Cognitive Robotics

Sensors and Perception

Hans-Dieter Burkhard
Rijeka 2018

Outline

Introduction

Sensors: General Considerations

Signals

Sensors: Special Types

Vision (introductory)

Camera Model

Image Processing (introductory)

Scene Interpretation (introductory)

Sensors

Sensus (lat.): the sense

- Recording information related to state or change of state (physical, chemical ...).

- Transformation between state/change of state
  by differentiation/integration
  e.g. distance – speed – acceleration
  drift problems over time (e.g. odometry)

- Conversion to internally processable information
  Technically: mostly electronic signals
  Nature: electrochemical processes

- Direct influence on actuators in case of sensor actor coupling

Human senses (more than 5)

see
listen
smell
taste
tactil

heat
pain
balance
hunger
thirst
muscle tension, joints, ...

Further senses in nature
e.g. magnetism, electricity

Processing of Sensations in Nature

- Stimulus excites a receptor
- Release of nerve impulses
- Forwarding to the spinal cord / brain:
  About 1 million receptor signals per second
  in the Central Nervous System
- Unconscious reflexes activated by spinal cord or brain
  "Sensor-Actor Coupling"
- Filtering in Thalamus:
  Only selected signals
  are consciously perceived in the cerebral cortex.

Problems in Perception

Humans can deal with incomplete and unreliable data
Humans use redundancies
Humans use world knowledge and experience
Humans can deal with high complexity

Recent machines are far from human performance
Useful results only in special cases
Missing robustness and reliability

Outline

Introduction

Sensors: General Considerations

Signals

Sensors: Special Types

Vision (introductory)

Camera Model

Image Processing (introductory)

Scene Interpretation (introductory)

Sensors of Nao (Academic Version)

4 Microphones (head)
2 CMOS digital cameras (head)
32 Hall effect sensors (joints)
1 Gyrometer 2 axis (torso)
1 Accelerometer 3 axis (torso)
2 Bumpers (feets)
2 Channel sonar (torso)
2 Infrared sensors (torso)
9 Touch sensors (head, hands)
8 Force Resistance Sensor (feets)

Sensors

- Passive sensors
  - record signals created in the environment
- Active sensors
  - send signals (sonar, laser, radar, infrared, ...) and measure the reflections
  - disadvantage: recognizable through their signals
- Proprioceptive sensors
  - bodily sensation

Sensors

Internal sensors:

„Proprioceptive sensors“ ("self")

- Position (body, joints)
- Motion
- Internal forces
- Temperature (inside)
- Resources
- Energy
- ...

External sensors:

("Environment")

- Light, Vision
- Sound
- Smell
- Distance
- External forces
- Temperature (outside)
- …

Sensors in Robotics

Exploit physical/chemical … features, e.g.

- Current - power - resistance - inductance - conductivity ...
- Wavelength - frequency - phase shift - echo - runtime ...
- Mass - force - speed - acceleration - inertia ...

Transformations by related mathematical/physical laws.
e.g. *State to velocity* by differentiation

Conversion into internal information (mostly electronic signals)

Sensor Model and Observation Model

s = state/feature of the world

o = observation: sensory data according to s

Sensor Model: „Forward model“

$$o = f_{sensor}(s)$$

Observation Model: „Backward model“

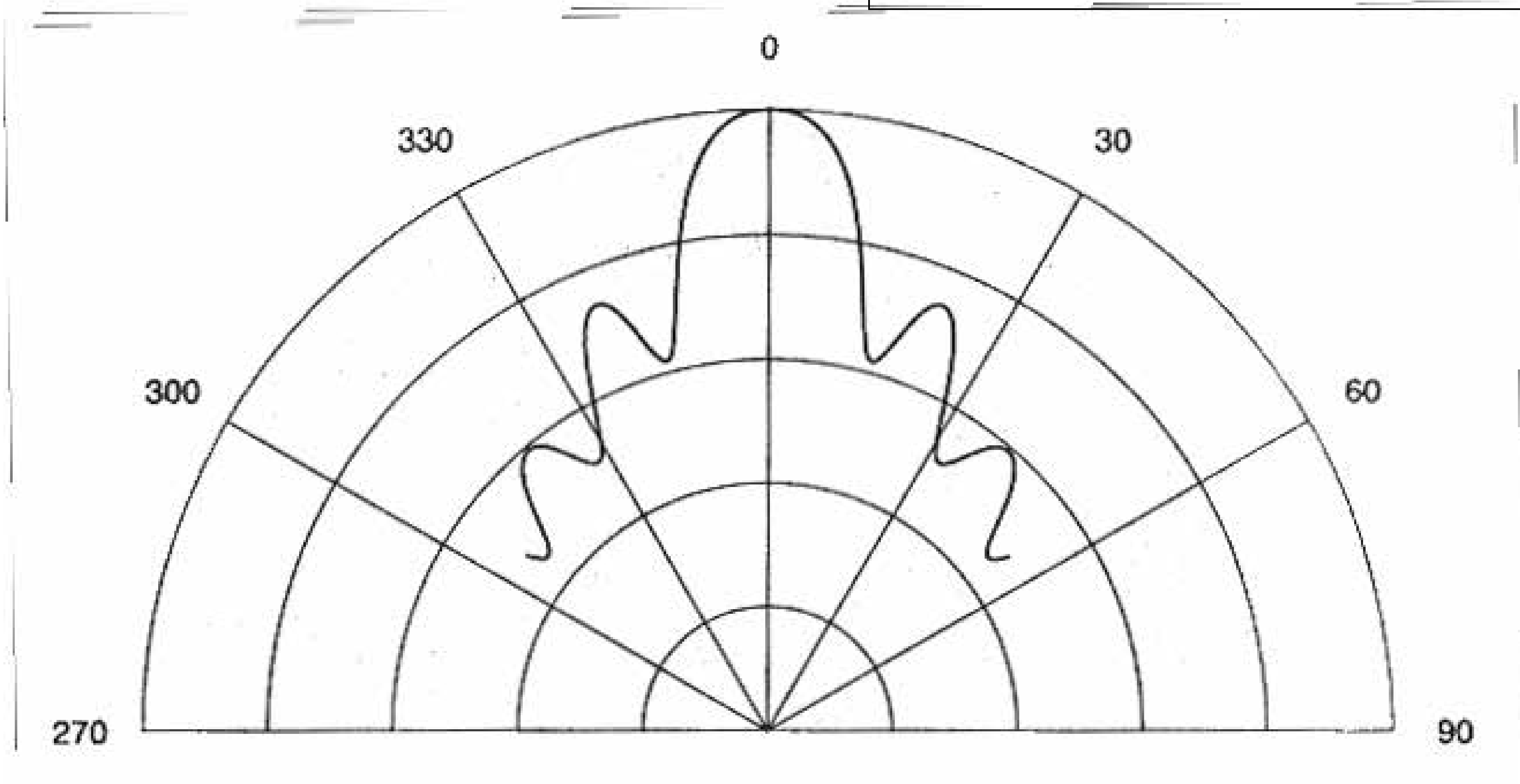
$$s = f^{-1}_{sensor}(o)$$

Sensor Model: Sonar

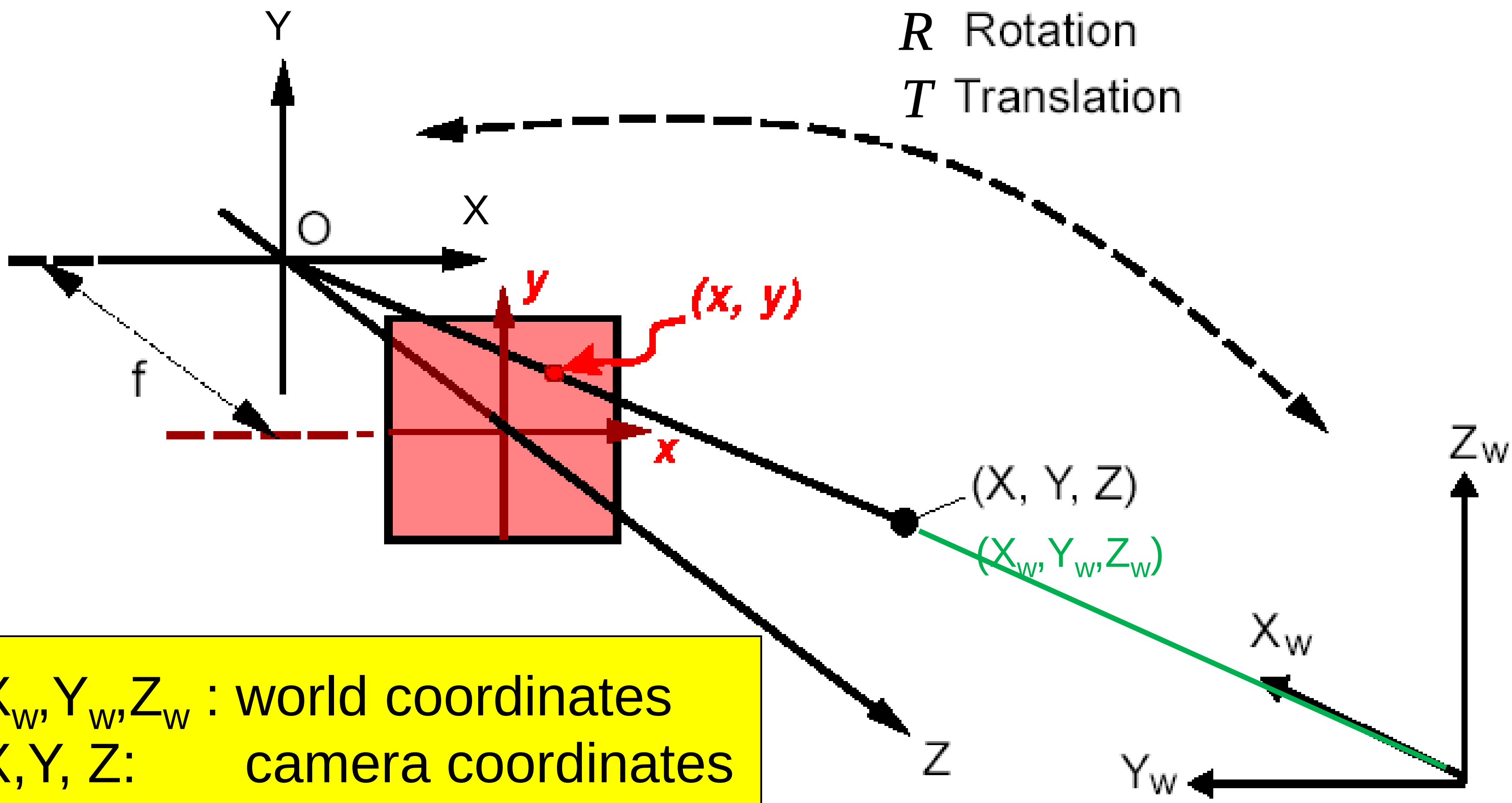
Acoustic propagation

(ca. 330 m/sec)

Image from “Where am I?" -- Systems and Methods for Mobile Robot Positioning by J. Borenstein, H. R. Everett, and L. Feng



Sensor Model: (Pinhole) Camera Projection

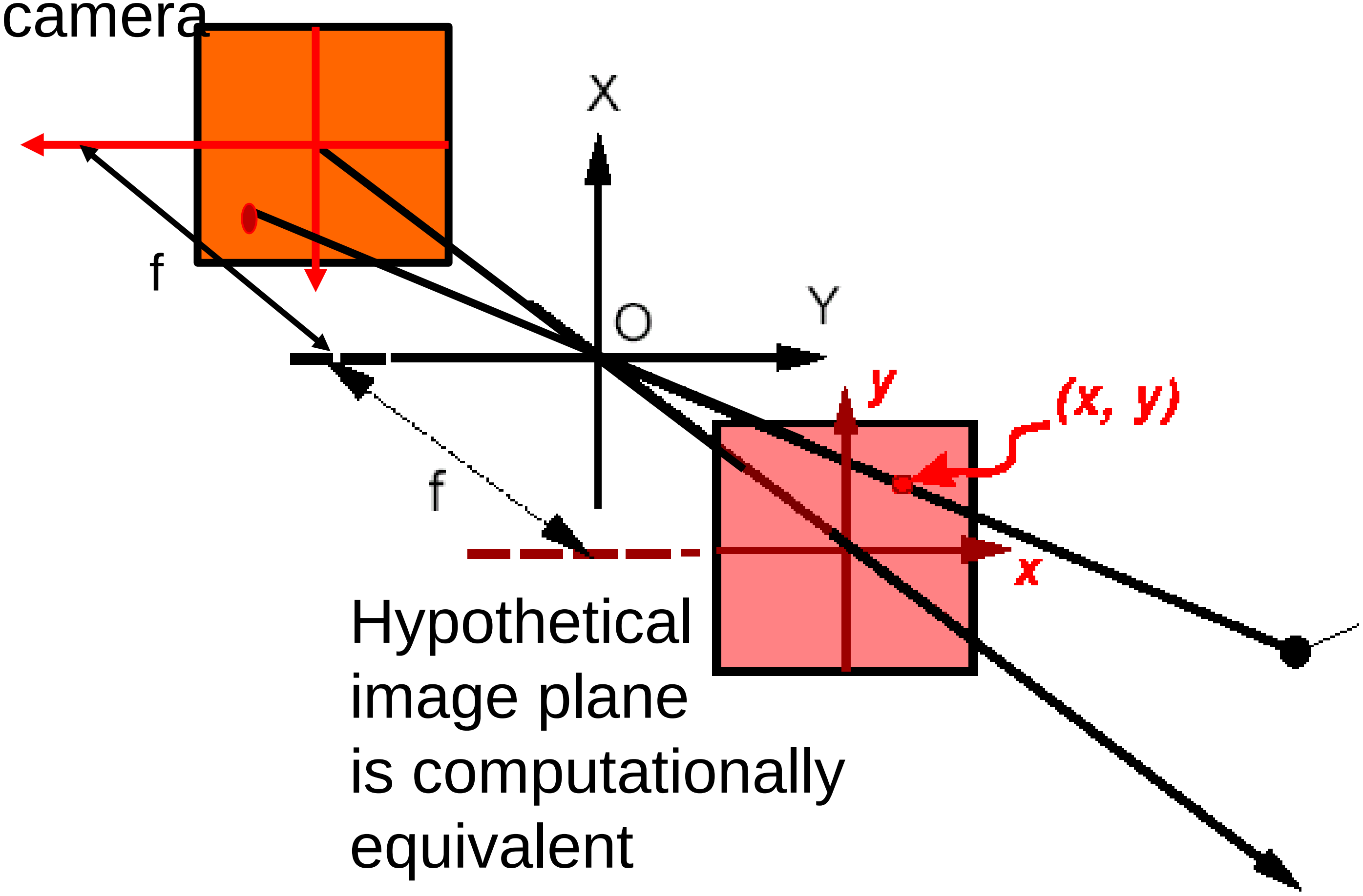


X_w, Y_w, Z_w : world coordinates
X,Y, Z: camera coordinates
x,y : image coordinates

Image from “Where am I?" -- Systems and Methods for Mobile Robot Positioning by J. Borenstein, H. R. Everett, and L. Feng

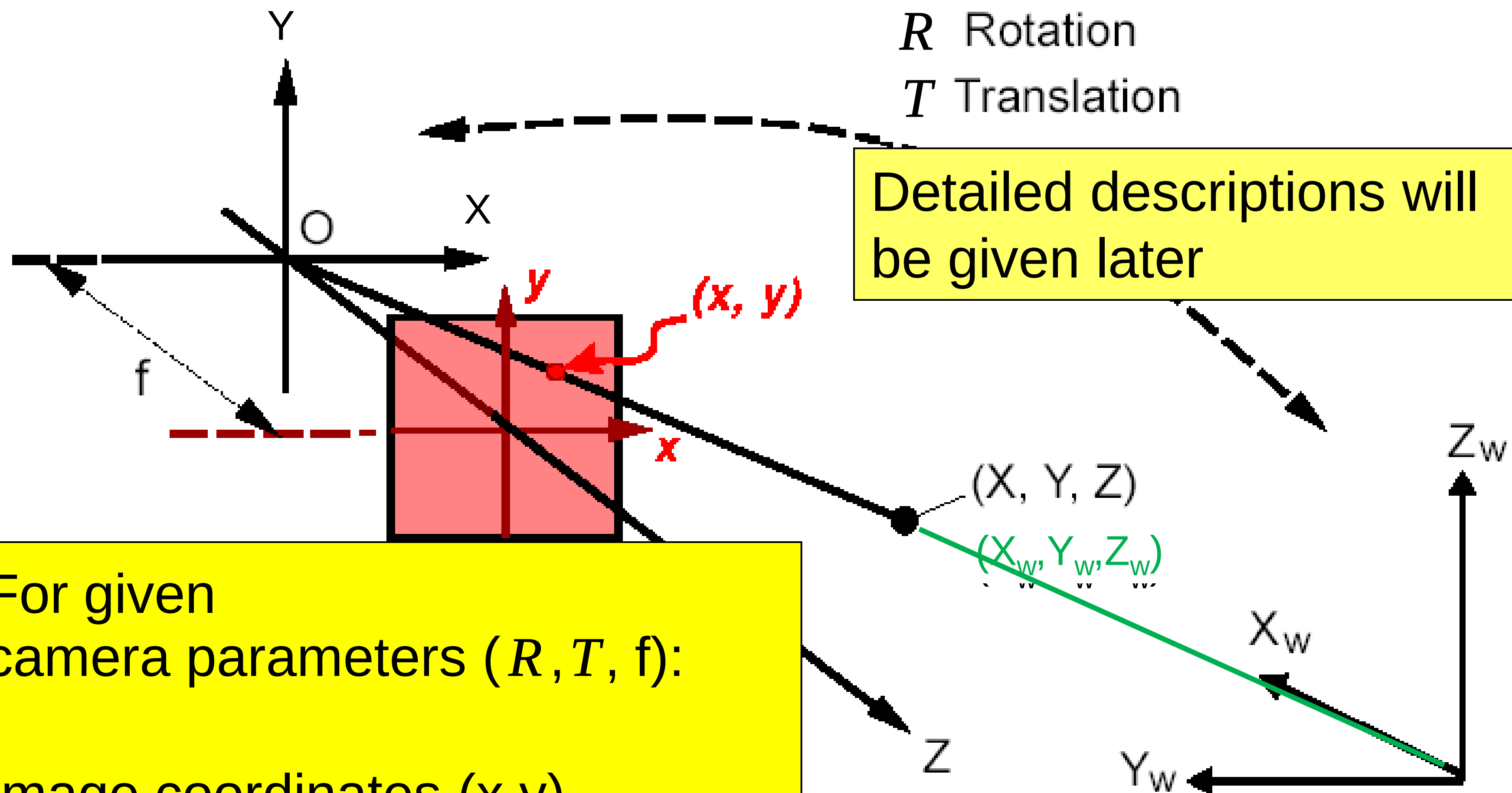
Sensor Model: (Pinhole) Camera Projection

Real image plane
of the camera



Hypothetical
image plane
is computationally
equivalent

Sensor Model: (Pinhole) Camera Projection



Detailed descriptions will be given later

For given
camera parameters (R, T, f):

Image coordinates (x,y)
are uniquely determined
by object coordinates (X_w, Y_w, Z_w)

Problems with Observation Model

- f_{sensor} often not bijective (f^{-1}_{sensor} not unique)

For given
camera parameters (R, T, f):

Object coordinates (X_w,Y_w,Z_w)
are **not** uniquely determined
by image coordinates (x,y) .

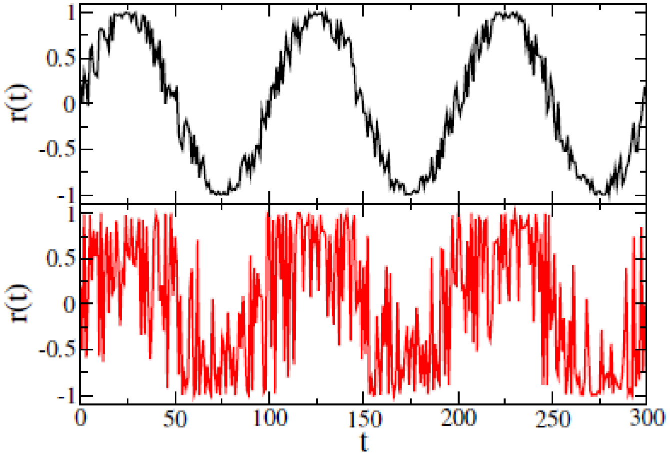
„Badly posted problem“

- noisy data: $o = f_{sensor}(s) + f_{noise}(s)$

Problems with Measurements

- Systematic errors (e.g. wrong position of sensors).

- Noise (caused by many inside and outside reasons):



different noise (PhD thesis J.N.E. Barrantos)

  - Modeling by noise models (often statistically).
  - Noise reduction by filtering.
  - Preprocessing in perception methods.

Outline

Introduction

Sensors: General Considerations

Signals

Sensors: Special Types

Vision (introductory)

Camera Model

Image Processing (introductory)

Scene Interpretation (introductory)

Signals

Information by

frequency, amplitude, pulse duration, ...

(Topics in Signal Processing)

Analog vs. discrete:

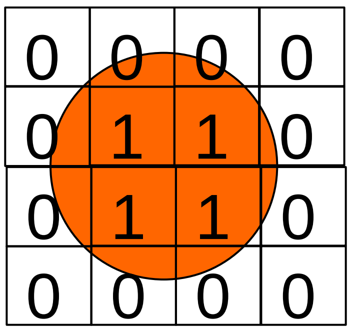
Depends on recording and processing

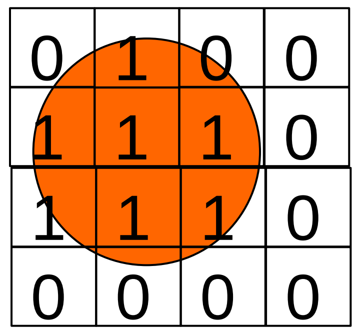
Conversion in both directions possible:

- Quantization
- Sampling
- Interpolation

Quantization

Discrete instead of continuous values (by rounding).

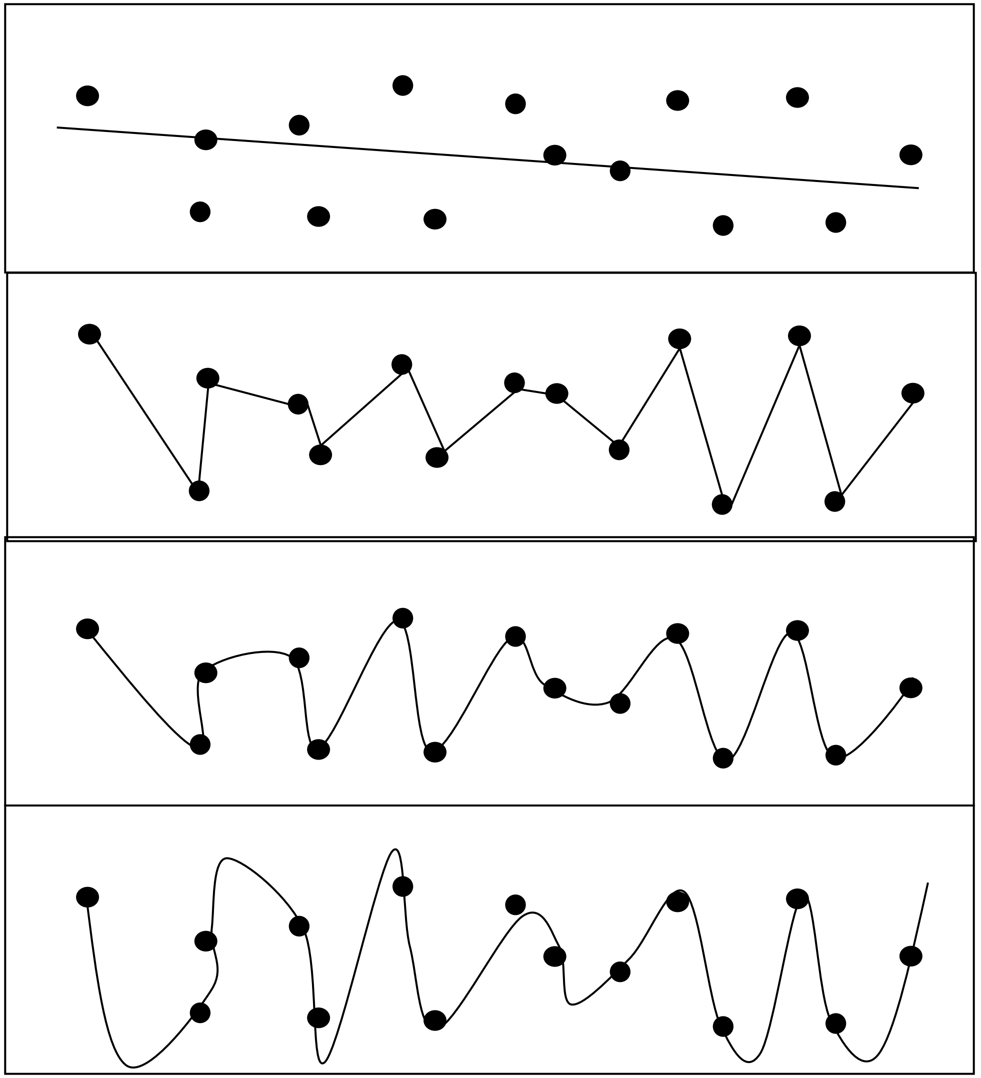




Small differences of continuous values can lead to larger differences of rounded values:

Noise: wrong values, Oscillations …

Interpolation

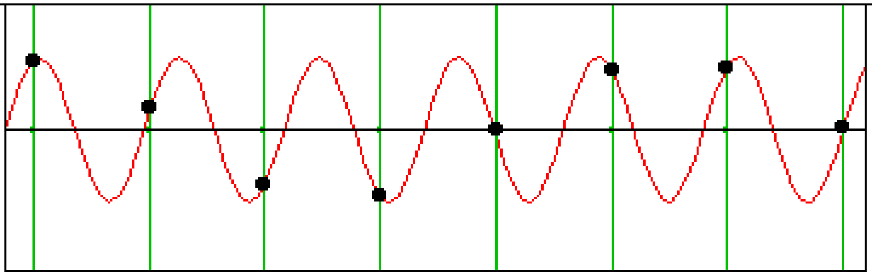


Find a curve which matches best given points.

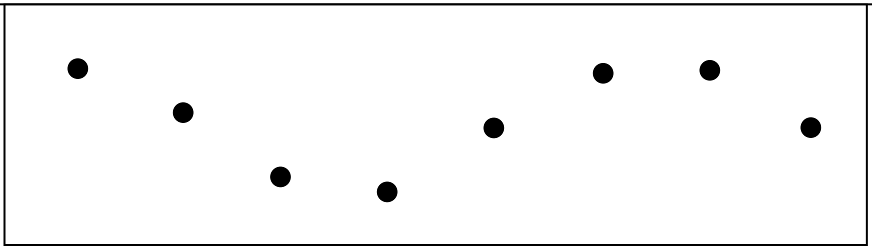
Depends on
- class of curve
  - o linear
  - o piecewise linear,
  - o quadratic, …)
- number of points
- error measure

Sampling Theorem (periodic functions)

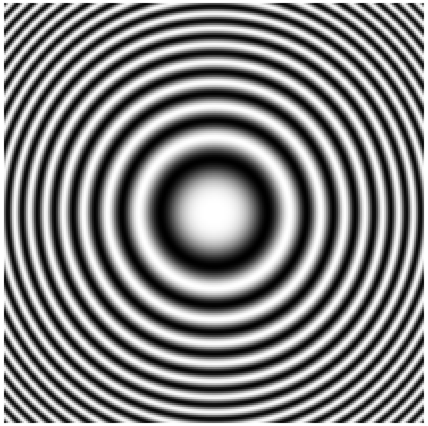
Problem:
The red curve is measured only at few points:
Only the black points are registered.



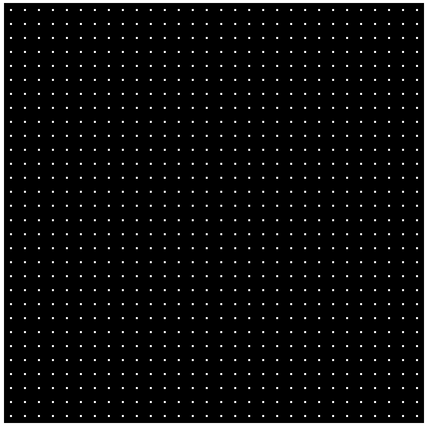
The black dots are interpreted as a lower frequency curve: "**Alias**"



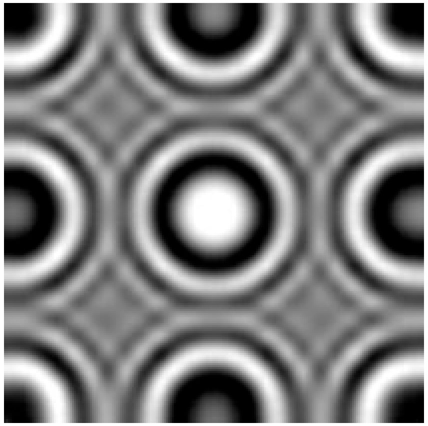
Aliasing



Original Image



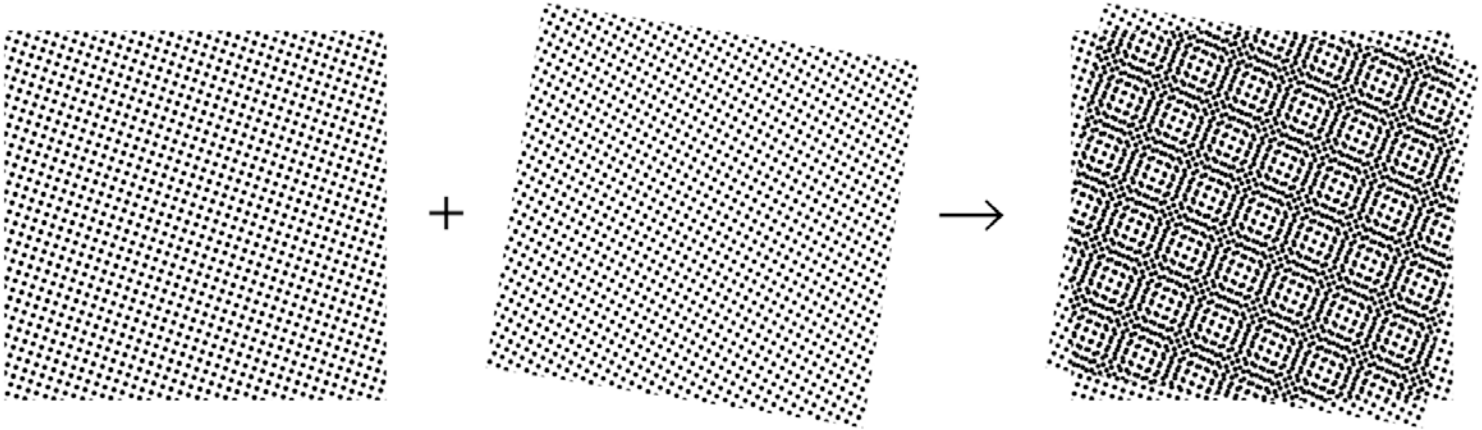
Sampling Points



Reconstructed Image

From Wikipedia, Author: Pemu

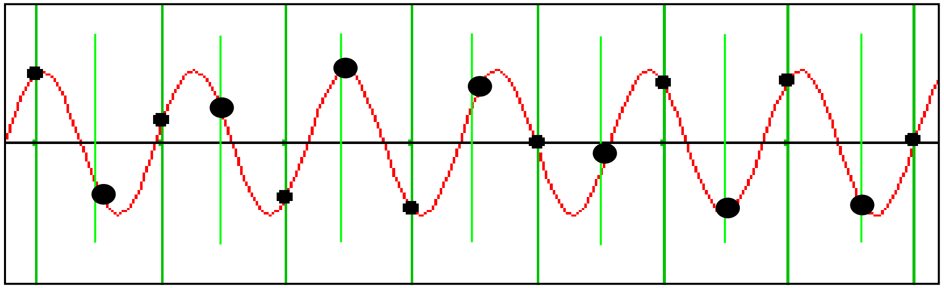
Moiré-Effects

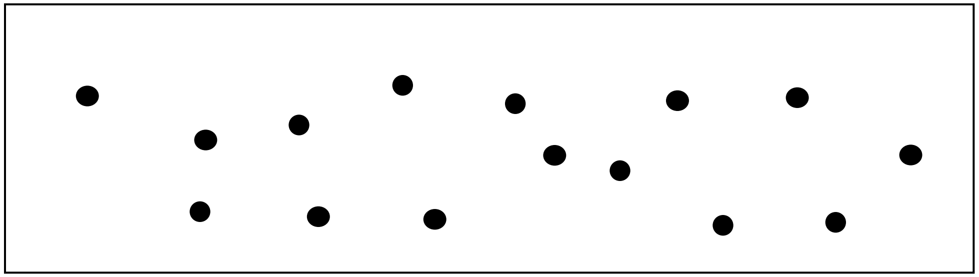


From Wikipedia, Author: Colin Pelka

Sampling Theorem

Example: Smaller intervals for measurements:
-- More points





How many measurements are needed?

