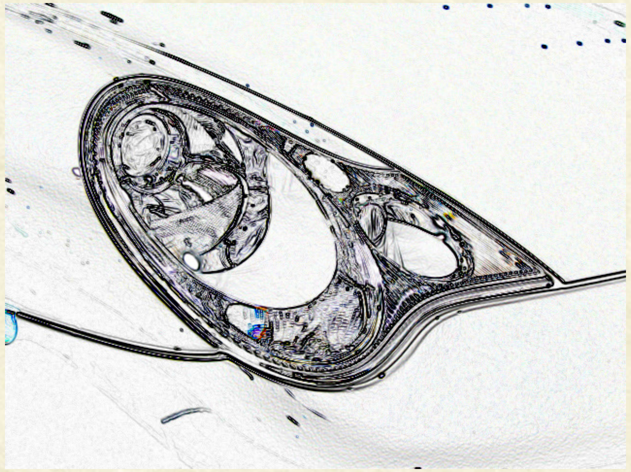
CS7.505: Computer Vision

Spring 2024: Active Illumination Methods



Anoop M. Namboodiri, Makarand Tapaswi

CVIT,

IIIT Hyderabad

Active vs. Passive Sensors

Passive Sensor

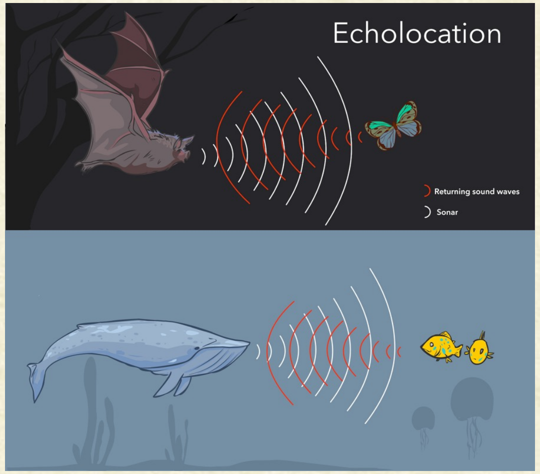
- Record the interaction between the scene and existing energy sources
- Does not change the world (other than sensor presence)
- Advantages:
  - Safety
  - Energy efficiency
  - Stealth
  - Multi-spectral imaging

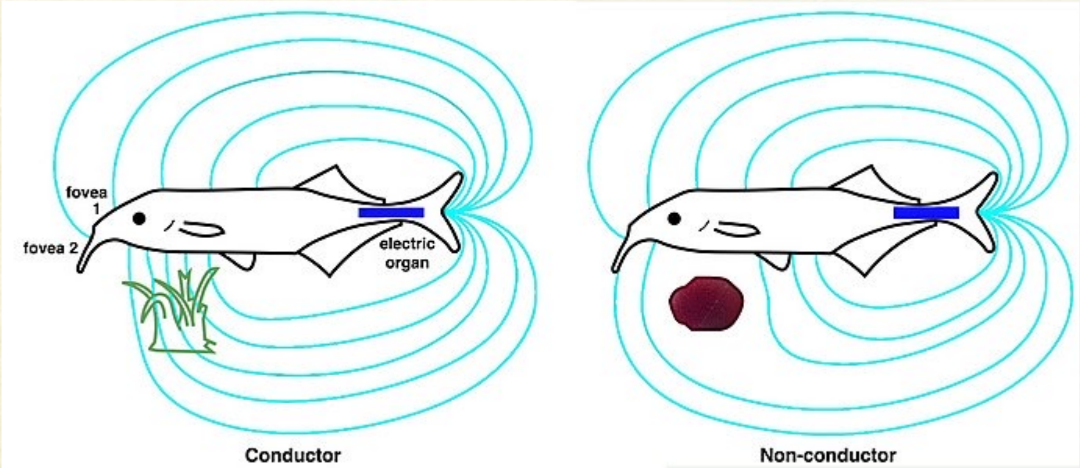
Active Sensors

- Send energy/probe out into the world and record its interaction with the world
- Modifies the world through introduction of probe
- Advantages:
  - Works in 'dark'
  - Accuracy
  - Reliability / Predictability
  - See through occlusions
  - Sense several scene characteristics

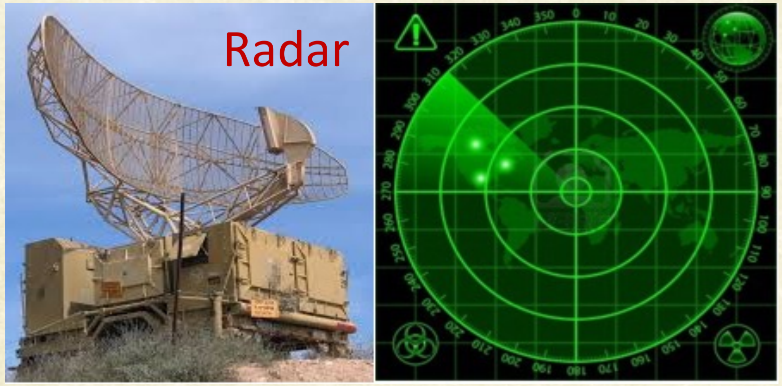
Active Sensors in Nature

- Echolocation
  - Bats, Whales
- Electrolocation
  - Electric eel
- Mechanical
  - Antennae
  - Whiskers
- Chemical
  - Slime Mold
- Bioluminesence
  - Firefly



