Lecture 4: CS677

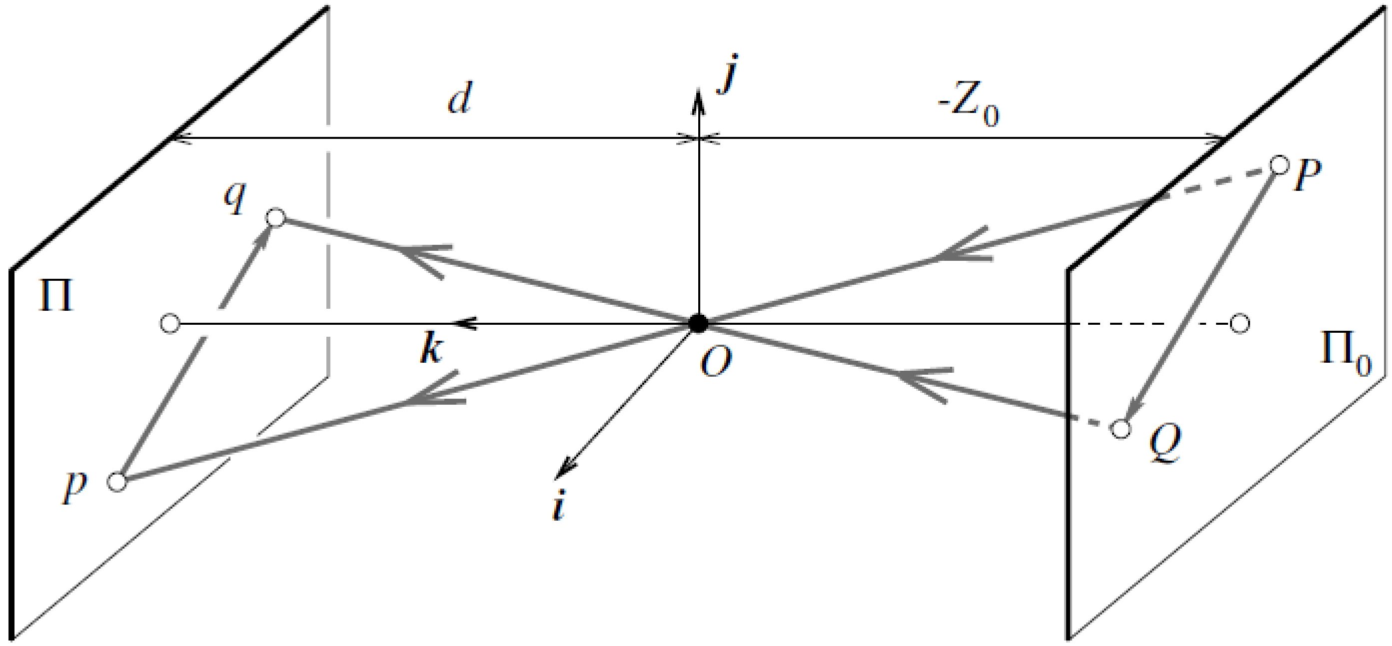
Aug 31, 2017

Review

- HW1 posted today; due September 12
- Previous class
  - Equations of projection
  - Homogeneous coordinates
  - Different coordinate systems
  - Intrinsic and extrinsic matrices
- Cloud computing: students can get a better personal account at: https://console.cloud.google.com/freetrial?_ga=2.228461851.-722665125.1503520492&page=1
- Today's objective
  - Weak perspective projection
  - Projective Geometry
  - Camera Calibration
  - Intro to Radiometry

Weak Perspective

Perspective projection but assume all points have the same z-value (object sizes small, compared to distance from camera)



$$x = -mX, \quad y = -mY, \quad \text{where} \quad m = -\frac{d}{Z_0}$$

Matrix form developed in next slide

Weak Perspective

- Equations become become simpler if we use homogeneous coordinates for P and non-homogeneous for image point p.
- Let Z_r be the distance of all points P; then, in normalized coord system, in matrix form

$$\begin{pmatrix} \hat{x} \\ \hat{y} \\ 1 \end{pmatrix} = \frac{1}{Z_r} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & Z_r \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