Sampling Theorem

Sampling Theorem

For correct reproduction we must have:

More than 2 sampling points per wavelength T,

i.e. sampling rate $Dx < T/2$ (Nyquist criterion)

or:

Sampling frequency must be more than twice as large as the highest occurring frequency.

Holds also for more dimensional signals (e.g. images).

Outline

Introduction

Sensors: General Considerations

Signals

Sensors: Special Types

Vision (introductory)

Camera Model

Image Processing (introductory)

Scene Interpretation (introductory)

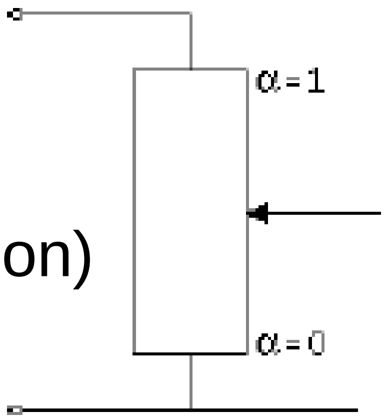
Example: Resistance Sensors

Potentiometer:

Voltage dependents on position on a resistor.

Sensor:

Transformation of mechanical values (e.g. position) into electrical signals.



Straingages:

Resistance depends on length (e.g. of meandering material)

Sensor:

Measurement of deformations.

Light Sensor / Infrared Sensor

Device with varying electronic properties (charge, resistance, …) depending on light intensity.

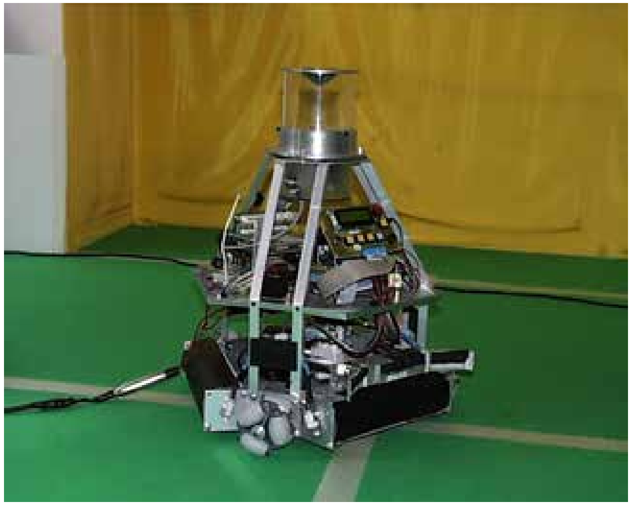
- Single sensor for measurement of brightness (cf. Braitenberg vehicle)
- Sensor fields (1D, 2D) with optics for cameras (visual sensor)
- Infrared sensor: measures temperature (alarm systems)
- Active infrared sensor for close distance measurements:
  - sends coded signals, measures reflected echo
  - similar to Sonar: no accurate measurement, cheap
  - arrangement as a ring: "non-contact bumpers“

Omnidirectional Camera

360 degrees of view

Can be realized by special (conic) mirror:

Different surface curvature for better resolution at close range



Omnidirectional Camera

Needs appropriate camera model and interpretation methods

From Wikipedia
Autor: Jahobr

Measurement of Distances

Many possibilities, e.g.

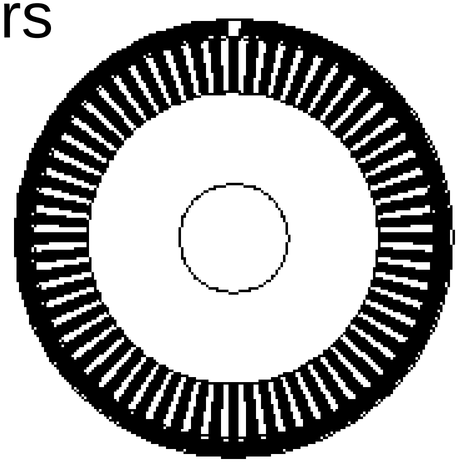
- Measure the performed path of a vehicle (wheel encoder)
- Send Signal, receive echo:
  - Time difference proportional to distance
  - Phase shift proportional to distance
- Image interpretation:
  - Size of objects reciprocally proportional to distance
  - Vertical view angle proportional to distance
  - Stereo vision: Shift proportional to distance

Sonar
Laser
Radar

Incremental Wheel Encoder

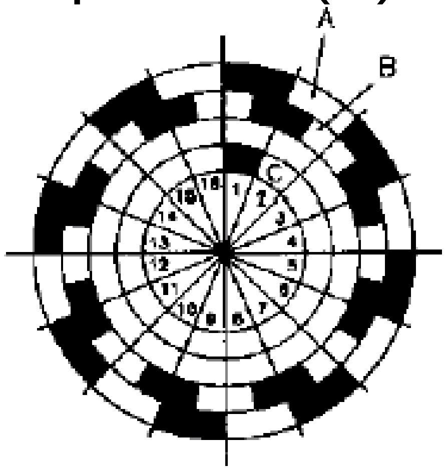
Measurement of rotation by identical markers

- speed (distance by integration)
- no wheel position, no direction
- Problem: error drifting



Multichannel Encoder:

- speed
- direction
- zero position (C)



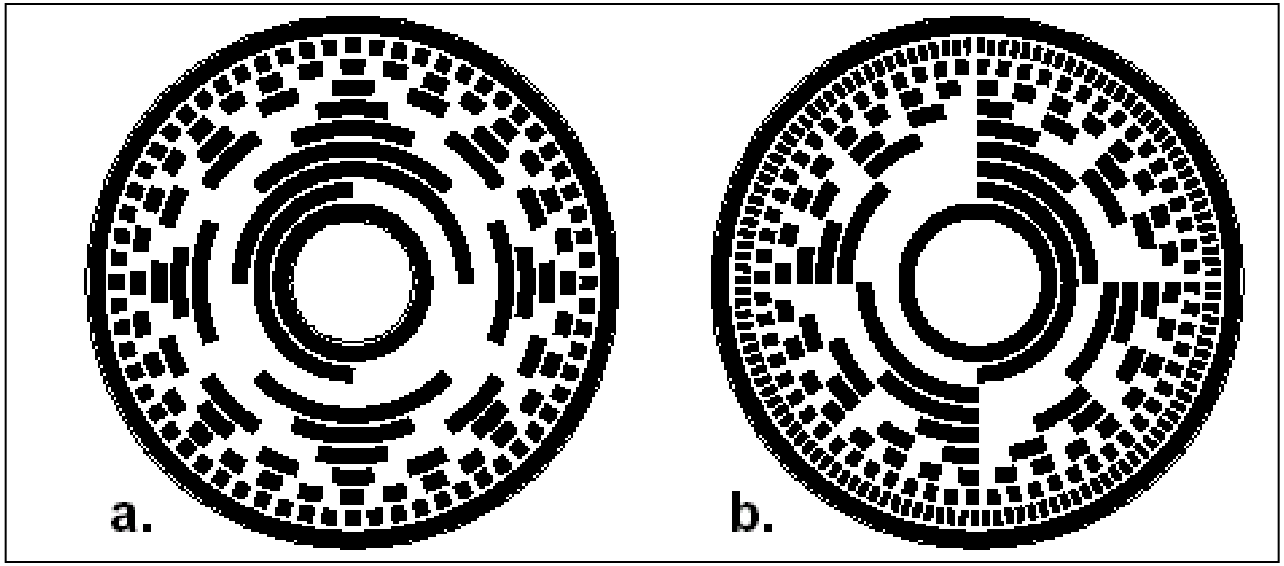
Images from “Where am I?" -- Systems and Methods for Mobile Robot Positioning by J. Borenstein, H. R. Everett, and L. Feng

Absolute Wheel Encoder

- Each position has individual word pattern

  Gray code (a),   BinaryCode (b)

- Disturbances without affecting
- 12 bits: 0.1 degree accuracy

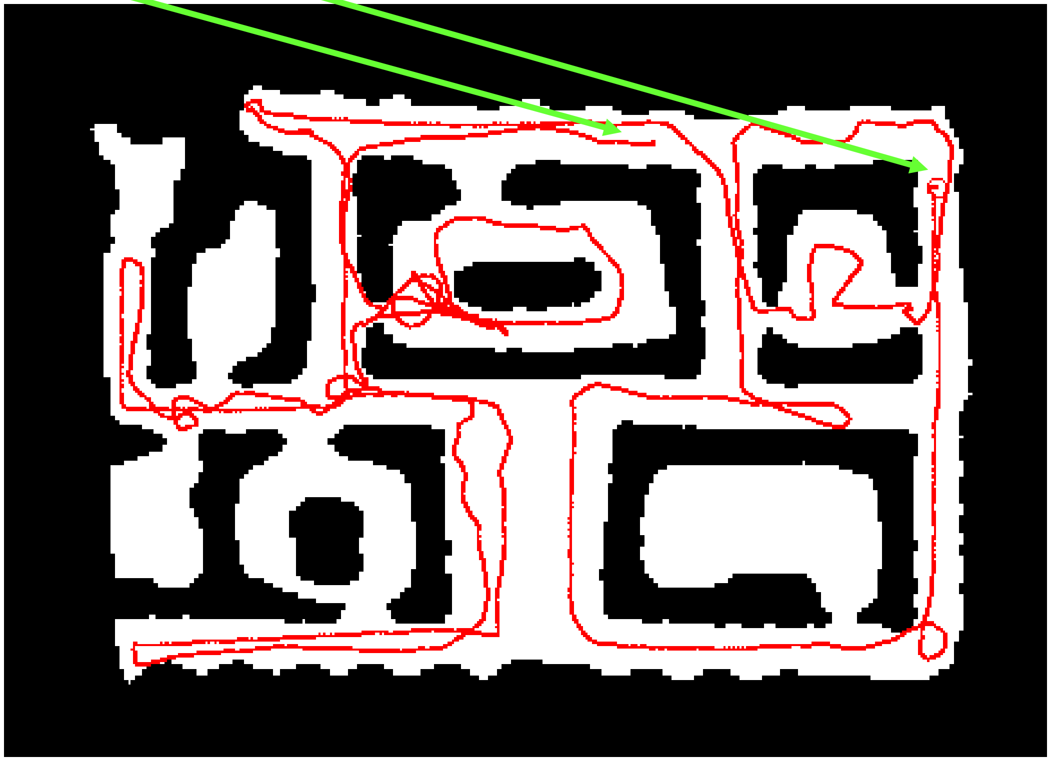


Odometry

Known start position

Actual position by measuremaent of pathes

- Wheel encoder
- Motion of legs
- Control
- Inertial sensors



Odometry: Measurement Errors

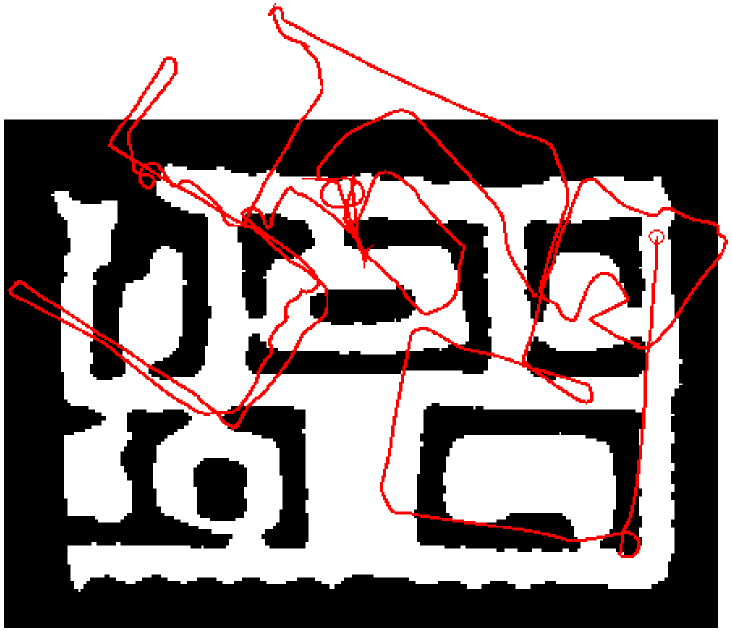
Systematic errors

- By sensors (e.g. wheel encoder)
- By controls (e.g. unsymmetric wheels)

Non-systematic errors

- Ground
- External forces (e.g. other robots)

Main problem:
Errors of direction

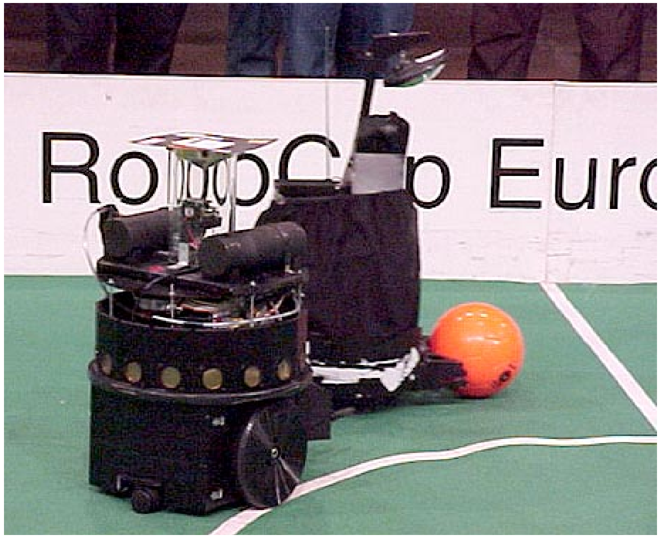


Sonar Sensors

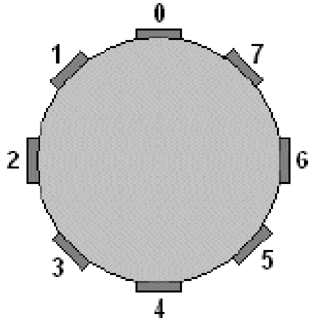
Sonar = sound navigation and ranging

Active ultrasonic sensor (> 20 kHz)

Cheap, but noisy and inaccurate



Arrangement as a ring for obstacle detection.



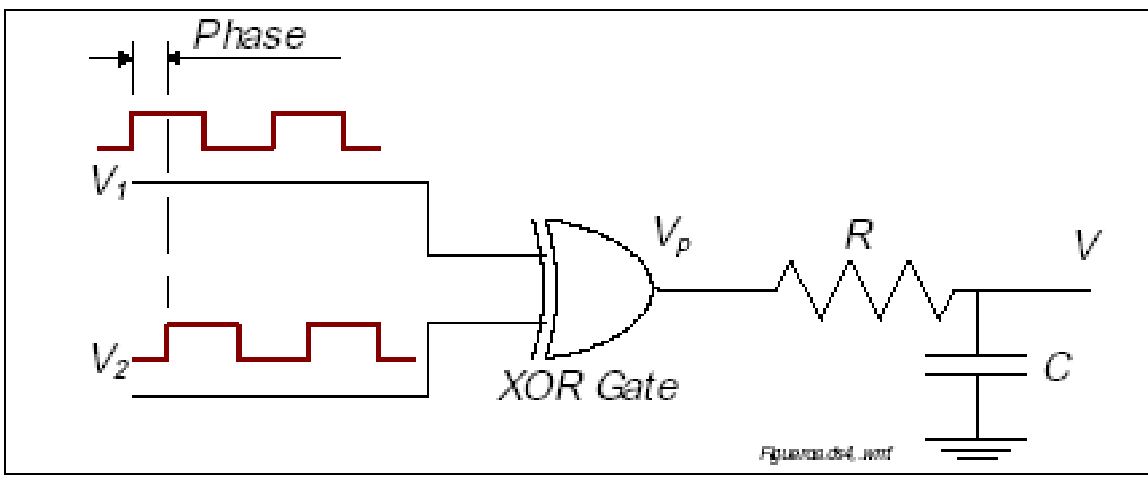
Sonar Sensors

Send pulse - receive echo:

- Time difference is proportional to the distance

alternatively:

- Phase shift proportional to the distance



V proportional to phase shift

Sonar Sensors

Sensor model:

Amplitude strength depends on the direction relative to the center of the signal

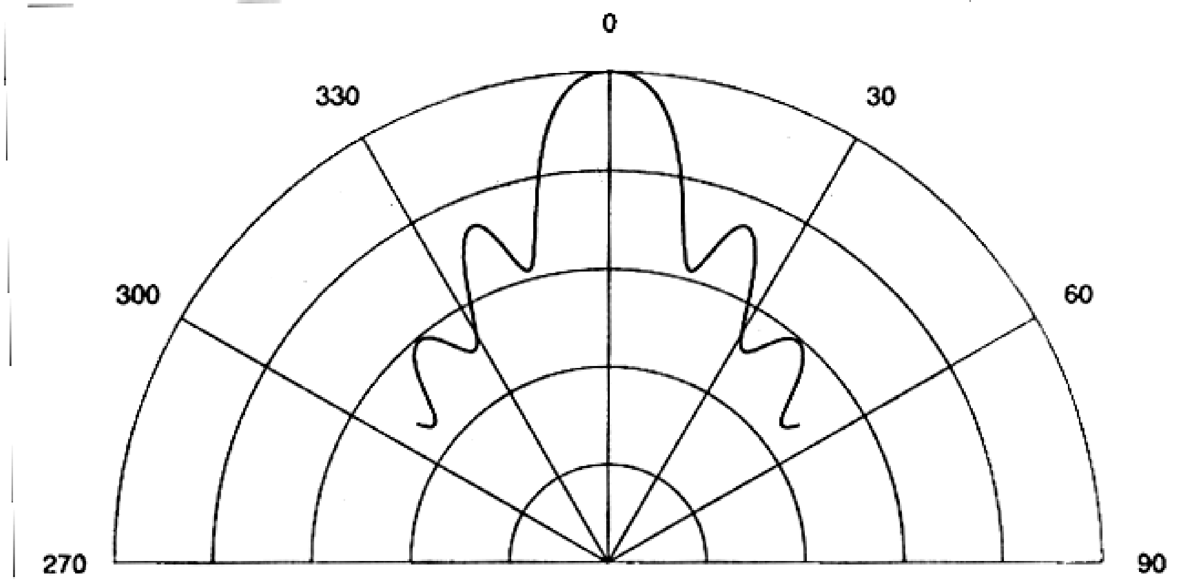
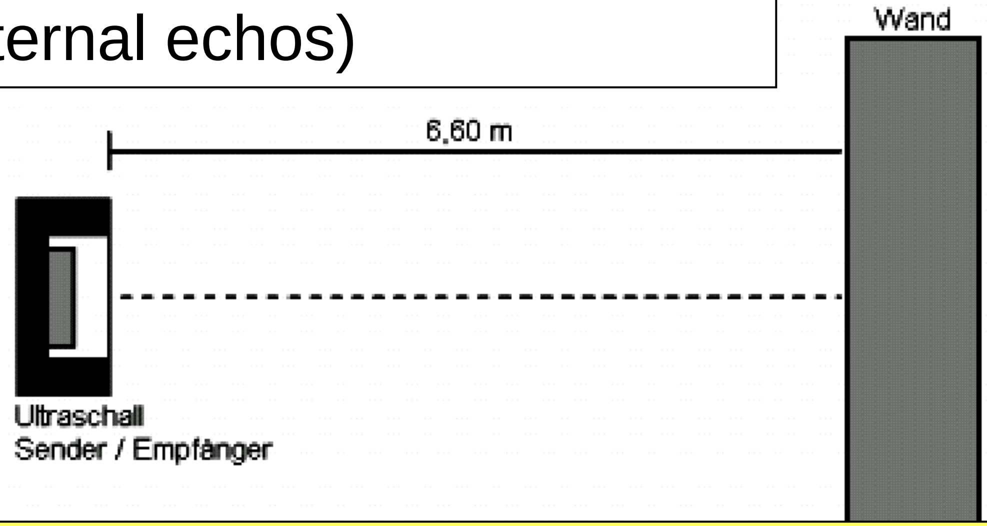


Image from “Where am I?" -- Systems and Methods for Mobile Robot Positioning by J. Borenstein, H. R. Everett, and L. Feng

Sonar Sensors

Device transmits a short sound,
then switch to work as microphon (receive echo).
No measurements in close distance (< 6cm)
(„blanking Intervall“: internal echos)



Distance (in m) $\boldsymbol{d = 0{,}5 \times c \times t}$ by echo runtime t (in s):

$$c = c_0 + 0{,}6T \; m/s$$

with $c_0 = 331\ m/s$, T=Temperature (Celsius)

Problems with Sonar Sensors

- „Crosstalk“

  Interference of reflexions:

  - Direct (a)
  - Indirect (b)

  To avoid:

  Use different frequencies and signals by the sensors.

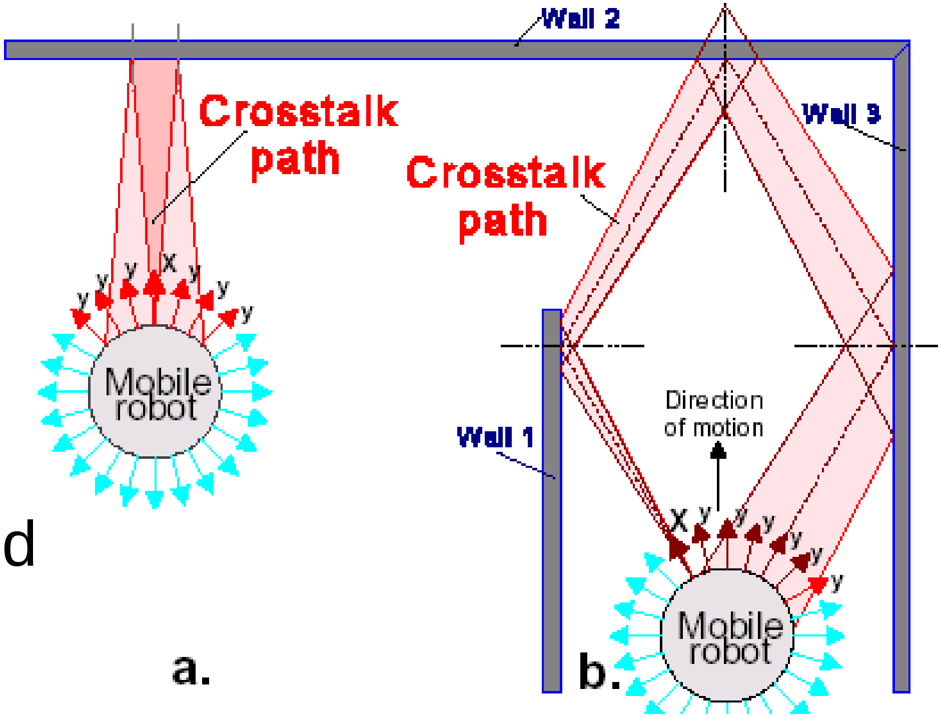
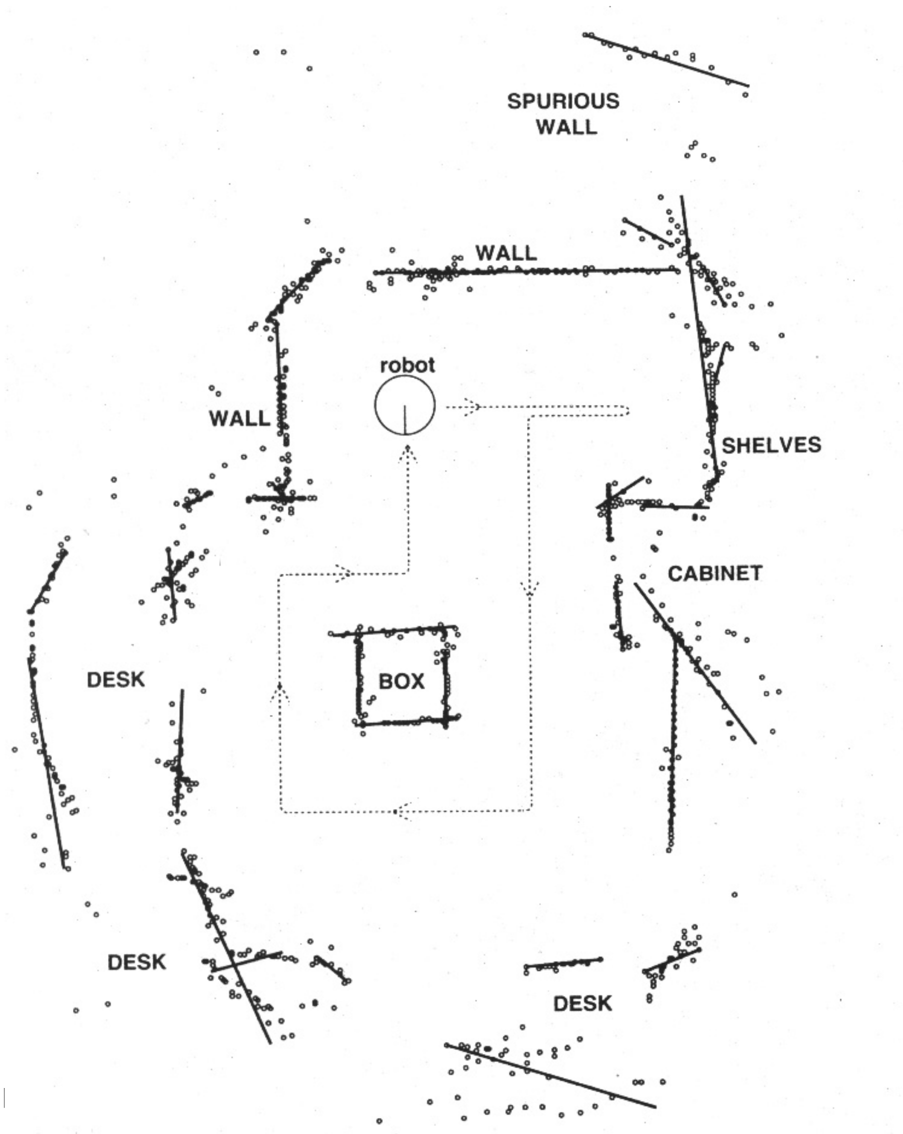


Image from “Where am I?" -- Systems and Methods for Mobile Robot Positioning by J. Borenstein, H. R. Everett, and L. Feng

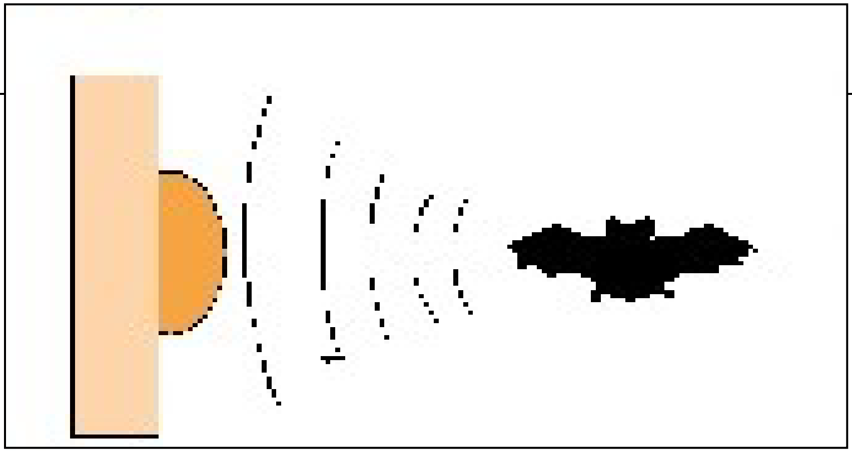
- Missing reflection
- Multiple reflection

Problems with Sonar Sensors

Sonar measurements of a robot driving on the depicted path



Sonar Sensors



Ultrasound organs in nature:

Dolphins,

Bats.

Bats use different frequencies and can identify flying insects.

- very complex skills
- not yet fully investigated

Laser Sensor

Active sensor using echo of laser impulses

Laser = **l**ight **a**mplification by **s**timulated **e**mission of **r**adiation

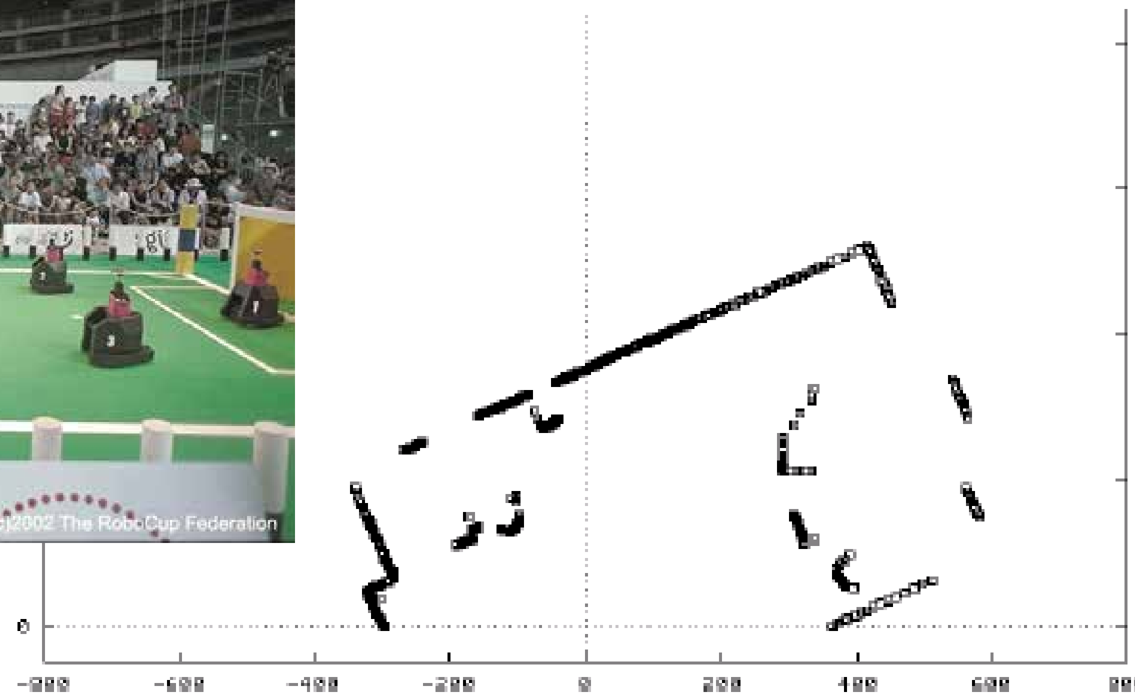
High intensity with short pulse

Different forms of production

- Very accurate distance measurement
- Very high range
- Short sampling time: even at high speeds
- More expensive devices

Laser Sensors

Detection of a RoboCup field by a Midsize League Robot



Laser Sensor

Different methods:

- Time of fly
- Phase shift
- Triangulation
- Blur

Problems:

- Multiple reflection
- No echo at transparent objects (glass)
- Eye sensitivity

Stereoscopy

2 camera images from different positions

(by 2 cameras or a moving camera)

Calculation of distances by different view angles of objects

Correspondence problem:

Which objects/pixels belong together?

