

Input Devices & Tracking

Hannes Kaufmann

Interactive Media Systems Group (IMS)

Institute of Visual Computing &
Human-Centered Technology

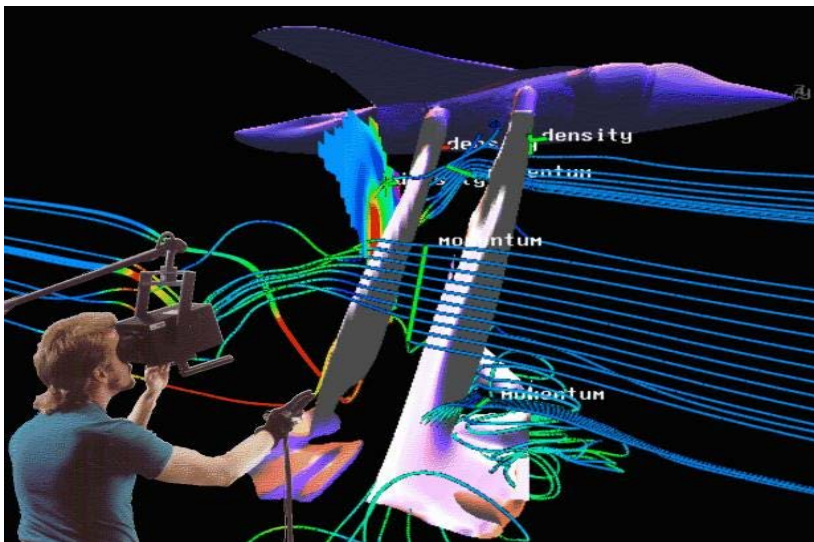
Based on material by Alexander Bornik, Dieter Schmalstieg, Oliver Bimber and Skip Rizzo

Motivation

- Last time: Various VR Application Areas
- VR/AR environment =
Hardware setup + VR software framework +
Application
- Detailed knowledge is needed about
 - Hardware: **Input Devices & Tracking**, Output Devices, 3D Graphics
 - Software: High-level programm., VR frameworks
 - Human Factors: Usability, Evaluations, Psychological Factors (Perception,...)

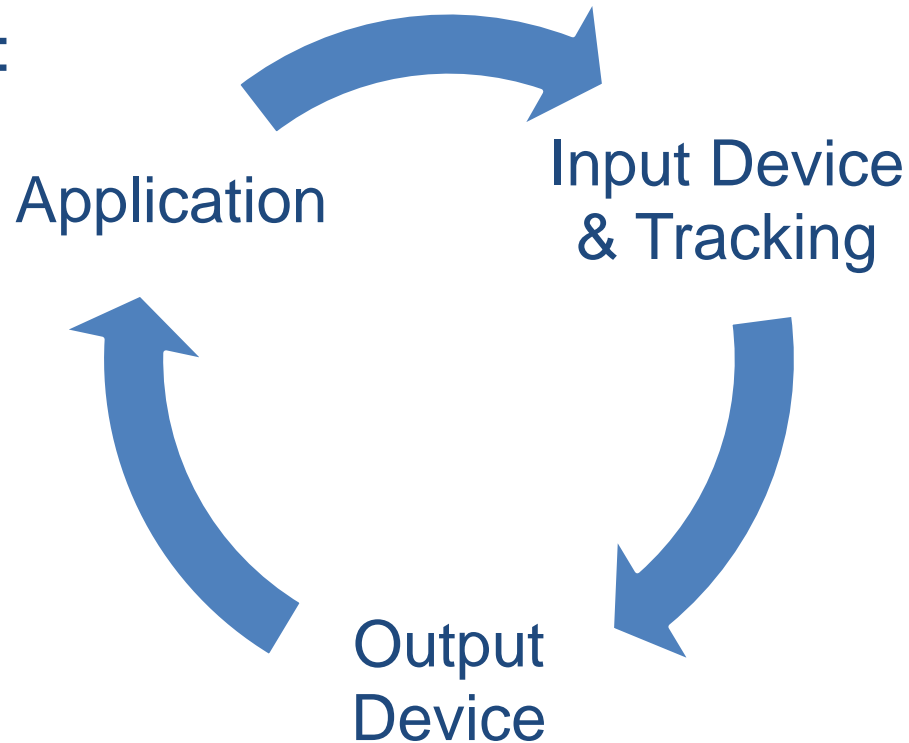
Input devices

- Hardware that allows the user to communicate with the system
- Input device vs. interaction technique (IT) (e.g. zoom)
- Single device can implement many ITs



The Interface Problem

Dependencies:



There is **not** a single ideal solution for all applications!

Examples: VR application for children - HMD?, data gloves?

CAD engineer – input by gestures?

How to find a good solution: **Know the possibilities!**

Input Device Classification

- Discrete – Continuous
- Degrees of Freedom
- Active – Passive
- Technological Principle



Degrees of Freedom (DOF)

- Number of independent values/measurements
- Commonly used:
 - (2DOF -> 2D e.g. mouse)
 - 3DOF -> position
 - 3DOF -> orientation (rotation relative to coordinate system axes): roll / pitch / yaw
 - 6DOF -> position + orientation

Discrete - Continuous

- Discrete / event-based
 - Generate one event at a time based on the user
 - Examples: Keyboard, Pinch Glove (see picture), Buttons
- Continuous / sampled
 - Continuously generate events in isolation (passive) or in response (active) to user action
 - Examples: Trackers, Datagloves
- Hybrids (e.g. mouse)
- Miscellaneous input
 - Speech
 - Gestures
 - Locomotion devices



Active / Passive Devices

- Purely active input device:
 - Requires user interaction
 - E.g. keyboard, mouse, joystick, dials, sliders
- Passive device:
 - Generate data continuously
 - No input required
 - E.g. motion tracker

Technological Classification

- Classification by Operating type
 - Architectural / Technological principle
 - E.g. Magnetic, Optical, Inertial,...
 - Regardless of actual implementation
 - How (tracking) data is acquired
- Classification by Implementation
 - Principle of (tracking) data acquisition
 - Physical principle
 - Types of sensors used

Overview

- Desktop Input
- Symbolic Input
- Tracking (6DOF)
 - Criteria
 - Technology
 - Mechanical, Magnetic, Optical, Inertial, Time of Flight, Hybrid
- Haptics
- Locomotion

For all devices:

- Technological Principle
- Technological Advantages & Disadvantages
- Human Factors Advantages & Disadvantages

Desktop Input Devices

Used in Desktop VR:

- Keyboards
- 2D Mice and trackballs
- Joysticks
- Pen based tablets
- 6-DOF devices for the desktop
- Haptics