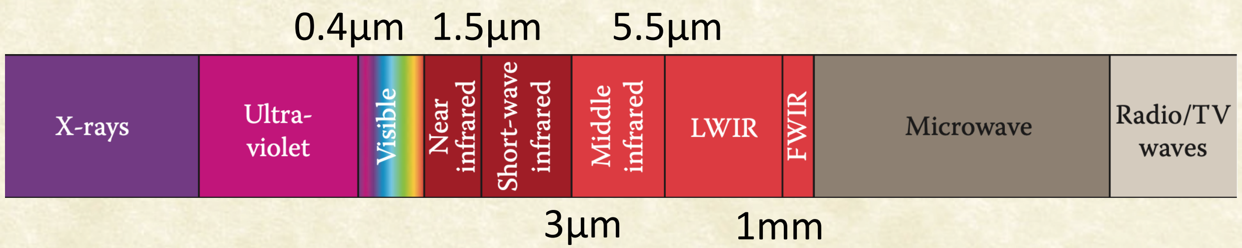


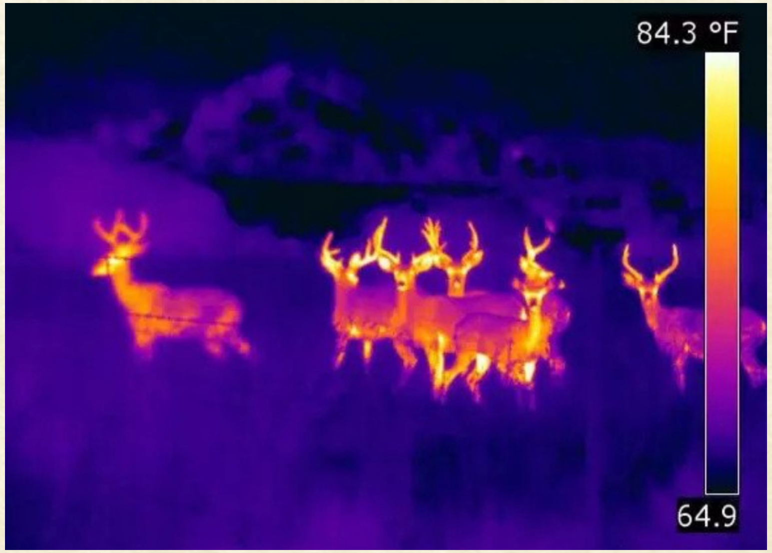
Active Sensing Devices



- Reflection
- Transmission
- Emission
- Induction
- Capacitance

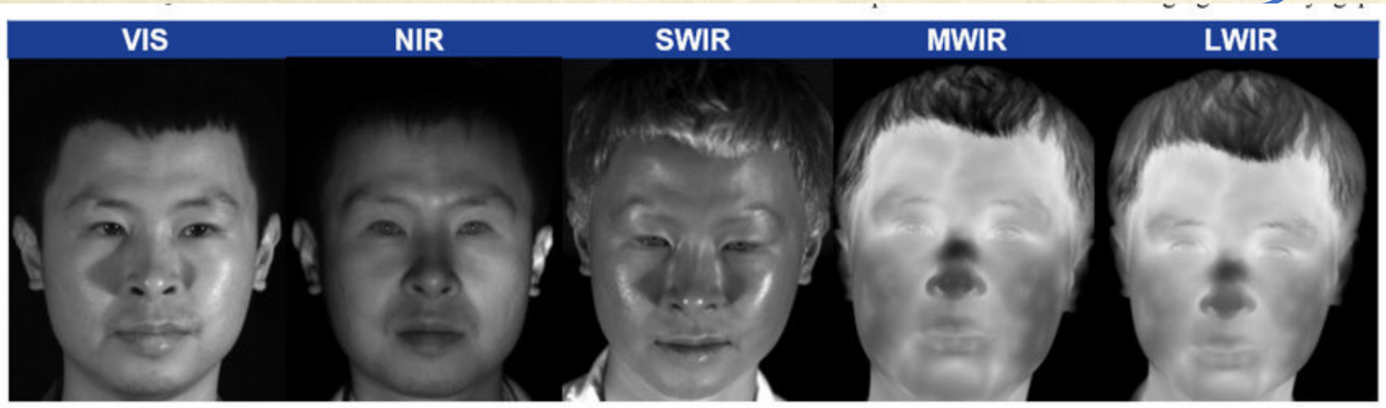
Infrared Cameras

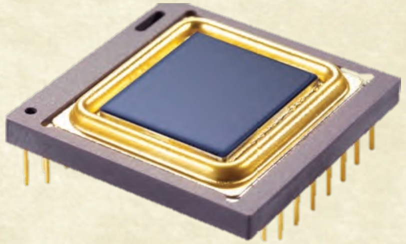




- Near Infrared (700 nm – 1.4 µm):
  - Most camera sensors are sensitive to this range (sans IR-cut filter)
- Short-Wave Infrared (1.4 – 3 µm):
- Mid-Wave Infrared (3 – 8 µm):
- Long-Wave Infrared (8 – 15 µm):
- Far-Infrared (15 – 1000 µm):

Thermal Camera





ULIS 640 × 480 17 µm

Photometric Stereo

Based on the Lectures:

First Principles of Computer Vision

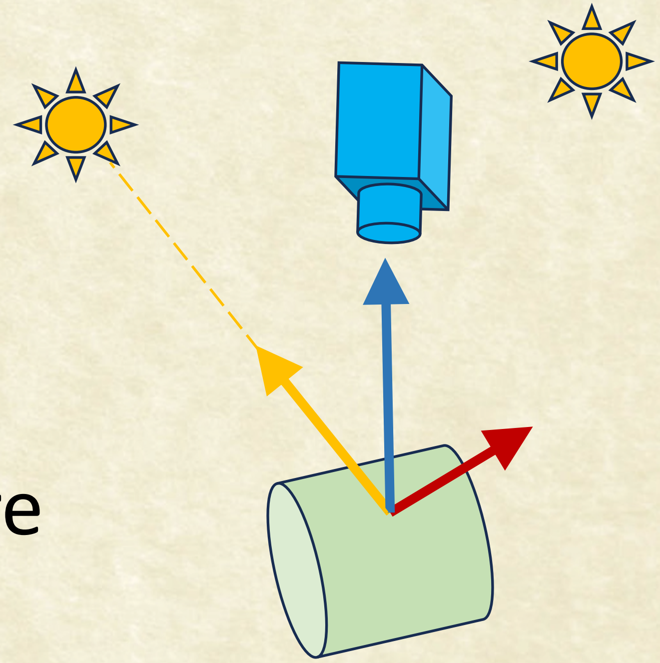
By Shree K Nayar

Photometric Stereo

- A method for recovering 3D shape from images measured using multiple known light sources.
- The image intensity, I, at a point is:

