General Principles of 3D Image Analysis

high-level interpretations
↕
objects
↕
scene elements
↕
image elements
↑
raw images

Extraction of 3D information from an image (sequence) is important for

- vision in general (= scene reconstruction)
- many tasks (e.g. robot grasping and navigation, traffic analysis)
- not all tasks (e.g. image retrieval, quality control, monitoring)

Recovery of 3D information is possible

- by multiple cameras (e.g. binocular stereo)
- by a monocular image sequence with motion + weak assumptions
- by a single image + strong assumptions or prior knowledge about the scene

1

Single Image 3D Analysis

Humans exploit various cues for a tentative (heuristic) depth analysis:

- size of known objects
- texture gradient
- occlusion
- colour intensities
- angle of observation
- continuity assumption
- generality assumption

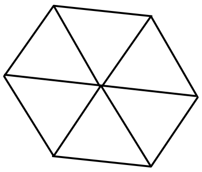
2

Generality Assumption

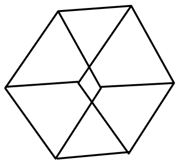
Assume that
- viewpoint
- illumination
- physical surface properties

are general, i.e. do not produce coincidental structures in the image.

<u>Example</u>: Do not interpret this figure as a 3D wireframe cube, because this view is not general.



General view:



The generality assumption is the basis for several specialized interpretation methods, e.g.
- shape from texture
- shape from shading

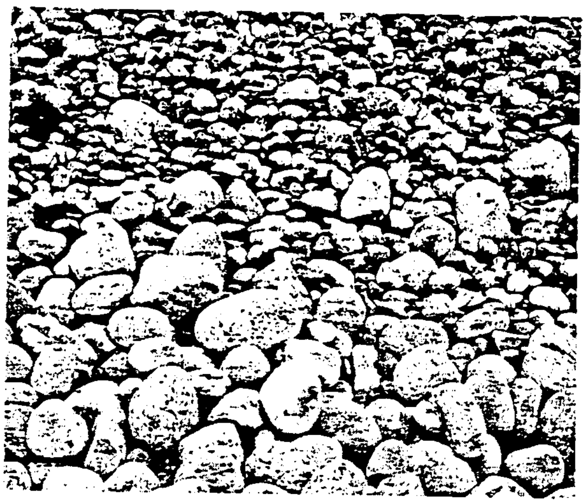
...
- "shape from X"

Texture Gradient

Assume that texture does not mimick projective effects

Interpret texture gradient as a 3D projection effect

(Witkin 81)

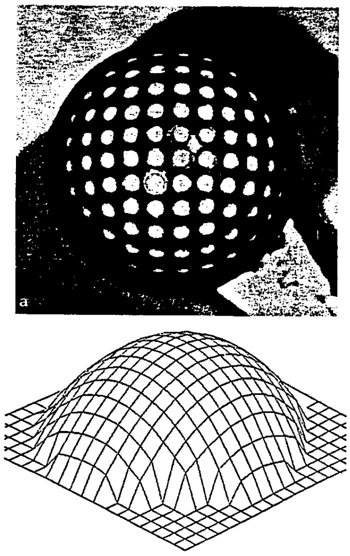


Shape from Texture

Assume
- homogeneous texture on 3D surface and
- 3D surface continuity

Reconstruct 3D shape from perspective texture variations

(Barrow and Tenenbaum 81)



5

Depth Cues from Colour Saturation

Humans interpret regions with less saturated colours as farther away.

hills in haze
nearby Graz

6

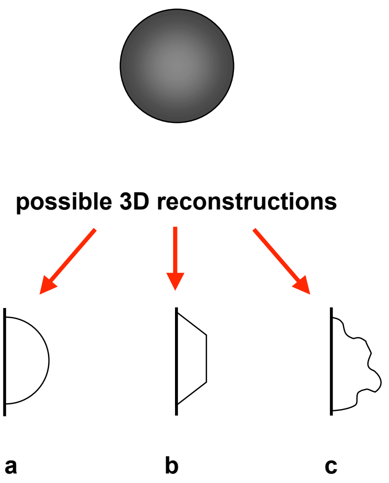
Surface Shape from Contour

Assume "non-special" illumination and surface properties

↓

3D surface shape maximizes probability of observed contours and minimizes probability of additional contours