Joysticks



- Isotonic Joystick
 - Movement continues until handle pushed back to neutral position
 - Rate control (not position control)
 - Haptics: force-feedback
 - Easy to use
- Isometric Joystick
 - Handle does not move
 - Output varies with force that is applied
 - E.g. laptop trackpoint



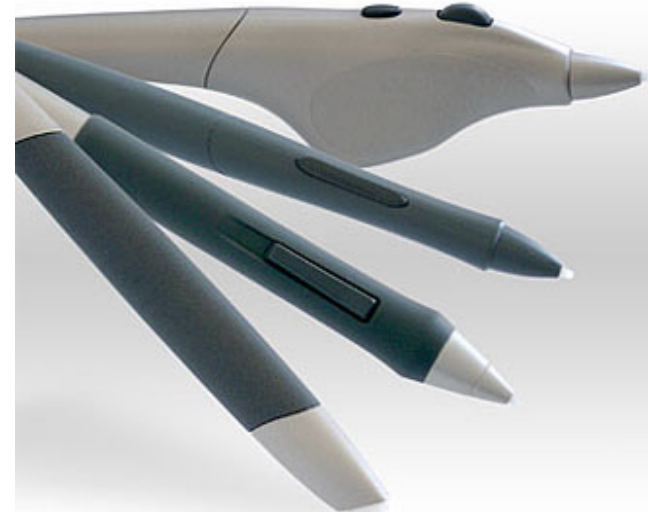
Pen based tablets



e.g. Wacom tablets

- Very precise input
- Absolute input!

Mice/Joystick: relative input



6-DOF Desktop Input Devices

3Dconnexion:

- SpaceMouse
- SpaceBall 5000



Developed for desktop 3D interaction

- CAD
- Used additionally to mouse
- Fine manipulation can be difficult
- Needs time to learn usage
- 3D experts very quick but needs practice
- Not ideal during mobile work



LEAP Motion

- 2 mono infrared (IR) cameras (300 fps), 3 IR LEDs
- LEDs generate a 3D pattern of dots of IR light
- Processing of point cloud data on PC
- <1 meter distance
- Advantage: High precision
- Disadvantage: Low range, small field of view



Haptic Input for Desktop VR

Phantom device:

- Accurate 6DOF input
- Mechanical Tracking
- Force Feedback (limited)
- Restricted Area
- Only **1-Point contact** !!!

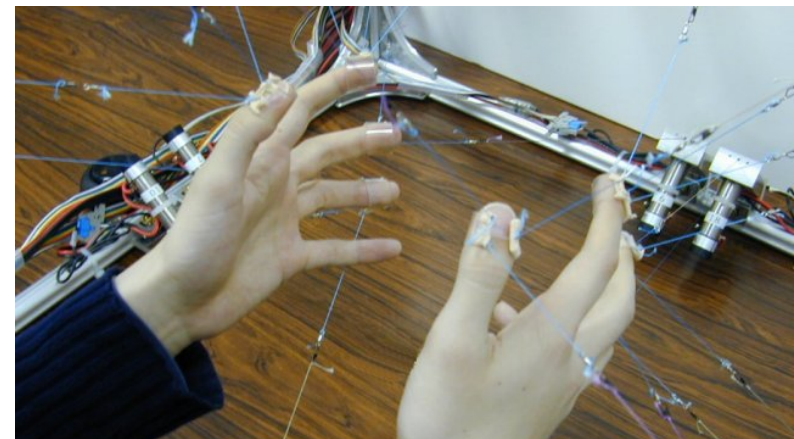


Other Desktop-bound Haptic Devices

- 6-DOF Delta device – spherical knob (Novint Falcon)
- Haptic feeling for all fingers holding the knob
- Not mobile
- Custom made research devices



Video



Symbolic Input Devices

Symbolic Input (Text, Numbers,...) in VR environments

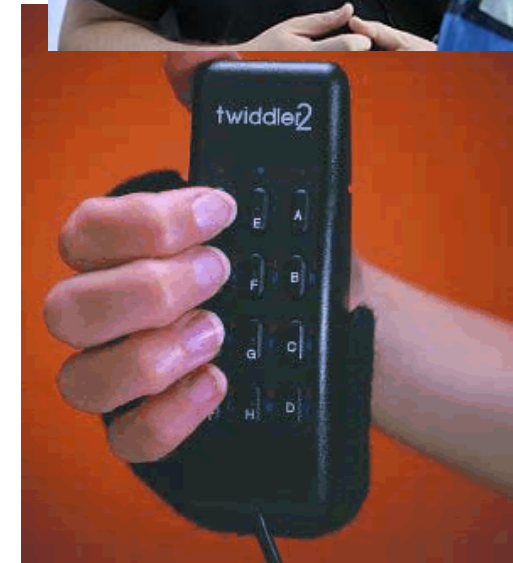
- Design annotation
- Filename entry
- Labeling
- Precise object manipulation
- Parameter setting
- Communication between users
- Markup (highlight, bold,...)

Features of Symbolic Input in 3D

- Desktop devices (e.g. keyboards) usually don't work in 3D:
 - Users are standing / physically moving
 - No surface to place a keyboard
 - Difficult or impossible to see keys in low-light environments or when wearing HMD
- Work-arounds possible
- Positive:
 - Symbolic input in 3D less frequent than in 2D

