} |f(i,j) - g(i,j)|$$

$$SSD = \sum_{[i,j] \in R} (f(i,j) - g(i,j))^2$$

$$C_{fg} = \sum_{[i,j] \in R} f(i,j)g(i,j)$$

} Most popular

Horst Bischof Computer Vision I SS11 48



Microsoft Kinect

3D DEPTH SENSORS
RGB CAMERA
MULTI-ARRAY MIC
MOTORIZED TILT
XBOX 360
XBOX LIVE

Horst Bischof Computer Vision I SS11

Triangulationprinzip

$$\frac{Z}{\sin \beta} = \frac{a}{\sin \alpha}$$

$$\frac{a}{\sin \alpha} = \frac{b}{\sin \gamma} \Rightarrow a = \frac{b \sin \alpha}{\sin \gamma}$$

$$\frac{Z}{\sin \beta} = \frac{b \sin \alpha}{\sin \gamma} \Rightarrow Z = \frac{b \sin \alpha \sin \beta}{\sin \gamma}$$

$$Z = \frac{b \sin \alpha \sin \beta}{\sin(180 - \alpha - \beta)}$$

Bekannte Parameter: α Winkel zwischen Basis und Lichtstrahl
 β Winkel zwischen Basis und Kameranormalen
 b Abstand zwischen Projektor und Kamera

Horst Bischof Computer Vision I SS11

Laser scanning

Object
Direction of travel
Laser sheet
CCD image plane
CCD
Cylindrical lens
Laser

Digital Michelangelo Project
<http://graphics.stanford.edu/projects/mich/>

- Optical triangulation
 - Project a single stripe of laser light
 - Scan it across the surface of the object
 - This is a very precise version of structured light scanning

Horst Bischof Computer Vision I SS11
Source: S. Seitz 57

Laser scanned models

The Digital Michelangelo Project, Levoy et al.

Horst Bischof Computer Vision I SS11
Source: S. Seitz 58

Laser scanned models

The Digital Michelangelo Project, Levoy et al.

Horst Bischof Computer Vision I SS11
Source: S. Seitz 59

Laser scanned models

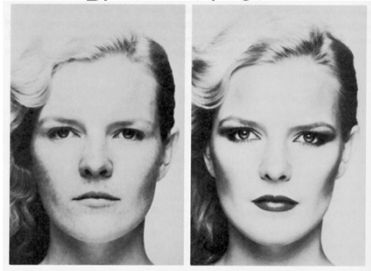
The Digital Michelangelo Project, Levoy et al.

Horst Bischof Computer Vision I SS11
Source: S. Seitz 60

Shape from x

Profilmessung eines Objektes
Shape from Shading
 Shape from Shadow
 Structured Light
 Shape from Focus
Photometric Stereo
 Interferometry
 Stereopsis
 Tomography

Horst Bischof Computer Vision I SS11 61



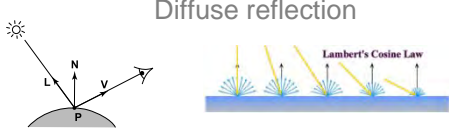
Merle Norman Cosmetics, Los Angeles

Readings

- R. Woodham, *Photometric Method for Determining Surface Orientation from Multiple Images*. Optical Engineering 19(1)139-144 (1980). [PDF](#)

Horst Bischof Computer Vision I SS11 62

Diffuse reflection



$$R_e = k_d \mathbf{N} \cdot \mathbf{L} R_i$$

image intensity of P $\rightarrow I = k_d \mathbf{N} \cdot \mathbf{L}$

Simplifying assumptions

- $I = R_e$: camera response function f is the identity function:
 - can always achieve this in practice by solving for f and applying f^{-1} to each pixel in the image
- $R_i = 1$: light source intensity is 1
 - can achieve this by dividing each pixel in the image by R_i

Horst Bischof Computer Vision I SS11 63

Shape from shading

- Suppose $k_d = 1$

$$\begin{aligned}
 I &= k_d \mathbf{N} \cdot \mathbf{L} \\
 &= \mathbf{N} \cdot \mathbf{L} \\
 &= \cos \theta_i
 \end{aligned}$$

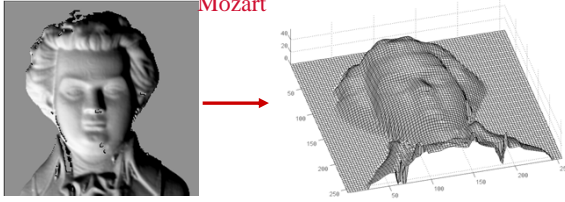
You can directly measure angle between normal and light source

- Not quite enough information to compute surface shape
- But can be if you add some additional info, for example
 - assume a few of the normals are known (e.g., along silhouette)
 - constraints on neighboring normals—"integrability"
 - smoothness
- Hard to get it to work well in practice
 - plus, how many real objects have constant albedo?

