Range Sensors (time of flight) (1)

- Large range distance measurement -> called range sensors
- Range information:
  - *key element for localization and environment modeling*
- Ultrasonic sensors as well as laser range sensors make use of propagation speed of sound or electromagnetic waves respectively. The traveled distance of a sound or electromagnetic wave is given by

$$d = c \cdot t$$

- Where
  - *d = distance traveled (usually round-trip)*
  - *c = speed of wave propagation*
  - *t = time of flight.*

Range Sensors (time of flight) (2)

- It is important to point out
  - *Propagation speed v of sound: 0.3 m/ms*
  - *Propagation speed v of of electromagnetic signals: 0.3 m/ns,*
    - *one million times faster.*
  - *3 meters*
    - *is 10 ms ultrasonic system*
    - *only 10 ns for a laser range sensor*
    - *time of flight t with electromagnetic signals is not an easy task*
    - *laser range sensors expensive and delicate*
- The quality of time of flight range sensors manly depends on:
  - *Uncertainties about the exact time of arrival of the reflected signal*
  - *Inaccuracies in the time of fight measure (laser range sensors)*
  - *Opening angle of transmitted beam (ultrasonic range sensors)*
  - *Interaction with the target (surface, specular reflections)*
  - *Variation of propagation speed*
  - *Speed of mobile robot and target (if not at stand still)*

Ultrasonic Sensor (time of flight, sound) (1)

- transmit a packet of (ultrasonic) pressure waves
- distance d of the echoing object can be calculated based on the propagation speed of sound c and the time of flight t.

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

- The speed of sound c (340 m/s) in air is given by

$$c = \sqrt{\gamma.R.T}$$

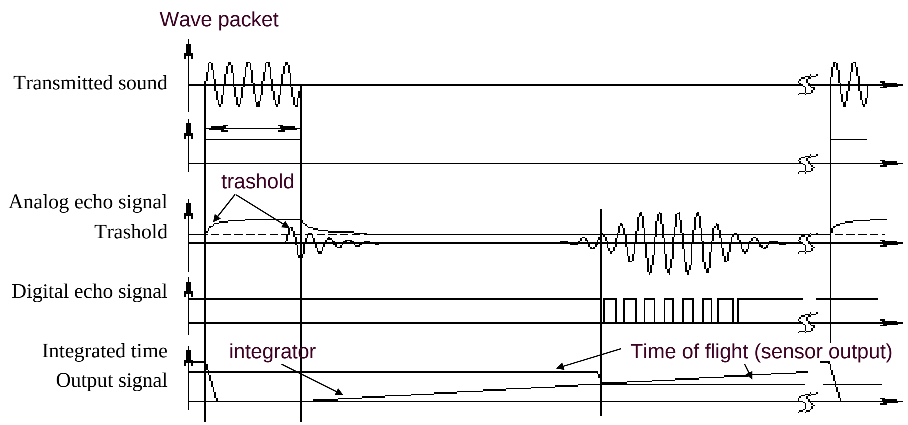
where

γ: ration of specific heats

R: gas constant

T: temperature in degree Kelvin

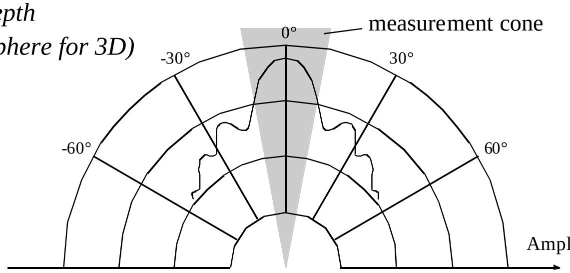
Ultrasonic Sensor (time of flight, sound) (2)



Signals of an ultrasonic sensor

Ultrasonic Sensor (time of flight, sound) (3)

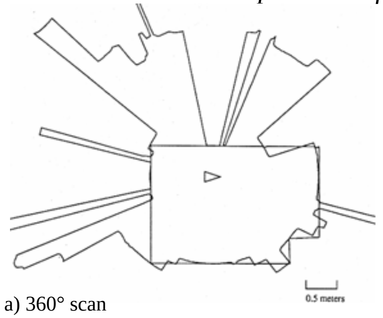
- typically a frequency: 40 - 180 kHz
- generation of sound wave: piezo transducer
  - *transmitter and receiver separated or not separated*
- sound beam propagates in a cone like manner
  - *opening angles around 20 to 40 degrees*
  - *regions of constant depth*
  - *segments of an arc (sphere for 3D)*