Symbolic Input: Keyboard Based Techniques

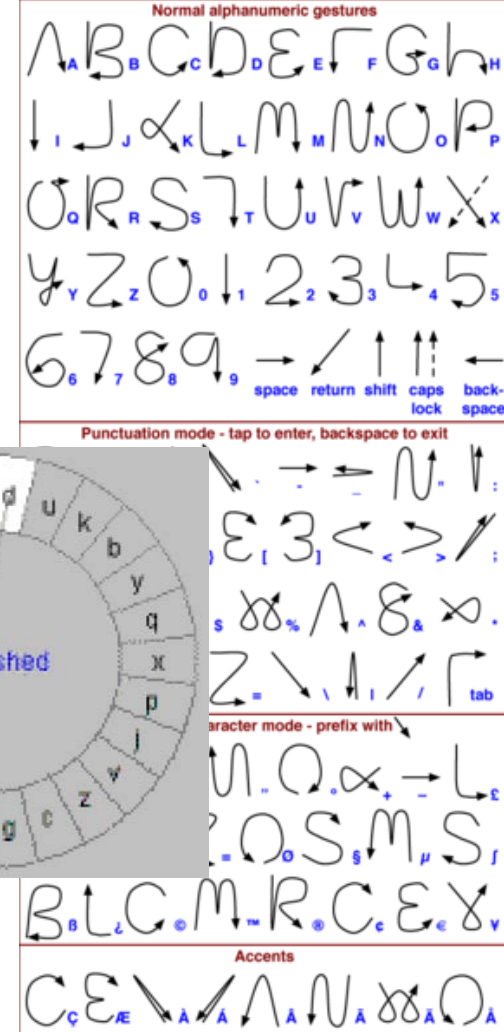
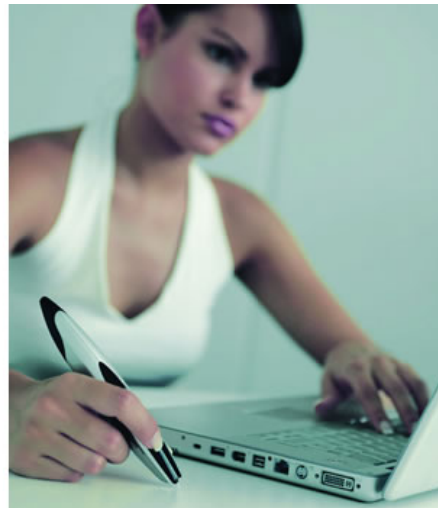
- Miniature keyboards
- Low key-count keyboards
 - e.g. mobile phone SMS; T9 system
- Chord keyboards (e.g. Twiddler)
 - Training required
- Pinch keyboard
 - Based on pinch glove
- Soft keyboards



Pen Based

- Pen-Stroke Gesture Recognition
 - PDA Graffiti Alphabet
 - Cirrin soft keyboard
- Pen Input (Digital Ink)
 - Recognition of handwriting

SCRIVO.1



Gesture-based Techniques

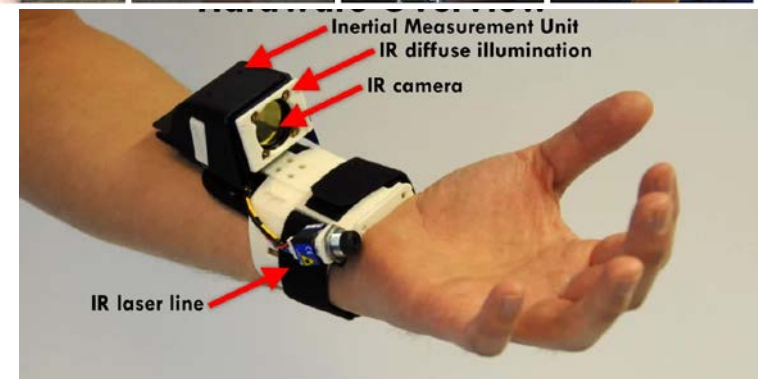
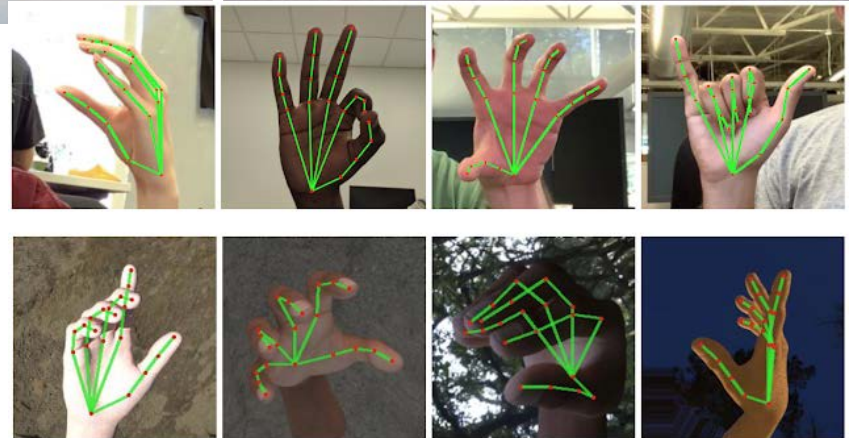
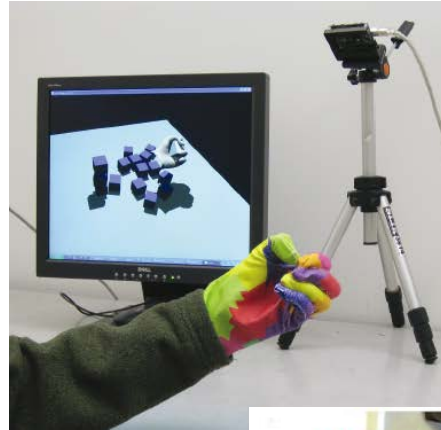
- Either by using gloves, computer vision or motion capture
 - Difficult: Fast Gesture recognition needed
 - Today more and more vision based approaches (with depth cameras)

3 Types:

- Numeric gestures
 - Use of fingers represent number
 - Both hands can be used
 - No application of this technique known
- Sign language gestures (only small percentage of population knows sign language)
- Instantaneous gestures

Sign Language Gestures

- Glove based
- 2D vision based
 - Difficult in real time
 - Occlusions
 - Using machine learning
- 3D Motion Capture solutions
 - Many approaches
 - See depth cameras & 3D User Interaction



Speech Input

- Complementary to other modes of interaction
- Got MUCH better due to machine learning approaches
- Issues to consider
 - continuous vs. one-time recognition
 - choice and placement of microphone
 - training vs. no training (= speaker independent)
 - handling of false positive recognition
 - surrounding noise interference
 - Language dependent



Speech Input - Hardware

- Headset:
 - wired or wireless (Bluetooth, RF)
 - surround sound (5.1, 7.1)
 - active noise cancelling



Tracking

Tracking systems

- Measure position and/or orientation of a sensor
 - 6 degrees of freedom in space
- Most VEs track the head and the hand
 - Correct viewing perspective
 - Interaction



Quality Factors & Criteria

- Degrees of freedom (3D pos.+orientation → 6DOF)
- Range or working volume
- Accuracy
 - Static
 - Dynamic
- Time for measurement - phase lag, “Real-time”
- Update rate (measures/sec)
- Signal to noise ratio
- Registration
- Sociability (Tracking should not hinder freedom)

Range, Working Volume

- Amount of space, where tracker works
- Range:
 - Distance from sensors
 - FOV of camera
- Working volume:
 - Volume resulting from ranges
- Time until stability degenerates
 - E.g. Drifting of inertial trackers

Static & Dynamic Accuracy

Static Accuracy: Maximum deviation from fixed tracker position to fixed reference value

- Influenced by:
 - Receiver sensitivity, Transmitter s/n ratio, A/D converter resolution, (Analog component noise tolerance levels), Environmental effects, Algorithmic errors, Installation errors