Comparison of image features

Correlation methods

Distance measurement with structured light

Patterns (e.g. stripes) are projected to the surface of an object.

The patterns are distorted by the geometry of the object.

Images are perceived by one or more cameras from the sides.

By related calculation, 3D point clouds are constructed.

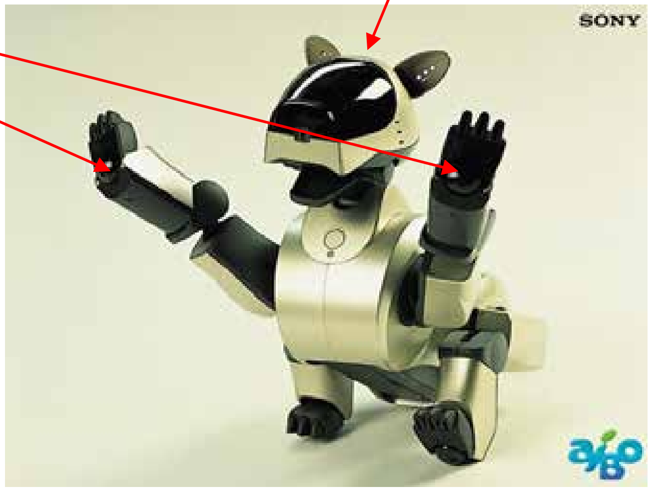
Kinect uses infra red.

Force Sensors

Transformation of force into electronical signals:

Change of electrical properties
(e.g. resistance, capacity, inductivity)
by mechanical deformation caused by forces

- Touch sensors (hand, feet, artificial skin …)
- Collision detection (bumper)
- Coupling with actuators



Inertialsensor: Accelerometer

Measures linear acceleration

- Inertial sensor (needs no contact with outside world).
- Must regard gravity.

Measurement of position by gravity.

Inertialsensor: Gyroscop

Measures rotation (forces caused by changing direction)

- Internal sensor (needs no contact with outside world).

Problems: Drift over time, Earth's rotation

Inertial System

Combination of accelerometer and gyroscop:

Measurement of linear and rotational motions.

Odometry using **only internal sensors:**

Calculation of the path from starting point

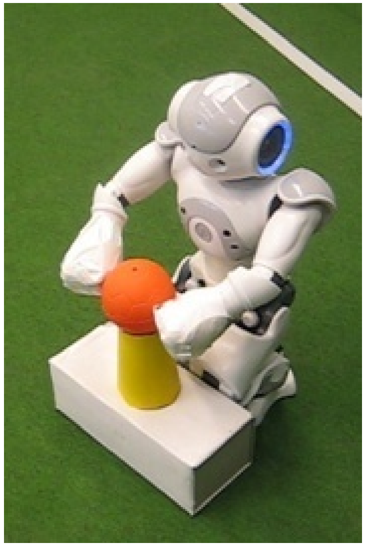
(Measurement of speed and path by integration over time)

Energy Consumption

Conclusion to external forces
by measuring the needed energy consumption (current)
or the generated heat.

Possibilities for feedback control.

Weight of objects
by current needed at
the shoulder joints



Acoustic Sensors: Microphone

Transformation of sound waves (forces) into electrical signals (e.g. membrane in magnetic field)

Time-dependent signals (limited polling frequency)

Noisy (internal, external noise)

Applications:

- noise detection
- speech recognition
- bearing (echo)

Language processing

"AI-hard": Turing Test.

Similar to image processing.

Complex process with many levels:

- Preprocessing
- Identification of sounds, syllables, words
- Identification of relationships
  (e.g. dereferencing pronouns)
- Interpretation: Identification of meanings/intentions

Requires knowledge about

- Sentence structure, grammar, syntax, ...
- Relationships, contexts, ...

Requires knowledge **about the world**

Outline

Introduction

Sensors: General Considerations

Signals

Sensors: Special Types

Vision (introductory)

Camera Model

Image Processing (introductory)

Scene Interpretation (introductory)

Vision

Magritte

Vision

Humans use about 50% of brain

for image processing and interpretation

Preprocessing already performed in the eyes

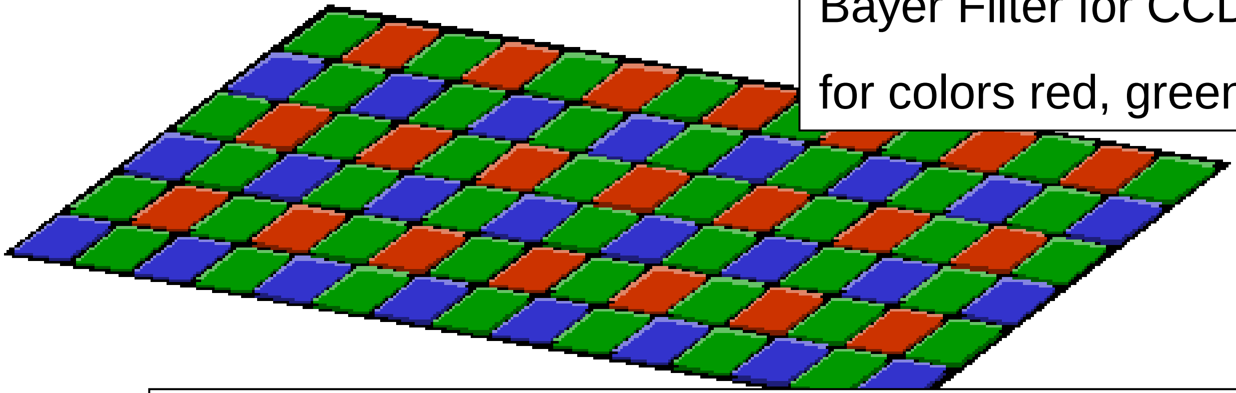
Optical Sensors

Light sensitive elements (e.g. CCD = Charge Coupled Device)

Arranged in form of a matrix with filters for different colors.

Result stored in a pixel matrix („frame“).

Short intervals: e.g. 30 frames per second (fps).



Bayer Filter for CCD (from Wikipedia)

for colors red, green and blue (RGB-system)

Brightness/contrast perception in human eyes especially by green (72%)

Optical Sensors

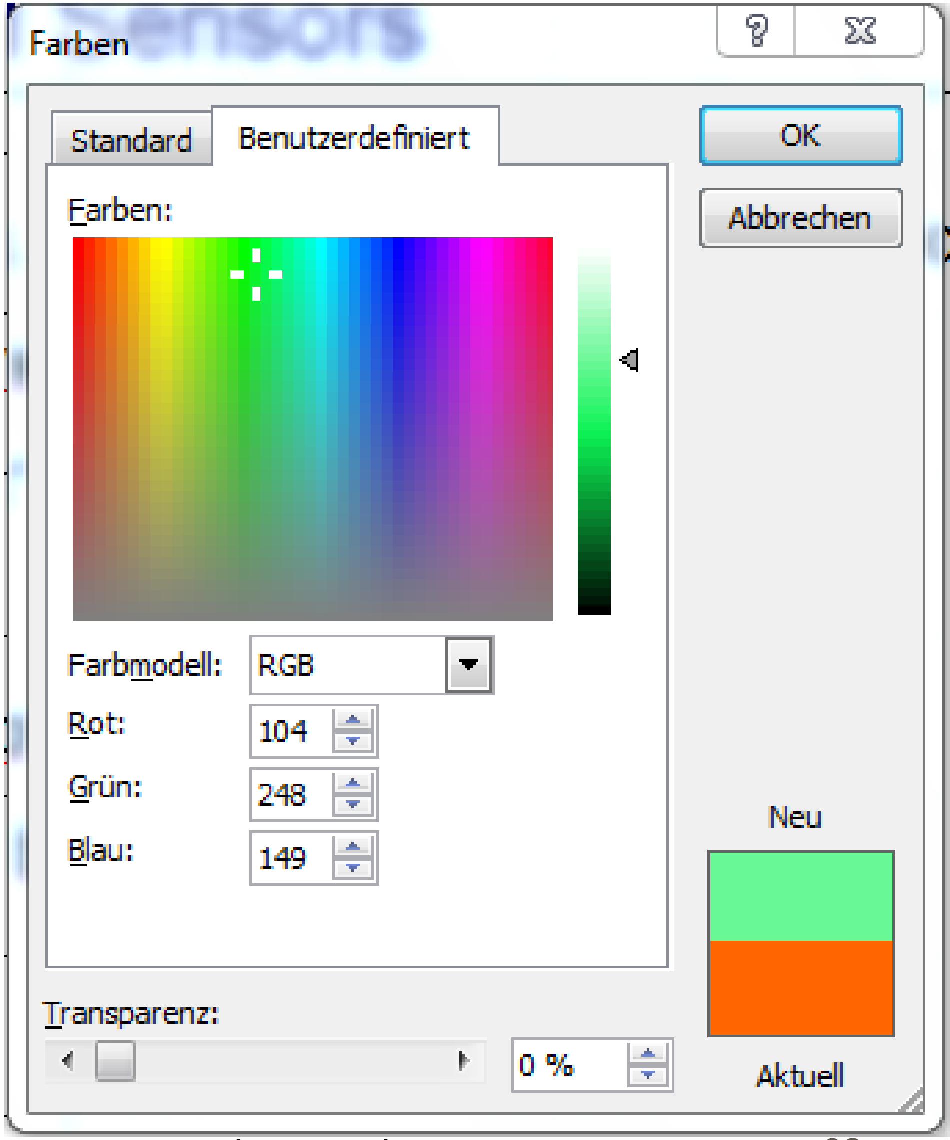
„Pixels“ = picture elements

Usually three color values

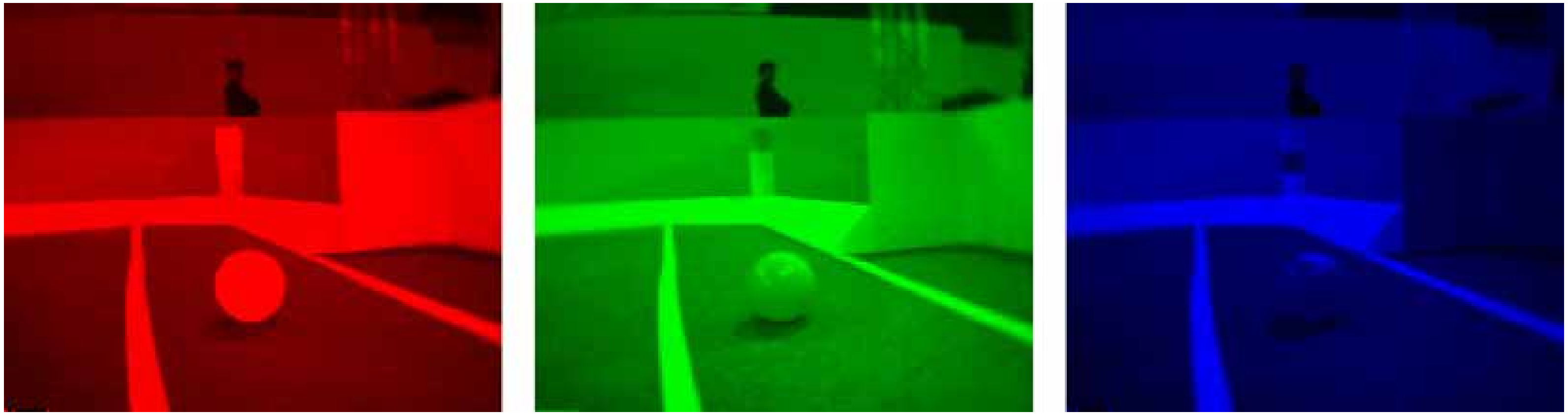
for intensity of **red, green, blue**

(higher value = more intensity)

RGB-system

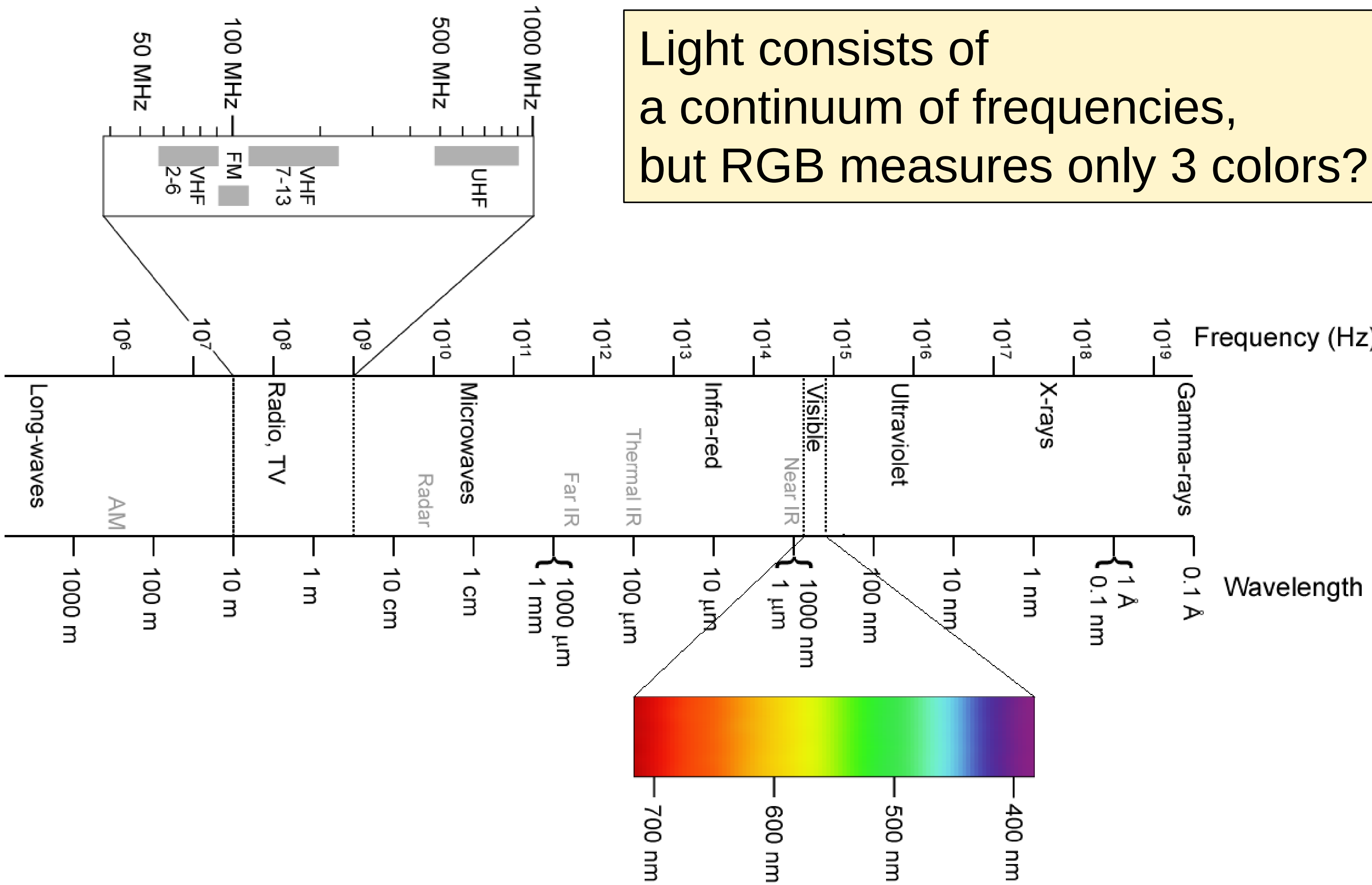


Color Channels in RGB-System

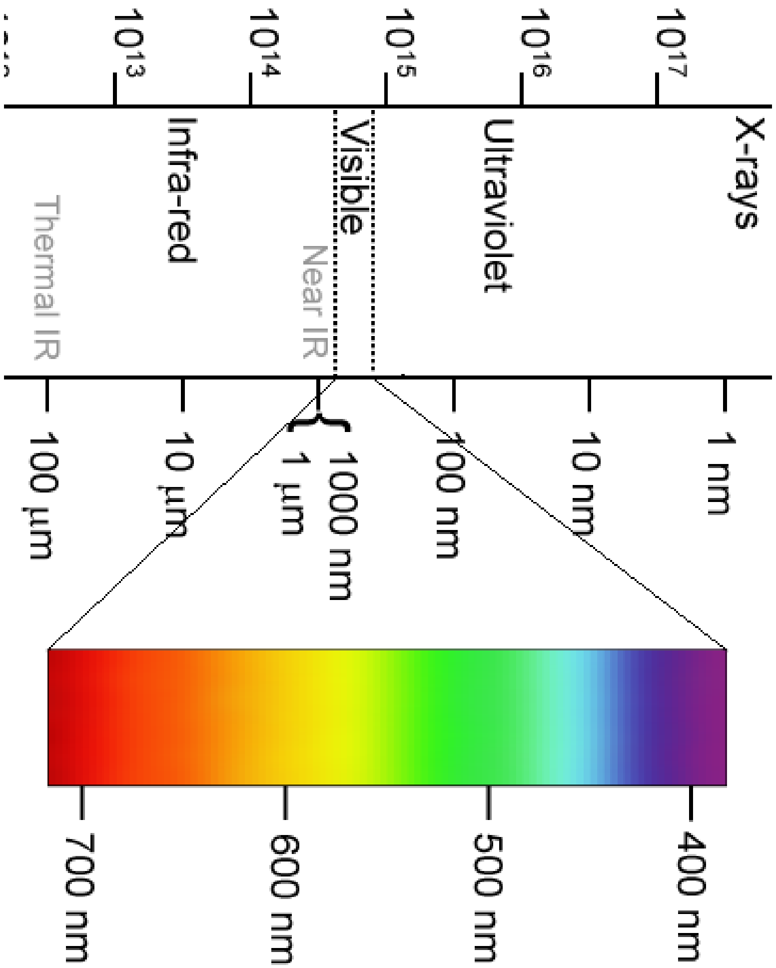


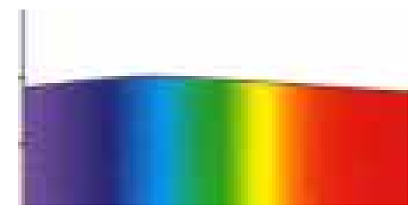
Spectral Colors vs. RGB

Light consists of
a continuum of frequencies,
but RGB measures only 3 colors?

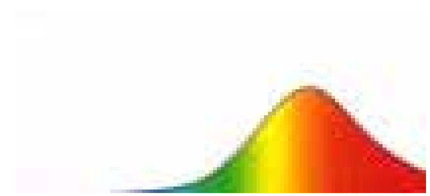


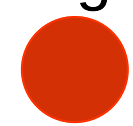
Spectral colors vs. RGB





Intensities of day light

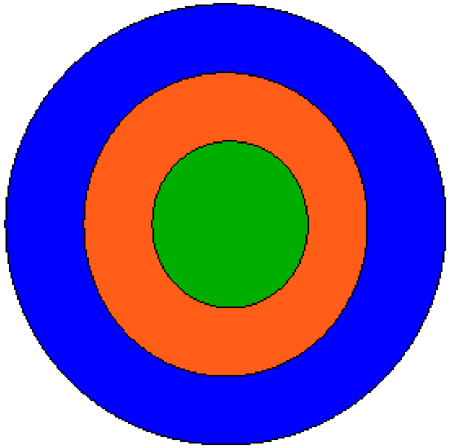


Intensities of a light that looks red 

Human eye …

… has only three types of color sensors („**cones**“)

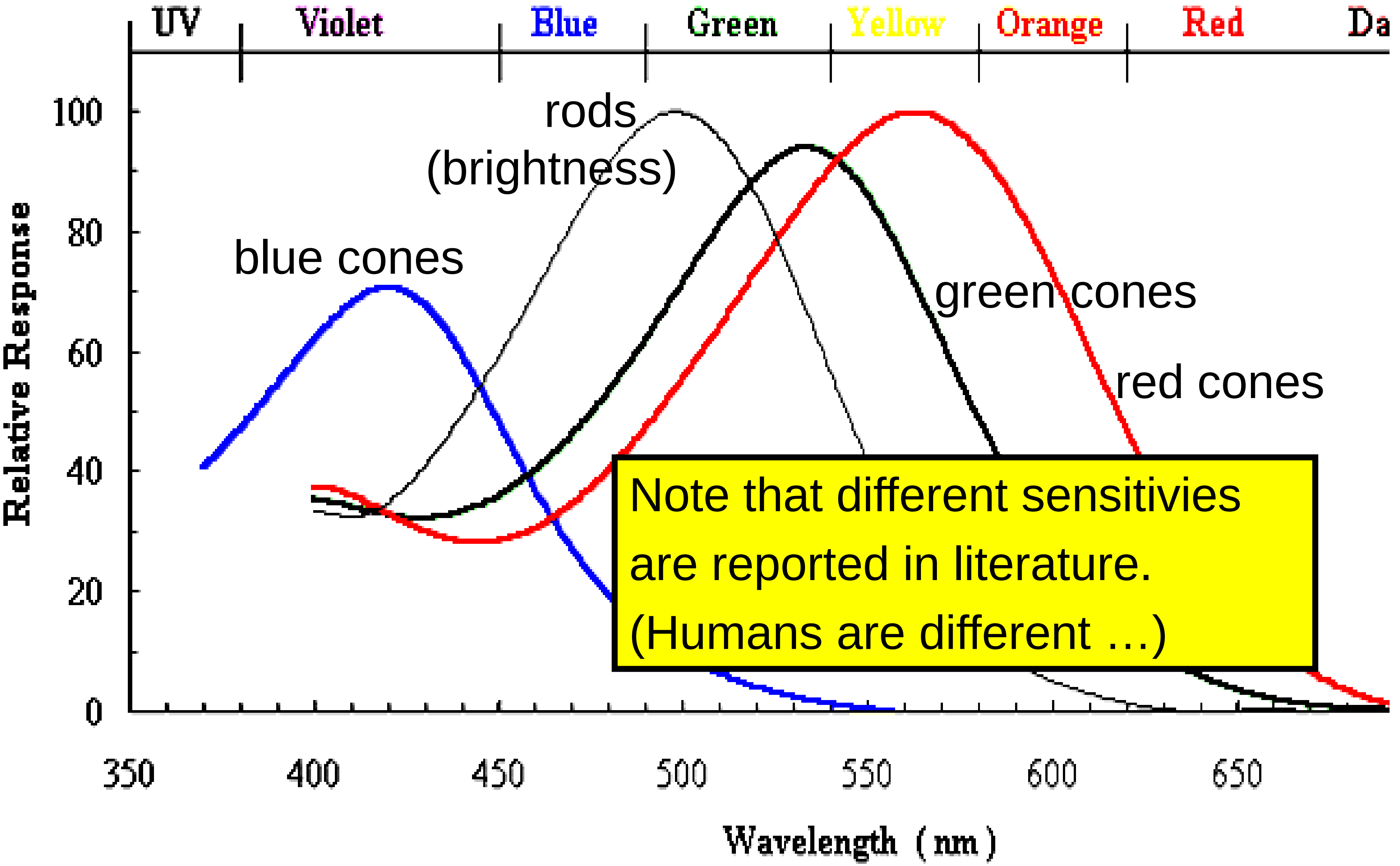
- red (64%) middle area
- green (32%) central area
- blue (4%) peripheral area



Blue text is exhausting to read

and additional light intensity sensors („**rods**“)

Sensivity of human eye sensors



Note that different sensitivies are reported in literature.
(Humans are different …)

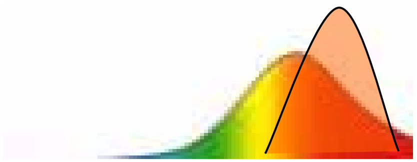
Response by Sensors

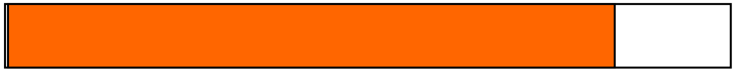
Sensor response *e* depends

on light intensity I and sensor sensitivity f for all frequencies λ :

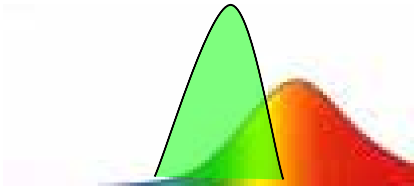
$$e = \int f(\lambda) I(\lambda) d\lambda$$

Example:

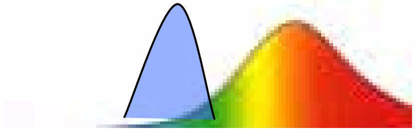




Response *e* for red sensitive sensor



Response *e* for green sensitive sensor



Response *e* for blue sensitive sensor

RGB-System

Different light may have identical RGB values:
Metamerism of light.

RGB tries to mimic human eyes
and therewith to produce acceptable rendering.

But:
Those colors don‘t „exist“ in nature,
they are only physiolocically grounded
(by individuals!).

Other color systems
for other applications
(YUV, CMYK etc.)

RGB-System

R=700 nm

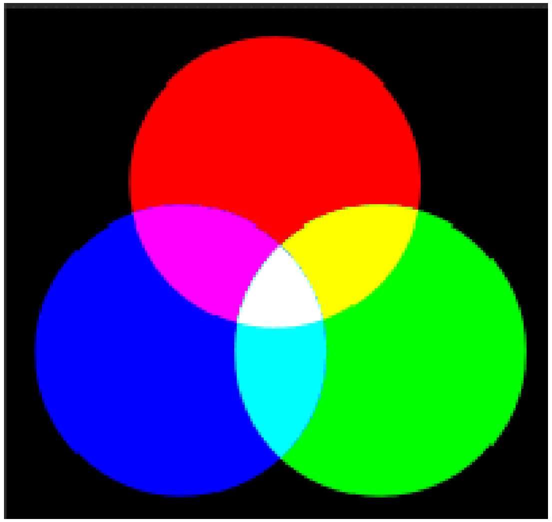
G=546,1 nm

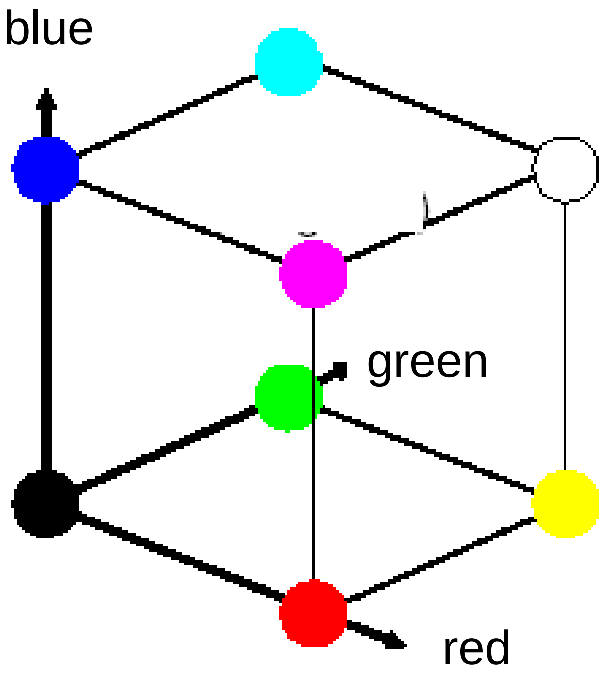
B=435,8 nm

Additive Model (3 dimensions)

Used for aktive media (e.g. displays)

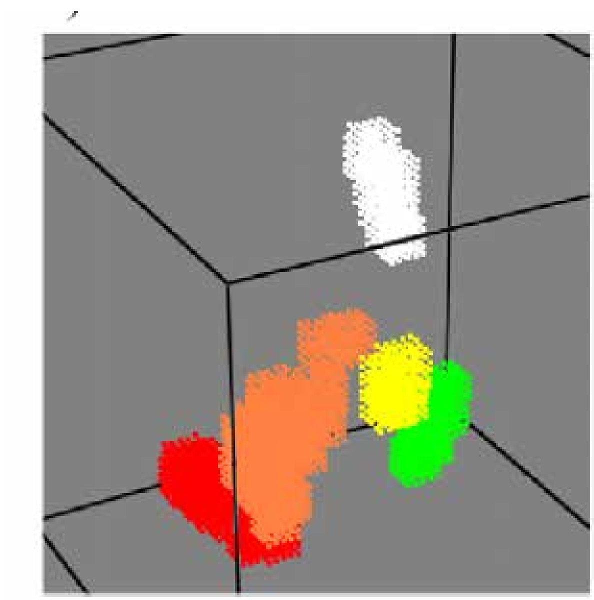
Spectral intensities are added

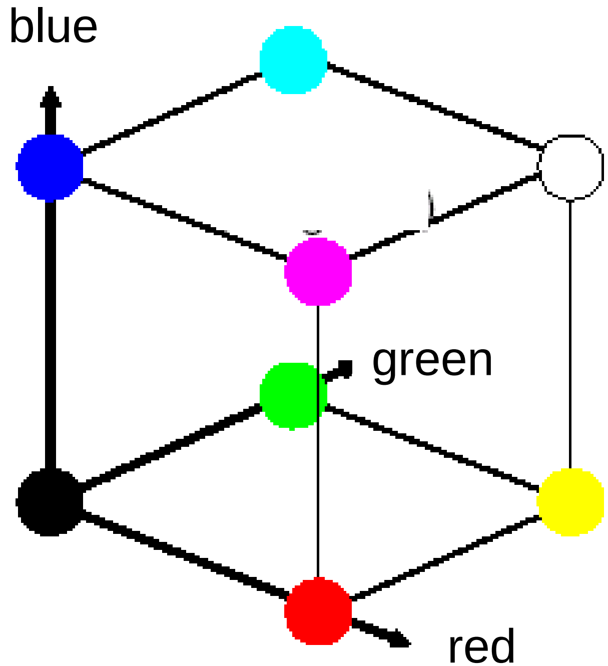




Color Classification

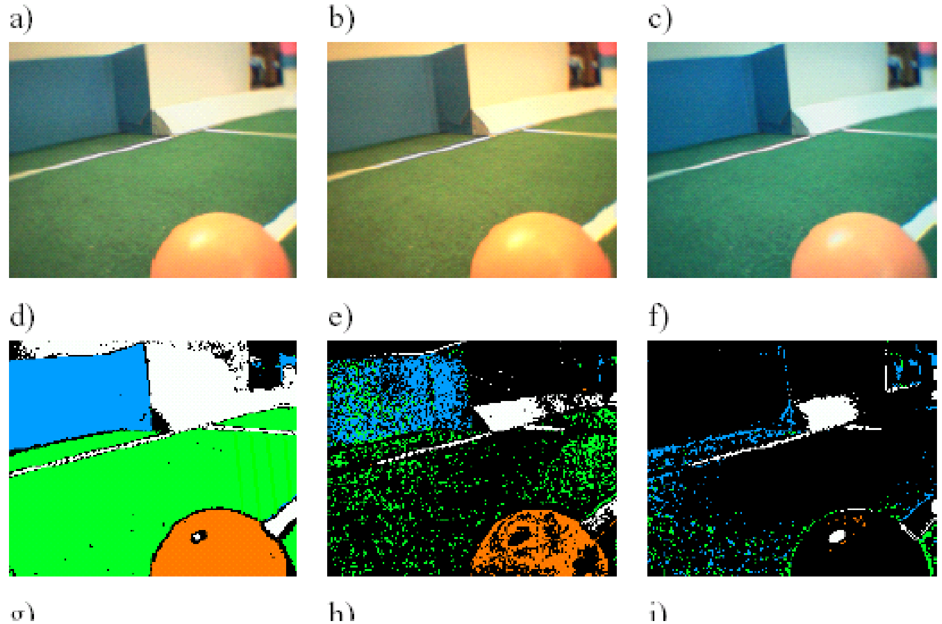
By regions in the RGB-Room





Problems with Colors

Distortion of colors by lighting and preprocessing in the camera



(a, b, c): images taken under different lighting conditions
(d, e, f): resulting color classifications by unique parameters
(from Diploma thesis Matthias Jüngel)

Adaptation/Calibration

Human color perception adapts to changing conditions.