Horst Bischof Computer Vision I SS11 64

SHAPE FROM SHADING

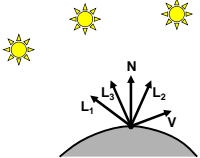
Mozart



Known reflection and illumination allow reconstruction of shape
„Inverse Rendering“

Horst Bischof Computer Vision I SS11 65

Photometric stereo



$$\begin{aligned}
 I_1 &= k_d \mathbf{N} \cdot \mathbf{L}_1 \\
 I_2 &= k_d \mathbf{N} \cdot \mathbf{L}_2 \\
 I_3 &= k_d \mathbf{N} \cdot \mathbf{L}_3
 \end{aligned}$$

Can write this as a matrix equation:

$$\begin{bmatrix} I_1 & I_2 & I_3 \end{bmatrix} = k_d \mathbf{N}^T \begin{bmatrix} \mathbf{L}_1 & \mathbf{L}_2 & \mathbf{L}_3 \end{bmatrix}$$

Horst Bischof Computer Vision I SS11 66

Solving the equations

$$\underbrace{\begin{bmatrix} I_1 & I_2 & I_3 \end{bmatrix}}_{\mathbf{I} \quad 1 \times 3} = k_d \underbrace{\mathbf{N}^T}_{\mathbf{G} \quad 1 \times 3} \underbrace{\begin{bmatrix} \mathbf{L}_1 & \mathbf{L}_2 & \mathbf{L}_3 \end{bmatrix}}_{\mathbf{L} \quad 3 \times 3}$$

$$\mathbf{G} = \mathbf{I} \mathbf{L}^{-1}$$

$$k_d = \|\mathbf{G}\|$$

$$\mathbf{N} = \frac{1}{k_d} \mathbf{G}$$

Horst Bischof Computer Vision I SS11 67

More than three lights

- Get better results by using more lights

$$\begin{bmatrix} I_1 & \dots & I_n \end{bmatrix} = k_d \mathbf{N}^T \begin{bmatrix} \mathbf{L}_1 & \dots & \mathbf{L}_n \end{bmatrix}$$

Least squares solution:

$$\mathbf{I} = \mathbf{G} \mathbf{L}$$

$$\mathbf{I} \mathbf{L}^T = \mathbf{G} \mathbf{L} \mathbf{L}^T$$

$$\mathbf{G} = (\mathbf{I} \mathbf{L}^T) (\mathbf{L} \mathbf{L}^T)^{-1}$$

Solve for N, k_d as before

What's the size of $\mathbf{L} \mathbf{L}^T$?

Horst Bischof Computer Vision I SS11 68

Results...



from Athos Georghiades
<http://cvc.yale.edu/people/Athos.html>

Horst Bischof Computer Vision I SS11 69

Limitations

- Big problems
 - doesn't work for shiny things, semi-translucent things
 - shadows, inter-reflections
- Smaller problems
 - camera and lights have to be distant
 - calibration requirements
 - measure light source directions, intensities
 - camera response function
- Newer work addresses some of these issues

Horst Bischof Computer Vision I SS11 70

Shape from x

Profilmessung eines Objektes

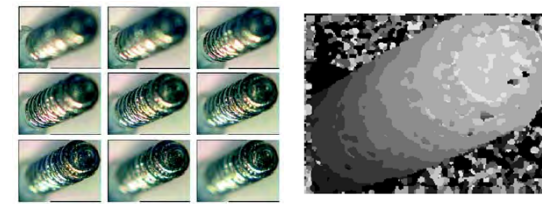
- Shape from Shading
- Shape from Shadow
- Structured Light
- Shape from Focus**
- Photometric Stereo
- Interferometry
- Stereopsis
- Tomography

Horst Bischof Computer Vision I SS11 71

Shape f. Focus/Defocus


Shape from Focus:
 Objekt bewegt entlang x/y/z, bis Oberfläche im Focus
 Wie Misst man Focus?

Shape from Defocus:
 Tiefeninformation durch Blurr (mindestens 2 Focus Einstellungen)



Horst Bischof Computer Vision I SS11 72

Form aus „Schärfe“ -
Shape-from-Focus



Microscopy

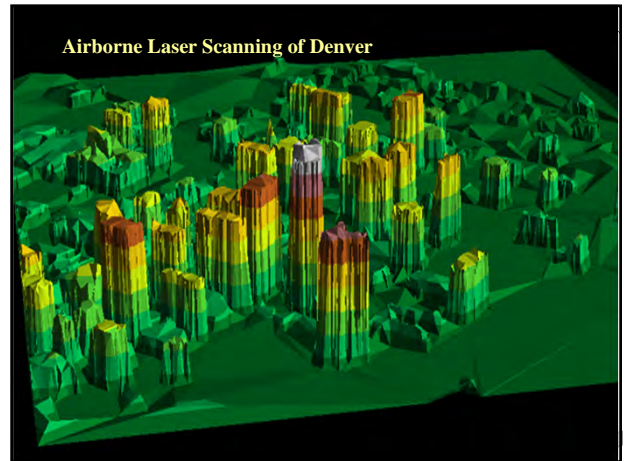
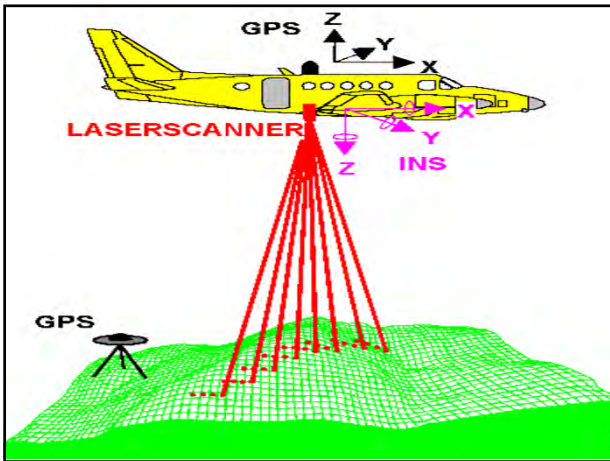
Horst Bischof Computer Vision I SS11 73

Shape from x

Profile (3D-Scanner...)


- Shape from Shading
- Shape from Shadow
- Structured Light
- Shape from Focus
- Photometric Stereo
- Interferometry
- Stereopsis
- Tomography (CAT)

Horst Bischof Computer Vision I SS11 74



Shape from Texture

Interpretation of Scene?
Depth Perception by Size of Squares
Texture contains 3D Information



Horst Bischof Computer Vision I SS11 77