Dynamic Accuracy: Accuracy as sensor is moved

- Dependent on static accuracy
- Influenced by:
 - Processor type, System architecture, Time dependent system components

Tracking System: Generic Architecture (Latency, Lag, Update Rate)

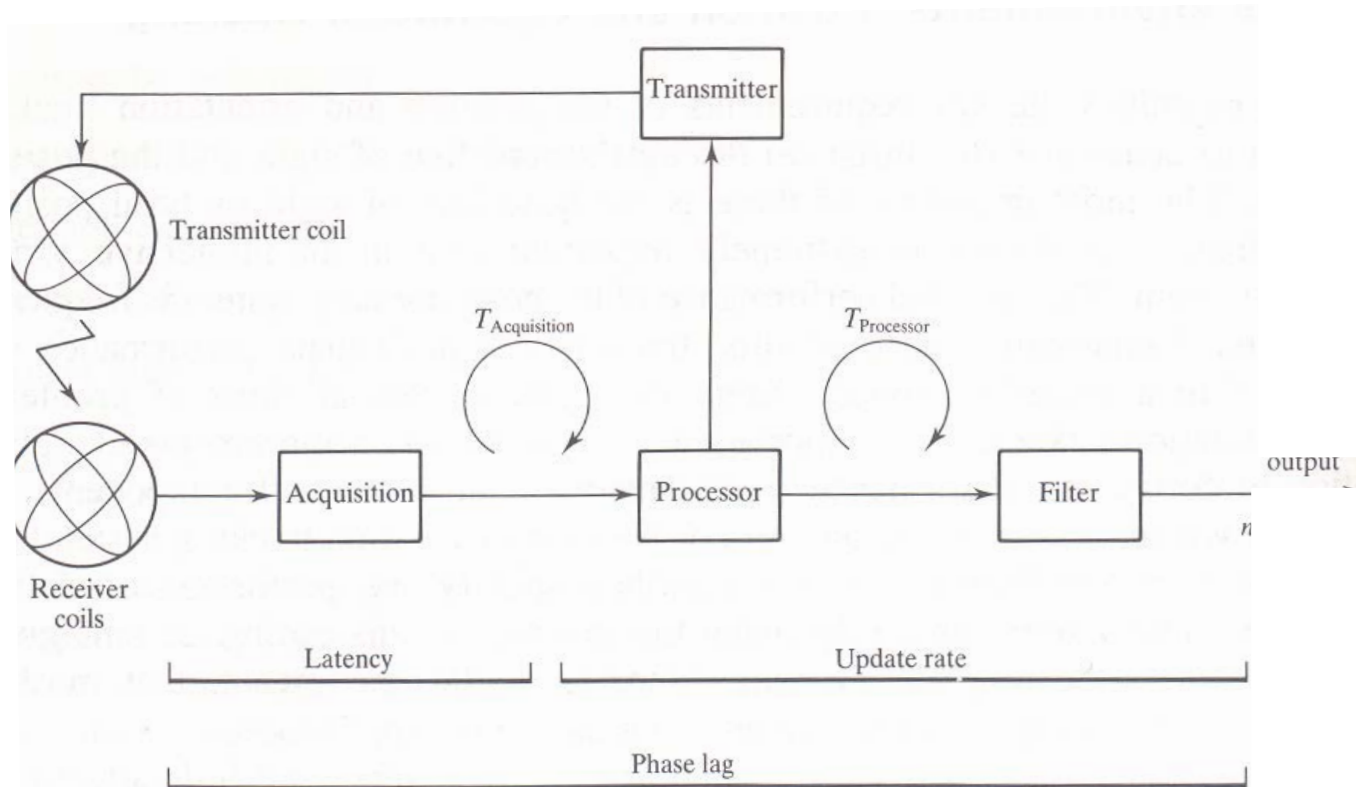


Figure 6.11 Generic architecture for a head tracker system.

Latency & Update Rate

Latency: Rate at which acquisition portion of system can acquire new data

- Aspects:
 - Hardware limitations
 - Time to sense change in receiver's position
- Example - optical tracking: Time to capture image

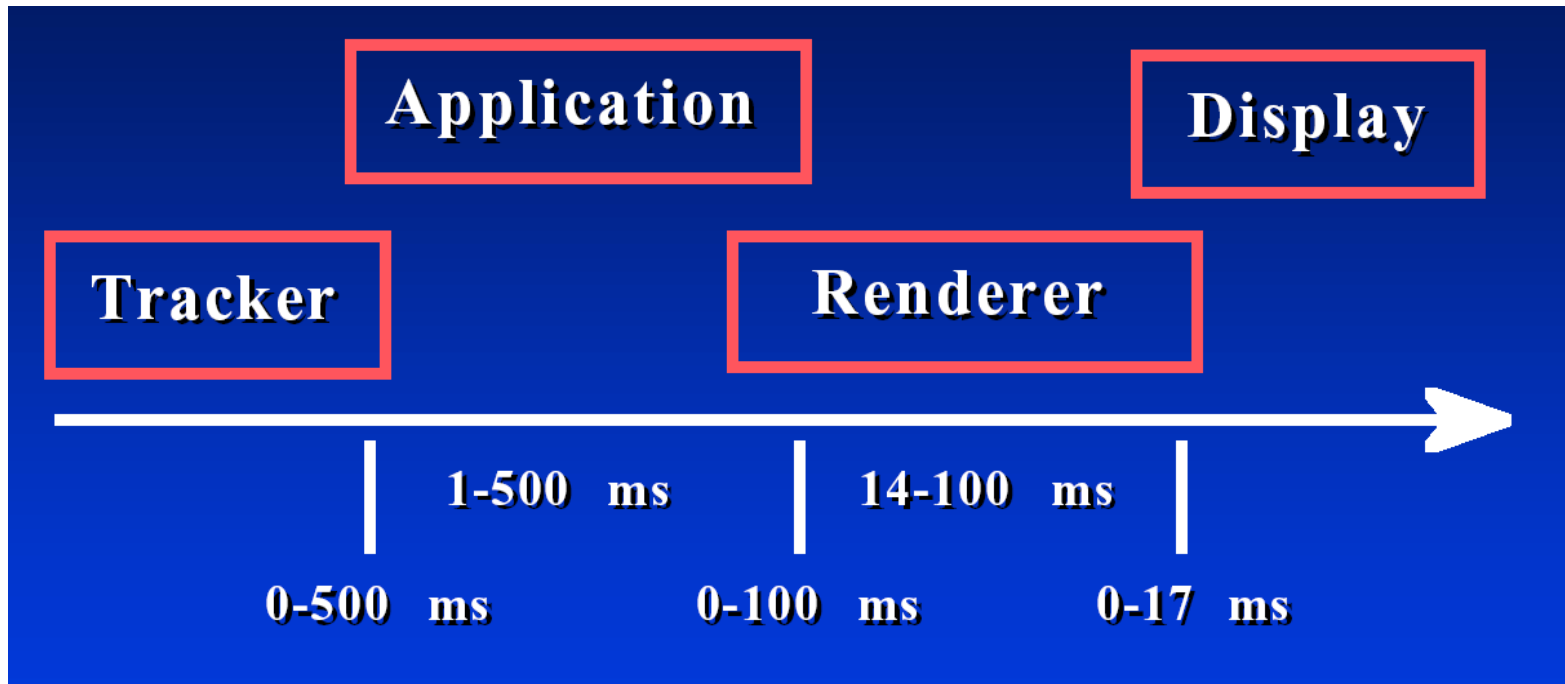
Update Rate: Trackers ability to output position and orientation data