- Including K, R and t

$$\boldsymbol{p} = \frac{1}{Z_r} \mathcal{K} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & Z_r \end{pmatrix} \begin{pmatrix} \mathcal{R} & \boldsymbol{t} \\ \mathbf{0}^T & 1 \end{pmatrix} \boldsymbol{P}$$

Rewrite $\mathcal{K} \stackrel{\text{def}}{=} \begin{pmatrix} \alpha & -\alpha \cot\theta & x_0 \\ 0 & \dfrac{\beta}{\sin\theta} & y_0 \\ 0 & 0 & 1 \end{pmatrix}$ as:

$$\mathcal{K} = \begin{pmatrix} \mathcal{K}_2 & \boldsymbol{p}_0 \\ \mathbf{0}^T & 1 \end{pmatrix}, \quad \text{where} \quad \mathcal{K}_2 \stackrel{\text{def}}{=} \begin{pmatrix} \alpha & -\alpha \cot\theta \\ 0 & \dfrac{\beta}{\sin\theta} \end{pmatrix} \quad \text{and} \quad \boldsymbol{p}_0 \stackrel{\text{def}}{=} \begin{pmatrix} x_0 \\ y_0 \end{pmatrix}$$

Weak Perspective (Continued)

- Rewrite weak perspective projection equation as:

$$p = \mathcal{M}P, \quad \text{where} \quad \mathcal{M} = \begin{pmatrix} \mathcal{A} & b \end{pmatrix}$$

p is a *non-homogeneous* coordinate vector here; M is 2x4

$$\mathcal{A} = \frac{1}{Z_r}\mathcal{K}_2\mathcal{R}_2 \quad \text{and} \quad b = \frac{1}{Z_r}\mathcal{K}_2 t_2 + p_0;$$

R_2 is the sub-matrix of R consisting of the first two rows; t_2 contains the first two terms of vector t.

Note t_3 does not appear in the projection equation.

- With further manipulation, we can derive:

$$\mathcal{M} = \frac{1}{Z_r}\begin{pmatrix} k & s \\ 0 & 1 \end{pmatrix}\begin{pmatrix} \mathcal{R}_2 & t_2 \end{pmatrix}$$

k denotes aspect ratio,
s denotes skew

- Some restrictions on A matrix for it to be a weak perspective projection matrix; details omitted for now.

Homogeneous Coordinates

- *Linearize* the entire image projection process, including various coordinate transformations
  - Allows us to work with a single 3 x 4 matrix transformation (for perspective projection) with at most 12 (11) parameters.
- Many geometrical entities (*e.g*. points, planes, *conics*) can be represented compactly
- Many relations between these entities and their mappings to image plane can also be represented compactly
- Allows points at infinity to be handled *homogeneously* with the other points
  - Consider $[a\ b\ c\ 0]^T$, we can still project it to a finite point in the image
- Studied as part of the field of *Projective Geometry*

Homogeneous Coords in 2-D

- Point X: normal coords (x,y); homogeneous $(x, y, 1)^T$
- Line l : ax + by + c = 0, homogeneous: $(a, b, c)^T$
- Point X is on line l *iff* $X^T l = 0$
- Lines l and l' intersect in $X = l \times l'$

$$\mathbf{u}\times\mathbf{v} = \begin{vmatrix} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \\ u_x & u_y & u_z \\ v_x & v_y & v_z \end{vmatrix},$$

$$= \hat{\mathbf{x}}\left(u_y v_z - u_z v_y\right) - \hat{\mathbf{y}}\left(u_x v_z - u_z v_x\right) + \hat{\mathbf{z}}\left(u_x v_y - u_y v_x\right)$$

- Intersection of parallel lines
  - Consider $(-1, 0, 1)^T$ $(-1, 0, 2)^T$; $X = (0, 1, 0)^T$, a *point at* ∞.
  - $(x_1, x_2, 0)^T$ set of all points at ∞, they all lie on line $(0, 0, 1)^T$, called the *line at infinity*. A point on this line corresponds to a *direction*.