2D image contour



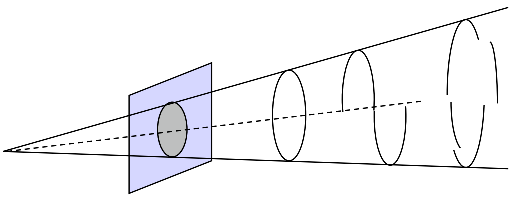
7

3D Line Shape from 2D Projections

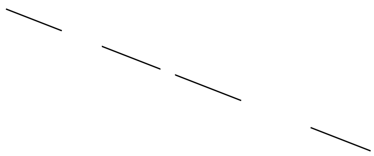
Assume that lines connected in 2D are also connected in 3D

↓

Reconstruct 3D line shape by minimizing spatial curvature and torsion



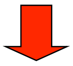
2D collinear lines are also 3D collinear



8

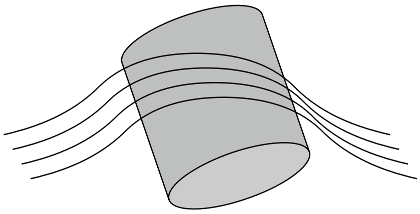
3D Shape from Multiple Lines

Assume that similar line shapes result from similar surface shapes



Parallel lines lie locally on a cylinder

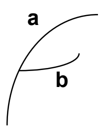
(Stevens 81)



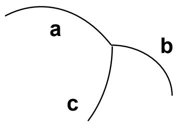
3D Junction Interpretation

rules for junctions of curved lines

(Binford 81)



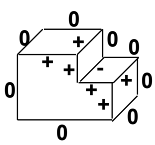
a not behind b



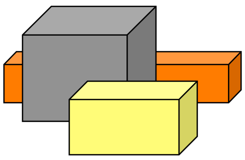
a, b and c meet

rules for blocks-world junctions

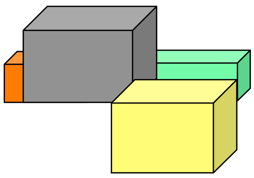
(Waltz 86)



edge labels



"general" ensemble

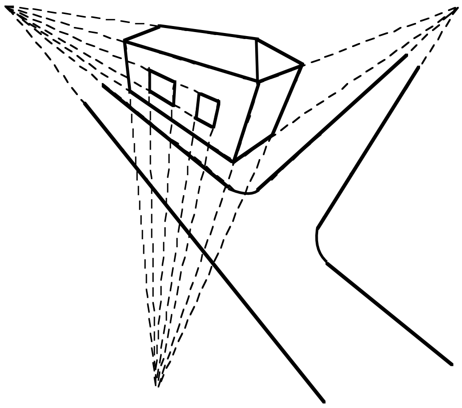


"special" ensemble

3D Line Orientation from Vanishing Points

From the laws of perspective projection:

The projections of 3D parallel straight lines intersect in a single point, the vanishing point.



Assume that more than 2 straight lines do not intersect in a single point by coincidence

If more than 2 straight lines intersect, assume that they are parallel in 3D

Obtaining 3D Shape from Shading Information

Under certain conditions, a 3D surface model may be reconstructed from the greyvalue variations of a monocular image.

From "Shape from Shading", B.K.P. Horn and M.J. Brooks (eds.), MIT Press 1989

Principle of Shape from Shading

See "Shape from Shading" (B.K.P. Horn, M.J. Brooks, eds.), MIT Press 1989

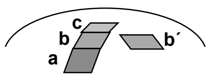
Physical surface properties, surface orientation, illumination and viewing direction determine the greyvalue of a surface patch in a sensor signal.

For a single object surface viewed in one image, greyvalue changes are mainly caused by surface orientation changes.

The reconstruction of arbitrary surface shapes is not possible because different surface orientations may give rise to identical greyvalues.

Surface shapes may be uniquely reconstructed from shading information if possible surface shapes are constrained by smoothness assumptions.