- Influenced by:
 - System processor
 - Algorithm
- Example: Time to process image & extract data

Phase Lag

- Latency + update rate
- Fast motions - tracker cannot keep track
- Contributors:
 - Architecture
 - Processor type
 - Algorithm
- Improving one of these can reduce lag

Visually coupled System – Still Interactive?



Tracker: phase lag

Application: time to process tracking data

Time for rendering

Interactivity: 15-20 FPS; Recommendation: 90 FPS

Low Latency – State of the Art

- On low latency in VR – see Michael Abrash's (Oculus VR) talk about

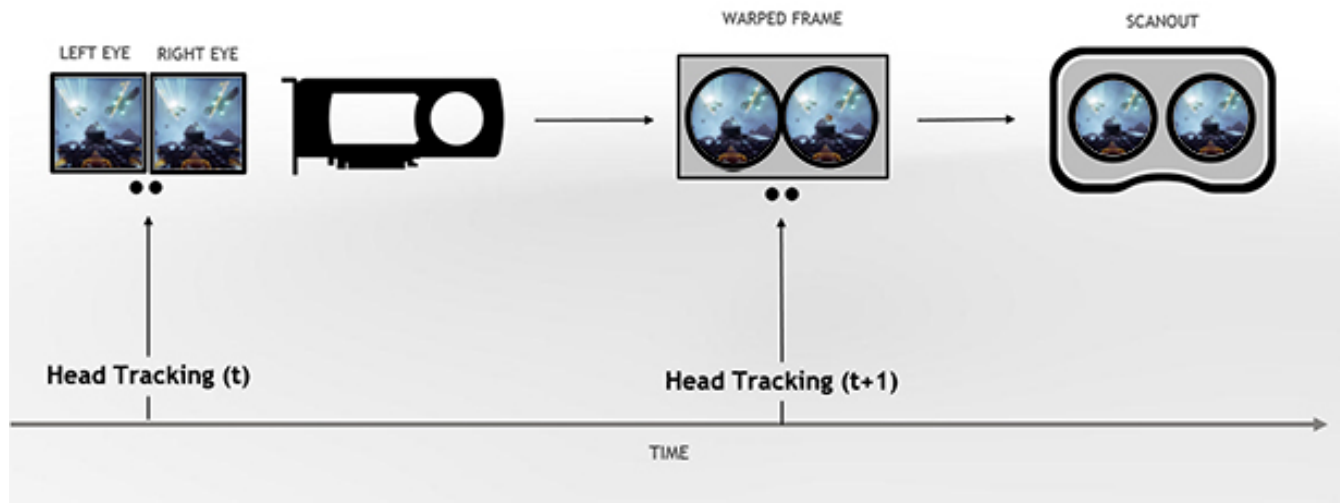
**What VR Could, Should, and Almost Certainly
Will Be within Two Years**

<http://youtu.be/G-2dQoeqVVo>

(Feb. 11, 2014)

Asynchronous Time Warp

CONTEXT PRIORITY FOR ASYNCHRONOUS TIME WARP
REDUCES HEAD TRACKING LATENCY



- Front Buffer Rendering
- Single Pass Stereo

Still 7ms difference between left/right eye on single display HMDs -> 2 displays used

Registration (esp. for AR)