Homogeneous Coords in 3-D

- Point X: normal coords $(x,y,z)^T$; homogeneous $(x, y, z,1)^T$
- Plane Π: ax + by + cz + d =0 is defined by $(a, b, c, d)^T$
- Point X is on plane *iff* $\Pi^T \boldsymbol{X} = 0$
- 3 points define a plane $(\boldsymbol{X}_1^T, \mathbf{X}_2^T, \mathbf{X}_3^T)^T \Pi = 0$
- 3 planes define a point $(\Pi_1^T, \Pi_2^T, \Pi_3^T)^T \boldsymbol{X} = 0$
- Plane at infinity: $\Pi_\infty = (0, 0, 0, 1)^T$.
  – Points on this plane, $(x_1, x_2, x_3, 0)^T$ or $(\boldsymbol{d}^T, 0)^T$ represent directions (of parallel lines)
  – Line of direction $\boldsymbol{d}$ intersects Π_∞ in $(\boldsymbol{d}^T, 0)^T$. Vanishing point is simply image of this point (= K $\boldsymbol{d}$)
  – Parallel planes intersect on a line in Π_∞. Suppose **n** is the direction of the normal to these planes. It can be shown that the vanishing line is given by $l = K^{-T} \mathbf{n}$.
- Lines in 3-D, harder to represent (more on this coming)
- Ref: Hartley- Zisserman book

Conics and Quadrics

- Consider a 2-D conic: $ax^2 + bxy + cy^2 + dx + ey + f = 0$
- Let $X = [x, y, 1]^T$ be a point on the conic, expressed in homogeneous coordinates, then equation of conic can be written as , $\mathbf{X}^T\mathbf{C}\mathbf{X} = 0$, where $\mathbf{C} = \begin{vmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{vmatrix}$

- Sphere or quadric surface represented as a simple vector equation (see eq. 2.3, FP book, first edition)

$$a_{200}x^2 + a_{110}xy + a_{020}y^2 + a_{011}yz + a_{002}z^2 + a_{101}xz + a_{100}x + a_{010}y + a_{001}z + a_{000} = 0.$$

$$P^T Q P = 0, \quad \text{where} \quad Q = \begin{pmatrix} a_{200} & \frac{1}{2}a_{110} & \frac{1}{2}a_{101} & \frac{1}{2}a_{100} \\ \frac{1}{2}a_{110} & a_{020} & \frac{1}{2}a_{011} & \frac{1}{2}a_{010} \\ \frac{1}{2}a_{101} & \frac{1}{2}a_{011} & a_{002} & \frac{1}{2}a_{001} \\ \frac{1}{2}a_{100} & \frac{1}{2}a_{010} & \frac{1}{2}a_{001} & a_{000} \end{pmatrix}.$$

3-D Lines in Homogeneous Coordinates

- Line Representation: more complex than for points and planes
- A line can be defined by two points on the line or the intersection of two planes
  - Say representation of line is concatenation of coordinates of two points
  - We can't project this new vector by using matrix M; if we project each point separately, we get two points in the image plane and then need to construct a line (which has only 3 parameters)
- Plücker coordinates define a neat 6 parameter representation of a line that can be operated on in similar ways as points and planes
- We will skip details of Plücker representation
  - For ref, see https://en.wikipedia.org/wiki/Pl%C3%BCcker_coordinates
    - Hartley-Zisserman book: Multi-view geometry
  - Details not necessary for 677 course