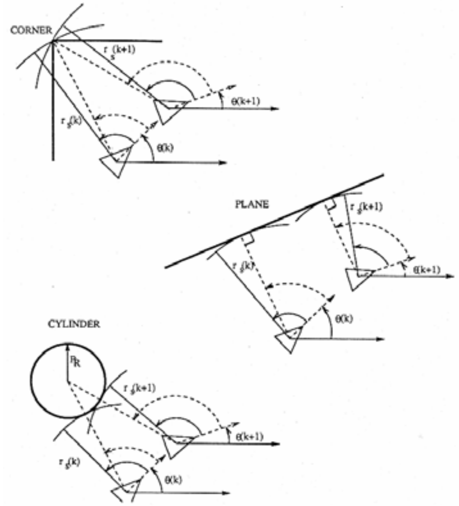
Typical intensity distribution of a ultrasonic sensor

Ultrasonic Sensor (time of flight, sound) (4)

- Other problems for ultrasonic sensors
  - *soft surfaces that absorb most of the sound energy*
  - *surfaces that are far from being perpendicular to the direction of the sound -> specular reflection*

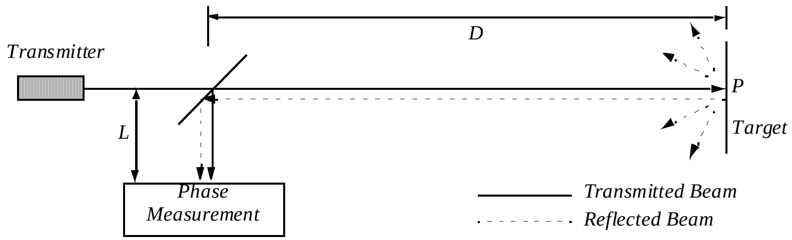


a) 360° scan



b) results from different geometric primitives

Laser Range Sensor (time of flight, electromagnetic) (1)



- Transmitted and received beams coaxial
- Transmitter illuminates a target with a collimated beam
- Received detects the time needed for round-trip
- A mechanical mechanism with a mirror sweeps
  - *2 or 3D measurement*

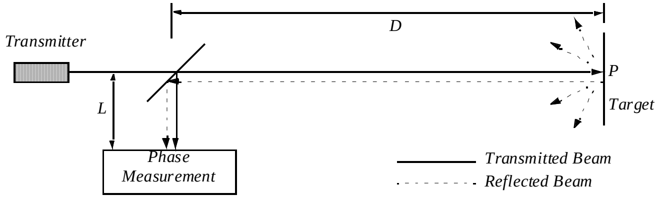
Laser Range Sensor (time of flight, electromagnetic) (2)

Time of flight measurement

- Pulsed laser
  - *measurement of elapsed time directly*
  - *resolving picoseconds*
- Beat frequency between a frequency modulated continuous wave and its received reflection
- Phase shift measurement to produce range estimation
  - *technically easier than the above two methods.*

Laser Range Sensor (time of flight, electromagnetic) (3)

- *Phase-Shift Measurement*



$\lambda = c/f \qquad D' = L + 2D = L + \frac{\theta}{2\pi}\lambda$

Where

c: is the speed of light; f the modulating frequency; D' covered by the emitted light is

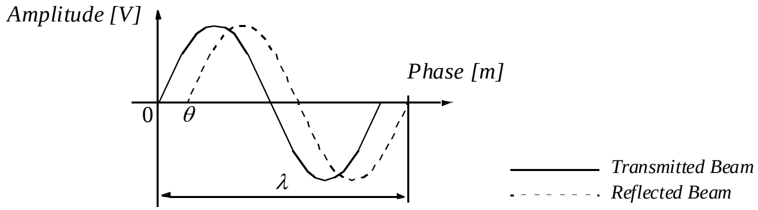
- *for f = 5 Mhz (as in the A.T&T. sensor), λ = 60 meters*

Laser Range Sensor (time of flight, electromagnetic) (4)

- Distance D, between the beam splitter and the target

$$D = \frac{\lambda}{4\pi}\theta \qquad (2.33)$$

- where
  - *θ: phase difference between the transmitted*
- Theoretically ambiguous range estimates
  - *since for example if λ = 60 meters, a target at a range of 5 meters = target at 65 meters* will result in the same phase shift



Laser Range Sensor (time of flight, electromagnetic) (5)

- Confidence in the range (phase estimate) is inversely proportional to the square of the received signal amplitude.