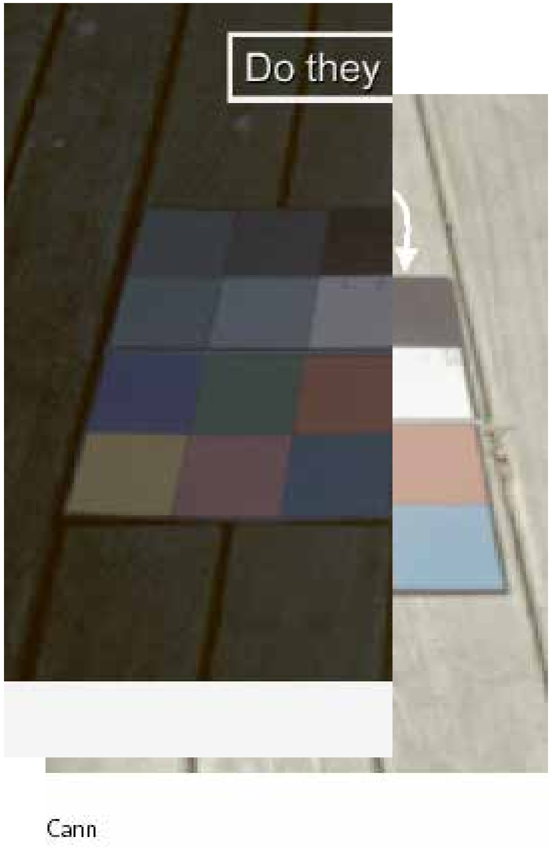
This may result in illusions.

Bild: John M. Canm

Adaptation/Calibration

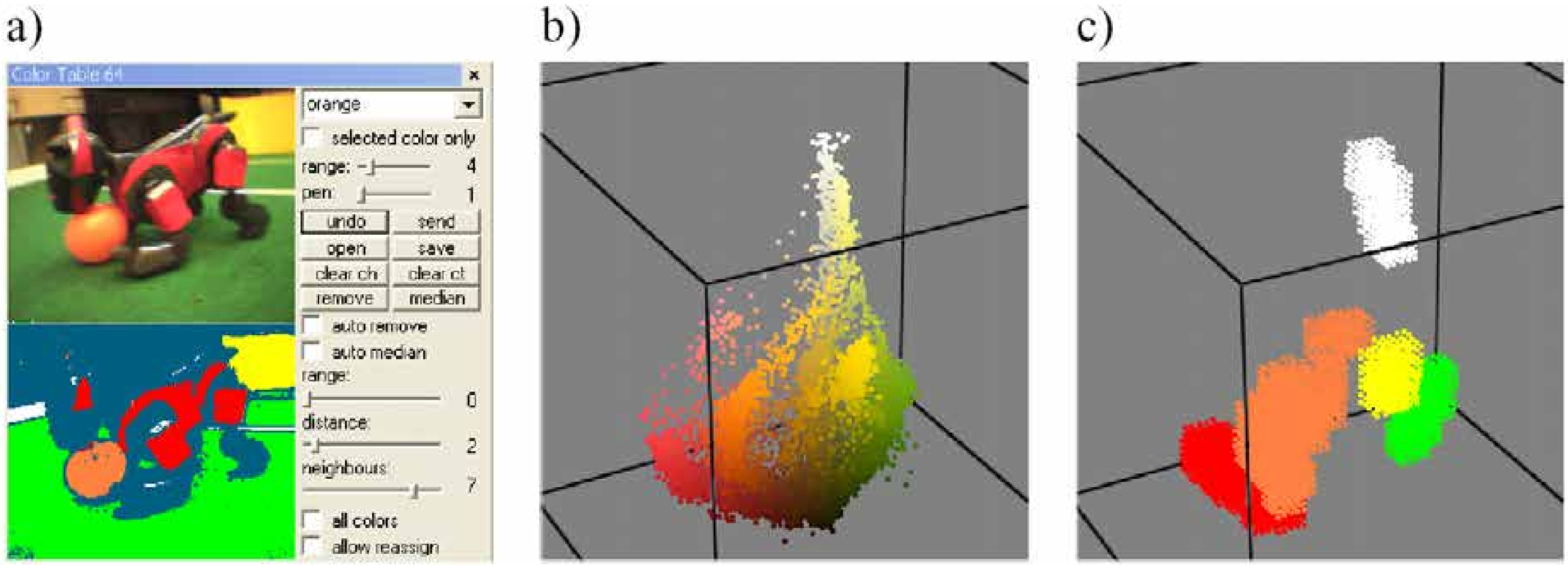
Human color perception adapts to changing conditions.

This may result in illusions.



Color Calibration

Tools for manual calibration



Different approaches for automatic calibration

Outline

Introduction

Sensors: General Considerations

Signals

Sensors: Special Types

Vision (introductory)

Camera Model

Image Processing (introductory)

Scene Interpretation (introductory)

Camera Model (Colors)

A complete color model had to regard:

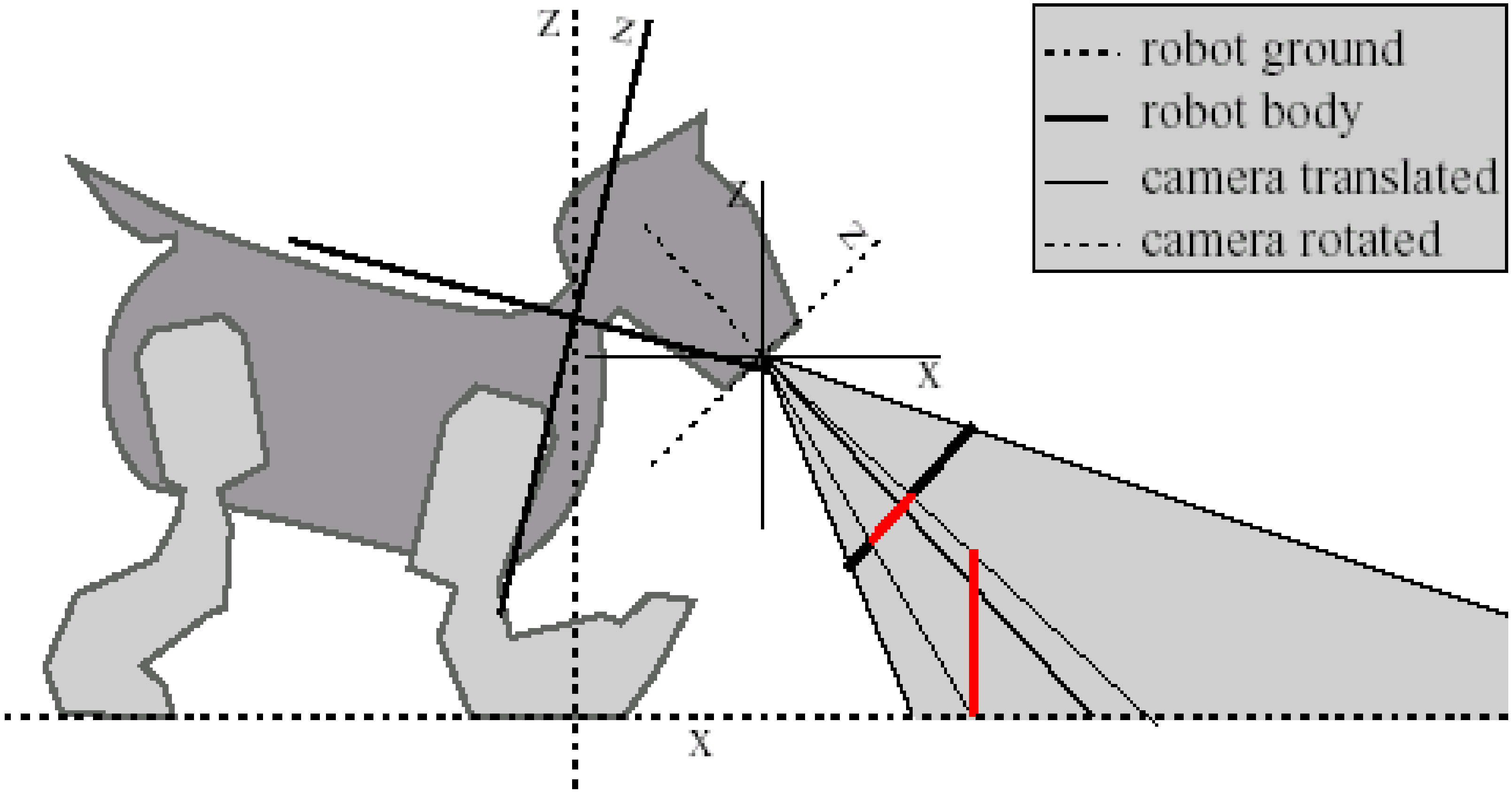
- The sources of illumination:
  - Their spectral characteristics (frequencies, intensities)
- The illuminated objects:
  - Their characteristics w.r.t. absorbance/reflection (directions, frequencies, intensities)
- The spatial relations between all sources/objects.

Very complex calculations: Only simplified models.

Color spaces like RGB are not exact models.

Difficulties in calibration.

Camera Model (Geometry)

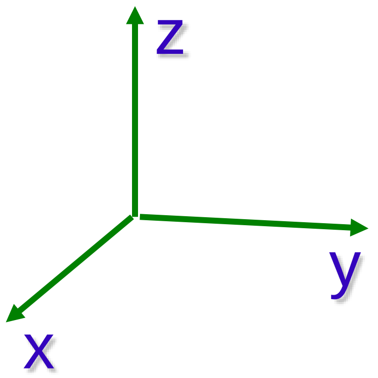


Diploma thesis Matthias Jüngel

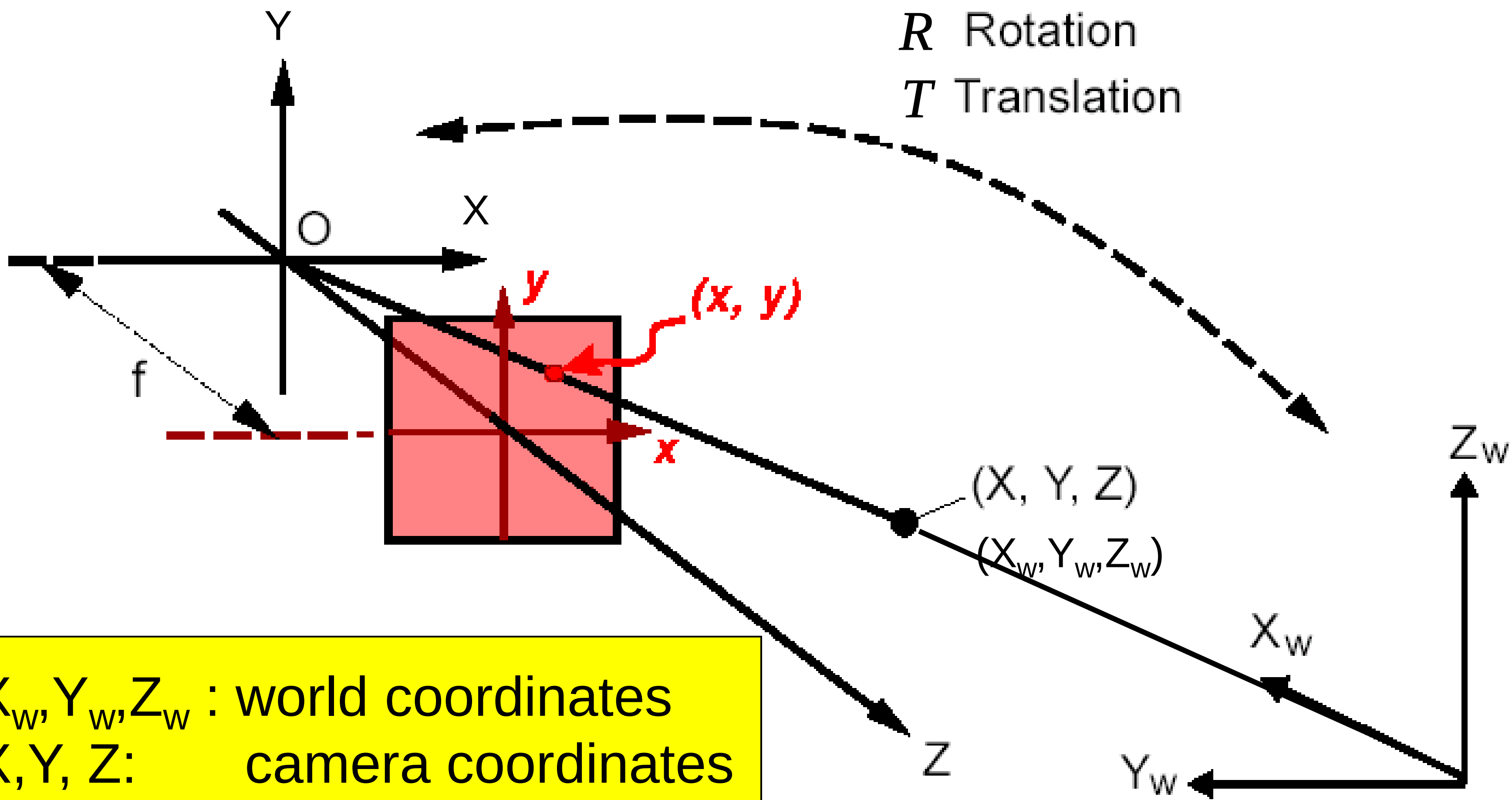
Conventions

Conventions:

- right hand coordinate systems
- angles are measured counter clockwise
- orthogonal matrices,
  hence $R^{-1} = R^{T}$



Camera Model



X_w,Y_w,Z_w : world coordinates
X,Y, Z: camera coordinates
x,y : image coordinates

Image from "Where am I?" -- Systems and Methods for Mobile Robot Positioning by J. Borenstein, H. R. Everett, and L. Feng

Camera Parameters („Camera Matrix“)

Extrinsic parameters:

Pose w.r.t. world coordinates X_w, Y_w, Z_w :

- Location of camera (camera center, 3 DOF)
- Orientation (3 DOF).

Camera Coordinates X,Y,Z with origin in camera center

direction of Z is optical axis

Intrinsic parameters:

Position of image plane (w.r.t. camera coordinates)

- Distance of image plan f (1 DOF)
- Intersection point of optical axis (2 DOF):

Image coordinates x, y

with origin at Z-axis and orientation parallel to XY plane

Perspective Projection (Central Perspective)

The image coordinates (x, y) are uniquely determined by Camera coordinates (X, Y, Z).

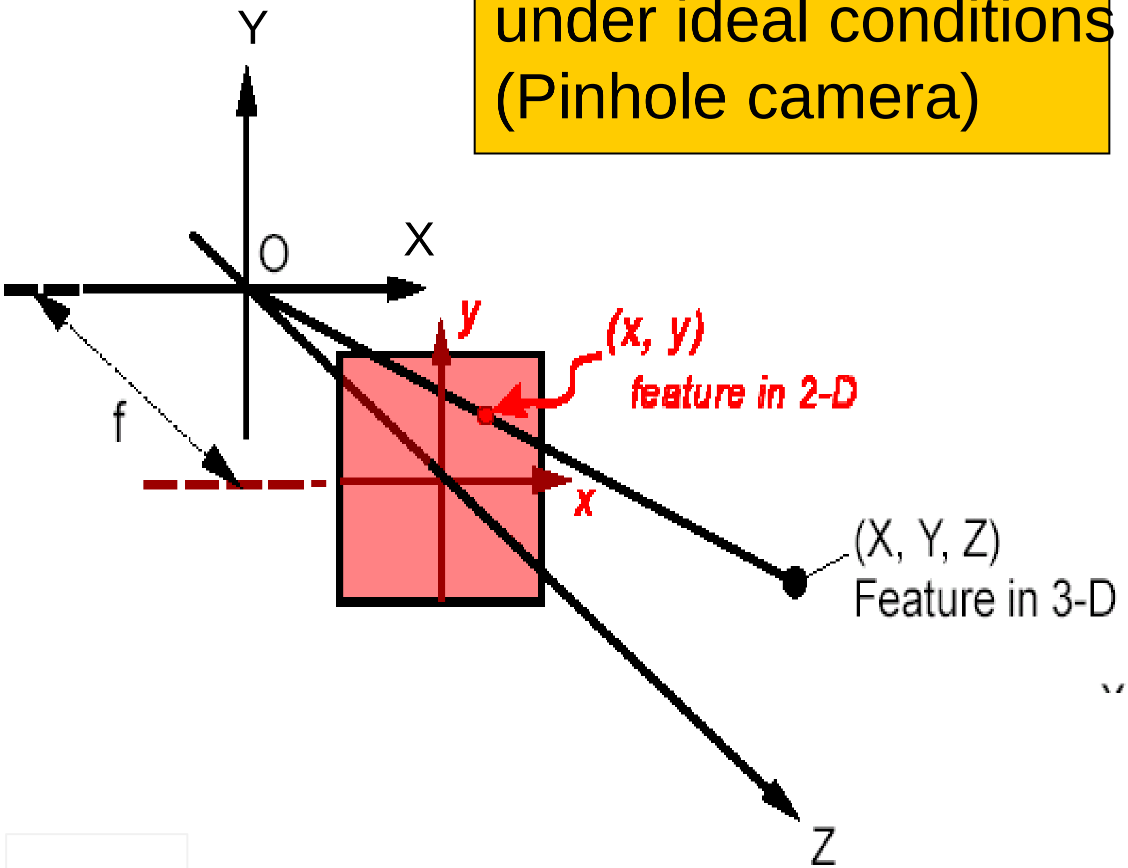
Exactly valid only under ideal conditions (Pinhole camera)

Intercept Theorem

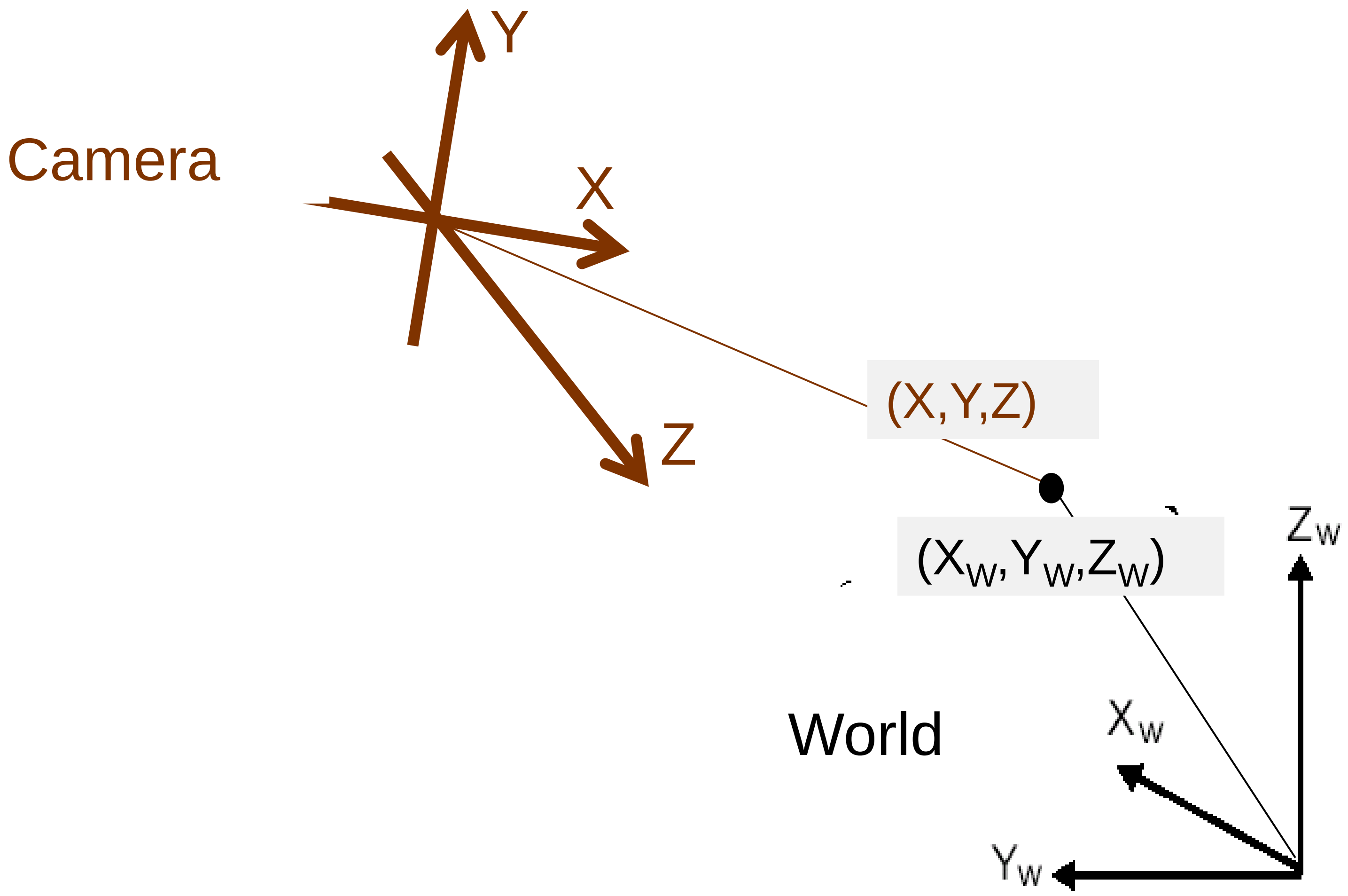
$Z : f = X : x = Y : y$

$x = f/Z \cdot X$

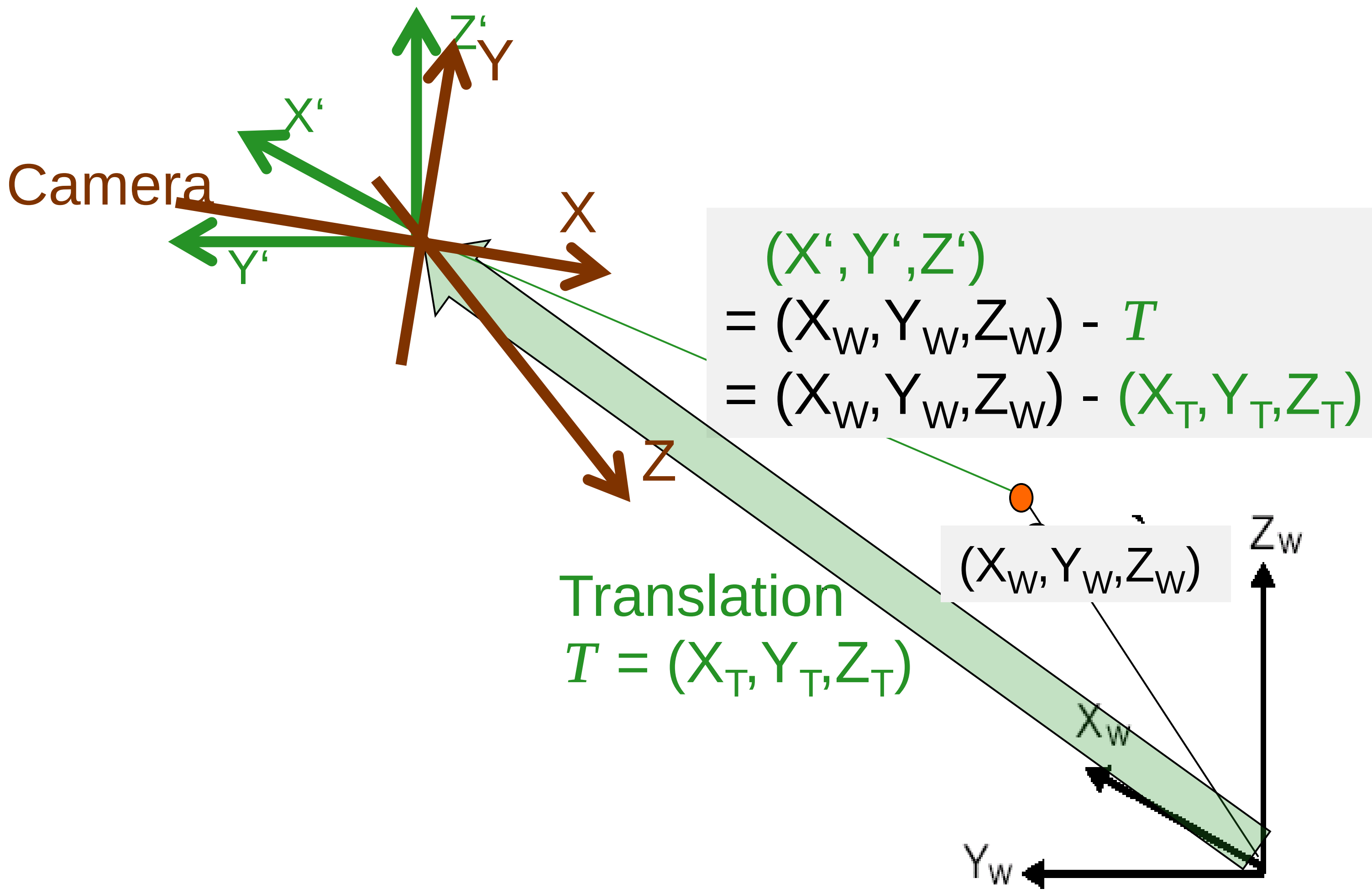
$y = f/Z \cdot Y$



From World to Camera Coordinates



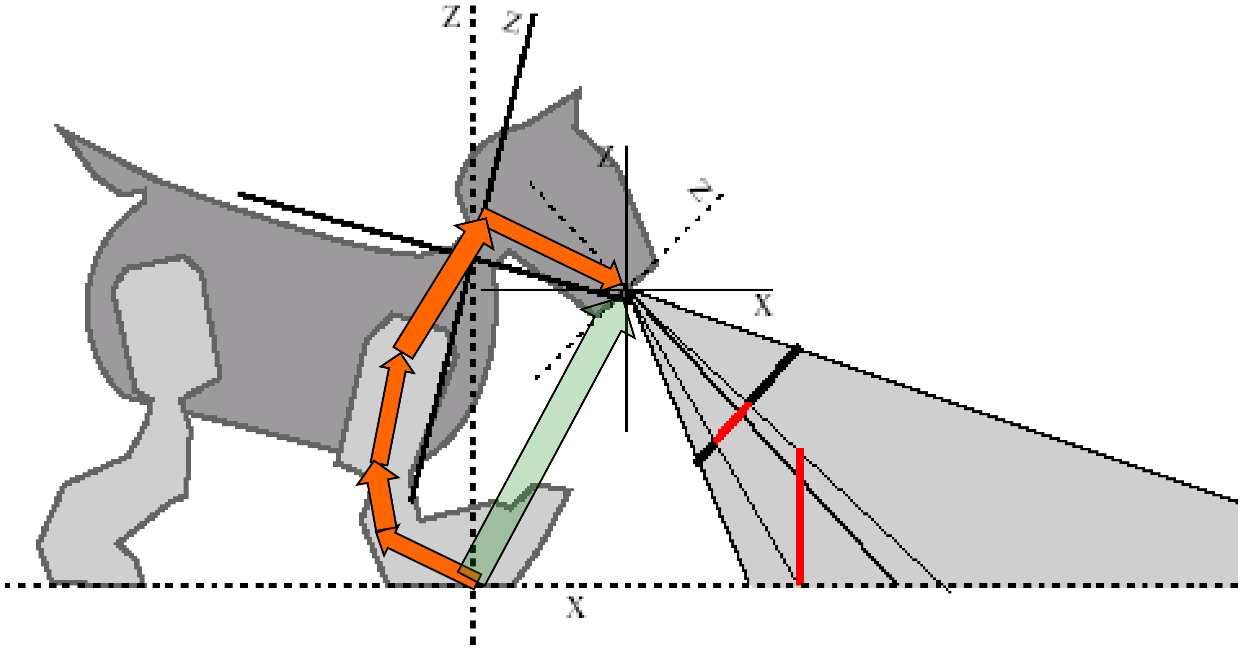
From World to Camera Coordinates: Translation



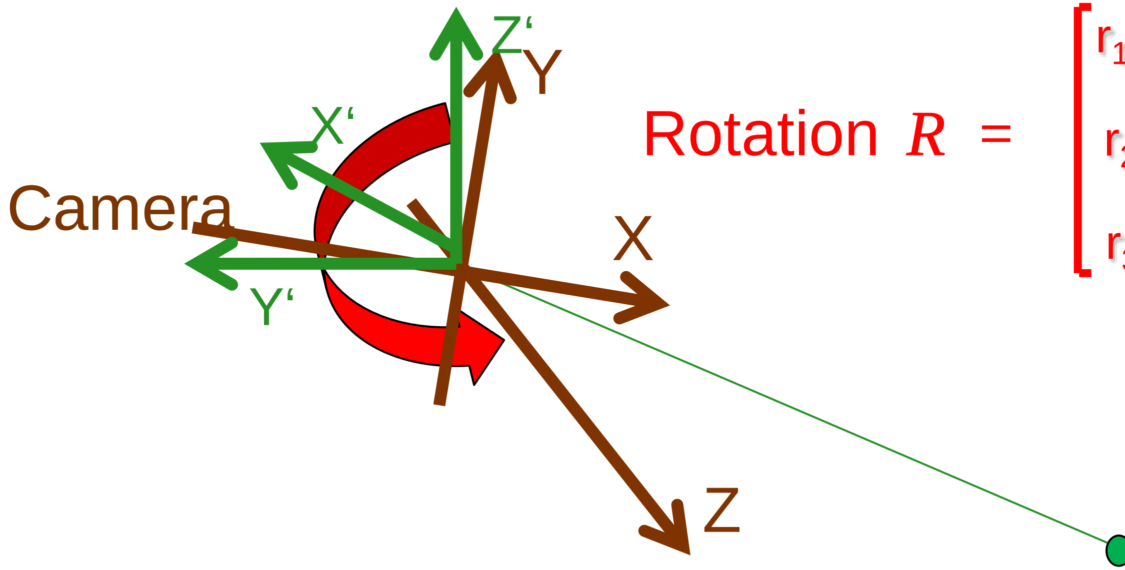
From World to Camera Coordinates: Translation

Usually, the translation vector T is not directly known.

It must be computed along the „Kinematic Chain“, i.e. with calculations by translations along limbs and rotations in joints.