Inverse Problem

- In graphics, object point(s) ***P*** and camera transformation matrix, *M*, are known, task is to compute the image ***p***
  - Matrix multiplication solves the problem
- Inverse problem
  - **p** is given, estimate **P**
    - *M* may or may not be known
  - Even if *M* is given, **P** is still not unique as *M* is not invertible; however, we can put some constraints on **P** (must lie on a specific line)
  - Given points in two images (say $\mathbf{p}_1$ and $\mathbf{p}_2$) and M_1 and M_2, we can solve for **P**
    - Stereo processing, requires finding corresponding points, $\mathbf{p}_1$ and $\mathbf{p}_2$
- Camera calibration problem
  - Given **p** and **P**, solve for *M*

Camera Calibration

- Find camera transform matrix ***M***
- Use a calibration object (cube, chessboard etc.)
  - 3-D positions of the points (set of $\mathbf{P}_i$) on the object are known (in an object centered coordinate system)
- Find correspondences between sets of image points ($\mathbf{p}_i$ = (x_i, y_i) and 3-D object points $\mathbf{P}_i$ (*manually* if necessary)
- Each correspondence provides two equations (relating $\mathbf{p}_i$ to $\mathbf{P}_i$ in terms of parameters of M).
- Given six matches, we can find an exact solution (ignoring degenerate cases)
- If we have more points, we can find a least mean squared error solution

Camera Calibration

- Given correspondences between sets of image points ($\mathbf{p}_i = (x_i, y_i)$) and 3-D object points $\mathbf{P}_i$

$$x_i = \frac{\boldsymbol{m}_1(\boldsymbol{\xi}) \cdot \boldsymbol{P}_i}{\boldsymbol{m}_3(\boldsymbol{\xi}) \cdot \boldsymbol{P}_i},$$
$$y_i = \frac{\boldsymbol{m}_2(\boldsymbol{\xi}) \cdot \boldsymbol{P}_i}{\boldsymbol{m}_3(\boldsymbol{\xi}) \cdot \boldsymbol{P}_i},$$

(ξ is the set of intrinsic and extrinsic parameters)

- Expanding and dropping ξ for simplicity:

$$\begin{aligned}
(\boldsymbol{m}_1 - x_i \boldsymbol{m}_3) \cdot \boldsymbol{P}_i &= \boldsymbol{P}_i^T \boldsymbol{m}_1 + \boldsymbol{0}^T \boldsymbol{m}_2 - x_i \boldsymbol{P}_i^T \boldsymbol{m}_3 &= 0, \\
(\boldsymbol{m}_2 - y_i \boldsymbol{m}_3) \cdot \boldsymbol{P}_i &= \boldsymbol{0}^T \boldsymbol{m}_1 + \boldsymbol{P}_i^T \boldsymbol{m}_2 - y_i \boldsymbol{P}_i^T \boldsymbol{m}_3 &= 0.
\end{aligned}$$

- Solve for m in $P\boldsymbol{m} = 0$ where

$$\mathcal{P} \stackrel{\text{def}}{=} \begin{pmatrix} \boldsymbol{P}_1^T & \boldsymbol{0}^T & -x_1 \boldsymbol{P}_1^T \\ \boldsymbol{0}^T & \boldsymbol{P}_1^T & -y_1 \boldsymbol{P}_1^T \\ \dots & \dots & \dots \\ \boldsymbol{P}_n^T & \boldsymbol{0}^T & -x_n \boldsymbol{P}_n^T \\ \boldsymbol{0}^T & \boldsymbol{P}_n^T & -y_n \boldsymbol{P}_n^T \end{pmatrix} \quad \text{and} \quad \boldsymbol{m} \stackrel{\text{def}}{=} \begin{pmatrix} \boldsymbol{m}_1 \\ \boldsymbol{m}_2 \\ \boldsymbol{m}_3 \end{pmatrix}$$

- Degenerate configurations: All points in same plane, 3 points in a line…

Computing Intrinsic and Extrinsic Parameters

- Matrix M combines intrinsic and extrinsic parameters in complex ways; seems difficult to separate them but this can be done by careful application of algebra (we omit details, see FP book)
- Non-linear calibration
  - Elements of M are not independent, hence linear solution is not completely accurate
  - Number of unknowns may be smaller than 12
    - *e.g.* skew is typically zero, aspect ratio is known
  - Three or four points may suffice (or two or three lines)
- Non-linear methods may be expensive and converge to locally optimum solutions
  - Many such solutions exist in the literature but will not be discussed further in our class (also not covered in FP book).