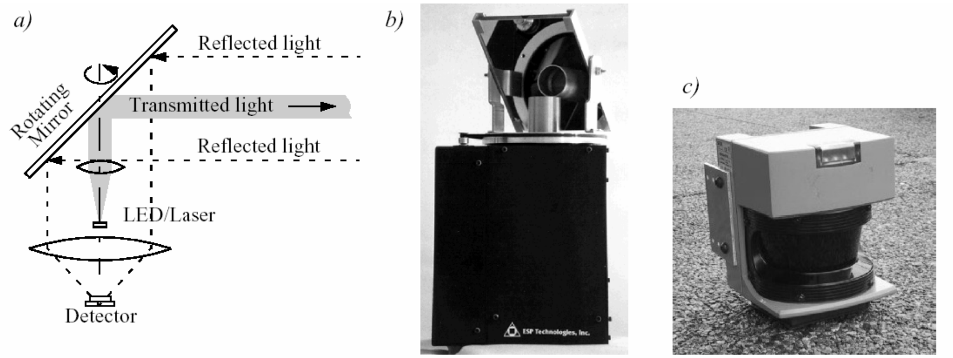
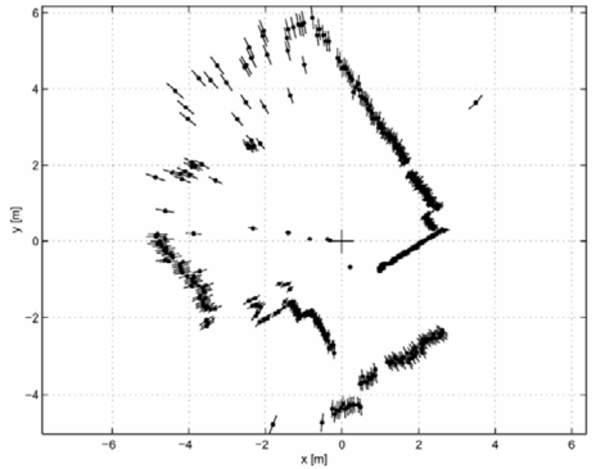


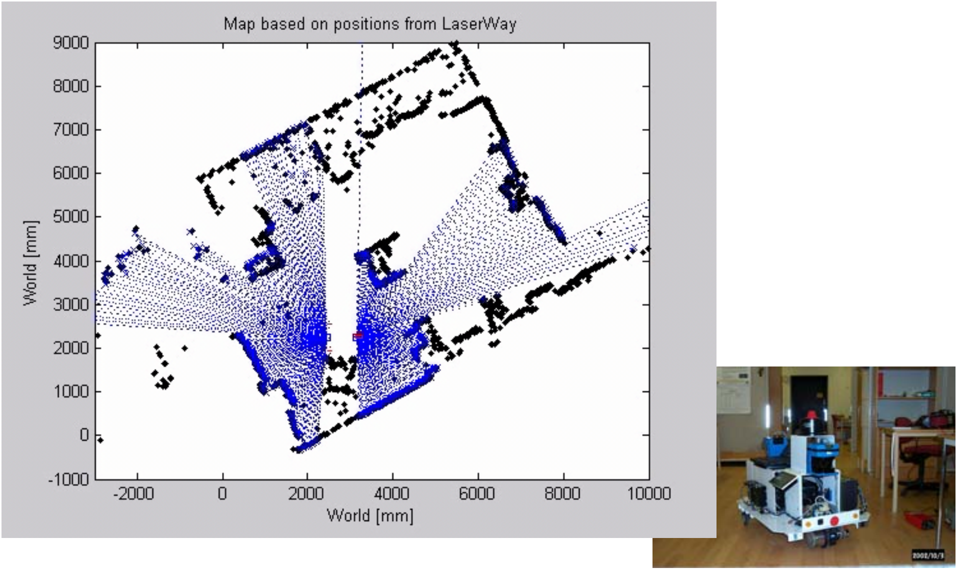
Figure 4.11
(a) Schematic drawing of laser range sensor with rotating mirror; (b) Scanning range sensor from EPS Technologies Inc.; (c) Industrial 180 degree laser range sensor from Sick Inc., Germany