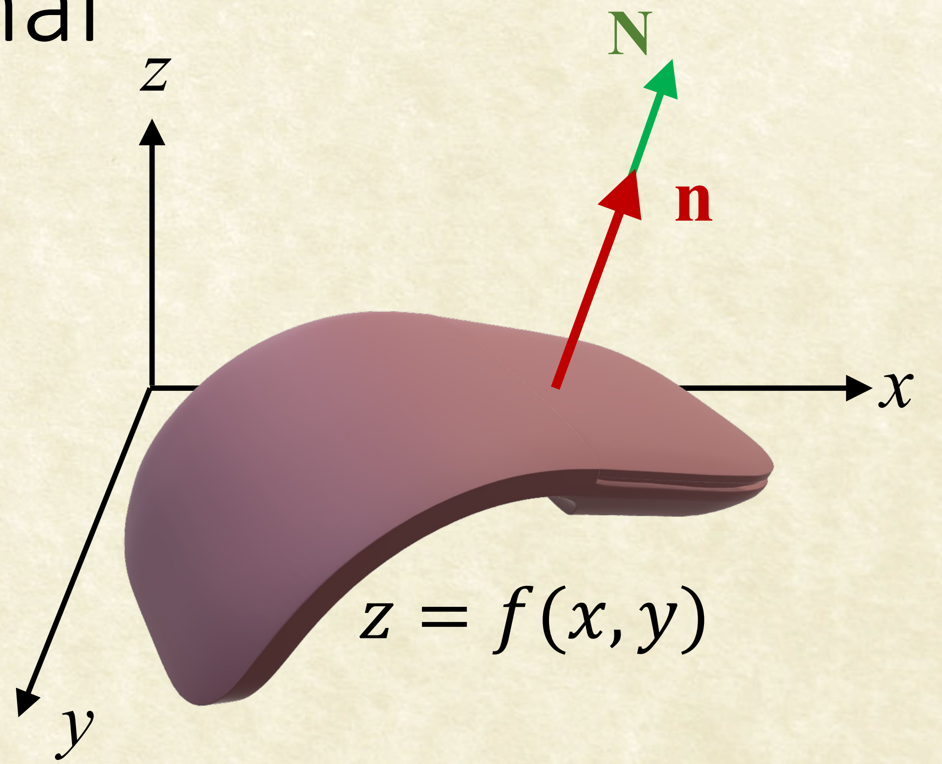
$$I = f(Source, Normal, Reflectance)$$

- Image intensities are measured
- Source directions, distances and intensities are known
- Reflectance is characterized by a known (often simplified) BRDF function



Surface Gradient and Normal

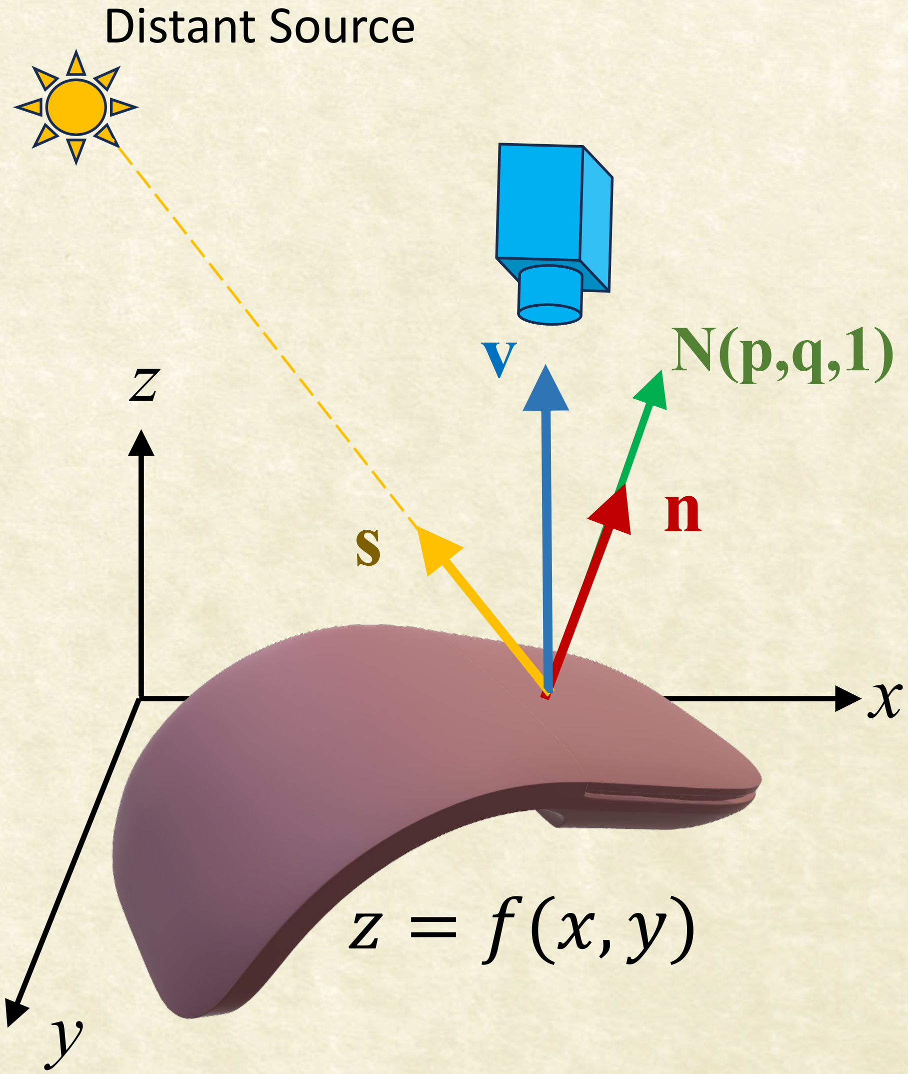
- Gradient: $\left(-\frac{dz}{dx}, -\frac{dz}{dy}\right) = (p, q)$
- Normal: $\left(-\frac{dz}{dx}, -\frac{dz}{dy}, 1\right) = (p, q, 1)$
- The surface normal, **N**, can be normalized to unit length to get: **n**.
- The surface normal is represented by two parameters: (p, q)
- Source direction can also be expressed in (p, q) space.