Radial Distortion

- Real lenses often have *radial distortion*, particularly wide-angle lenses
  - Scaling is proportional to distance from the center
  - Projections of straight lines near the edges of image appear curved
  - Radial distortion can be accounted for, details may be found in section 1.3.2 but not required for our class

Self-calibration

- What can we infer about the camera without a calibration object?
  - Typical images don't come with calibration objects or transformation matrices?
- Inherent ambiguities
  - Primarily that of scale
- Use of vanishing points
  - "Vertical" vanishing point provides "tilt" angle
  - Three orthogonal vanishing points define a triangle
    - Orthocenter of this triangle gives the principal point
    - Horizon line gives "roll" angle directly
    - Size of triangle provides focal length
    - Height and scale can be estimated if an object of known size can be seen in the image
- Details omitted; self-calibration from single camera not included in exams.

Example

A street scene

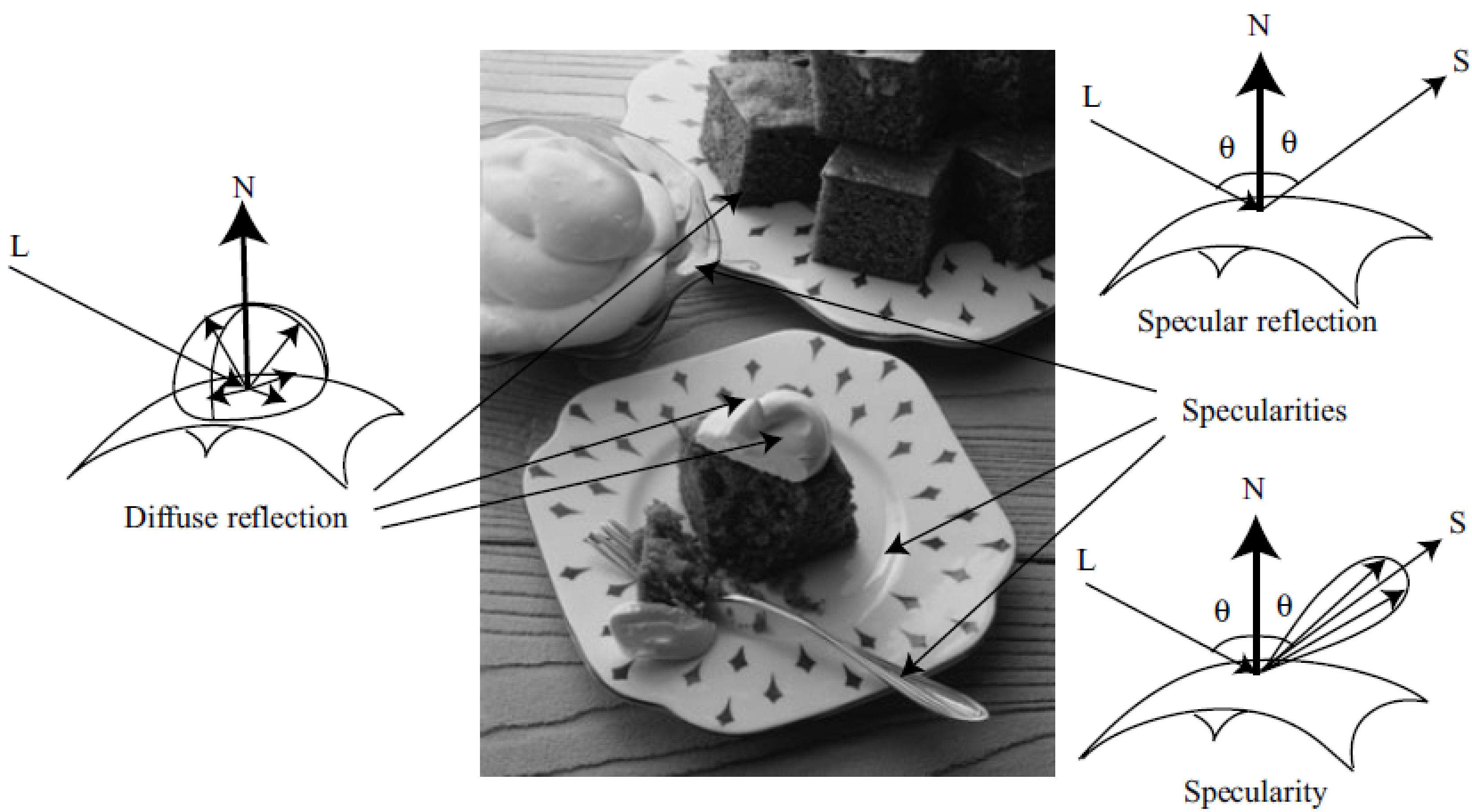
- Where are the vanishing points?
- Where is the horizon?