Laser Range Sensor (time of flight, electromagnetic)

- Typical range image of a 2D laser range sensor with a rotating mirror. The length of the lines through the measurement points indicate the uncertainties.



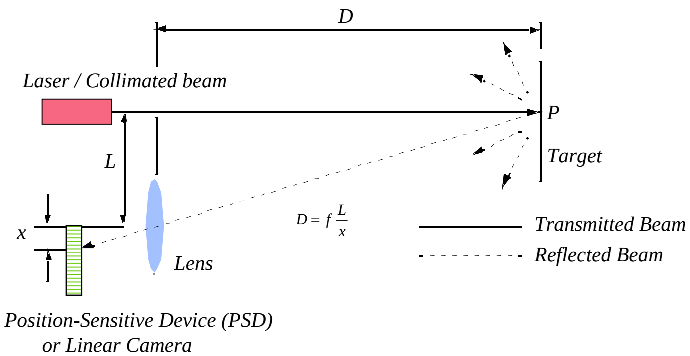
Examples of laser scans perceived by a time-of-flight sensor



Triangulation Ranging

- geometrical properties of the image to establish a distance measurement
- e.g. project a well defined light pattern (e.g. point, line) onto the environment.
  - *reflected light is than captured by a photo-sensitive line or matrix (camera) sensor device*
  - *simple triangulation allows to establish a distance.*
- e.g. size of an captured object is precisely known
  - *triangulation without light projecting*

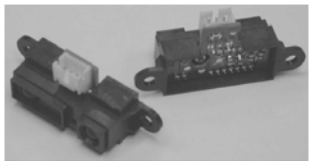
Laser Triangulation (1D)



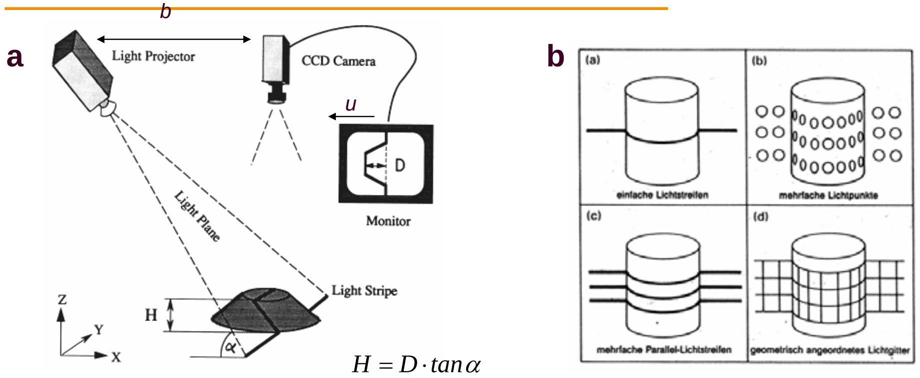
Principle of 1D laser triangulation.

- *distance is proportional to the 1/x*

$$D = f\frac{L}{x}$$



Structured Light (vision, 2 or 3D)

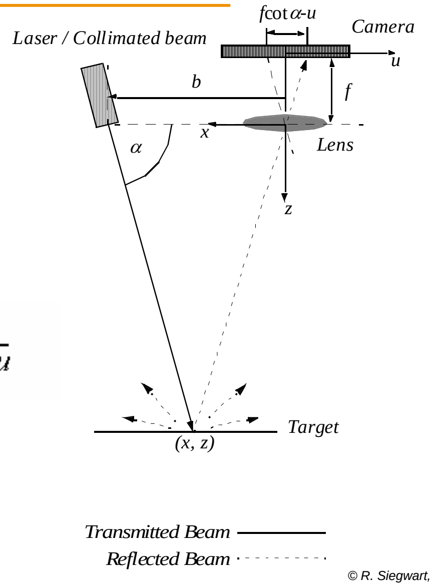


- Eliminate the correspondence problem by projecting structured light on the scene.
- Slits of light or emit collimated light (possibly laser) by means of a rotating mirror.
- Light perceived by camera
- Range to an illuminated point can then be determined from simple geometry.