From Translated World to Camera Coordinates: Rotation



Rotation $R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$

(X,Y,Z)

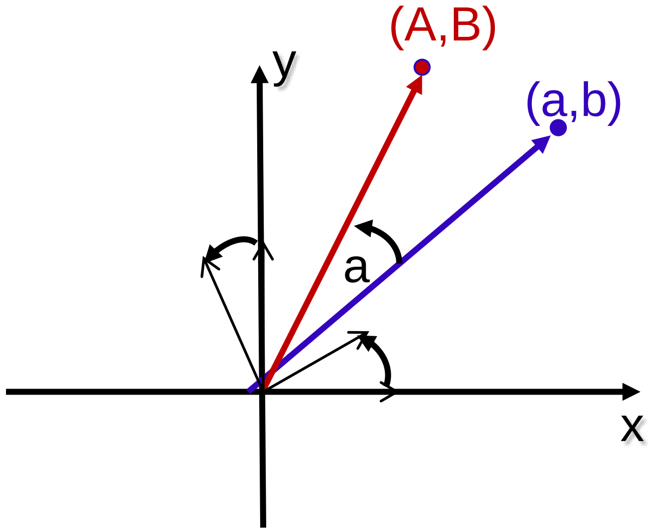
$= R \times (X',Y',Z')$

$= R \times ((X_W,Y_W,Z_W) - T)$

$= R \times ((X_W,Y_W,Z_W) - (X_T,Y_T,Z_T))$

Rotation in 2D Euclidean Space

Rotation of a point by angle a from (a,b) to (A,B)



Rotation matrix

$$R = \begin{bmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{bmatrix}$$

$$\begin{bmatrix} a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ changes to $\begin{bmatrix} \cos a \\ \sin a \end{bmatrix}$

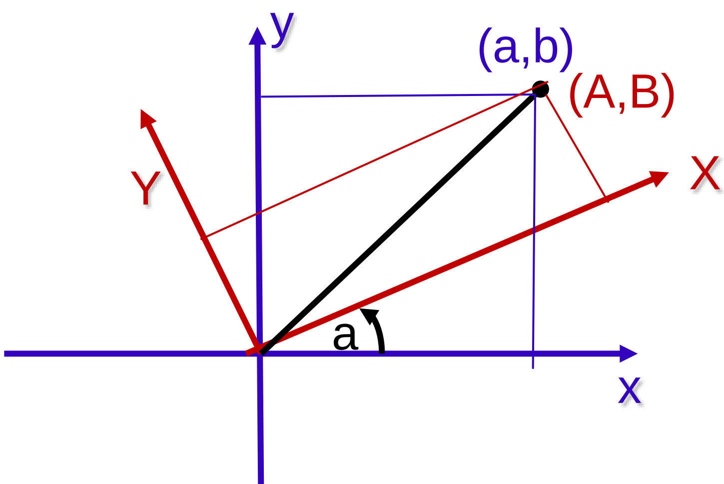
$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ changes to $\begin{bmatrix} -\sin a \\ \cos a \end{bmatrix}$

$\begin{bmatrix} a \\ b \end{bmatrix}$ changes to $\begin{bmatrix} a \cos a - b \sin a \\ a \sin a + b \cos a \end{bmatrix}$

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \cos a & -\sin a \\ \sin a & \cos a \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

Rotation in 2D Euclidean Space

Rotation of a the old coordinate system (lower letters x,y, blue) by angle a into new rotated coordinate system (capital letters X,Y, red).
It changes coordinates of a point from (a,b) to (A,B)



Corresponds to rotation of the point (a,b) by inverse rotation R^{-1}

Corresponds to rotation of the point (a,b) by angle - a

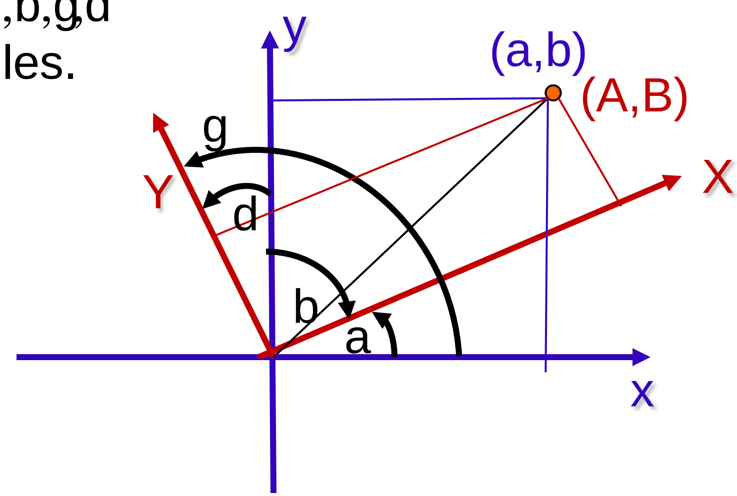
$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} \cos a & \sin a \\ -\sin a & \cos a \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

Rotation matrix

$$R = \begin{bmatrix} \cos a & \sin a \\ -\sin a & \cos a \end{bmatrix}$$

Rotation in 2D Euclidean Space

Angles between old and new axis: a,b,g,d
Direction cosine: cosine of those angles.



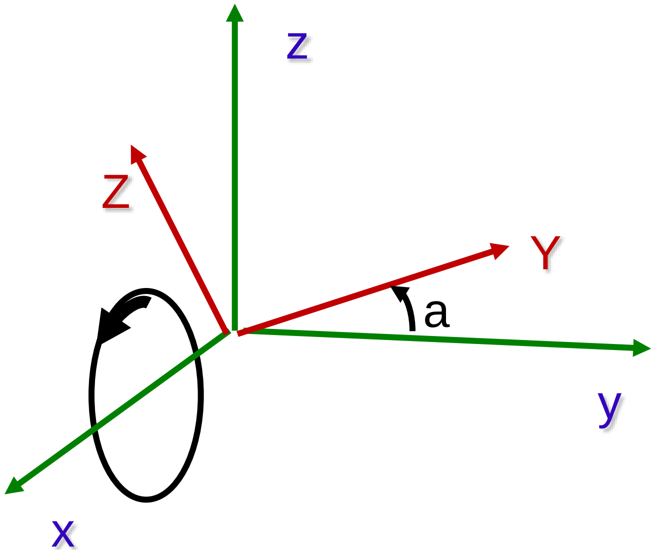
b = p/2- a
g= p/2+a
d = a

cos b = - sin a
cos g = sin a

$$R = \begin{bmatrix} \cos a & \sin a \\ -\sin a & \cos a \end{bmatrix} = \begin{bmatrix} \cos a & \cos g \\ \cos b & \cos d \end{bmatrix}$$

Rotation matrix with „direction cosine“

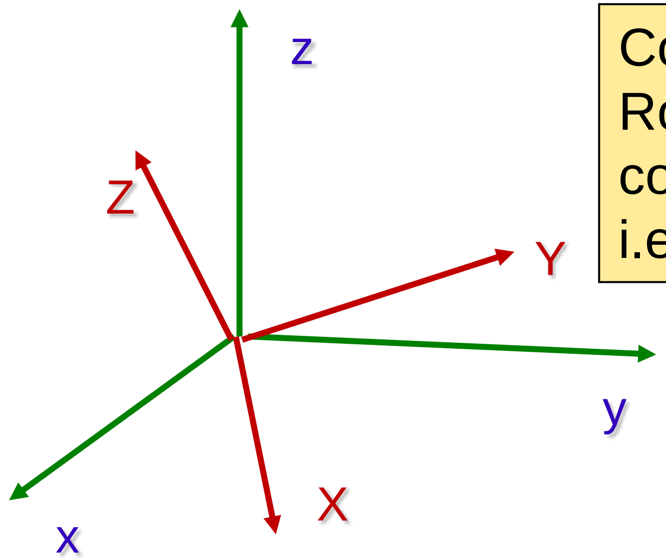
Rotation Around a Single Axis in 3D



Rotation around old x-axis rotates coordinates in y-z-plane from y-z to Y-Z

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos a & \sin a \\ 0 & -\sin a & \cos a \end{bmatrix}$$

Rotation Around a Single Axis in 3D



Convention:
Rotation around an axis are oriented counter clockwise when looking „from above", i.e. against orientation of the rotation axis.

Rotation by a around x-axis

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos a & \sin a \\ 0 & -\sin a & \cos a \end{bmatrix}$$

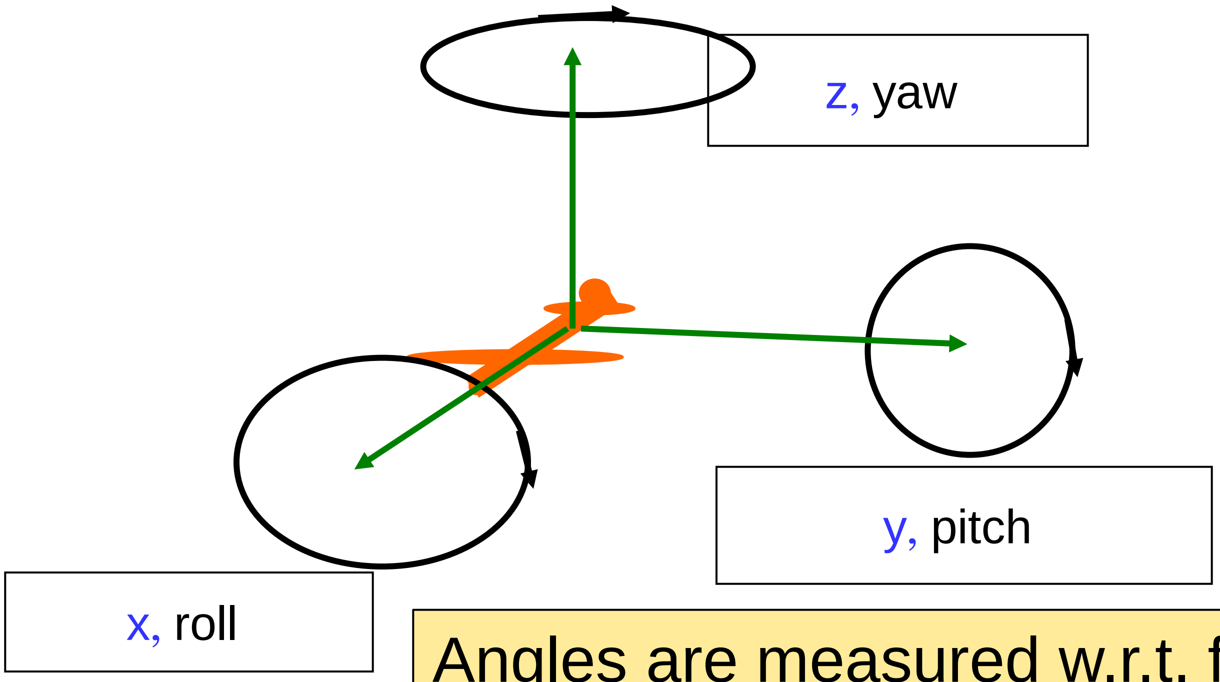
Rotation by a around y-axis

$$\begin{bmatrix} \cos a & 0 & -\sin a \\ 0 & 1 & 0 \\ \sin a & 0 & \cos a \end{bmatrix}$$

Rotation by a around z-axis

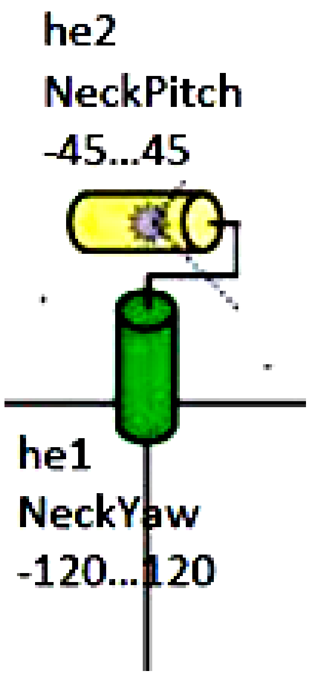
$$\begin{bmatrix} \cos a & \sin a & 0 \\ -\sin a & \cos a & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

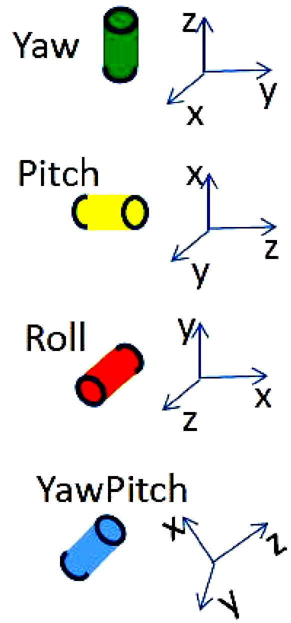
Yaw, Pitch, Roll in Aviation and Nautics

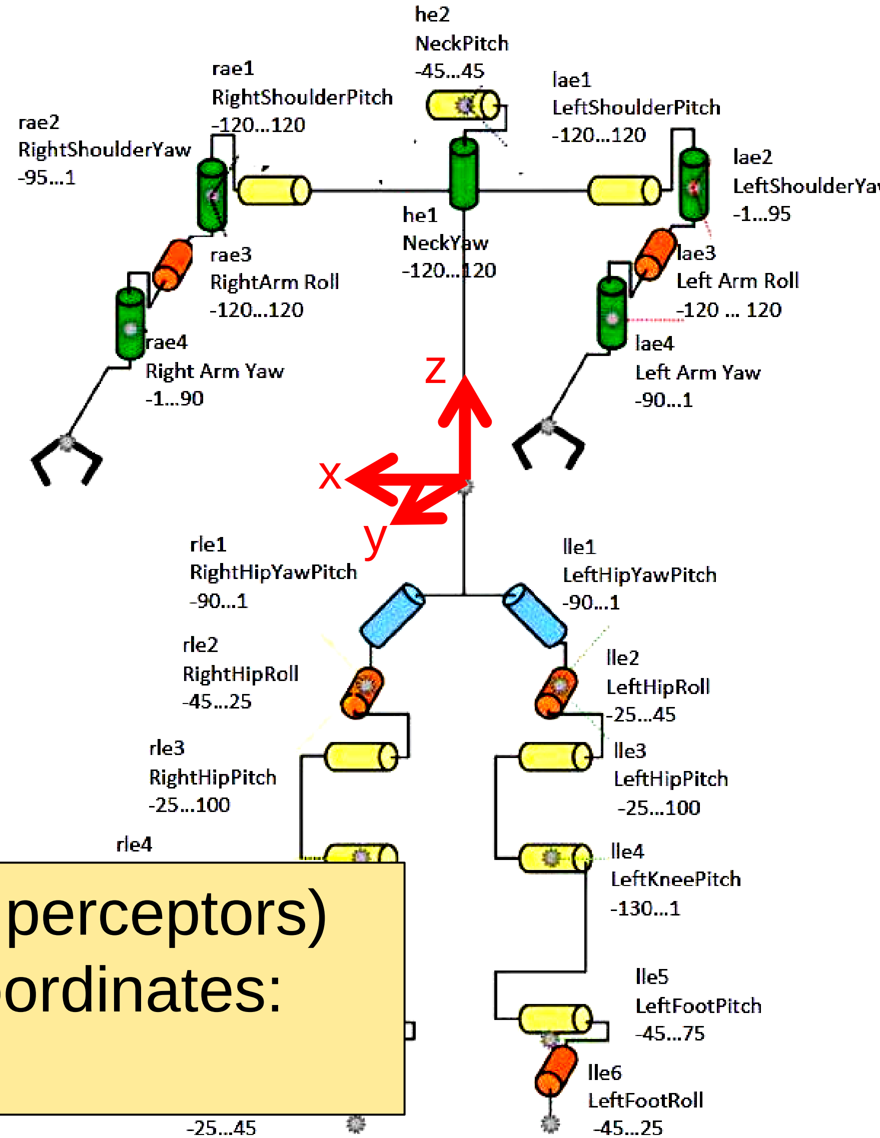


Angles are measured w.r.t. fixed coordinates: *Extrinsic rotations*.

Yaw, Pitch, Roll in Robotics

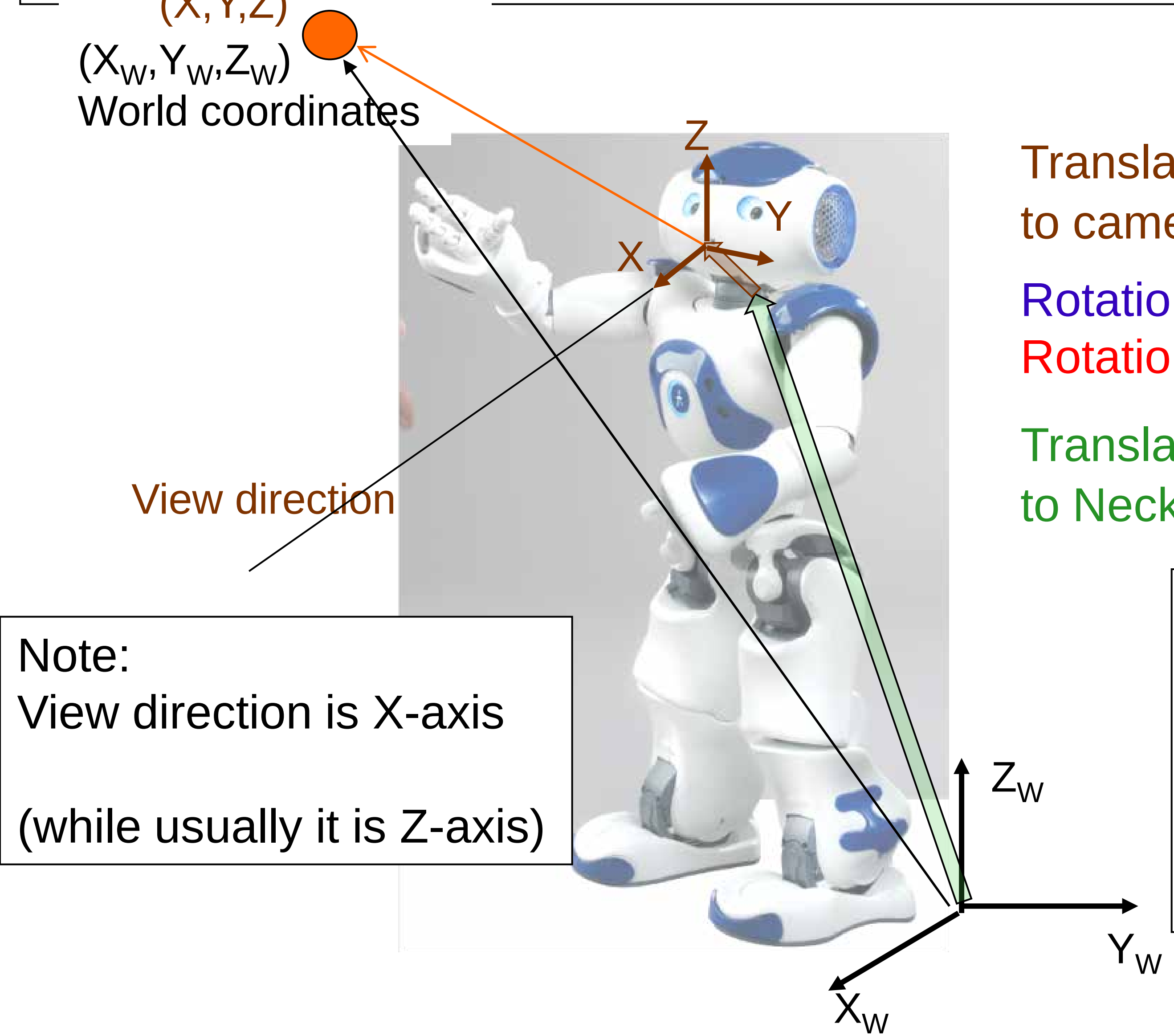






Angles are measured (effectors, perceptors) with respect to changing local coordinates: *Intrinsic rotations.*

Camera Model Nao



Translation T_F =(X_F,Y_F,Z_F) to camera center

Rotation Neck Yaw a

Rotation Neck Pitch b

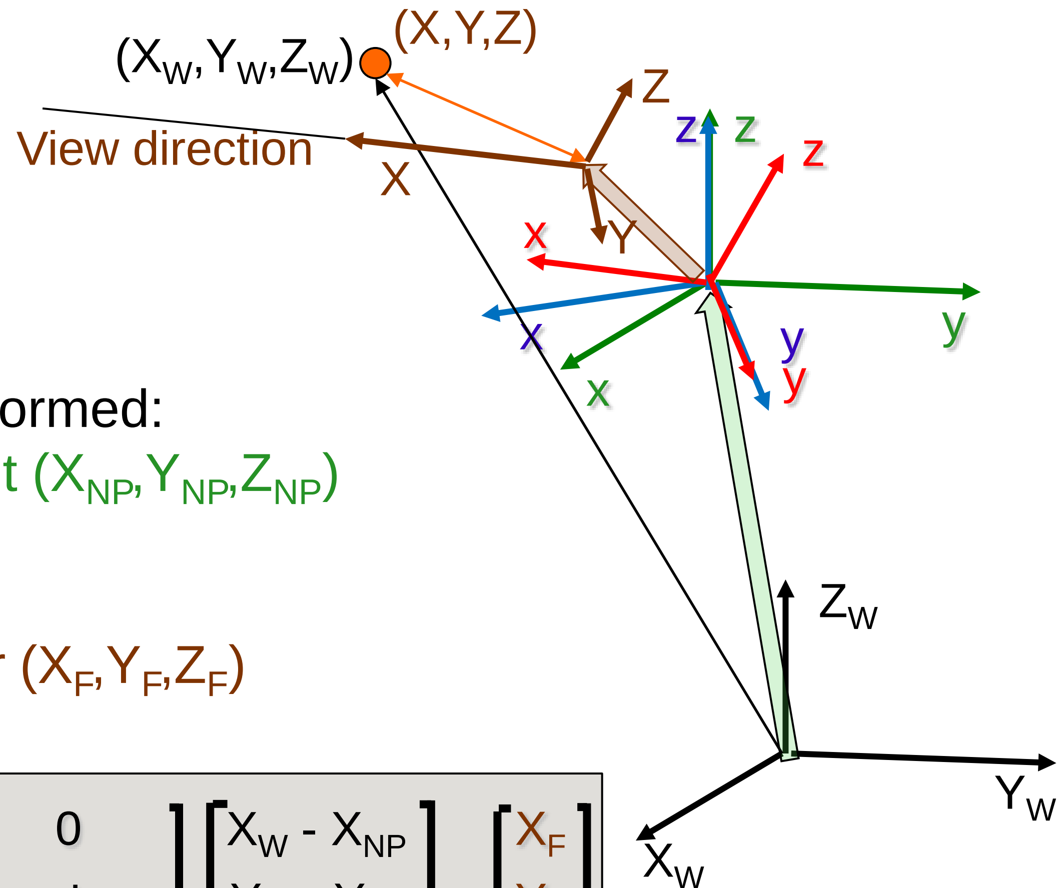
Translation T_{NP} =(X_{NP},Y_{NP},Z_{NP}) to Neck Pitch joint

Note:
View direction is X-axis

(while usually it is Z-axis)

Neck Pitch is located at axis of Neck Yaw: No translation needed between these joints using translation from world to NeckPitch.

Camera Model Nao



World coordinates are transformed:

Translation to neck pitch joint (X_{NP}, Y_{NP}, Z_{NP})

Rotation Neck Yaw a

Rotation Neck Pitch b

Translation to camera center (X_F, Y_F, Z_F)

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} \cos b & 0 & -\sin b \\ 0 & 1 & 0 \\ \sin b & 0 & \cos b \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos a & \sin a \\ 0 & -\sin a & \cos a \end{bmatrix} \begin{bmatrix} X_W - X_{NP} \\ Y_W - Y_{NP} \\ Z_W - Z_{NP} \end{bmatrix} - \begin{bmatrix} X_F \\ Y_F \\ Z_F \end{bmatrix}$$

Perspective Projection (Central Perspective)

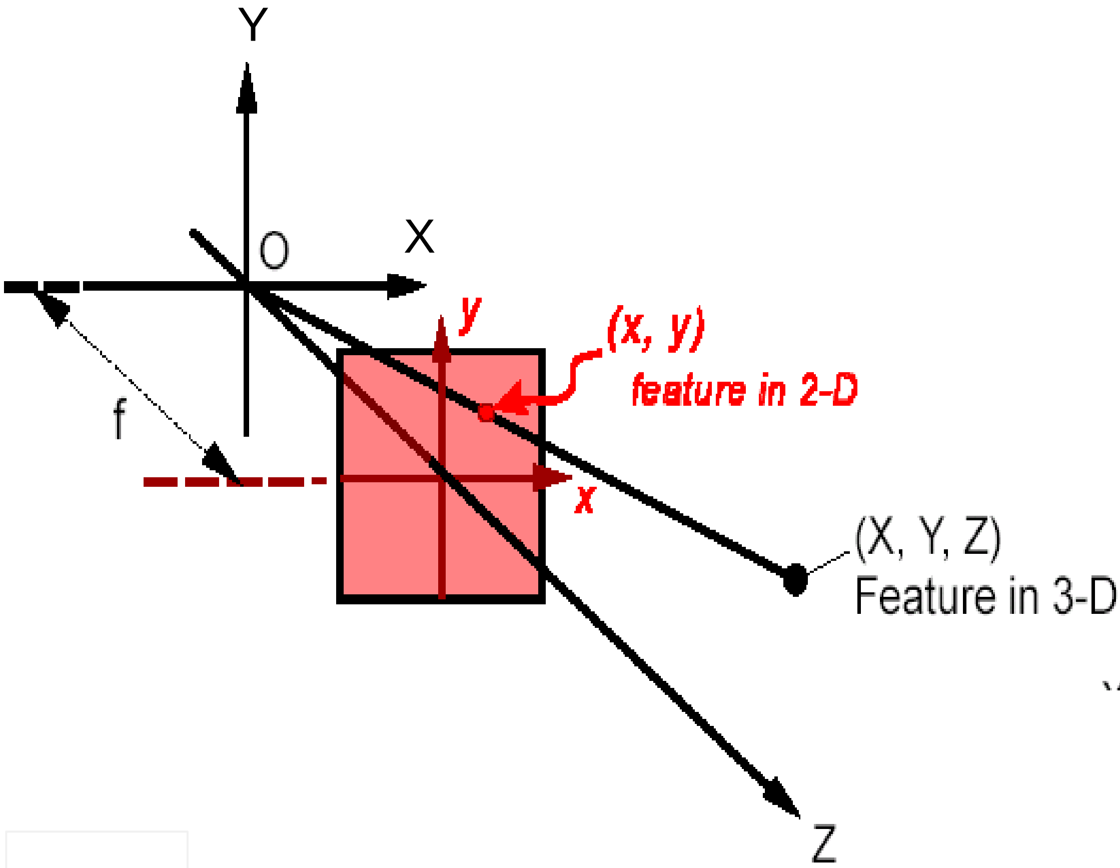
The image coordinates (x, y) are uniquely determined by Camera coordinates (X, Y, Z).

Intercept Theorem

$Z : f \;=\; X : x \;=\; Y : y$

$x = f/Z \bullet X$

$y = f/Z \bullet Y$



Perspective Projection (Central Perspective)

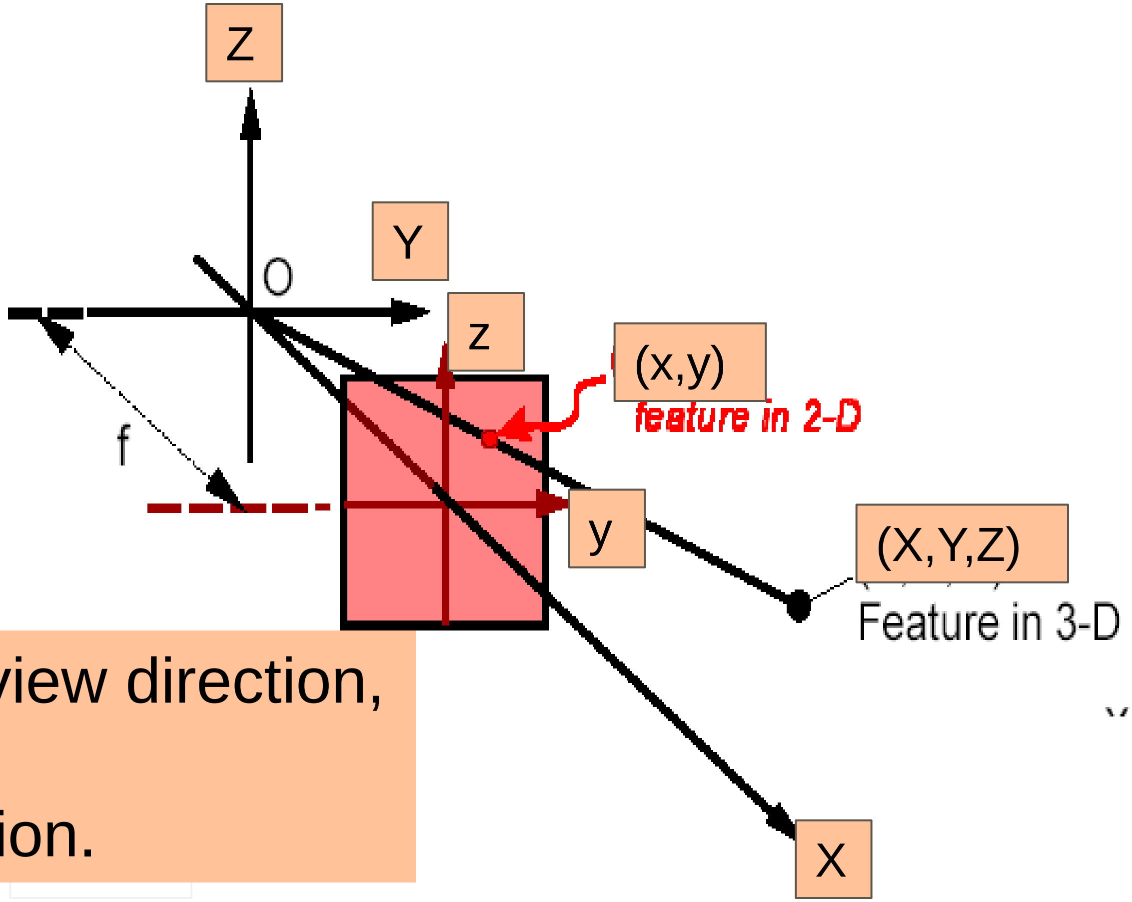
The image coordinates (x, y) are uniquely determined by Camera coordinates (X, Y, Z).

Intercept Theorem

X : f = Y : y = Z : z

y = f/X • Y

z = f/X • Z



Usually Z-axis points in view direction, while for Simulated Nao, view direction is X-direction.