Reflectance Map: $\boldsymbol{R(p,q)}$

- The Reflectance Map, $\boldsymbol{R(p,q)}$, gives the image intensity at any point (x,y), computed as a function of the source direction, surface gradient (p,q), and surface reflectance (BRDF).
- For a Lambertian (matte) surface, the image intensity is independent of viewing direction.

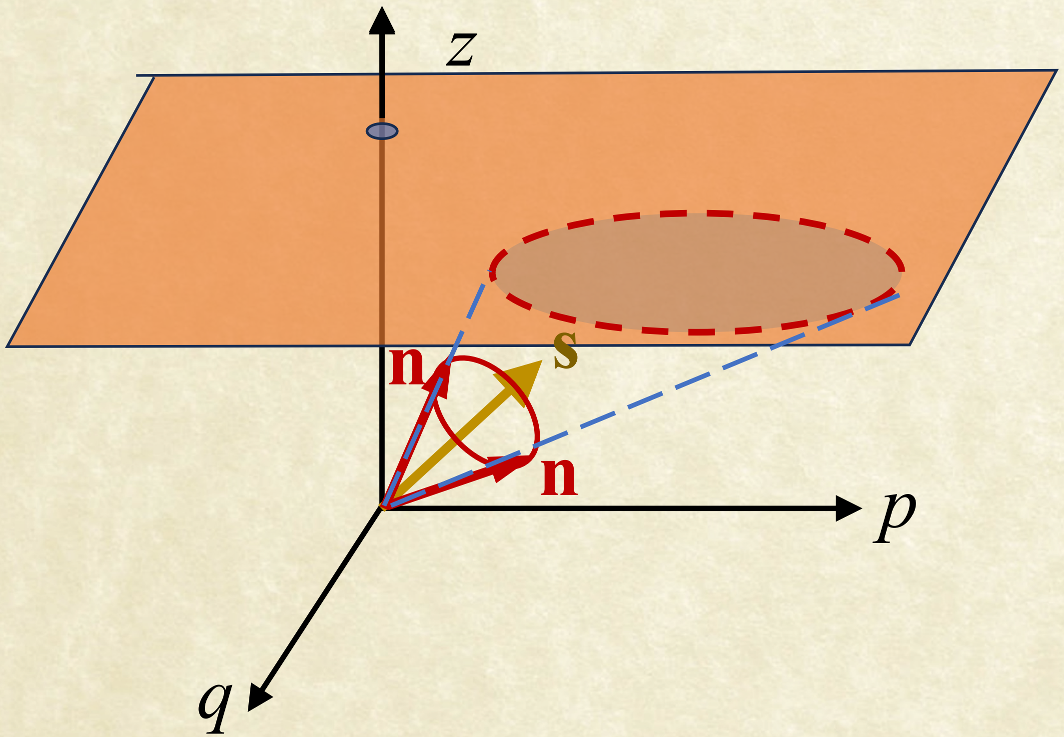
$$I = c\frac{\rho}{\pi}\frac{J}{r^2}\cos\theta_i = c\frac{\rho}{\pi}k(\mathbf{n}.\mathbf{s}) = \mathbf{n}.\mathbf{s}$$