Structured Light (vision, 2 or 3D)

- One dimensional schematic of the principle
- From the figure, simple geometry shows that:

$$x = \frac{b \cdot u}{f\cot\alpha - u} \; ; \quad z = \frac{b \cdot f}{f\cot\alpha - u}$$



Structured Light (vision, 2 or 3D)

- Range resolution is defined as the triangulation gain G_p :

$$\frac{\partial u}{\partial z} = G_p = \frac{b \cdot f}{z^2}$$

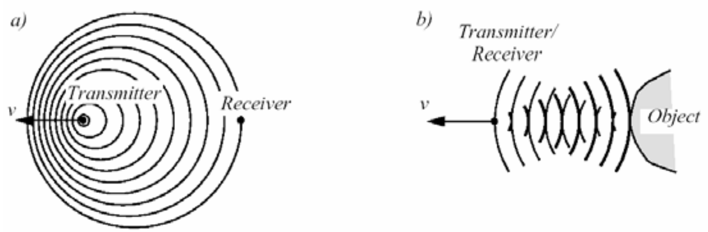
- Influence of α:

$$\frac{\partial \alpha}{\partial z} = G_\alpha = \frac{b \sin \alpha^2}{z^2}$$

- Baseline length *b*:
  - *the smaller b is the more compact the sensor can be.*
  - *the larger b is the better the range resolution is.*

  Note: for large b, the chance that an illuminated point is not visible to the receiver increases.
- Focal length f:
  - *larger focal length f can provide*
    - *either a larger field of view*
    - *or an improved range resolution*
  - *however, large focal length means a larger sensor head*

Doppler Effect Based (Radar or Sound)



a) between two moving objects b) between a moving and a stationary object

$f_r = f_t(1 + v/c)$ if transmitter is moving $\qquad f_r = f_t \frac{1}{1 + v/c}$ if receiver is moving

$\Delta f = f_t - f_r = \frac{2 f_t v \cos\theta}{c}$ Doppler frequency shift $\qquad v = \frac{\Delta f \cdot c}{2 f_t \cos\theta}$ relative speed

- *Sound waves: e.g. industrial process control, security, fish finding, measure of ground speed*
- *Electromagnetic waves: e.g. vibration measurement, radar systems, object tracking*

Vision-based Sensors: Hardware

- CCD (*light-sensitive, discharging capacitors of 5 to 25 micron*)

2048 x 2048 CCD array — Orangemicro iBOT Firewire — Sony DFW-X700 — Cannon IXUS 300

- CMOS (*Complementary Metal Oxide Semiconductor technology*)

Vision in General

Vision is our most powerful sense. It provides us with an enormous amount of information about our environment and enables us to interact intelligently with the environment, all without direct physical contact. It is therefore not surprising that an enormous amount of effort has occurred to give machines a sense of vision (almost since the beginning of digital computer technology!)

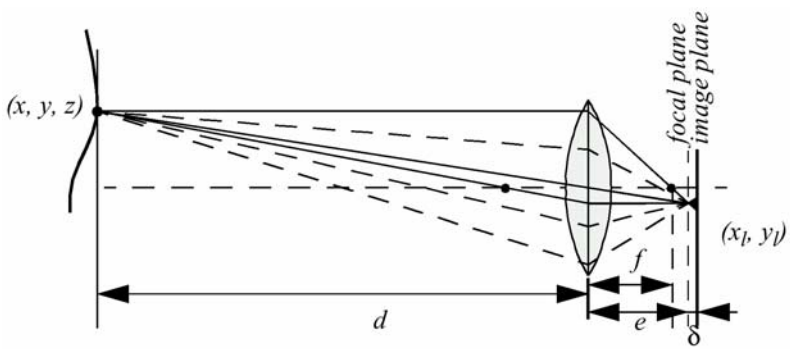
Vision is also our most complicated sense. Whilst we can reconstruct views with high resolution on photographic paper, the next step of understanding how the brain processes the information from our eyes is still in its infancy.