RADIOMETRY

- Goal: Determine the *brightness* of an image point
- Depends on:
  - Intensity and direction of incoming light
  - Surface reflection properties
    - May be direction dependent
      - Specular (mirror), all light reflected in one direction
      - Lambertian (diffused), looks equally bright from all directions
      - Most objects can be modeled as a combination of the two
  - Sensor response
- We will focus on a *local* shading model only; intensity variations essentially depend on the local surface orientation
- Image is digitized (sampled at discrete points) and quantized (values are integers)

Fig 2.2: Diffuse and Specular Reflections

- Fig 2.33 RS book



Distant Point Light Source

- Some parts of scene don't get any direct light: "cast shadows"
- Inter-reflection
- Lambertian + Specular model
- Let I(***x***) be *intensity* of point ***x*** in the image;

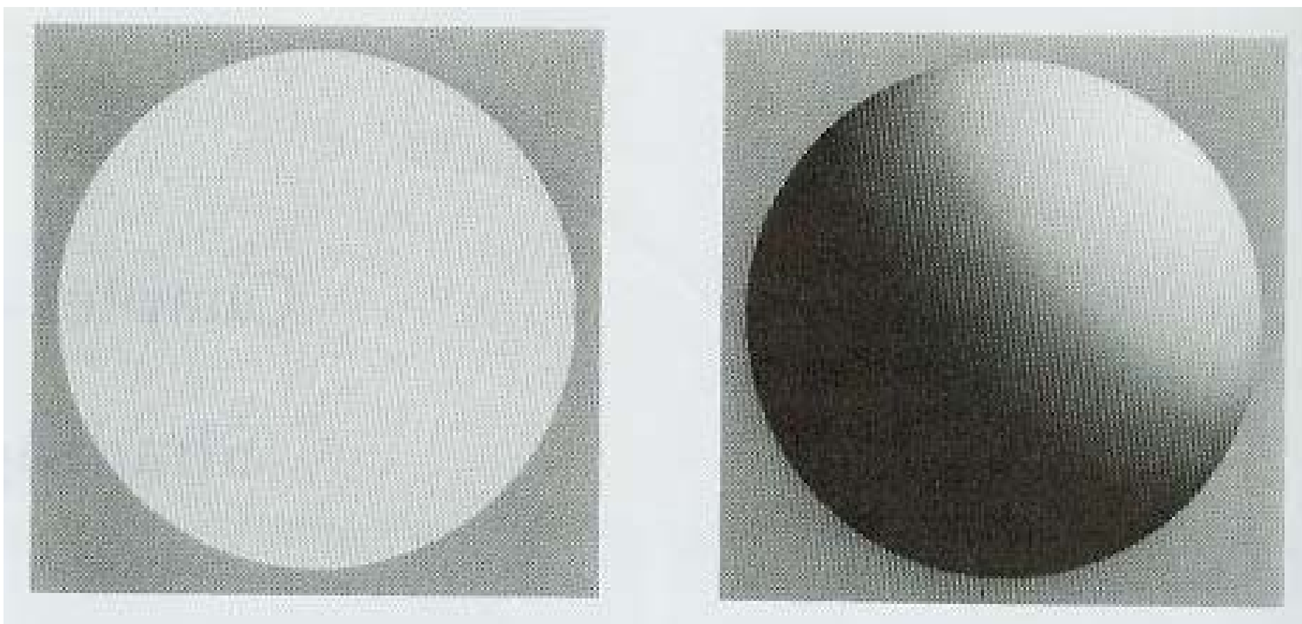
  N(*x*) be surface normal at ***x***;