Principle of incremental procedure for surface shape reconstruction:

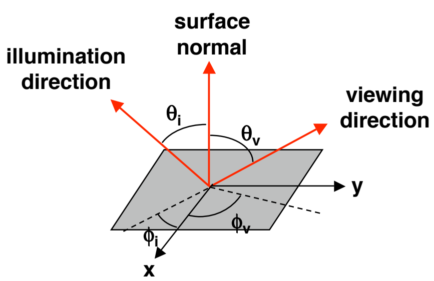


a: patch with known orientation

b, c: neighbouring patches with similar orientations

b´: radical different orientation may not be neighbour of a

Photometric Surface Properties



θ_i, θ_v polar (zenith) angles

ϕ_i, ϕ_v azimuth angles

In general, the ability of a surface to reflect light is given by the Bi-directional Reflectance Distribution Function (BRDF) r:

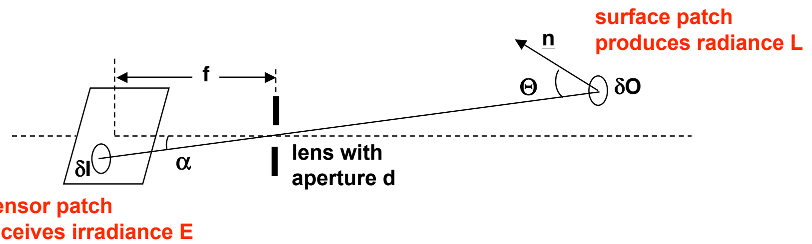
$$r(\theta_i, \phi_i; \theta_v, \phi_v) = \frac{\delta L(\theta_v, \phi_v)}{\delta E(\theta_i, \phi_i)}$$

δE = irradiance of light source received by the surface patch

δL = radiance of surface patch towards viewer

For many materials the reflectance properties are rotation invariant, in this case the BRDF depends on θ_i, θ_v, ϕ, where $\phi = \phi_i - \phi_v$.

Irradiance of Imaging Device



irradiance = light energy falling on unit patch of imaging sensor,
sensor signal is proportional to irradiance

$$E = L\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4\alpha$$

sensor signal depends on span-off angle α of surface element ("vignetting")

➡ off-center pixels in wide-angle images are darker

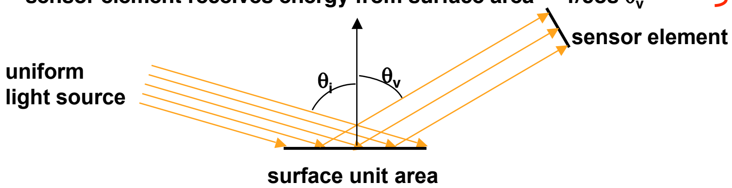
15

Lambertian Surfaces

A Lambertian surface is an ideally matte surface which looks equally bright from all viewing directions under uniform or collimated illumination. Its brightness is proportional to the cosine of the illumination angle.

- surface receives energy per unit area $\sim \cos\theta_i$
- surface reflects energy $\sim \cos\theta_v$ due to matte reflectance properties
- sensor element receives energy from surface area $\sim 1/\cos\theta_v$

} cancel out



$r_{Lambert}(\theta_i, \theta_v, \phi) = \rho(\lambda)/\pi$

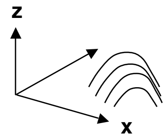
$$\rho(\lambda) = \frac{\int_\Omega L\,\partial\Omega}{E_i}$$

"albedo" = proportion of incident energy reflected back into half space Ω above surface

16

Surface Gradients

For 3D reconstruction of surfaces, it is useful to represent reflectance properties as a function of surface orientation.



$z(x, y)$	surface
$p = \delta z/\delta x$	x-component of surface gradient
$q = \delta z/\delta y$	y-component of surface gradient

$\begin{bmatrix} 1 \\ 0 \\ p \end{bmatrix}$ tangent vector in x direction

$\begin{bmatrix} 0 \\ 1 \\ q \end{bmatrix}$ tangent vector in y direction

$\begin{bmatrix} -p \\ -q \\ 1 \end{bmatrix}$ vector in surface normal direction

$\underline{n} = \frac{1}{\sqrt{1+p^2+q^2}} \begin{bmatrix} -p \\ -q \\ 1 \end{bmatrix}$ surface normal vector

If the z-axis is chosen to coincide with the viewer direction, we have

$$\cos\theta_v = \frac{1}{\sqrt{1+p^2+q^2}} \qquad \cos\theta_i = \frac{1+p_i p+q_i q}{\sqrt{1+p^2+q^2}\sqrt{1+p_i^2+q_i^2}} \qquad \cos\varphi = \frac{1}{\sqrt{1+p_i^2+q_i^2}}$$

The dependency of the BRDF on θ_i, θ_v and ϕ may be expressed in terms of p and q (with p_i and q_i for the light source direction).