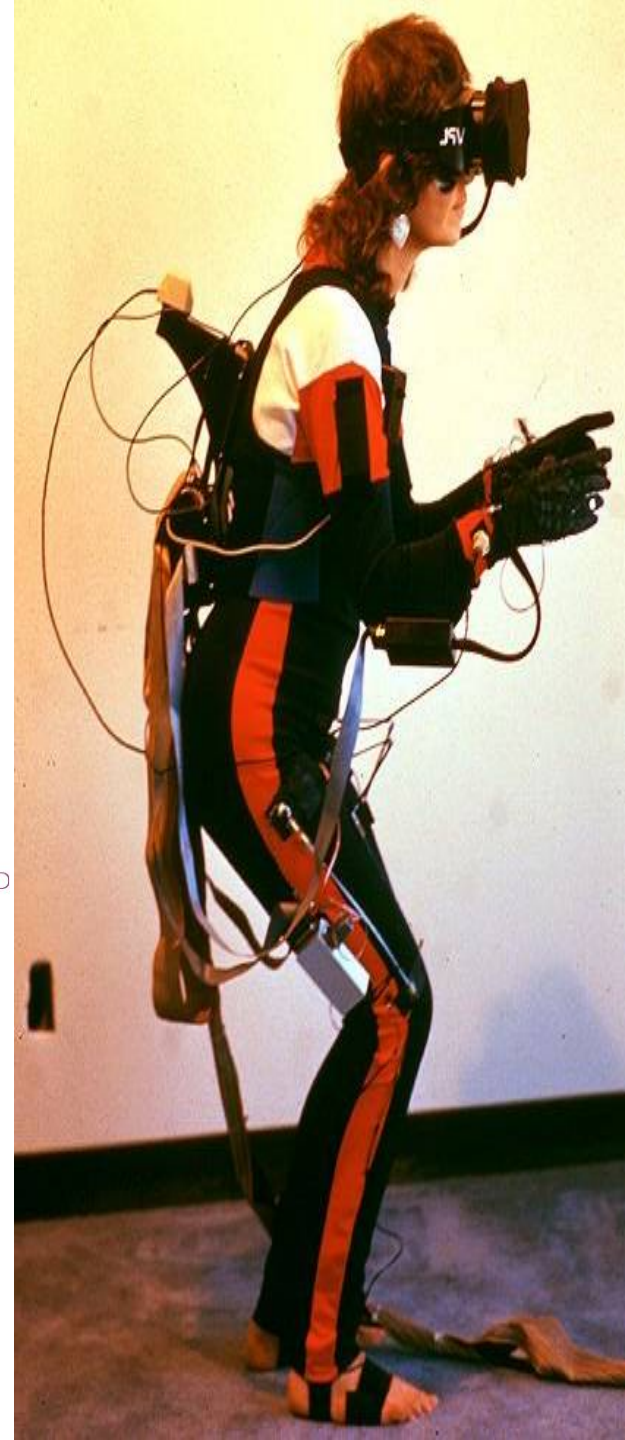
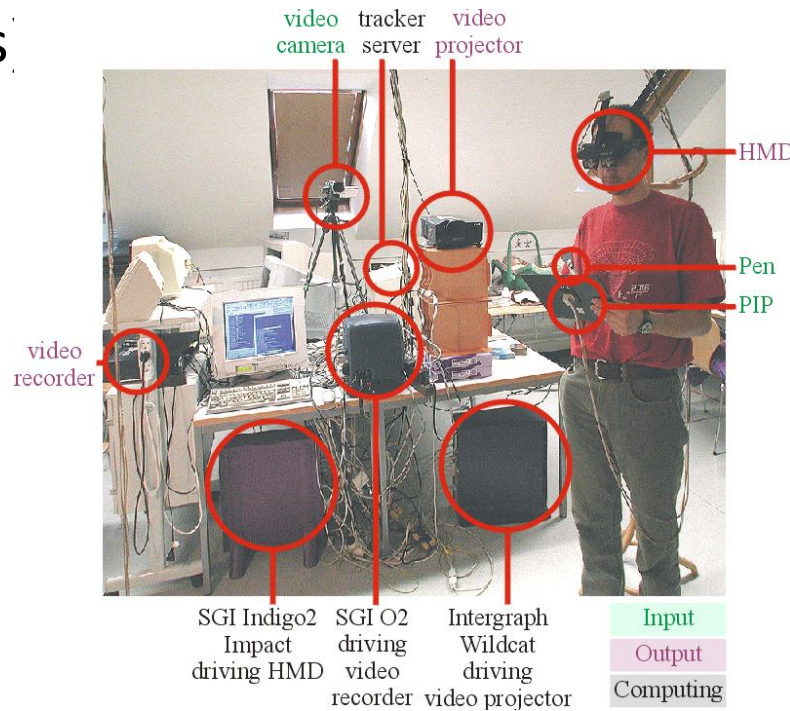
- Correspondence:
 - $\text{Actual position / orientation} = \text{Reported position / orientation}$
- Good accuracy/resolution - does not imply good registration
- Important for more than one sensor
- Important usability factor – can cause bad side effects (Cybersickness)

Effects of Inaccurate Tracking

- Objects appear where they are not
- Proprioceptive conflicts
 - Static: limb location conflicts
 - Dynamic: visual delay (lag)
 - Limb jitter or oscillation
- Misregistration
 - Constant OK (immersive, non see-through)
 - Changing can hurt
- User forced to adapt
- Simulator sickness (Cybersickness)

Sociability

- Can it be used by multiple users?
- Is interaction possible then?
- Restrictions by:
 - Operating principle (number of sensors, occlusions)
 - Weight
 - Size
 - Phase lag
 - Wires
 - ...



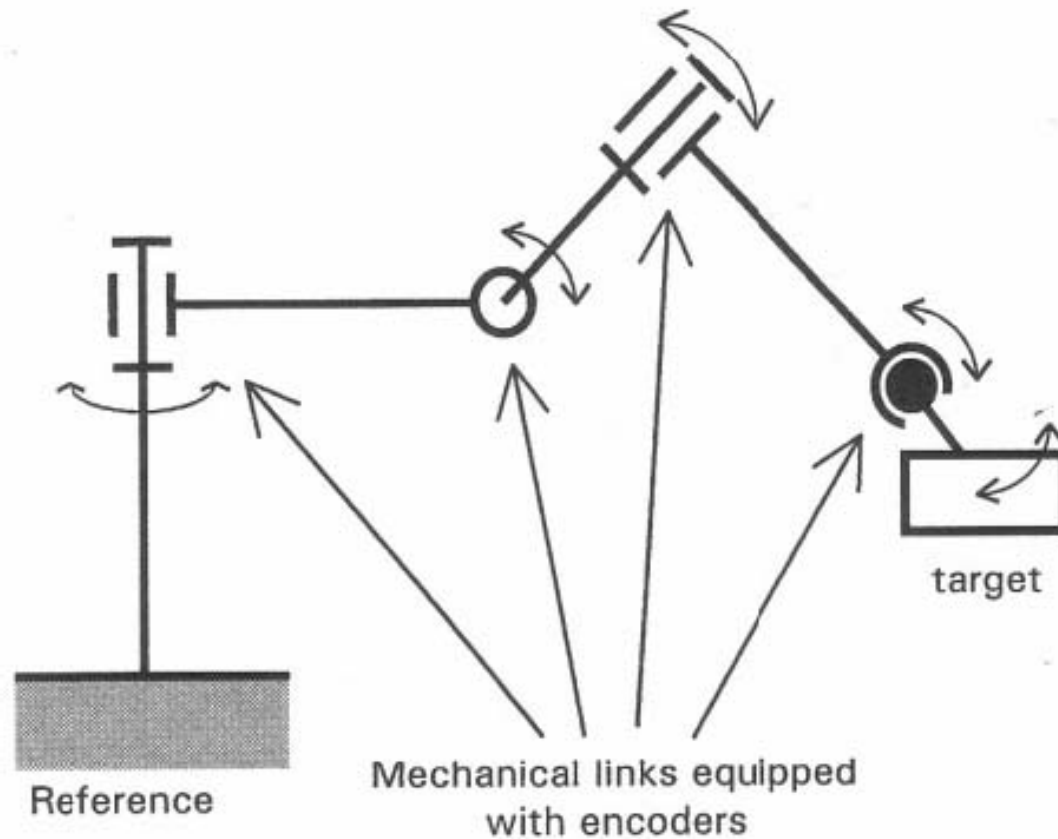
Classification by Operation Principle

- **Mechanical**
- **Magnetic** (AC/DC/passive)
- **Optical**
 - Marker based
 - Natural feature; Vision (Gestures)
 - Motion capture
- **Inertial** (gravity, acceleration)
- **Time-of-Flight & Frequency Measuring**
 - (Radio Waves) GPS ...
 - (Sound waves) Acoustic
- **Hybrids** (combination of multiple)