  S(***x***) be the direction vector towards source (source is at ∞);

  – $I(\boldsymbol{x}) = \rho(\boldsymbol{x})\,(N(\boldsymbol{x}).\,S(\boldsymbol{x})) + \rho(\boldsymbol{x})\,A + M$
  – $\rho(\boldsymbol{x})$ includes effects of surface *albedo*, sensor response, and illumination intensity
  – Second term accounts for ambient light, last for specular term
- Area source: sum of points sources; complex in general

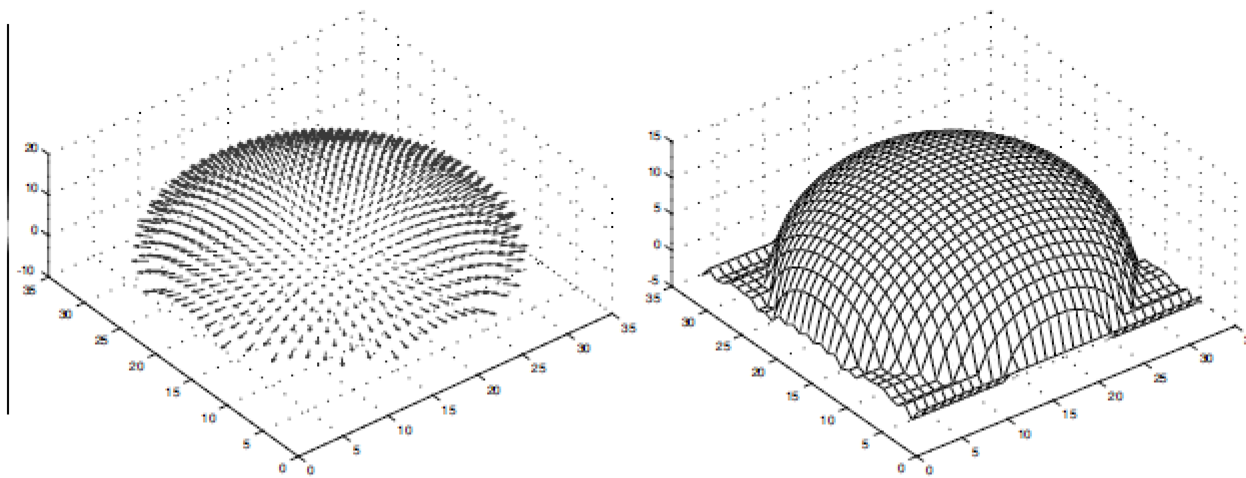
Some Examples

- Examples
  - Constant for a plane, source far away
  - Disk vs sphere (same geometric shape), (from Nalwa fig 5.2)



Normal Field vs Height Field

- Surface normals can be derived from height (depth) by computing partial derivatives
- Height (depth) can be derived from surface normals by integration



Next Class

- FP: Chapter 2, section 2.2.1, Chapter 3, sections 3.1, 3.2 and 3.3