When an image is recorded through a camera, a 3 dimensional scene is projected onto a 2 dimensional plane (the film or a light sensitive photo sensitive array). In order to try and recover some "useful information" from the scene, usually edge detectors are used to find the contours of the objects. From these edges or edge fragments, much research time has to been spent attempting to produce fool proof algorithms which can provide all the necessary information required to reconstruct the 3-D scene which produced the 2-D image. Even in this simple situation, the edge fragments found are not perfect, and will require careful processing if they are to be integrated into a clean line drawing representing the edges of objects. The interpretation of 3-D scenes from 2-D images is not a trivial task. However, using stereo imaging or triangulation methods, vision can become a powerful tool for environment capturing.

Vision-based Sensors: Sensing

- Visual Range Sensors
  - *Depth from focus*
  - *Stereo vision*
- Motion and Optical Flow
- Color Tracking Sensors

Depth from Focus (1)



$$\frac{1}{f} = \frac{1}{d} + \frac{1}{e}$$

$$R = \frac{L\delta}{2e}$$

Depth from Focus (2)

- Measure of sub-image gradient $sharpness_1 = \sum_{x,y} |I(x,y) - I(x-1,y)|$

$$sharpness_2 = \sum_{x,y} (I(x,y) - I(x-2,y-2))^2$$

Stereo Vision => Image Analysis II

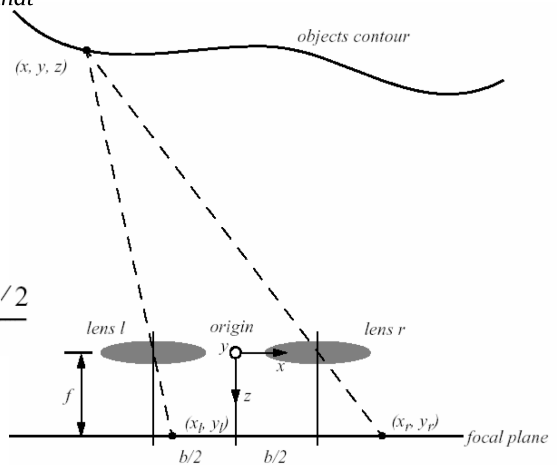
- Idealized camera geometry for stereo vision
  - *Disparity between two images -> Computing of depth*
  - *From the figure it can be seen that*

$$\frac{x_l}{f} = \frac{x + b/2}{z} \text{ and } \frac{x_r}{f} = \frac{x - b/2}{z}$$

$$\frac{x_l - x_r}{f} = \frac{b}{z}$$

$$x = b\frac{(x_l + x_r)/2}{x_l - x_r} \; ; \; y = b\frac{(y_l + y_r)/2}{x_l - x_r}$$

$$z = b\frac{f}{x_l - x_r}$$



Stereo Vision

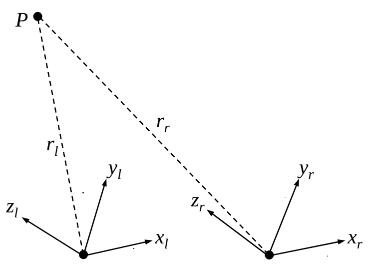
1. Distance is inversely proportional to disparity
   - *closer objects can be measured more accurately*
2. Disparity is proportional to b.
   - *For a given disparity error, the accuracy of the depth estimate increases with increasing baseline b.*
   - *However, as b is increased, some objects may appear in one camera, but not in the other.*
3. A point visible from both cameras produces a conjugate pair.
   - *Conjugate pairs lie on epipolar line (parallel to the x-axis for the arrangement in the figure above)*

Stereo Vision – the general case

- The same point P is measured differently in the left camera image :

$$r'_r = R \cdot r'_l + r_0 \qquad \begin{bmatrix} x'_r \\ y'_r \\ z'_r \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{21} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x'_l \\ y'_l \\ z'_l \end{bmatrix} + \begin{bmatrix} r_{01} \\ r_{02} \\ r_{03} \end{bmatrix}$$