17

Simplified Image Irradiance Equation

Assume that

- the object has uniform reflecting properties,
- the light sources are distant so that the irradiation is approximately constant and equally oriented,
- the viewer is distant so that the received radiance does not depend on the distance but only on the orientation towards the surface.

With these simplifications the sensor greyvalues depend only on the surface gradient components p and q.

$$E(x,y) = R(p(x,y), q(x,y)) = R(\frac{\partial z}{\partial x}, \frac{\partial z}{\partial y})$$

"Simplified Image Irradiance Equation"

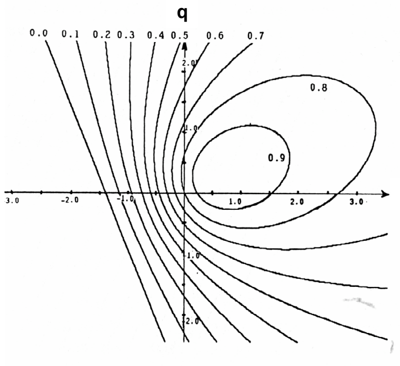
R(p, q) is the reflectance function for a particular illumination geometry. E(x, y) is the sensor greyvalue measured at (x, y). Based on this equation and a smoothness constraint, shape-from-shading methods recover surface orientations.

18

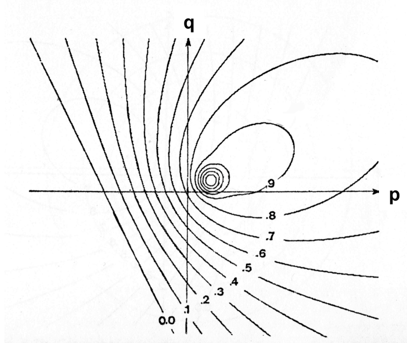
Reflectance Maps

R(p, q) may be plotted as a reflectance map with iso-brightness contours.

Reflectance map for Lambertian surface illuminated from $p_i = 0.7$ and $q_i = 0.3$



Reflectance map for matte surface with specular component



19

Characteristic Strip Method

Given a surface point (x, y, z) with known height z, orientation p and q, and second derivatives $r = z_{xx}$, $s = z_{xy} = z_{yx}$, $t = z_{yy}$, the height z+δz and orientation p+δp, q+δq in a neighbourhood x+δx, y+δy can be calculated from the image irradiance equation E(x, y) = R(p, q).

Infinitesimal change of height:

$\delta z = p\,\delta x + q\,\delta y$

Changes of p and q for a step δx, δy:

$\delta p = r\,\delta x + s\,\delta y \quad \delta q = s\,\delta x + t\,\delta y$

Differentiation of image irradiance equation w.r.t. x and y gives

$E_x = r\,R_p + s\,R_q \qquad E_y = s\,R_p + t\,R_q$

Choose step δξ in gradient direction of of the reflectance map ("characteristic strip"):

$\delta x = R_p\,\delta\xi \quad \delta y = R_q\,\delta\xi$

For this direction the image irradiance equation can be replaced by

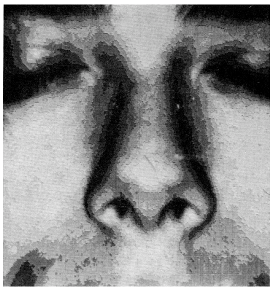
$\delta x/\delta\xi = R_p \quad \delta y/\delta\xi = R_q \quad \delta z/\delta\xi = p\,R_p + q\,R_q \quad \delta p/\delta\xi = E_x \quad \delta q/\delta\xi = E_y$

Boundary conditions and initial points may be given by
- occluding contours with surface normal perpendicular to viewing direction
- singular points with surface normal towards light source.