Camera Model Simulated Nao

$$y = f/X \bullet Y$$
$$z = f/X \bullet Z$$

f is the distance of image plan, and X,Y, Z are calulated by

$$\begin{bmatrix} Y \\ Z \\ X \end{bmatrix} = \begin{bmatrix} \cos b & 0 & -\sin b \\ 0 & 1 & 0 \\ \sin b & 0 & \cos b \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos a & \sin a \\ 0 & -\sin a & \cos a \end{bmatrix} \begin{bmatrix} X_W - X_{NP} \\ Y_W - Y_{NP} \\ Z_W - Z_{NP} \end{bmatrix} - \begin{bmatrix} X_F \\ Y_F \\ Z_F \end{bmatrix}$$

Rotation matrix
by multiplication of matrices

Camera Model

$$X = R\,(X_W - T_{NP}) - T_F$$

Inverse Camera Model Nao

Camera Model

$$X = R\,(X_W - T_{NP}) - T_F$$

x = f/Z • X
y = f/Z • Y

Inverse Camera Model

$$X_W = R^{-1}(X + T_F) + T$$

X = x • Z / f
Y = y • Z / f

X= (X,Y,Z) can not be completely reconstructed from x, y only

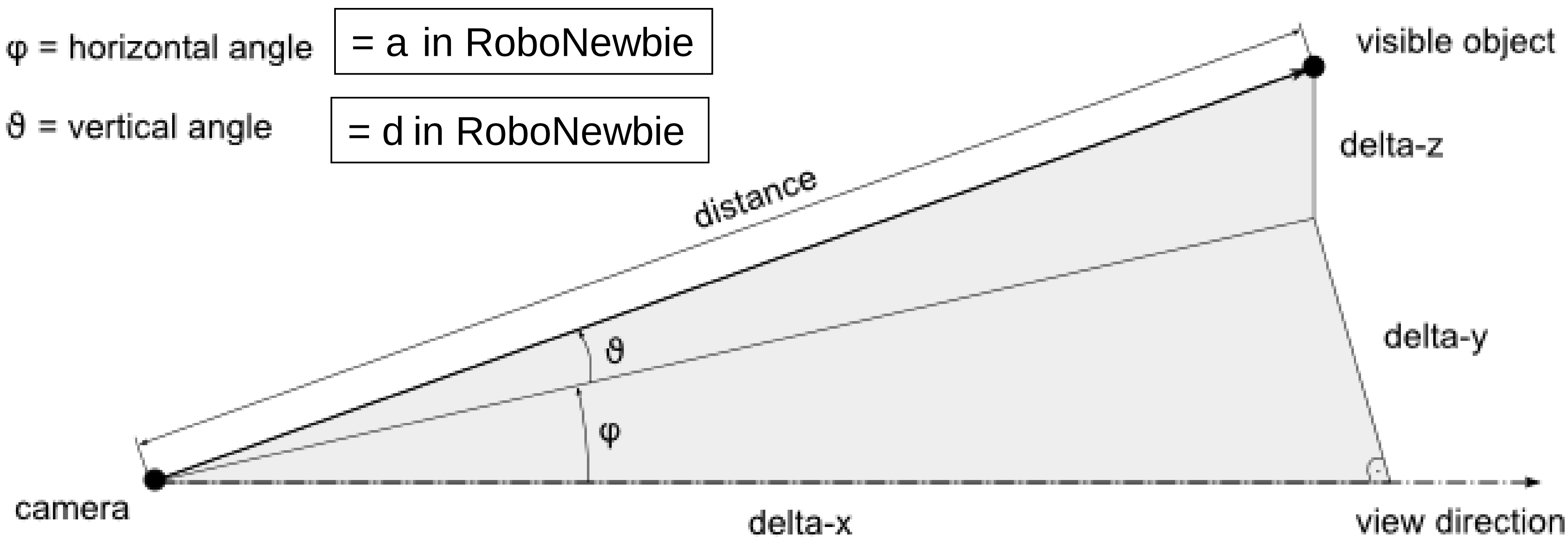
Additional information is needed, e.g.

- Distance Z
- Size of an object
- Location on ground

Change for Simulated Nao as before:
View direction is X-direction.

Vision Perceptor of Simulated Nao

Provides *polar coordinates* relatively to the camera



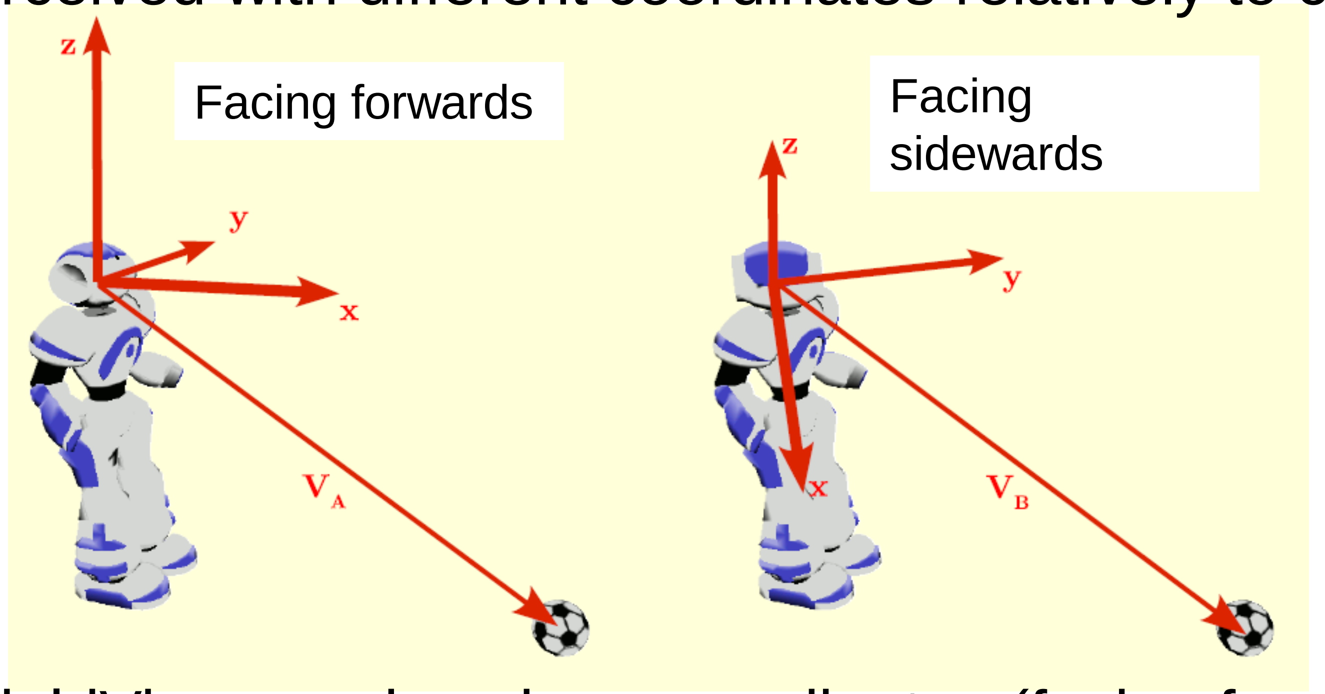
Preprocessing for Perception in RoboNewbie

LookAroundMotion moves the head (the camera) continuously:

Turns down to 40°, back to upright position,

then left to 60°, right to -60° and back to initial position.

Objects perceived with different coordinates relatively to camera.



But LocalFieldView needs unique coordinates (facing forwards).

Simplification in RoboNewbie

The vision perceptor collects visual data while moving the head.

The position of an object is described by polar coordinates (d, a, d) with distance d, horicontal angle a and vertical angle d.

Direction of the head (camera) by LookAroundMotion is:

1. in horizontal direction (yaw y) while vertical angle (pitch f) is 0.
2. in vertical direction (pitch f) while horizontal angle (yaw y) is 0.

LocalFieldView is to provide transformed data (d’, a’, d’) according to the coordinate system when facing forward.

Simplification in RoboNewbie

The distance d remains unchanged, i.e. d' = d,

but angles a' and d' need to be calculated from a, d, y, f .

Correct calculation need transformations as described before.

Instead, a simple approximation is performed by RoboNewbie:

a' and d' are calculated using the offsets y resp. f .

$$a' = a + y \qquad d' = d + f$$

The result is correct

- for vertical angle d' .
- for horizontal angle a' as long as f = 0 .

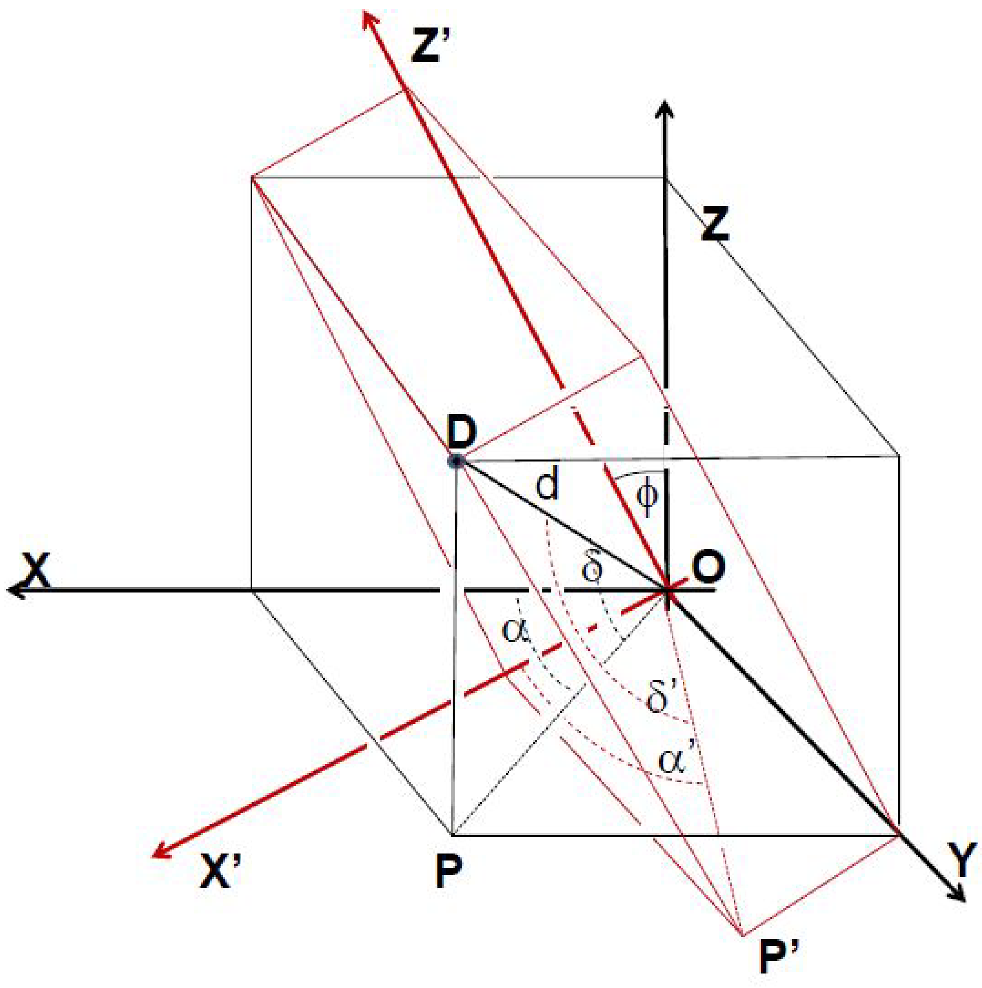
It is only an approximation for angle a' if f ≠ 0 (head tilded)

Simplification in RoboNewbie

The angles d and a of perception change according to the change from XY-plane to X'Y-plane (tilded head).

Correct transformations would need complex geometrical calculations.

Drawback of simplified calculation: Deviations of position for near objects.



Rotation Matrix for Intrinsic Rotations

Intrinsic Rotations:

Rotations are given w.r.t. recent object coordinates.

If A, B, C are successive intrinsic rotations,
then the resulting rotation is described by $R = C\ B\ A$

Rotation Matrix for Extrinsic Rotations

Extrinsic Rotations:

Rotations are given w.r.t. a fixed coordinate system

(e.g. yaw-pitch-roll in Aviation/Nautics).

If A, B, C are extrinsic rotations,
then the resulting rotation is described by $R = A\ B\ C$

result of the first rotation is given by A

result of the first two rotations is given by

intrinsic rotations $(A\,B\,A^{-1})$ and A resulting in $(A\,B\,A^{-1})A = A\,B$

result of all three rotations is then given by

intrinsic rotations $((A\,B)C(A\,B)^{-1})$ and $(A\,B)$

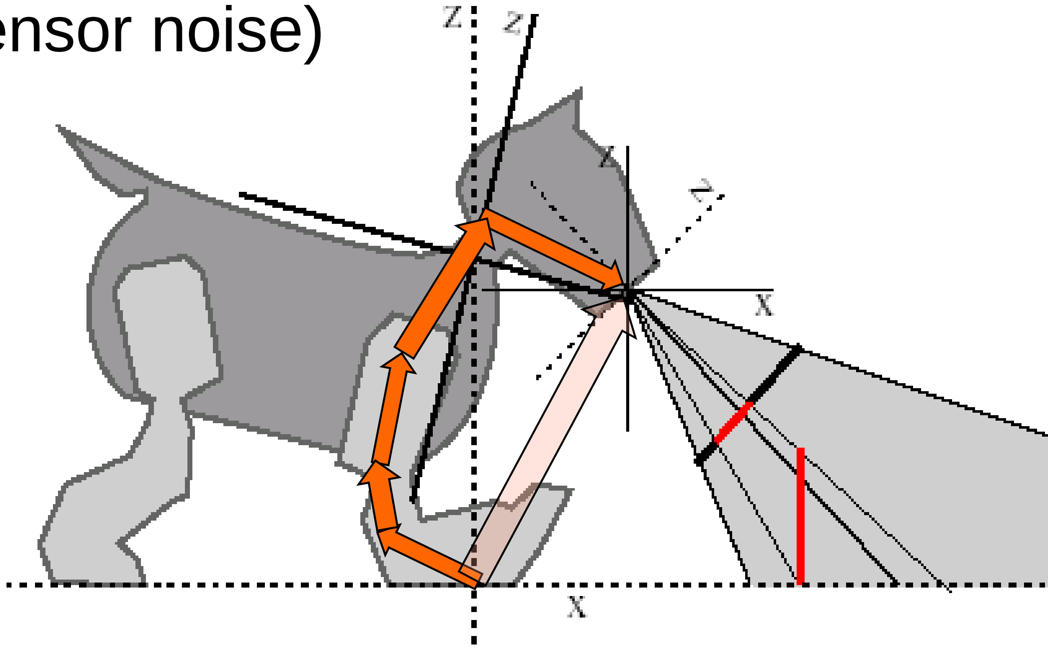
resulting in $((A\,B)C(A\,B)^{-1})\ (A\,B) = A\,B\,C$

Problems with Camera Model

Position of camera (extrinsic parameters):

Errors in the kinematic chain:

- joint angles (backslash, sensor noise)
- distortion during motion

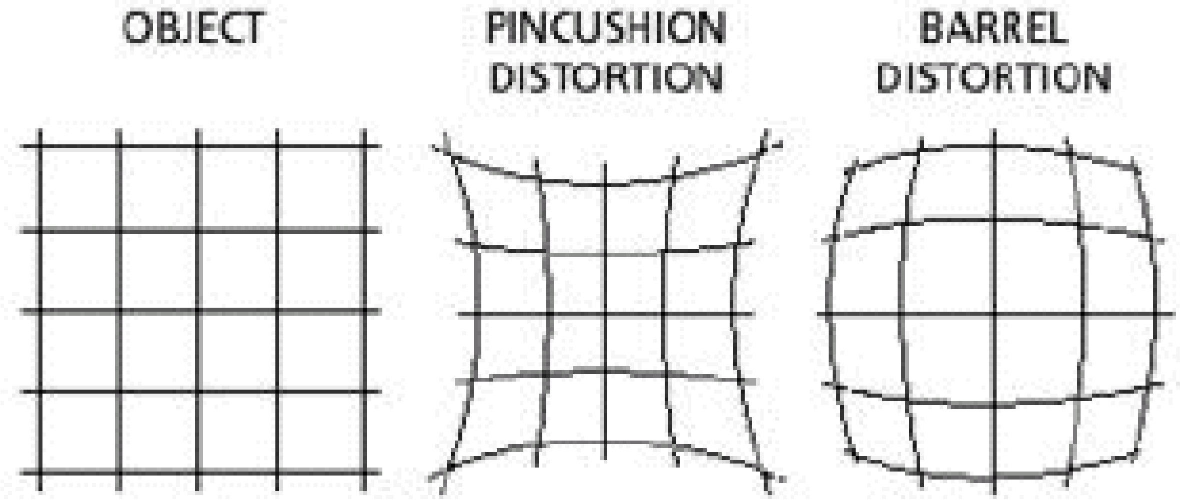


Can be determined/corrected by known landmarks

(cf. localization methods: later)

Problems with Camera Model

Geometrical distortion by optics (intrinsic parameters) by refraction of the light at the inlet and outlet from the media

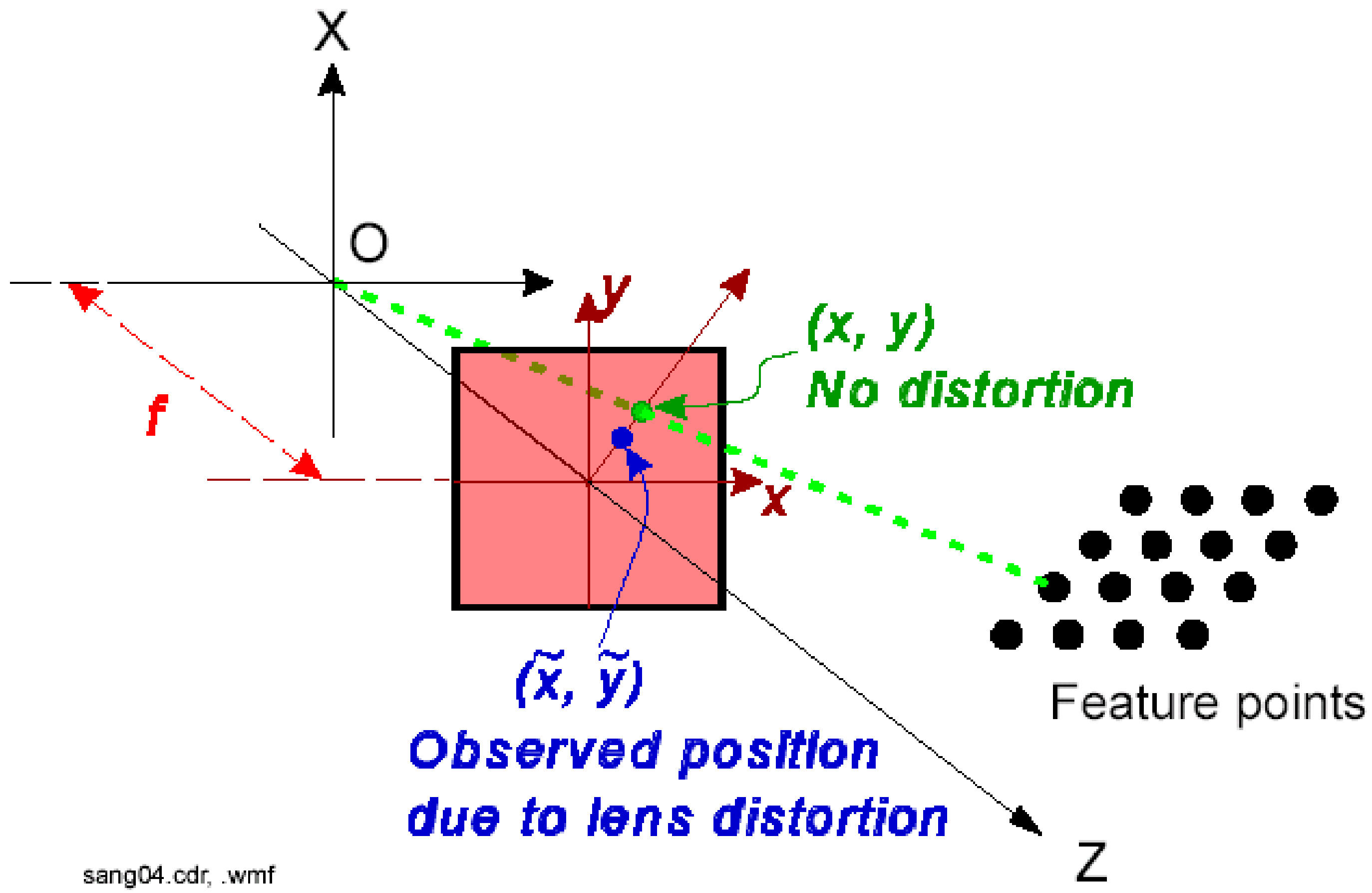


Problems with Camera Model

Distortions determined/corrected/calibrated by experiments:

Imaging parameters determined by

corresonding points in reality and in image.



Problems with Camera Model

Example for calibration:
Ceiling camera in Small Size League
(FU-Fighters Berlin)

Fig. 2. Barrel distortion of the field with a 4.2 mm lens

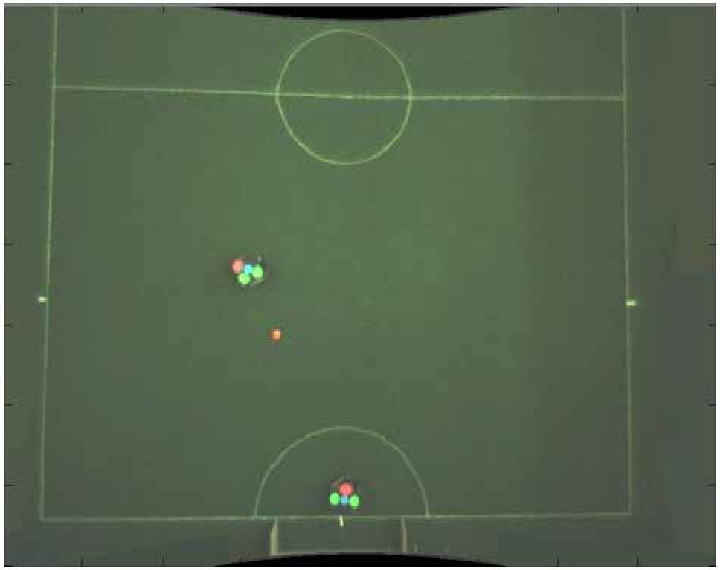
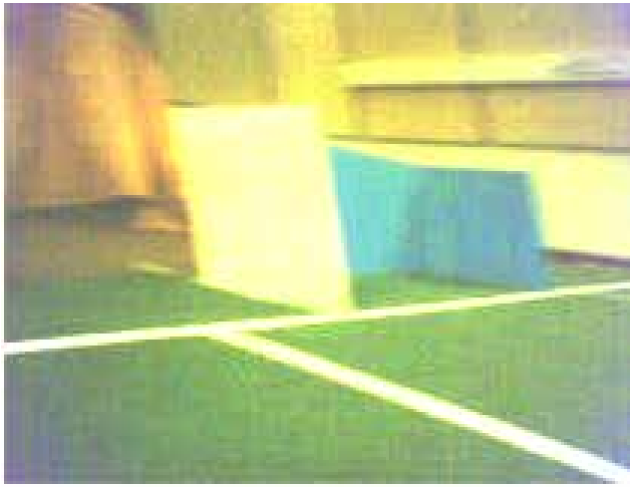


Fig. 6. Corrected field image

Problems with Camera Model

Motion Blur (delays while reading pixels during motion)



You can also see color distortion (blue in the corners)

Outline

Introduction

Sensors: General Considerations

Signals

Sensors: Special Types

Vision (introductory)

Camera Model

Image Processing (introductory)

Scene Interpretation (introductory)

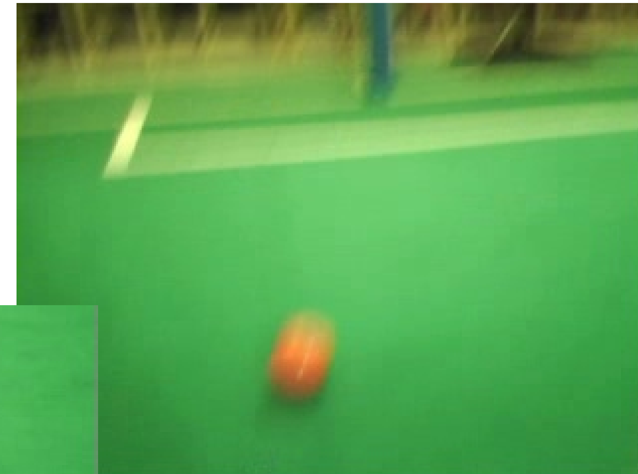
Image Processing

Given a pixel matrix: what is the content of the image?

Can include many processes:

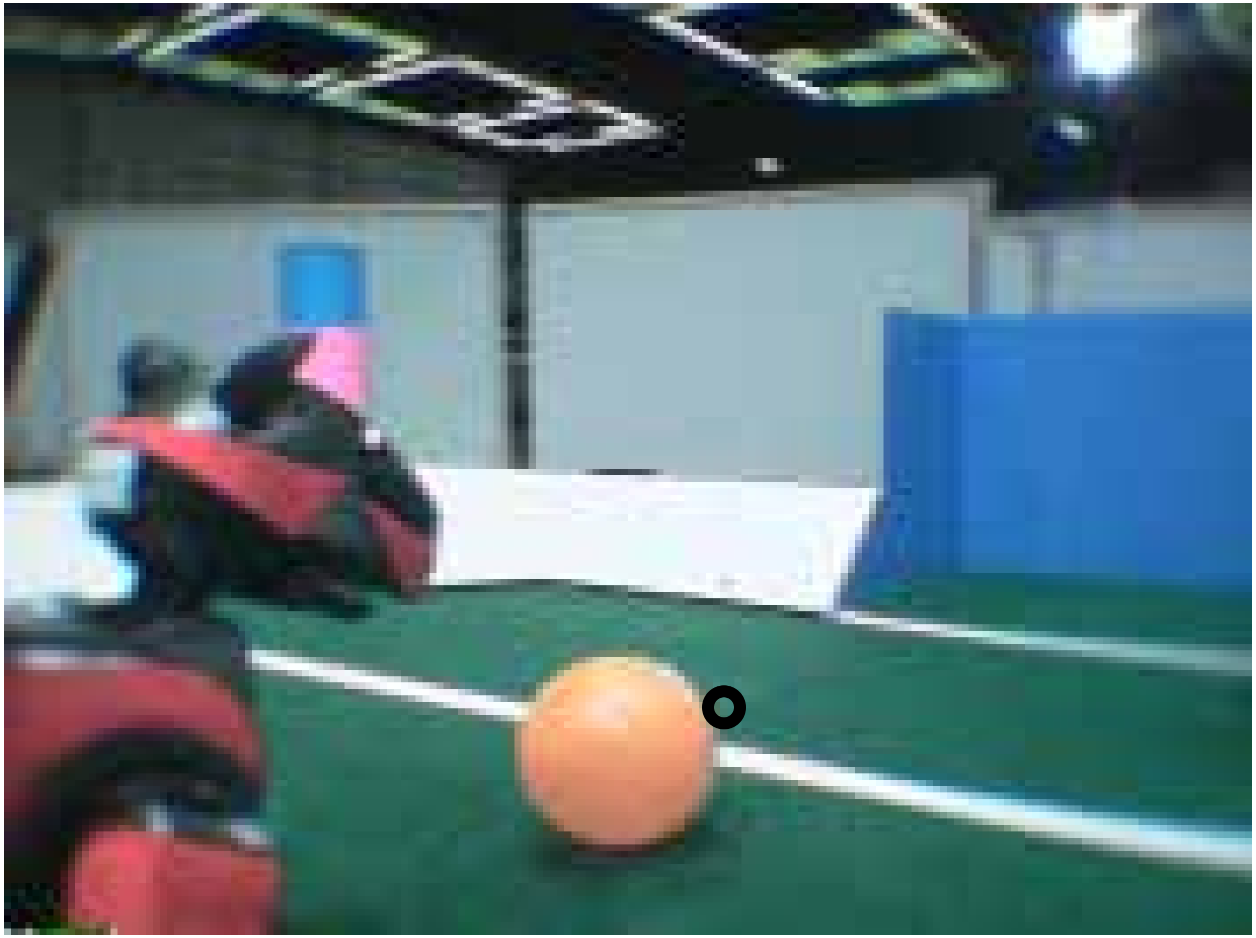
- Signal processing (noise reduction, …)
- Low level identification (line detection, color detection,…)
- Object identification
- Relation between objects
- Scene reconstruction (3D-model)
- Scene interpretation

Visual Information for Real Nao: Images

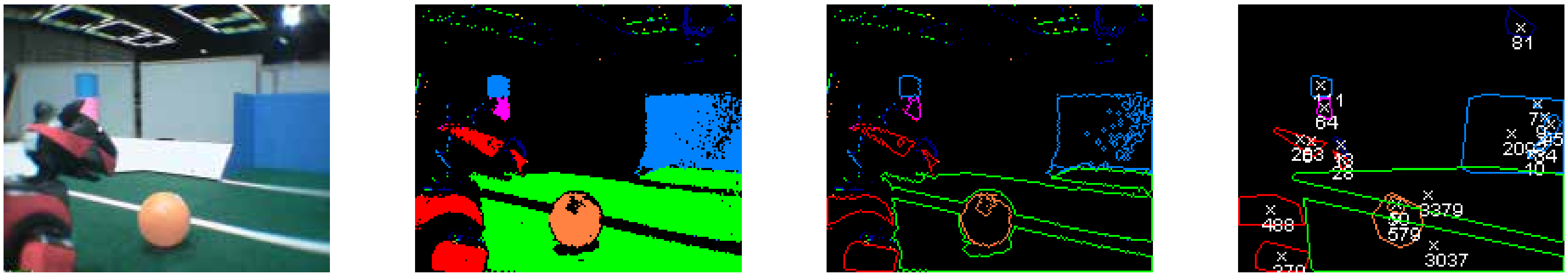


Interpretation needs complex image processing.

It is possible to provide synthetic images for simulation, but standard in 3D league are already preprocessed data.



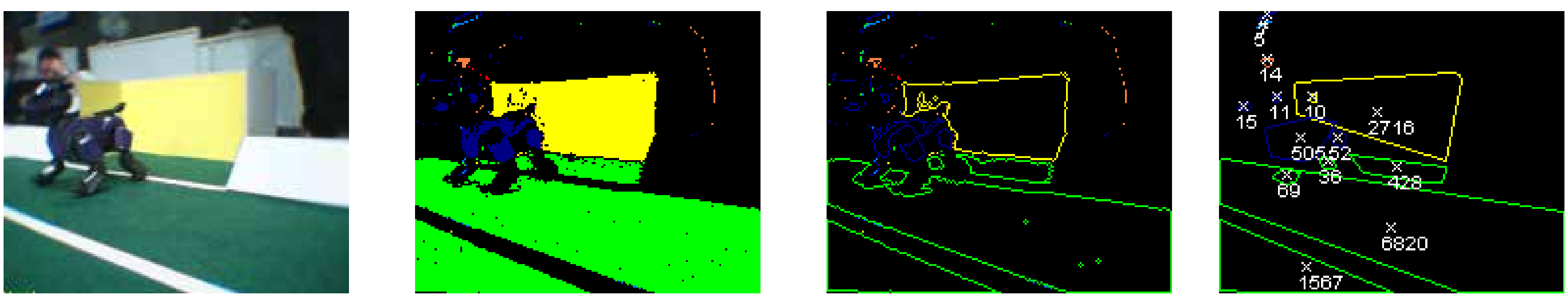
Preprocessing of Images



Color classification

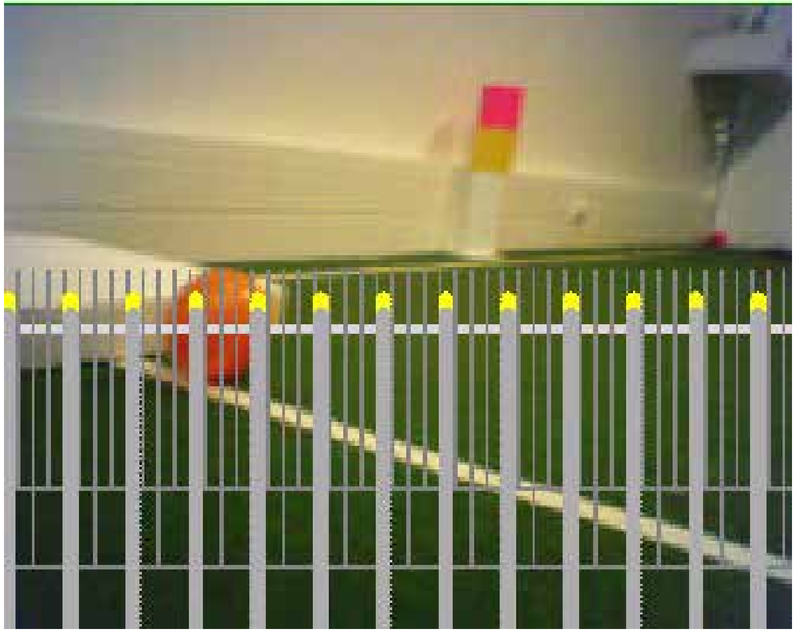
Boundary of objects

Identification of objects



Ball detection by region growing

Scan lines for minimal search

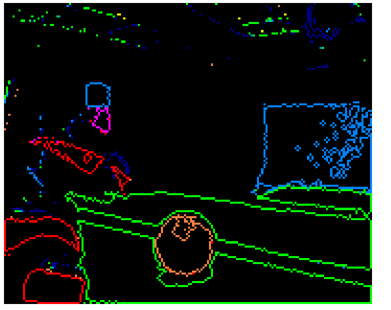


Identification

Identification of an individual object:

Based on known features.