- where
  - R *is a 3 x 3 rotation matrix*
  - r_0 *= offset translation matrix*



left camera coordinate system

right camera coordinate system

- The above equations have two uses:
  - *We can find r_r if we knew R and r_l and r_0. Note: For perfectly aligned cameras R=I (unity matrix)*
  - *We can calibrate the system and find r_{11}, r_{12} … given corresponding values of x_l, y_l, z_l, x_r, y_r and z_r.*
- We have 12 unknowns and require 12 equations:
  - *we require 4 conjugate points for a complete calibration.*
  - *Note: Additionally there is a optical distortion of the image*

Stereo Vision

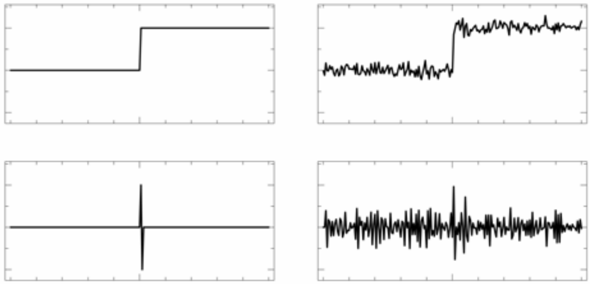
Calculation of Depth
- *The key problem in stereo is now how do we solve the correspondence problem?*

Gray-Level Matching
- *match gray-level wave forms on corresponding epipolar lines*
- *“brightness” = image irradiance I(x,y)*
- *Zero Crossing of Laplacian of Gaussian is a widely used approach for identifying feature in the left and right image*

Zero Crossing of Laplacian of Gaussian

- Identification of features that are stable and match well
- Laplacian of intensity image $L(x, y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}$
- Convolution with P: $L = P \otimes I$ $\quad P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$
- Step / Edge Detection in Noisy Image
  - ➢ *filtering through Gaussian smoothing*

$$\begin{bmatrix} \frac{1}{16} & \frac{2}{16} & \frac{1}{16} \\ \frac{2}{16} & \frac{4}{16} & \frac{2}{16} \\ \frac{1}{16} & \frac{2}{16} & \frac{1}{16} \end{bmatrix}$$



Stereo Vision Example

- Extracting depth information from a stereo image
  - ➢ *a1 and a2: left and right image*
  - ➢ *b1 and b2: vertical edge filtered left and right image; filter = [1 2 4 -2 -10 -2 4 2 1]*
  - ➢ *c: confidence image: bright = high confidence* (good texture)
  - ➢ *d: depth image: bright = close; dark = far*

Optical Flow (1)

- *E (x, y, t)* = irradiance at time *t* at the image point *(x, y)*.
- *u (x, y)* and *v (x, y)* = optical flow vector at that point
  - ➢ *find a new image for a point where the irradiance will be the same at time t+ δt*

$$E(x + u\delta t, y + v\delta t, t + \delta t) = E(x, y, t)$$

- If brightness varies smoothly with *x, y* and *t* we can expand the left hand side as a Taylor series to obtain:

$$E(x, y, t) + \delta x\frac{\partial E}{\partial x} + \delta y\frac{\partial E}{\partial y} + \delta t\frac{\partial E}{\partial t} + e = E(x, y, t)$$

- *e* = second and higher order terms in δx …
  - ➢ *With δt -> 0*

$$\frac{\partial E}{\partial x}\frac{dx}{dt} + \frac{\partial E}{\partial y}\frac{dy}{dt} + \frac{\partial E}{\partial t} = 0 \qquad u = \frac{dx}{dt} \; ; \; v = \frac{dy}{dt}$$

Optical Flow (2)

- from which we can abbreviate:

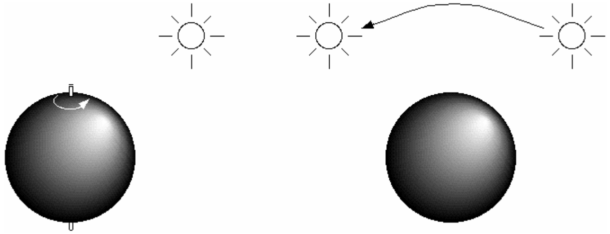
$$E_x u + E_y v + E_t = 0$$

optical flow constraint equation

  - ➢ *The derivatives Ex, Ey and Et are estimated from the image.*
- From this equation we can only get the direction of the velocity (*u, v*) and not unique values for *u* and *v*.
  - ➢ *One therefore introduces additional constraint, smoothness of optical flow (see lecture notes)*