Structure – Tracking Hardware

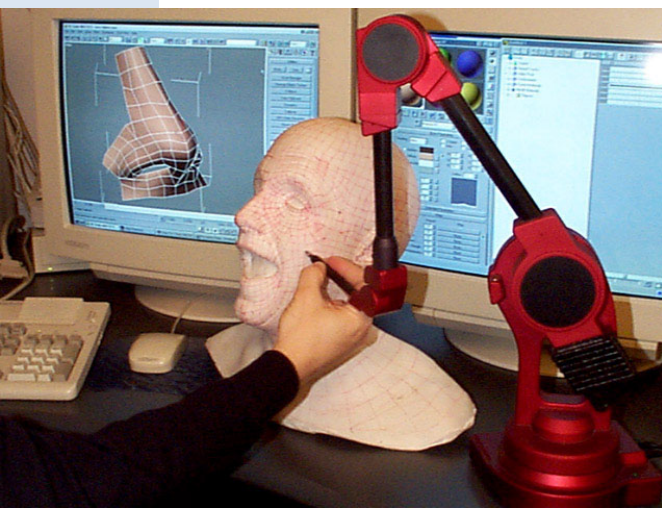
- Technological Principle
- Technological Advantages & Disadvantages
- Human Factors Advantages & Disadvantages
- Examples

Mechanical Tracking



Mechanical Tracking

- Oldest tracking technology
- Measure angle of human joints
- Potentiometers for angular measurements
- Position reconstructed using kinematics



Mechanical Tracking

■ Advantages

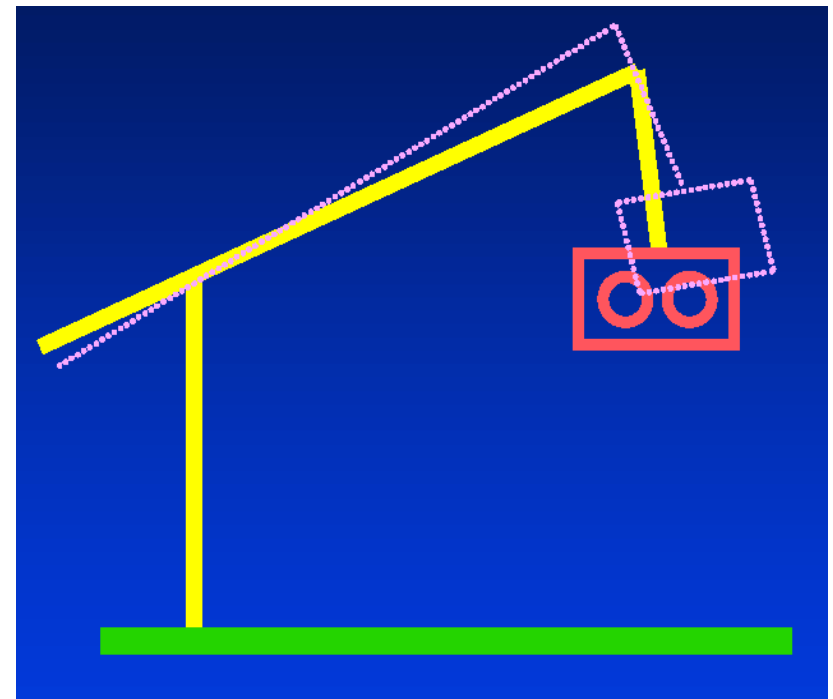
- Technology well known
- Tracks multiple users in real-time

■ Disadvantages

- Ground-referenced
 - Limited working volume
- Uncomfortable
- Exoskeleton
 - affects movement
 - Need to adjust per person, can be ill-fitting
 - No global movement tracking

Mechanical

- Pros
 - High precision
 - Very fast
 - High S/N
- Cons
 - Infrastructure
 - Compounded measurement error
 - Restricted working volume
 - Restricted motion
 - Not very sociable



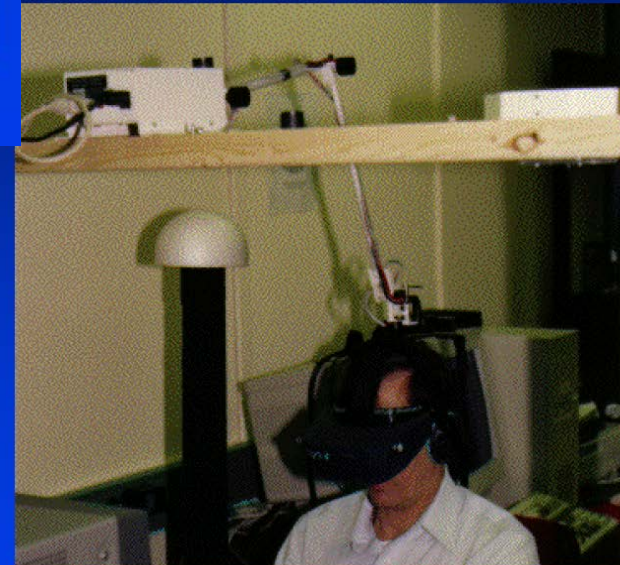
Fakespace

Boom
3C/HF/Push



Shooting Star
Technologies

ADL-1



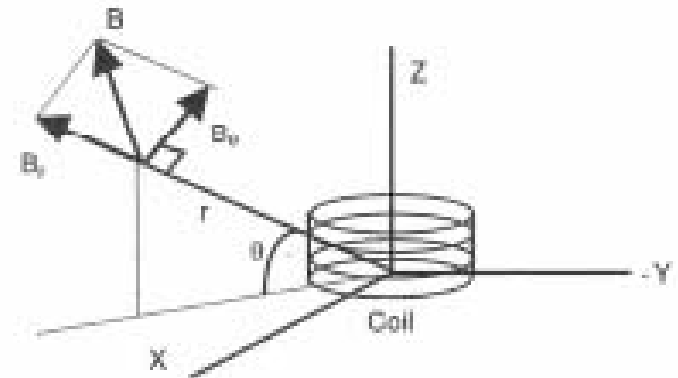
CyberForce

- Most natural interface - the human hand
- Immersion Corp. Products:
 - CyberGlove
 - CyberTouch
 - CyberGrasp
 - **CyberForce**
- 6DOF tracking
- Usability problem