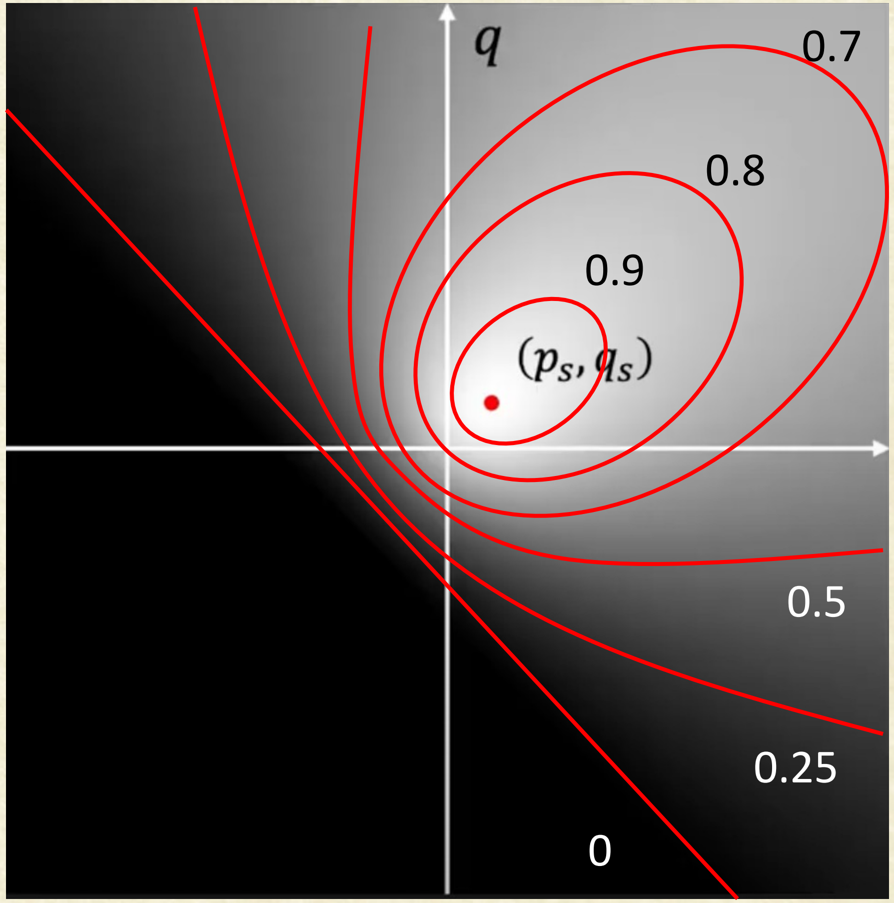


$$= \frac{pp_s + qq_s + 1}{\|\boldsymbol{N}\|\|\boldsymbol{S}\|} = R(p,q)$$

From Intensities to Surface Normals

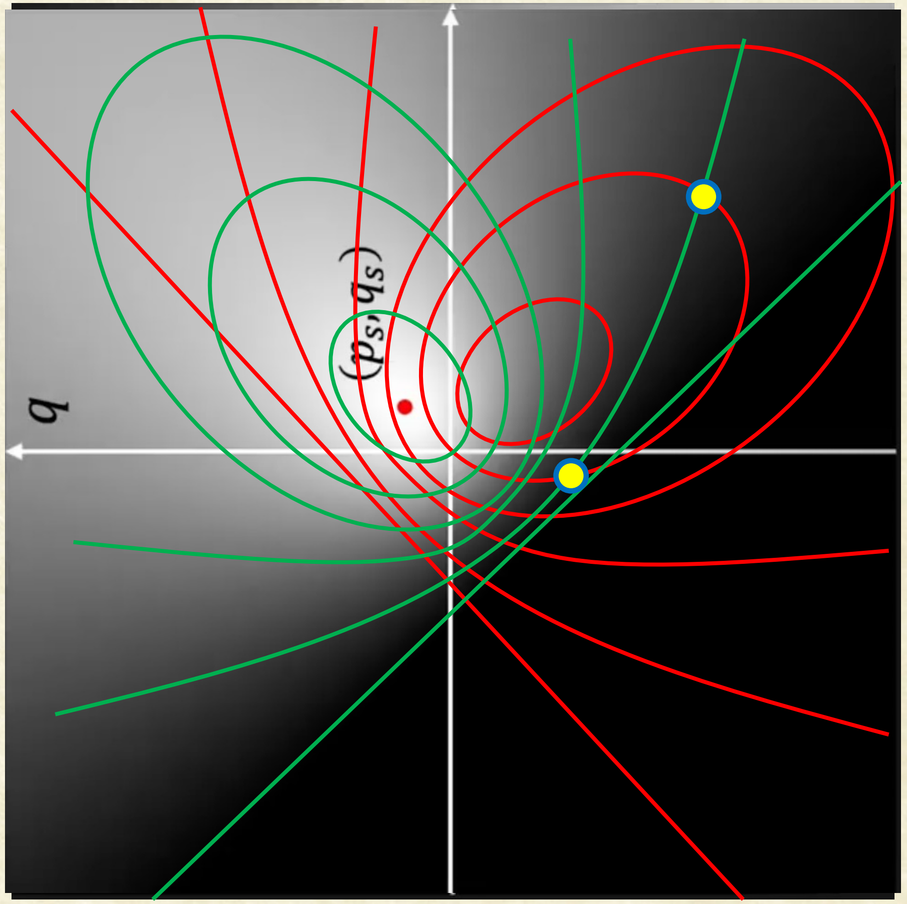
- Given a Reflectance map for a point source, the iso-intensity contours are conics
- Given an intensity there are several possible normals





Computing Unique Normals

- Use multiple light sources
  - Note: We assume the viewing direction and source directions are same for all points
- The correct normal should satisfy all reflectance maps
- Can be uniquely solved by three light sources (or more)
  - In-fact can compute albedo also with 3 sources



Photometric Stereo: A Naïve Algorithm

1. Capture k images with k known light sources
2. Using the known source details and BRDF, compute the reflectance map for each source.: $R_1(p,q),\ldots,R_k(p,q)$.
3. Quantize the R-Map and for each value of (p,q), enter the k intensities and corresponding normal into a lookup table.
4. For each pixel (x,y), lookup table row that best matches the intensities: $\boldsymbol{I}_1, \boldsymbol{I}_2, \boldsymbol{I}_3, , \boldsymbol{I}_{\mathrm{k}}$ and choose the corresponding (p,q).

Notes:

- All light sources may not illuminate all points
- We assume the surface is Lambertian and Isotropic.
- What happens if the above are not satisfied?

Computing Normals: Lambertian Case

- $I_1 = \frac{\rho}{\pi} n.s_1, \qquad I_2 = \frac{\rho}{\pi} n.s_2\,, \qquad I_3 = \frac{\rho}{\pi} n.s_3,$

where $n = \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix},$ and $s_i = \begin{bmatrix} s_{xi} \\ s_{yi} \\ s_{zi} \end{bmatrix}$

- We can write the three equations together in matrix form:

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \frac{\rho}{\pi} \begin{bmatrix} s_{x1} & s_{y1} & s_{z1} \\ s_{x2} & s_{y2} & s_{z2} \\ s_{x3} & s_{y3} & s_{z3} \end{bmatrix} n$$

$\boldsymbol{I} = \frac{\rho}{\pi} \boldsymbol{S}.\boldsymbol{n};$ or $\boldsymbol{I} = \boldsymbol{S}.\boldsymbol{N}$ and $N = S^{-1} I;$

Normal: $n = {}^{N}/_{\|N\|};$ Albedo: $\rho = \pi \|N\|$

Extending to k Light Sources

$$\begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_k \end{bmatrix} = \frac{\rho}{\pi} \begin{bmatrix} s_{x1} & s_{y1} & s_{z1} \\ s_{x2} & s_{y2} & s_{z2} \\ \vdots & \vdots & \vdots \\ s_{xk} & s_{yk} & s_{zk} \end{bmatrix} n$$