Problems with Optical Flow

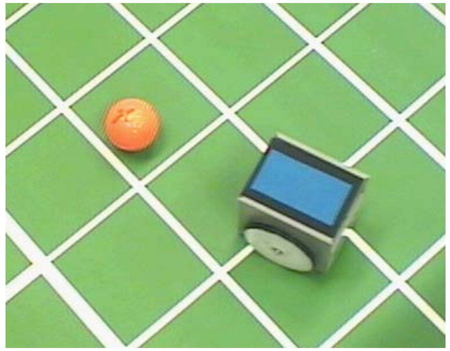
- Motion of the sphere or the light source here demonstrates that optical flow is not always the same as the motion field.

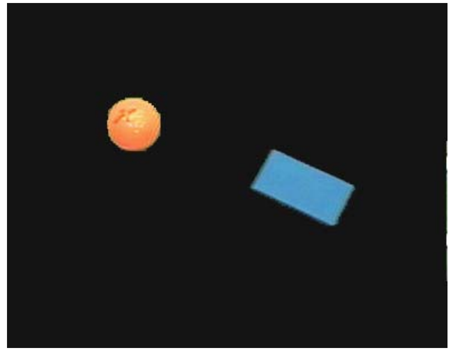


- Left: Discontinuities in Optical Flow
  - *silhouettes (one object occluding another)*
    - *discontinuities in optical flow*
  - *find these points*
    - *stop joining with smooth solution.*
- Right: Motion of sphere, moving light sources

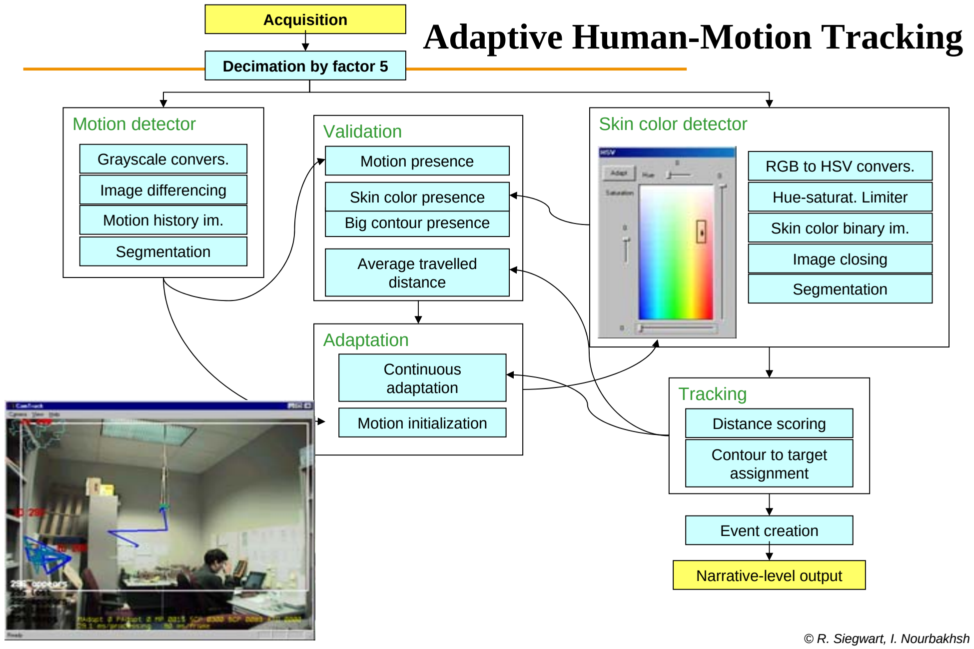
Color Tracking Sensors

- Motion estimation of ball and robot for soccer playing using color tracking





Adaptive Human-Motion Tracking



Adaptive Human-Motion Tracking