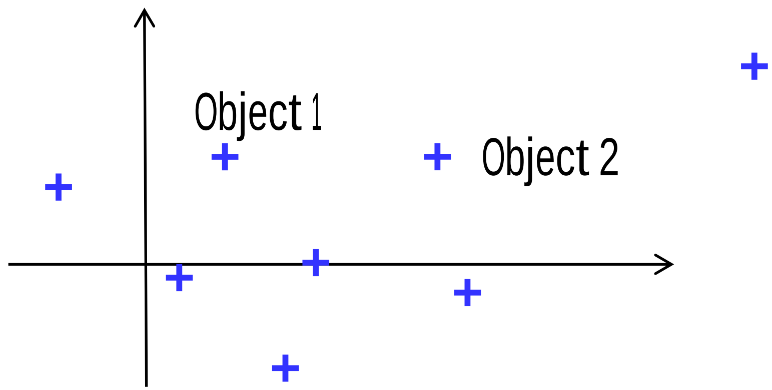
Features should be invariant against
rotation and scaling.



Features computed e.g. from
- Colors
- Shapes
- Size
- Statistics in useful regions (SIFT, SURF)
- Relations between points

Identification

Each object has a (*high dimensional)* feature vector ("signature")



In simple cases, the objects can be identified using explicit world knowledge (e.g. "the ball is orenge").

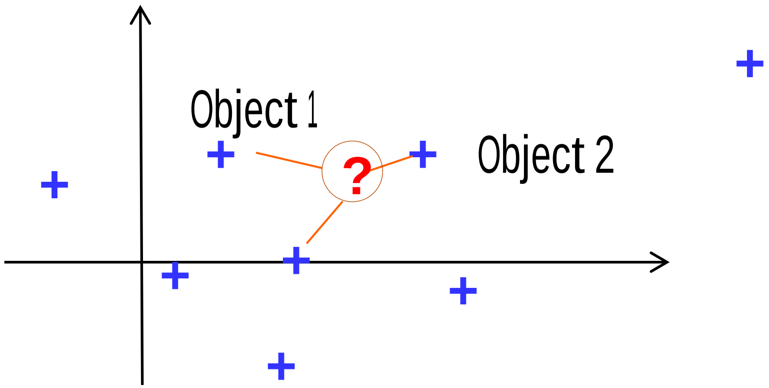
In general, the world is more complex.

Identification

Nearest Neighbor Method

Compare observed object by similarity to known objects

Choose most similar object (or reject)



Example: Face Recognition
1. Identify related regions
2. Identify (biometric) features
3. Compare with database

Available by commercial products

Works well with frontal faces
Depends on available resources

Classification

Classify objects

Based on known features, properties, relations

Problems with diversity of objects in the same class

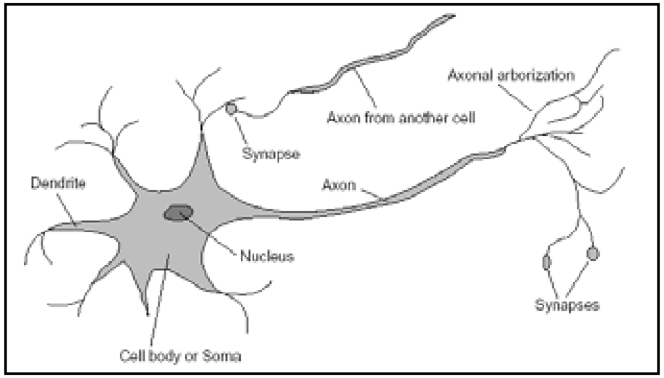
Classification

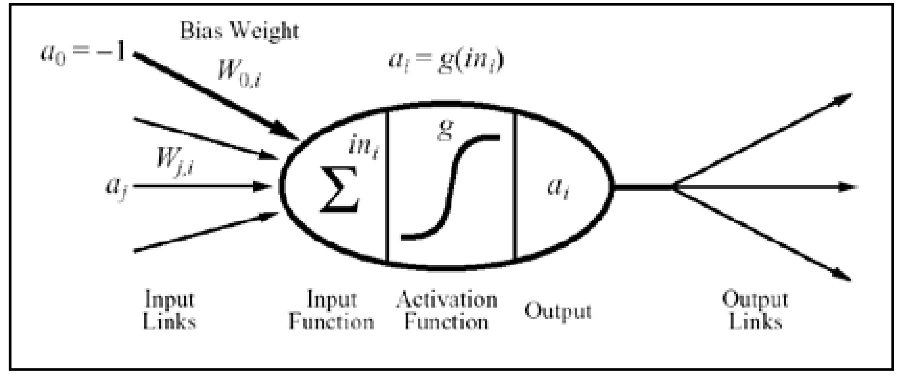
Classify objects

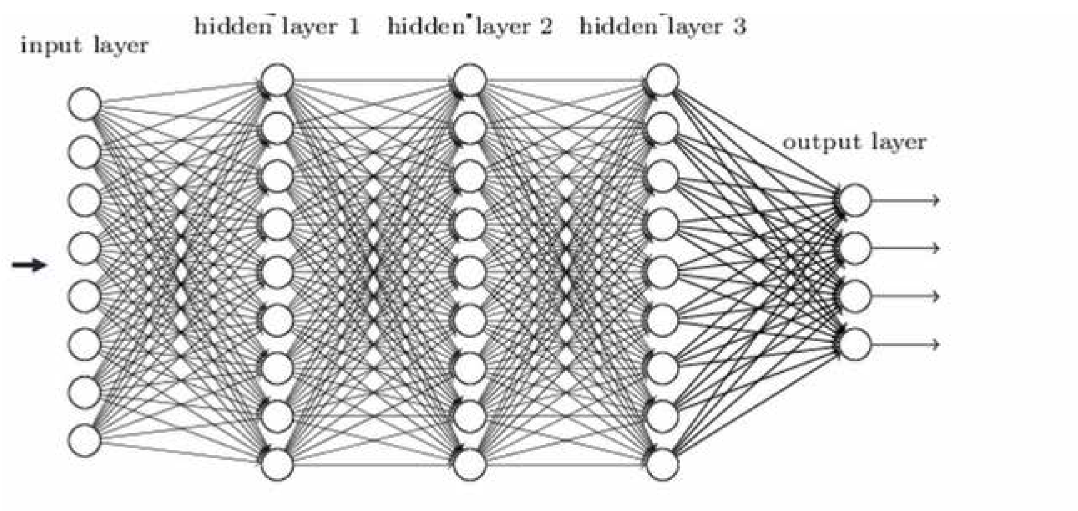
Based on known features, properties, relations

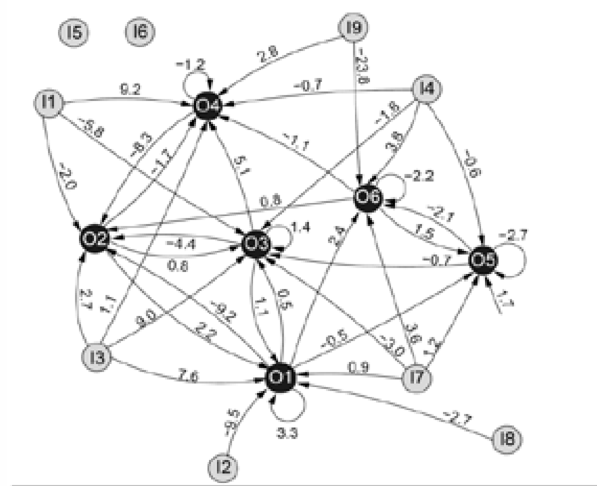
Problems with diversity of objects in the same class

Biologically Inspired Processing: Neural Networks

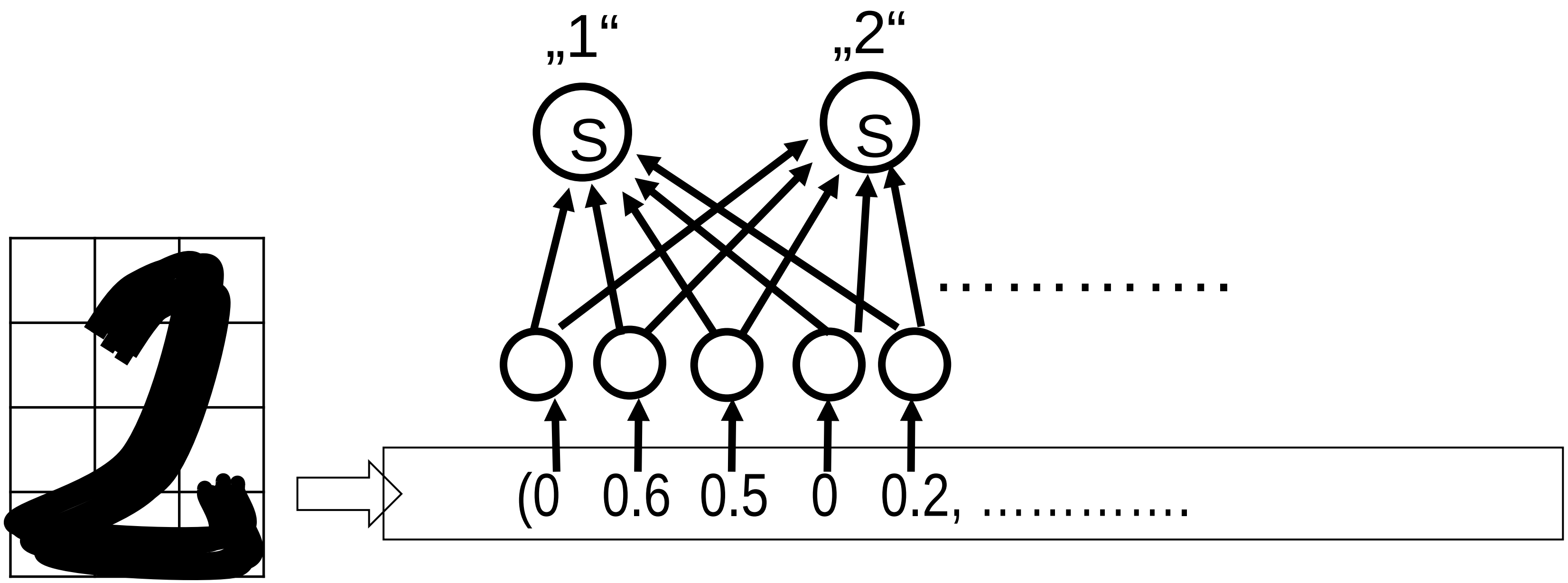


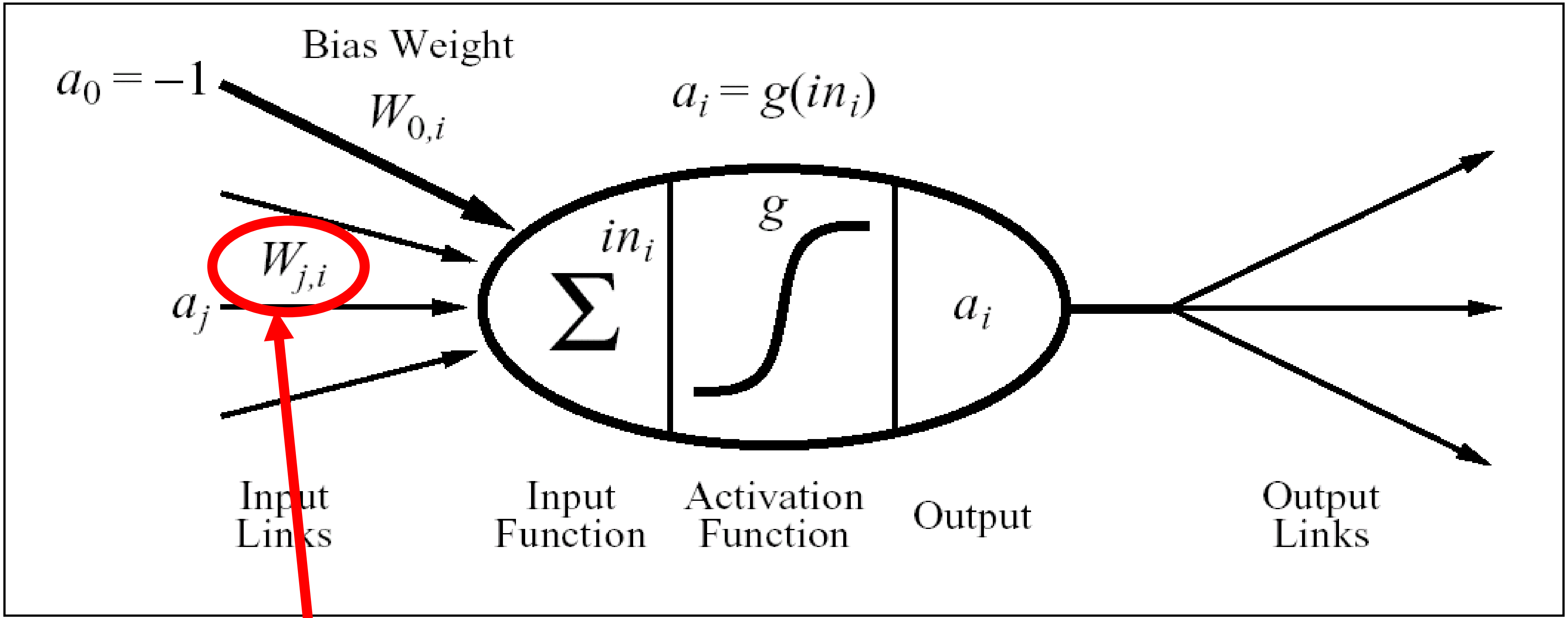




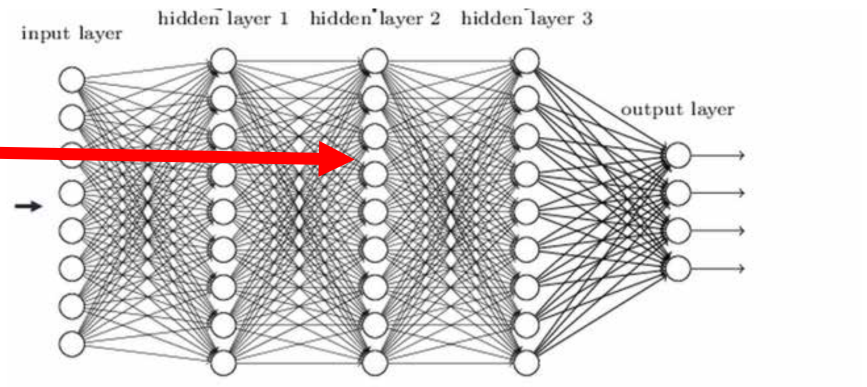


Processing in Neural Networks : Propagate Activations
(„Voting“: Maximal Votes Win)





Define appropriate weights for each neuron.

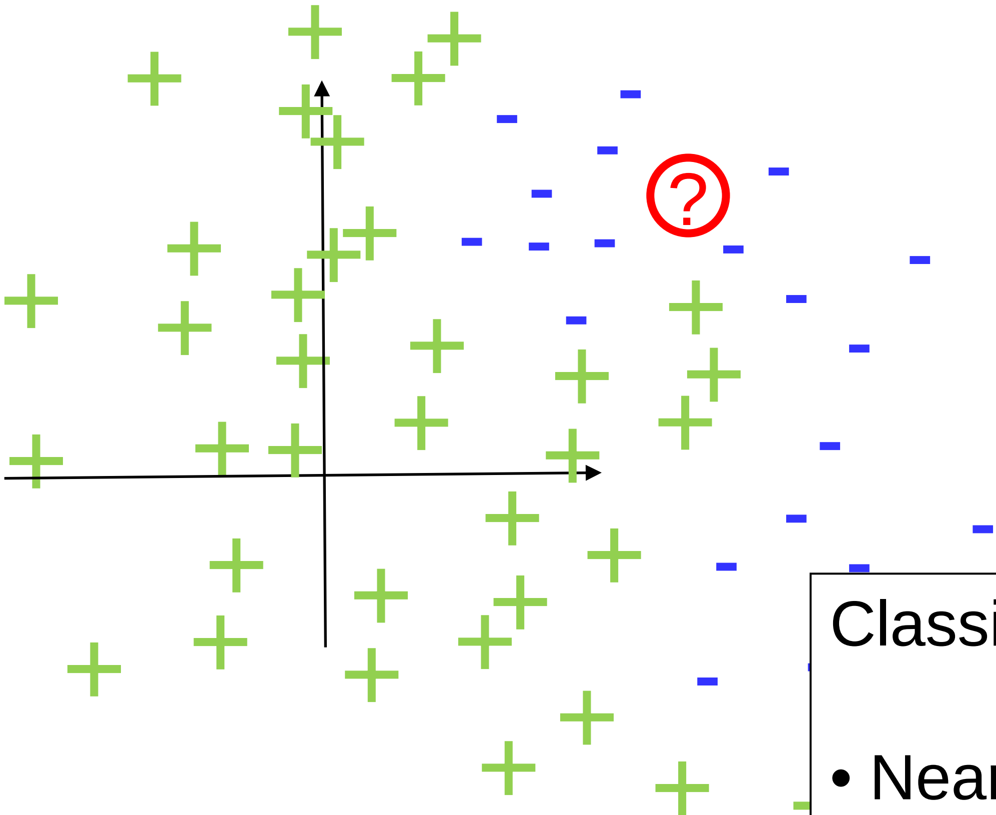


Machine Learning for Neural Networks

Adjustment of weights is done by "Training":

1. Classify examples with the network
2. Evaluate correctness
3. Change weights to reduce errors (gradient descent of the error function)

Classification (good +, bad -) based on similar fesatures

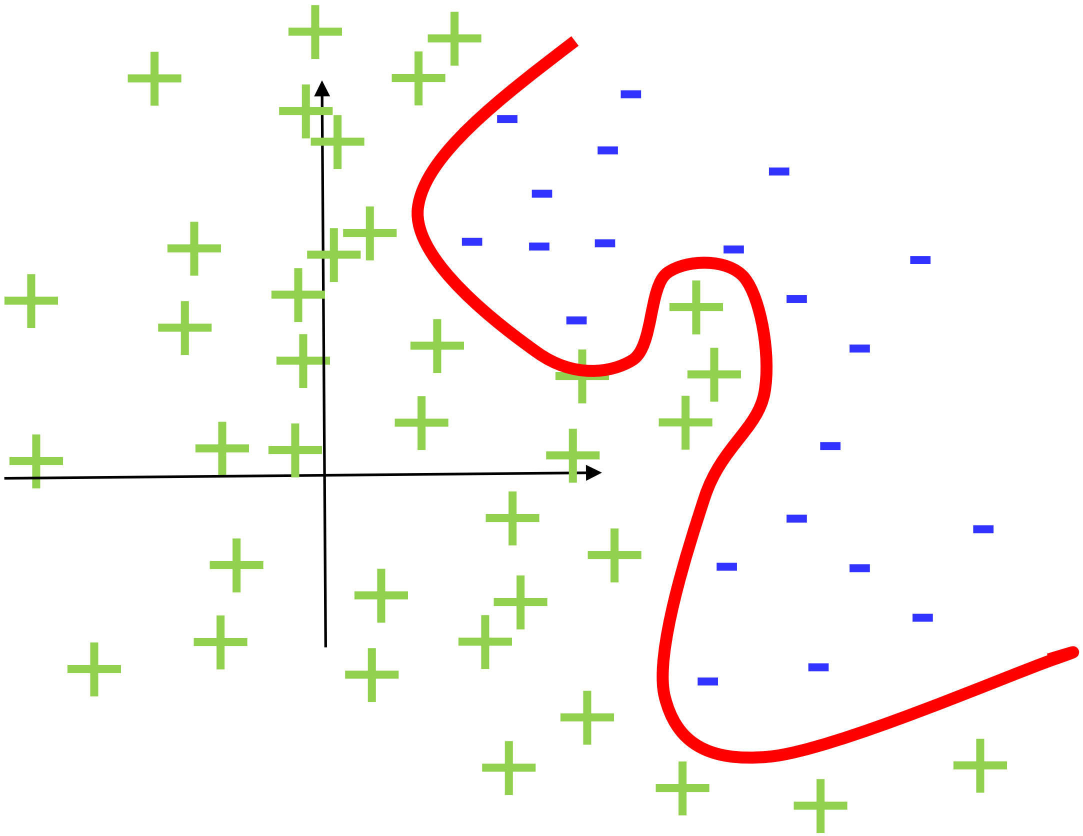


Problems:
- High Dimensionality

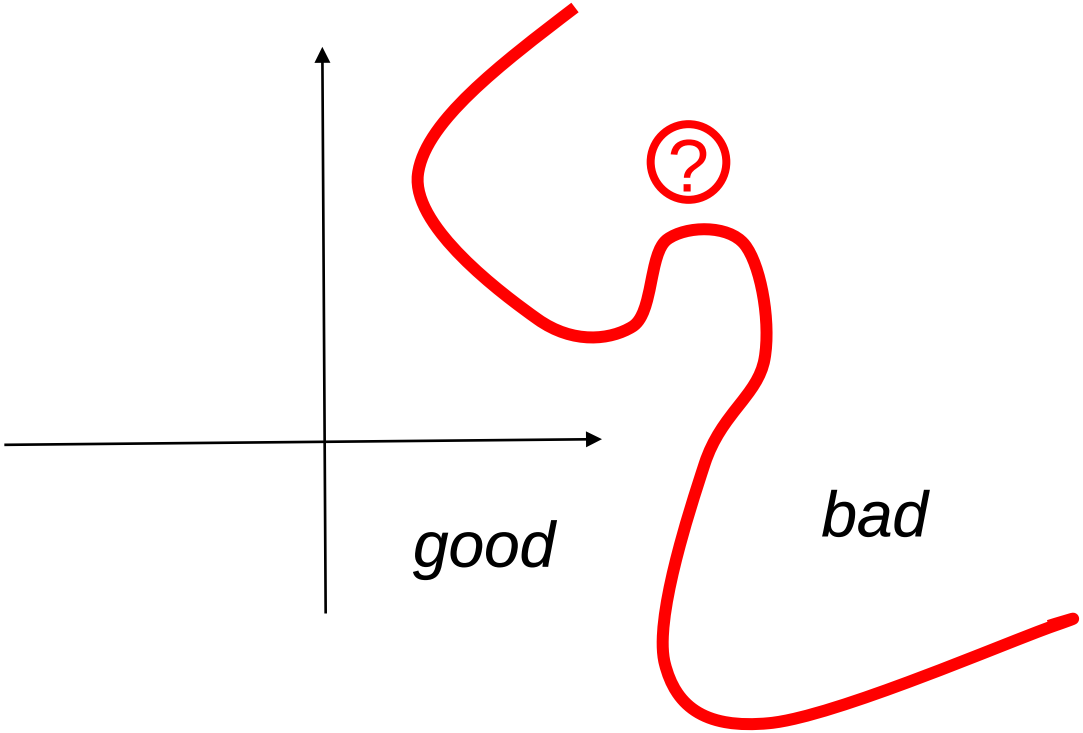
Classification methods e.g.

- Nearest Neighbor
- Decision tree
- Neural Network
- Support Vector Machine (SVM)

Neural Network learns a Partition Line

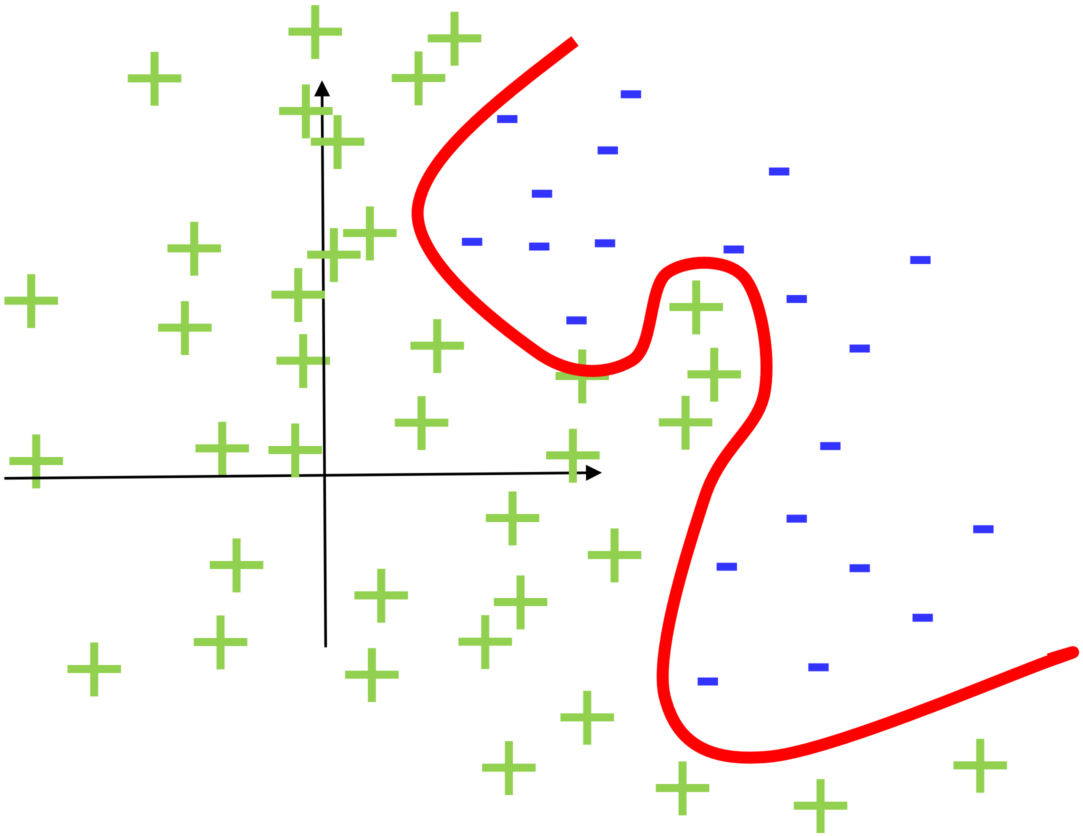


Classification by Partition Line (Left: good, Right: bad)

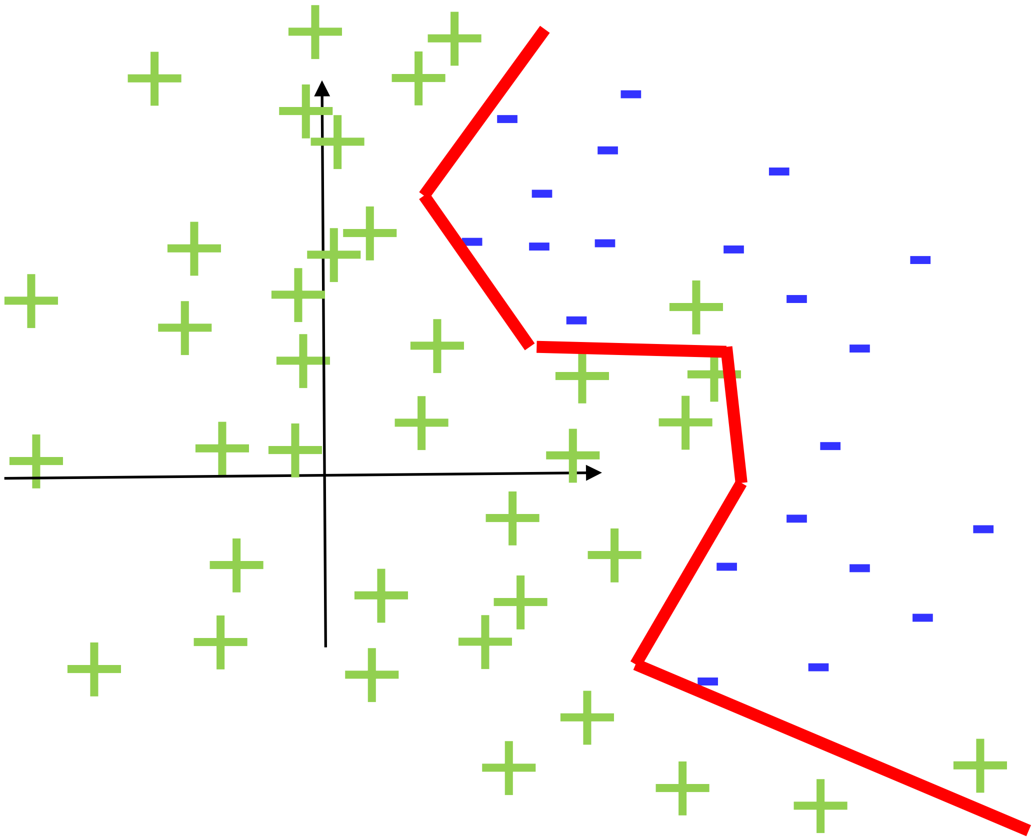


Support Vector Machines (SVM):
Construction of a partition line from examples

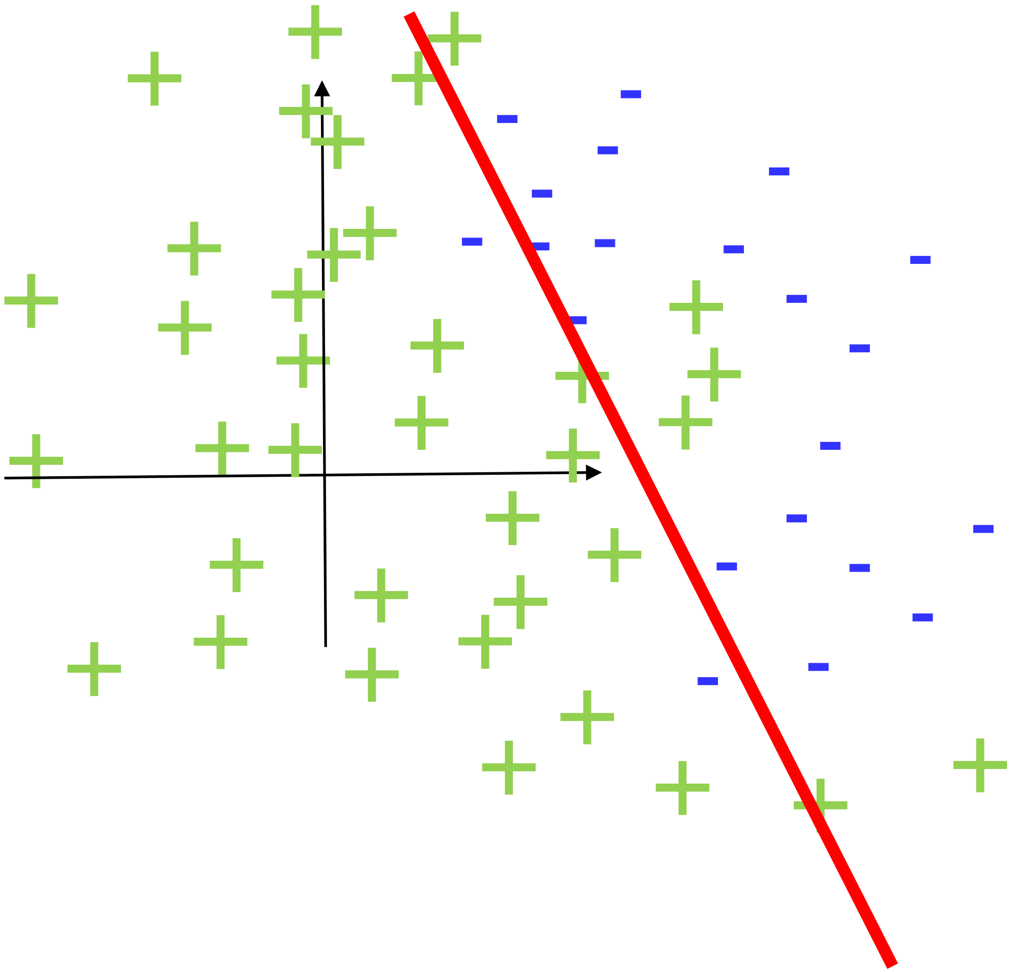
Problem:
Which line is the best?