- There are k equations in 3 unknowns
  - Find the least-squared-error solution
- Mathematically,
  - S is a kx3 matrix, which cannot be inverted.
- Solution:
  - Pre-multiply both sides by S^{-1} and solve.
  - The method of pseudo-inverse

Other Considerations

- Dealing with non-point light sources
- Dealing with non-Lambertian surfaces
- Dealing with color objects
- Dealing with inter-reflections
- Other errors in assumption

By Calibration

Computing Depth from Normals

- Find a set of z values that minimizes the error between measured normals and normals computed from z values.
  - Naïve approach: Incrementally compute z from its derivatives ($\boldsymbol{n}$)
  - Does not work well in presence of noise
- Minimize the squared error between the two. i.e, find z that minimizes:

$$D = \iint\limits_{Image} \left(\frac{dz}{dx} + p\right)^2 + \left(\frac{dz}{dy} + q\right)^2 \, dx\, dy$$

Note, Gradient: $\left(-\frac{dz}{dx}, -\frac{dz}{dx}\right) = (p, q)$

From Normals to Depth: Frankot & Chellappa

- Key Idea: Minimize D in Fourier domain
- Let $Z(u, v)$, $P(u, v)$ and $Q(u, v)$ be the Fourier transforms of $z(x, y)$, $p(x, y)$ and $q(x, y)$ respectively.
- Find $Z(u, v)$ that minimizes D by setting: $\frac{\partial D}{\partial z} = 0$.
- Solution:

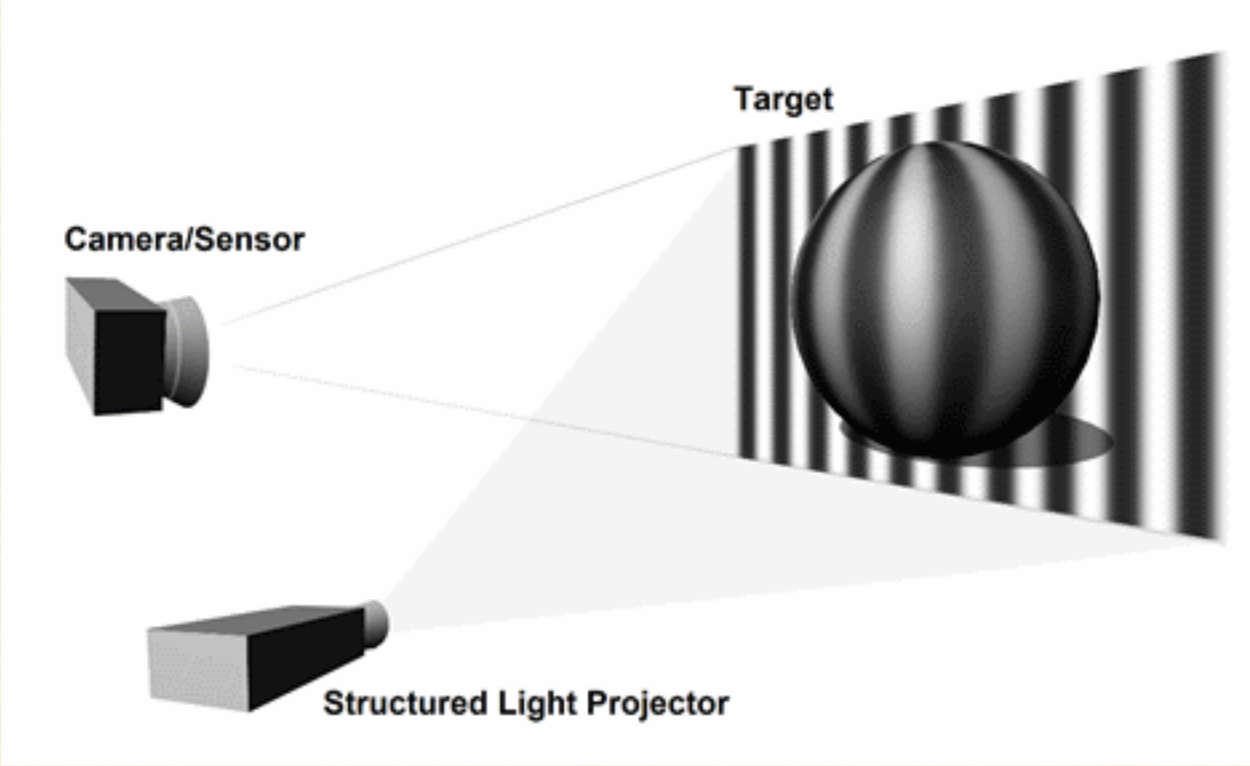
$$\tilde{Z}(u, v) = \frac{iuP(u, v) + ivQ(u, v)}{u^2 + v^2}$$

- Compute $\mathcal{F}^{-1}\left(\tilde{Z}(u, v)\right)$ to get $\tilde{z}(x, y)$.