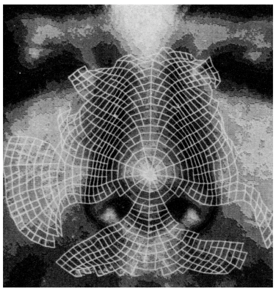
20

Recovery of the Shape of a Nose

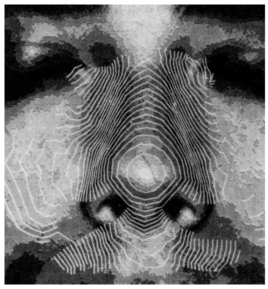
Pictures from B.K.P. Horn "Robot Vision", MIT Press,1986, p. 255



nose with crudely quantized greyvalues



superimposed characteristic curves



superimposed elevations at characteristic curves

Nose has been powdered to provide Lambertian reflectance map

21

Shape from Shading by Global Optimization

Given a monocular image and a known image irradiance equation, surface orientations are ambiguously constrained. Disambiguation may be achieved by optimizing a global smoothness criterion.

Minimize $$D(x,y) = \underbrace{\left[E(x,y) - R(p,q)\right]^2}_{\text{violation of reflectance constraint}} + \underbrace{\lambda}_{\text{Lagrange multiplier}} \underbrace{\left[\left(\nabla^2 p\right)^2 + \left(\nabla^2 q\right)^2\right]}_{\text{violation of smoothness constraint}}$$

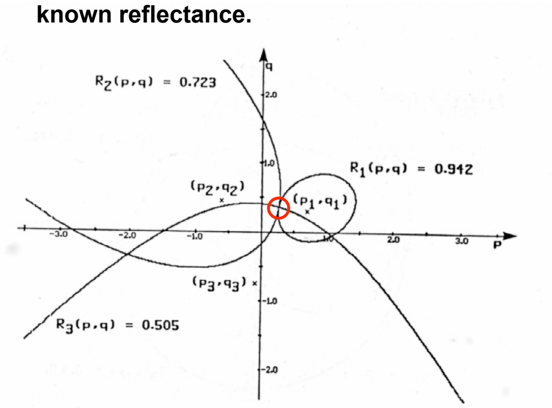
There exist standard techniques for solving this minimization problem iteratively. In general, the solution may not be unique.

Due to several uncertain assumptions (illumination, reflectance function, smoothness of surface) solutions may not be reliable.

22

Principle of Photometric Stereo

In photometric stereo, several images with different known light source orientations are used to uniquely recover 3D orientation of a surface with known reflectance.



- The reflectance maps $R_1(p, q)$, $R_2(p, q)$, $R_3(p, q)$ specify the possible surface orientations of each pixel in terms of iso-brightness contours ("isophotes").
- The intersection of the isophotes corresponding to the 3 brightness values measured for a pixel (x, y) uniquely determines the surface orientation (p(x, y), q(x, y)).

From "Shape from Shading", B.K.P. Horn and M.J. Brooks (eds.), MIT Press 1989

Analytical Solution for Photometric Stereo

For a Lambertian surface:

$$E(x, y) = R(p, q) = \rho \cos(\theta_i) = \rho\, \underline{i}^T \underline{n}$$

$\underline{i}$ = light source direction, $\underline{n}$ = surface normal, ρ = constant

If K images are taken with K different light sources $\underline{i}_k$, k = 1 ... K, there are K brightness measurements E_k for each image position (x, y):

$$E_k(x, y) = \rho\, \underline{i}_k^T \underline{n}$$

In matrix notation:

$$\underline{E}(x, y) = \rho\, L\, \underline{n} \quad \text{where } L = \begin{bmatrix} \underline{i}_1^T \\ \vdots \\ \underline{i}_K^T \end{bmatrix}$$

For K=3, L may be inverted, hence

$$\underline{n}(x,y) = \frac{L^{-1}\underline{E}(x,y)}{\left\| L^{-1}\underline{E}(x,y) \right\|}$$

In general, the pseudo-inverse must be computed:

$$\underline{n}(x,y) = \frac{\left(L^T L\right)^{-1} L^T \underline{E}(x,y)}{\left\| \left(L^T L\right)^{-1} L^T \underline{E}(x,y) \right\|}$$