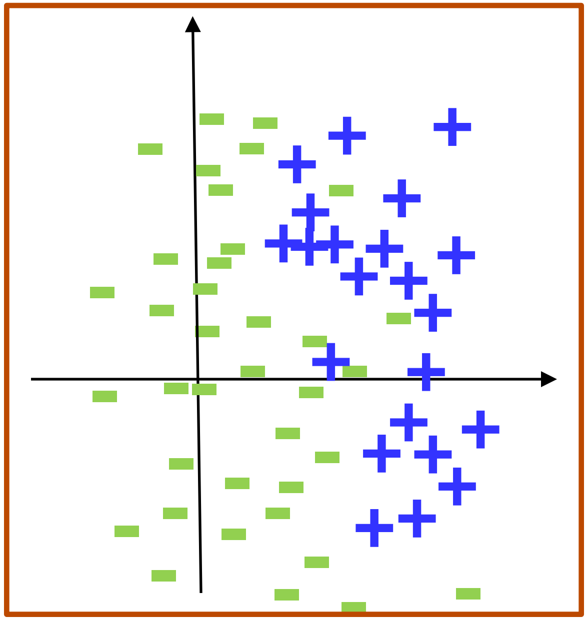
Problem:
Which line is the best?



Problem:
Which line is the best?

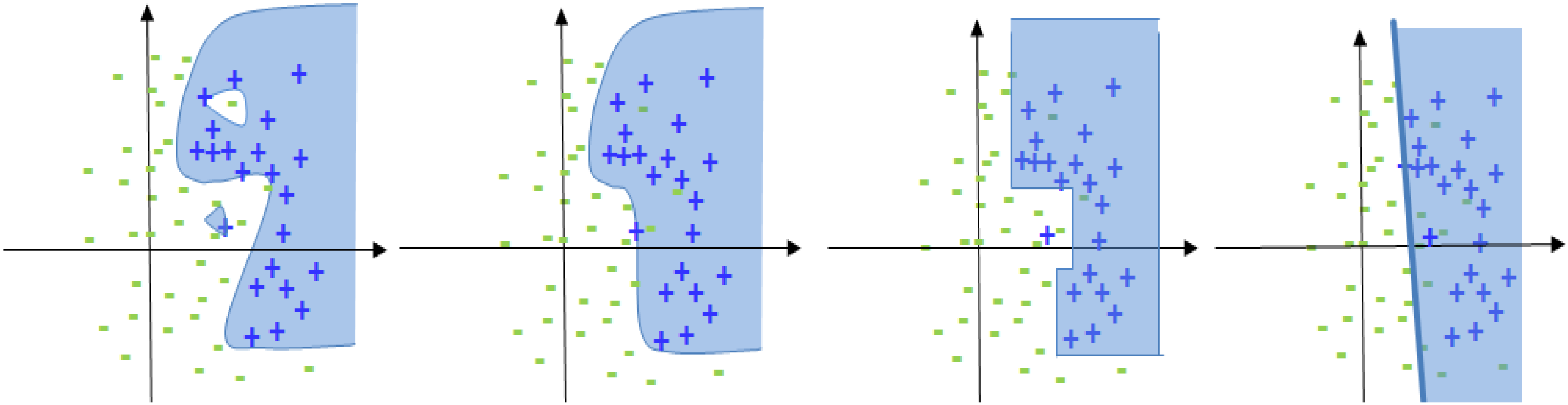


Classification



Generalization Problem:

The classification of new objects depends on the choice of the learning method (“inductive bias”)



Neural Networks for Image Classification

Deep Learning: Better hardware and new methods allow more layers of the net

Outline

Introduction

Sensors: General Considerations

Signals

Sensors: Special Types

Vision (introductory)

Camera Model

Image Processing (introductory)

Scene Interpretation (introductory)

What the Robot Sees

Scene Interpretion

Badly posed problem:

Reconstruction of a 3D scene from 2D image

M.C.Escher

Scene Interpretion

There are many available informations

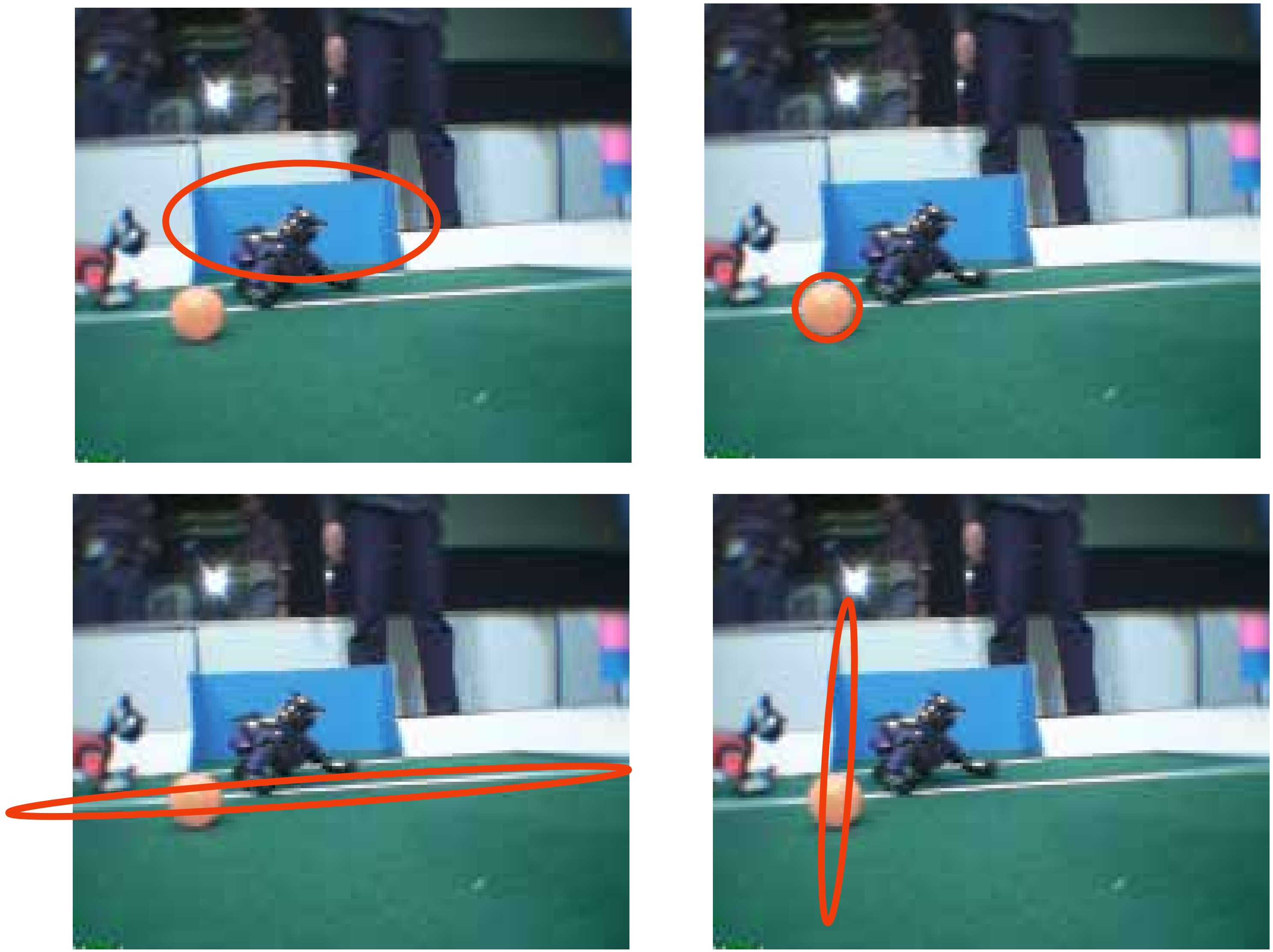
- i.g. enough to reconstruct a scene even from 2D images by using world knowledge.
- i.g. redundant for dealing with noise.

But: It is hard to compute.

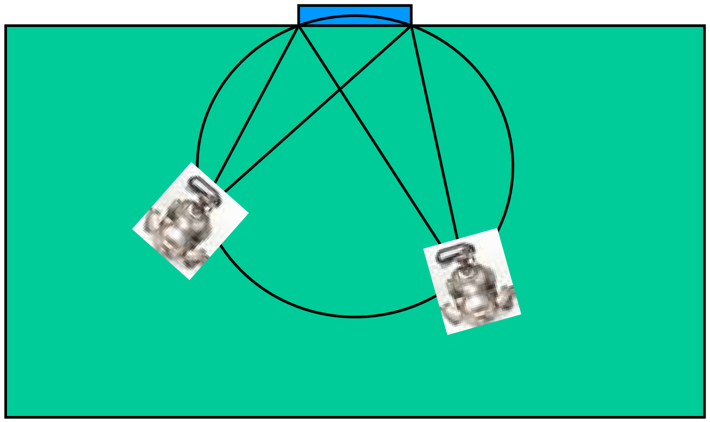
Exploiting Redundancy

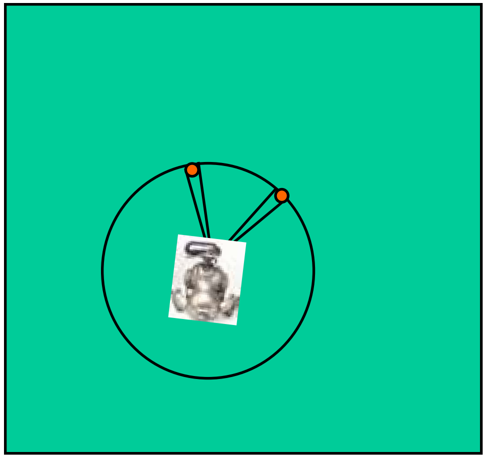
Where am I ?
Where is the ball ?

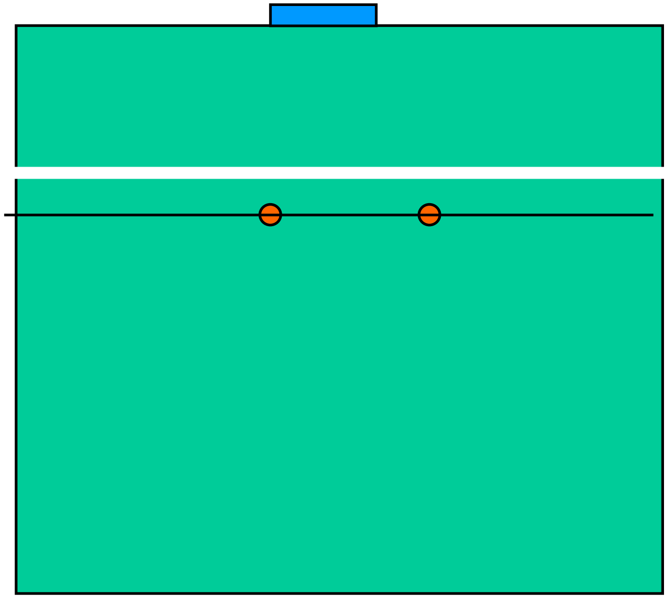
Exploiting Redundancy

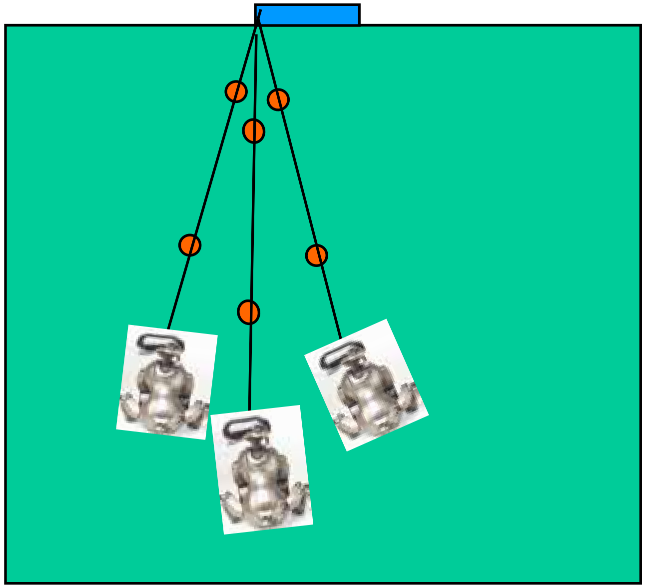


Exploiting Redundancy



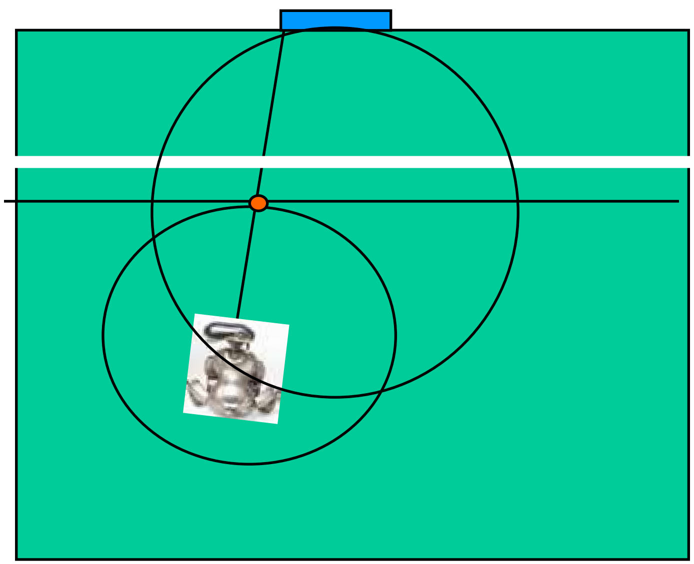


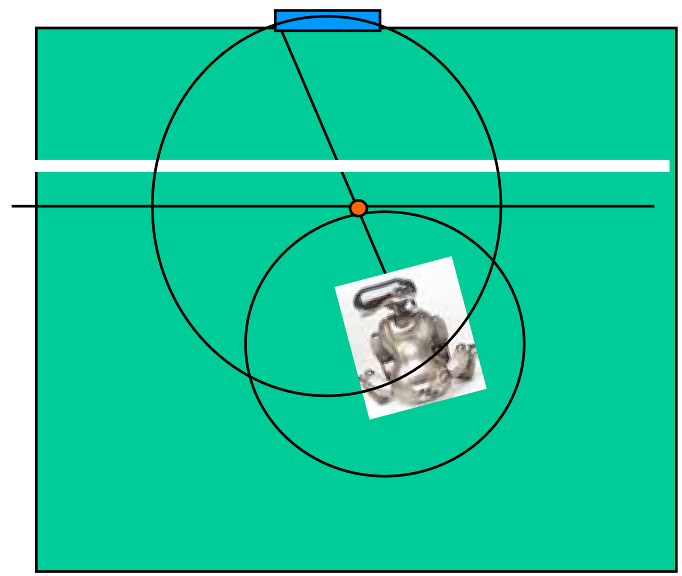




Exploiting Redundancy

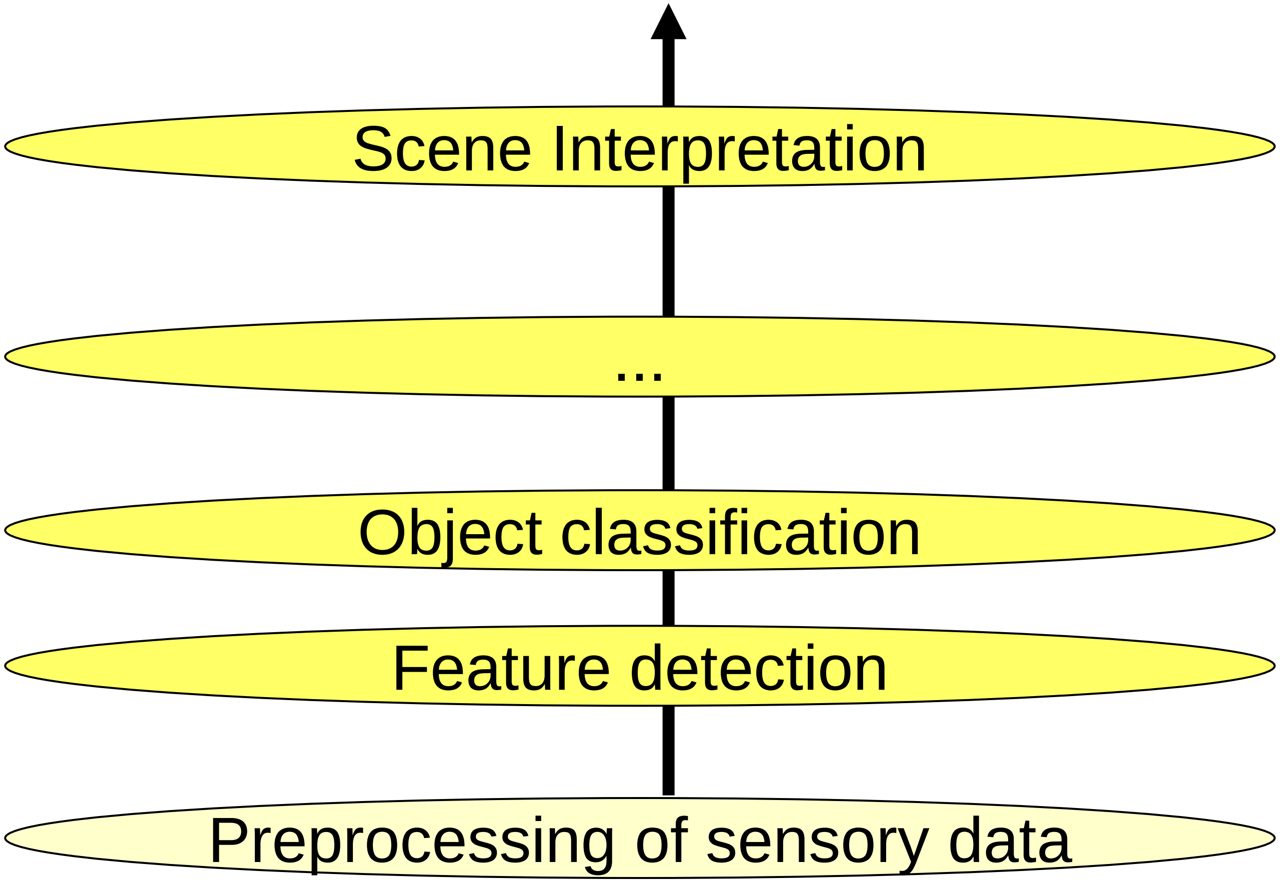
Combination yields 2 possible positions





Integration of Information

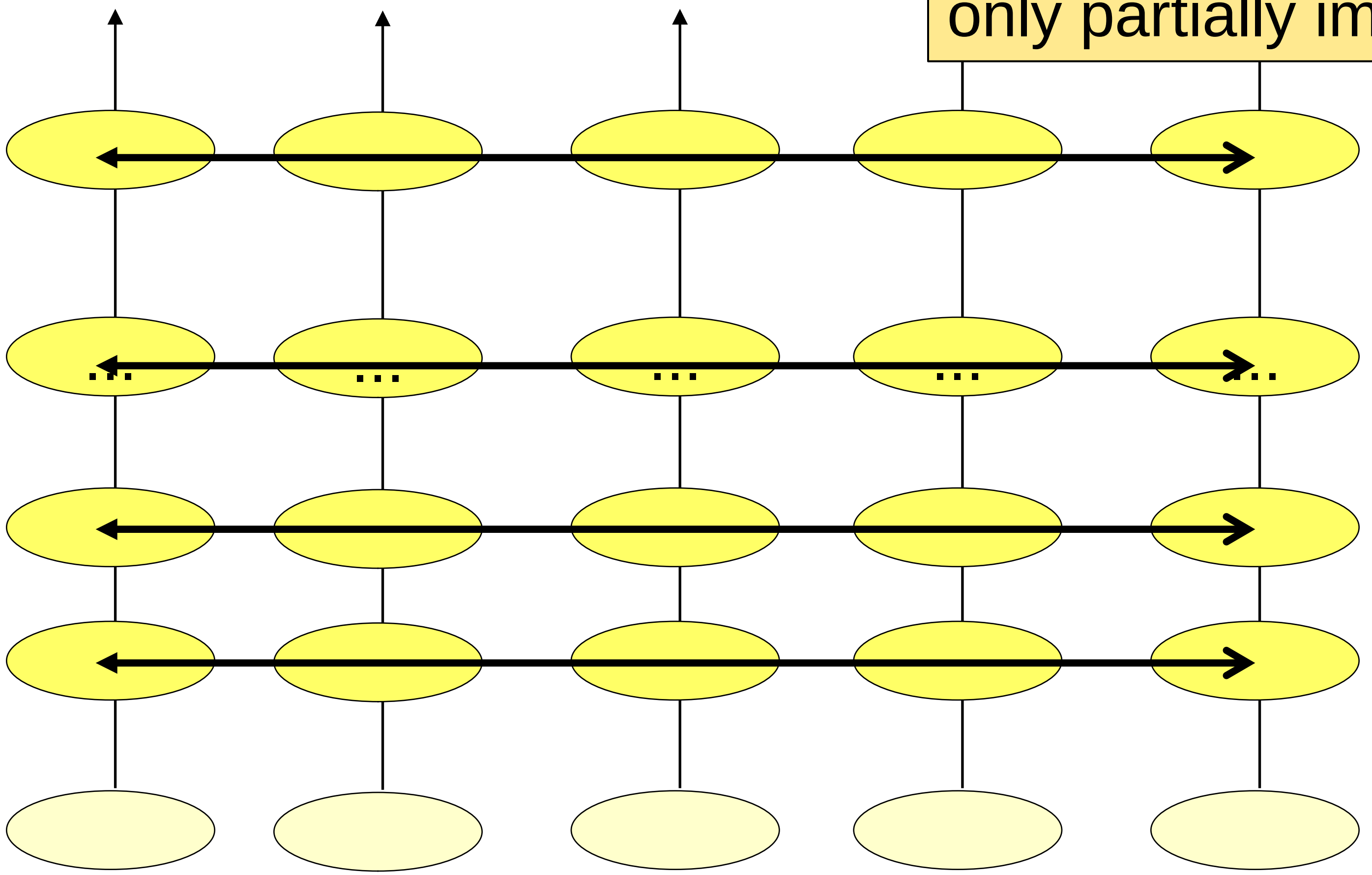
Processing on different levels



Integration of Information

Integration of different sensors

In most systems
only partially implemented

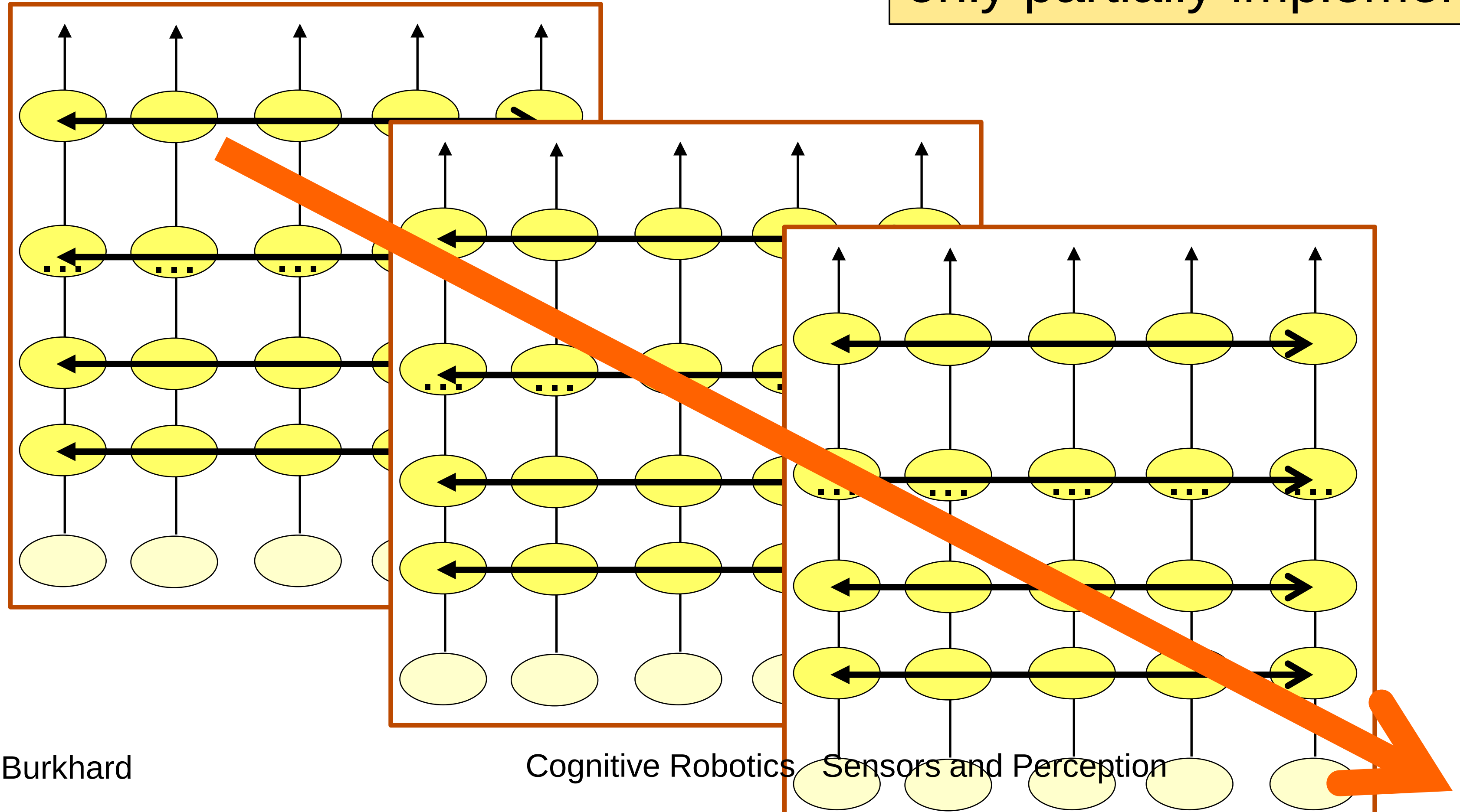


Integration of Information

Integration over time

Attendance/Focussing

In most systems
only partially implemented



World Model

World Model is called “belief”.
Because it needs not to be correct!

Objectives behind:

Keep perceived information because

- Environment only partially observable
- Observations are unreliable and noisy

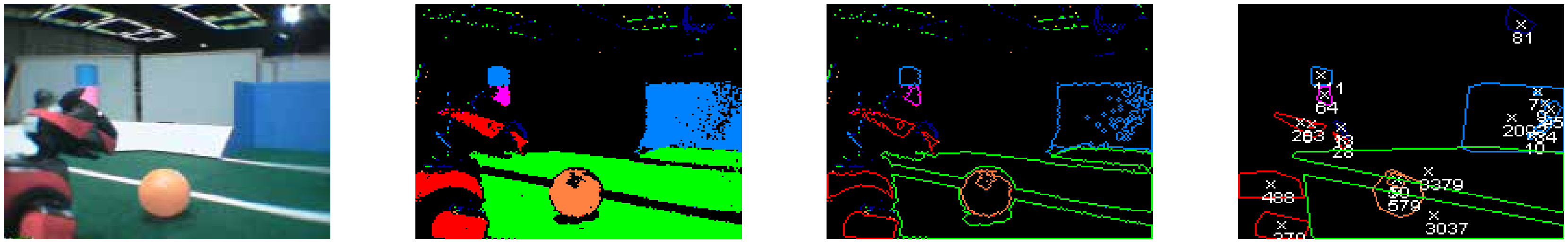
New belief= old belief + new sensory data

Using

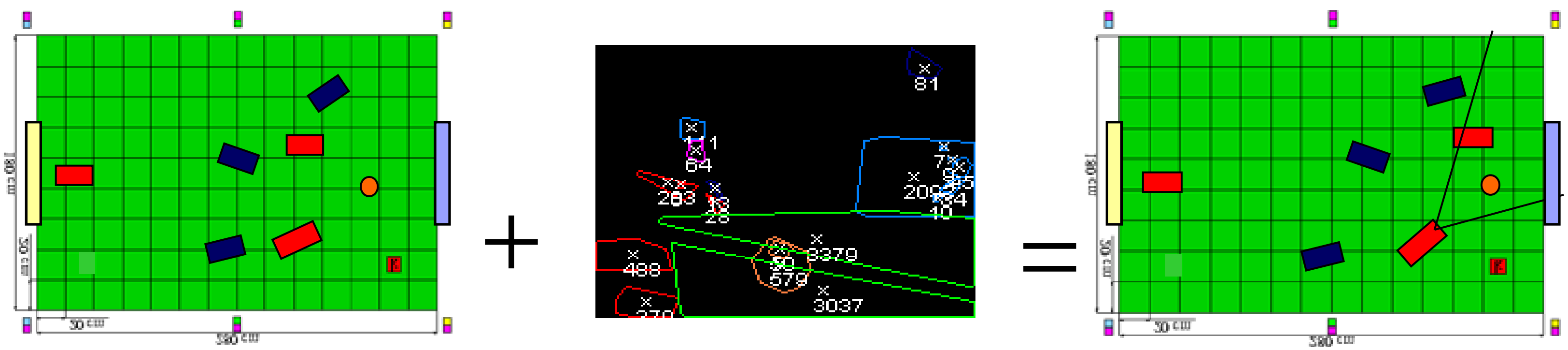
- Knowledge (e.g. maps) about the world
- Tracking of objects over time

World model

Update of belief



new perception from recent image



Belief_new := update (Perception, **Belief_old**);

Scene Interpretation

Calculcate spatial model from geometrical/topological data using

- maps
- perceived objects
- relations between objects

Usually by statistical methods, e.g. Bayesian methods

Probability to be at location s given an observation z:

$$P(s|z) = P(z|s) \cdot P(s) / P(z)$$

Where am I?

Scene Interpretation

Calculate mental attitudes of other actors using

- communication
- observation
- behavior patterns

What will he do?