Structured Light Sensing

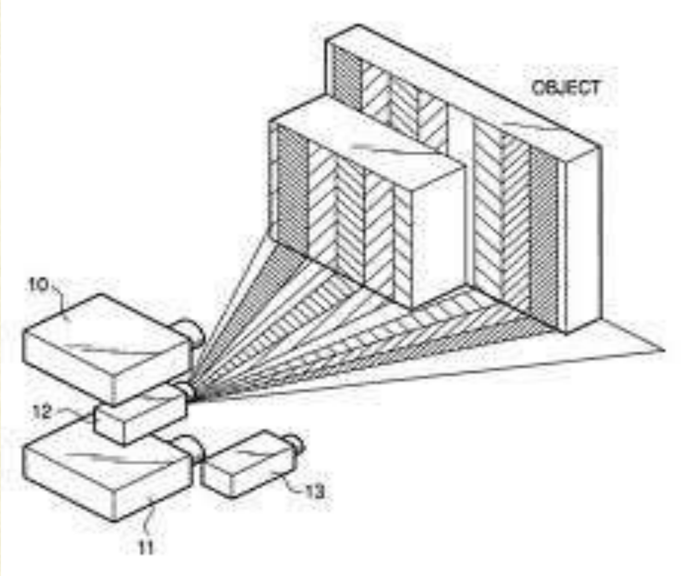
Projector-Camera Systems

- Projector is a camera
- Projector and camera forms a stereo pair.
- We can control what we project
  - Single point at a time
  - Binary Coding
  - Vertical/Horizontal stripes



Structured Lighting with Stereo Cameras

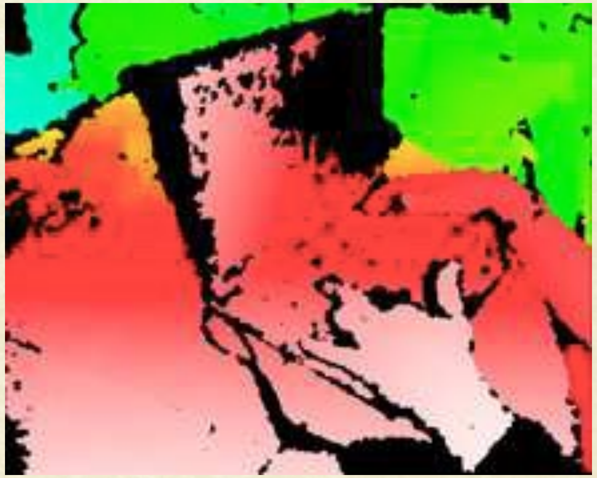
- Finding correspondences is hard by itself
- Can we help it by projecting patterns onto the world?
- Light-strip range finders, etc.
- Combination of sinusoids sometimes to get dense matches
- Active vision, as it changes the appearance
- The light projected need not be in the visible spectrum



Xbox Kinect

IR-based range sensor for Xbox

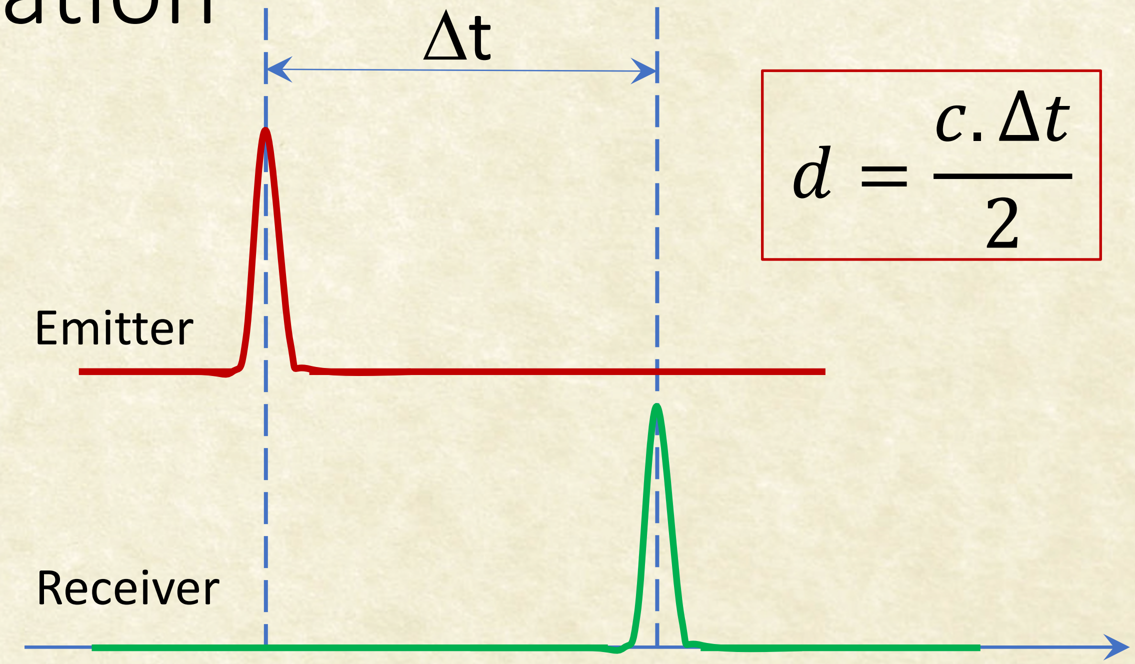
- Aligned depth and RGB images at 640 × 480
- Original goal: Interact with games in full 3D
- Computer vision happy with real-time depth and image
    - Games, HCI, etc
    - Action recognition
    - Image based modelling of dynamic scenes
- Fastest selling electronic appliance ever!!
- Other products that use PrimeSense sensor



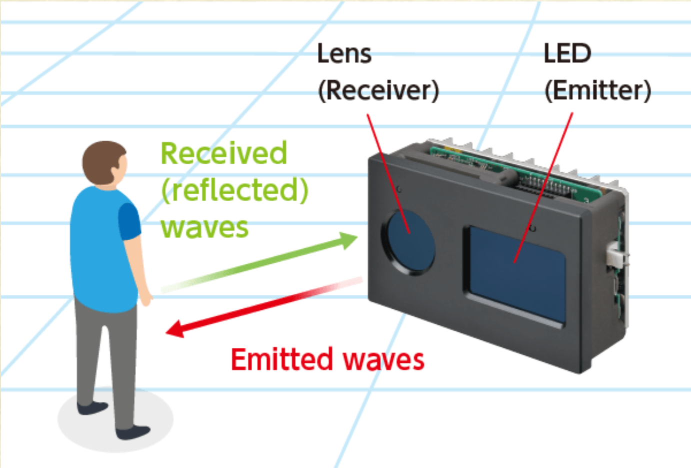
Optical Time-of-Flight (ToF) Sensors

Direct ToF or Pulse Modulation

- Send a short light pulse
- Measure the time taken for reflection
- Compute distance based on time elapsed



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

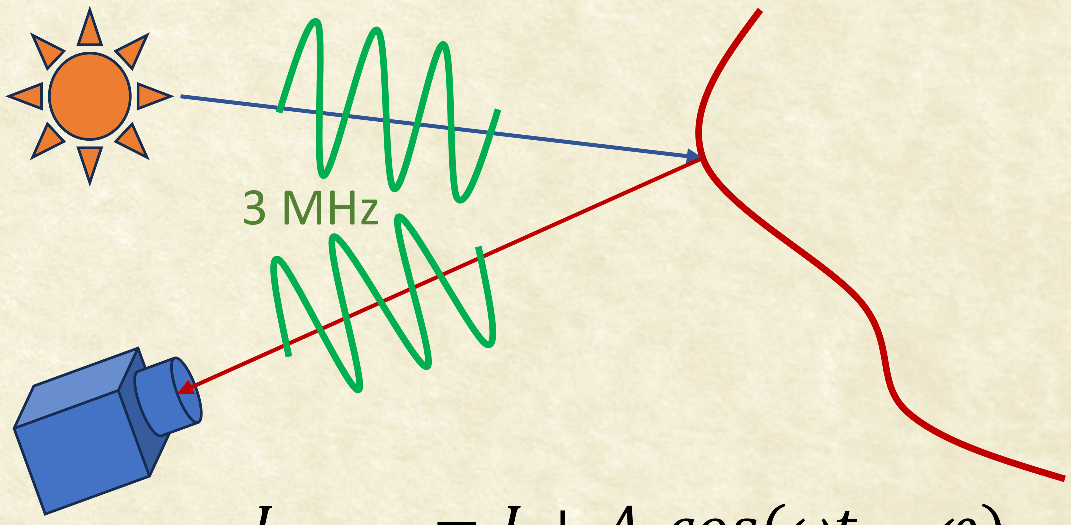


- Need very short pulse for accuracy
- Shorter pulses should be very bright for effective sensing (referred to as the flash method)

Indirect ToF or Continuous Modulation

- Use a temporally modulated light source
- Find the phase difference of the reflected light
- Distance is proportional the phase difference
- Terminology:
  - ω: Modulation Frequency
  - I: Ambient Illumination
  - φ: Phase difference

$$L_{emit} = cos(\omega t)$$



$$L_{scene} = I + A.cos(\omega t - \varphi)$$

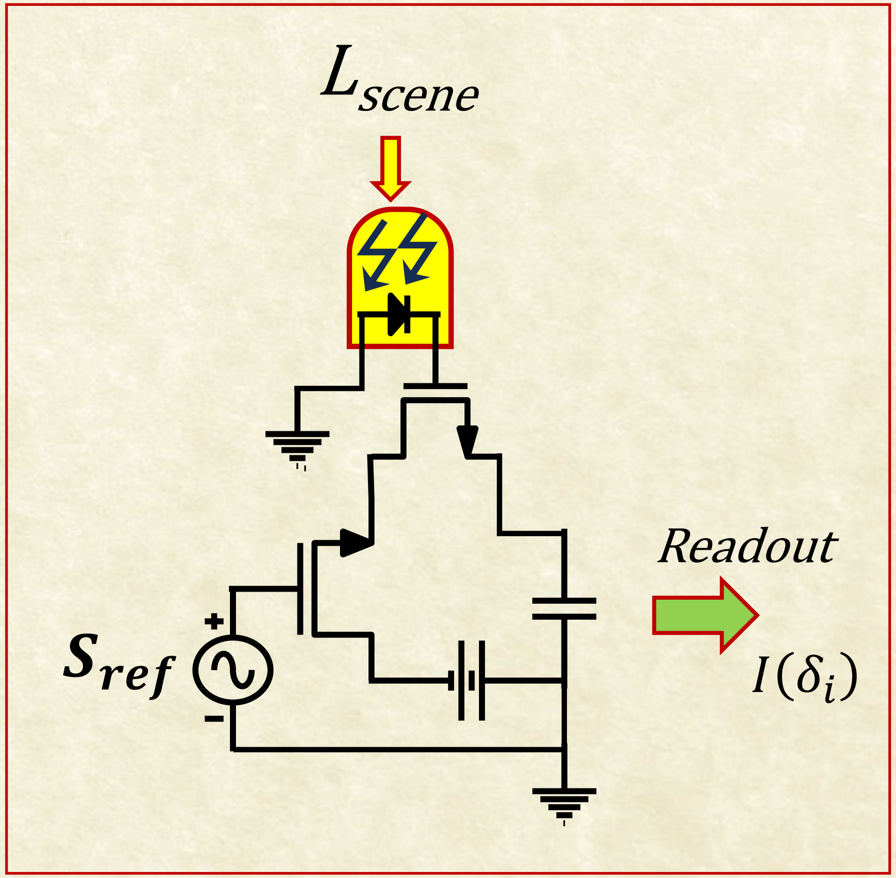
Computing Depth

- $L_{scene} = I + A.cos(\omega t - \varphi)$
- $S_{ref} = cos(\omega t - \delta_i)$
- $I(\delta_i) = \int_0^t L_{scene}(t) \times S_{ref}(t, \delta_i)\, dt$

Substituting and simplifying, we get:

$$I(\delta_i) = \boldsymbol{P} + \boldsymbol{Q}\, cos(\delta_i - \boldsymbol{\varphi})$$

- Measure $I(\delta_i)$ for three different δ_i to solve for the unknowns and get $\boldsymbol{\varphi}$.
- E.g., $d(f = 3MHz, \varphi = \pi) = 25m$



$$\boxed{d(f, \varphi) = \frac{c.\varphi}{4\pi f}}$$

$f = {}^{\omega}/_{2\pi}$

Questions?