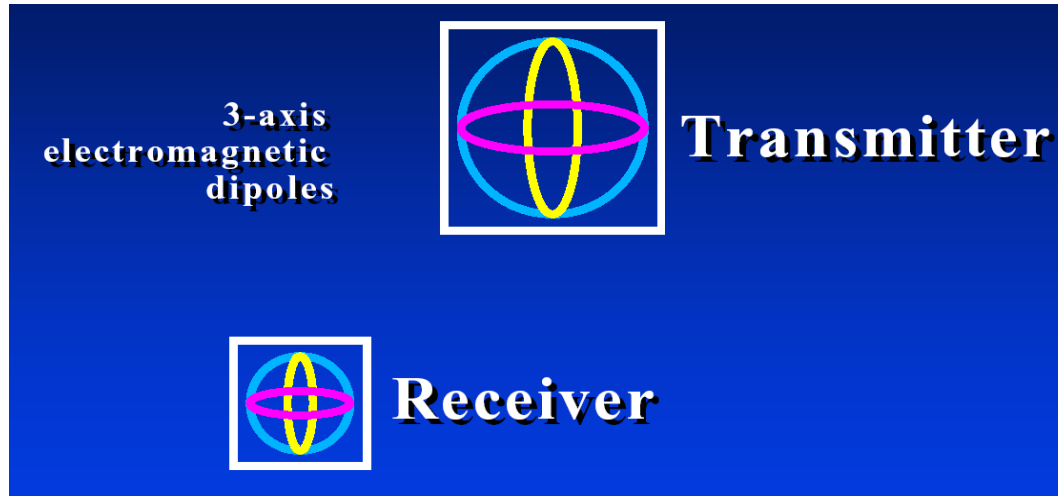
Magnetic Tracking

- Circulating electric current in coil \rightarrow magnetic field is generated
- At distance r , the field has polar coordinates B_θ and B_r



- Magnetic flux is created in receiver (magnetic field sensor)
- Magnetic flux: Function of distance and orientation relative to coil
- To measure position and orientation of receiver in space, emitters consist of 3 orthogonal coils and receivers of 3 sensors \rightarrow combination of three elementary orthogonal directions

Magnetic Tracking - Alternating Current (AC)

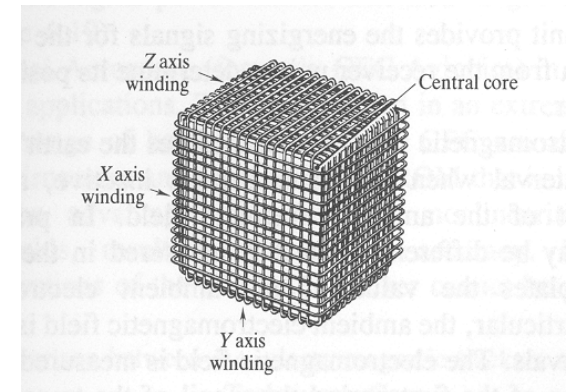


- Transmitter: electromagnetic dipoles \rightarrow AC field
- Receiver: induced voltage measured
- Voltage dependent on:
 - Distance transmitter \leftrightarrow receiver \rightarrow position
 - Orientation of coils in magnetic fields \rightarrow orientation

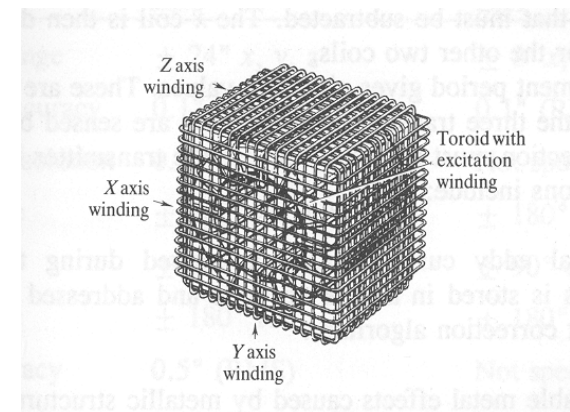
DC Magnetic - Overview

- Similar operation to AC
- Short DC pulses vs. contin. AC
- Transmitter
 - 3 orthogonal coils, cubic core
 - Mounted rigidly to reference structure
 - Driven sequentially
- Receiver
 - 3 orthogonal coils, cubic core
 - Additional energizing coil
 - All 3 measures concurrent
- 4 phase measurement (4.: earth magn. Field – subtract out)

Transmitter



Receiver

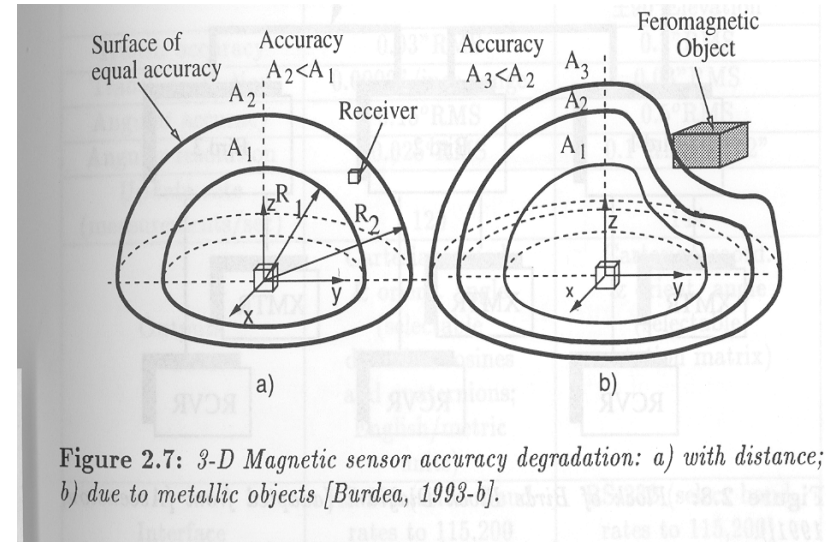


Magnetic: Advantages

- No line of sight restrictions – continuous data
- Small/lightweight sensors
- Wireless versions available
- Off the shelf available, robust
- High update rate (200 Hz or more)
- Price depends on tracking range – can be relatively inexpensive

Magnetic: Disadvantages

- External noise
 - At metal, power wires,...
 - Unwanted eddy currents (Field distortions)
 - conductive material will distort the magnetic field (monitors)
- Field strength $\sim 1/d^3$
 - Jitter at boundaries (Filtering increases phase lag)
 - Cannot increase electromagnetic field strength -> might have consequences on humans
- Not the best accuracy
- Limited working volume (max. 3-5m radius)



AC Magnetic Products

- Polhemus (Fasttrak, InsideTrak, LongRanger ...)



Longranger:
sphere ~ 50 cm
with orthogonal coils



DC Magnetic - Ascension

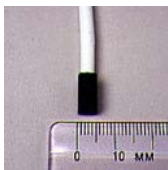
[Video of our former VR lab](#)



MotionStar



PciBird



MiniBird

AC/DC Magnetic - Specifications

AC

Model	Latency	Max. # of sensor	Range	Accuracy	Costs
FastTrak	4 ms	4	~ 3 m	2mm, 0,15 degree	€ 7.000
IsoTrak	20 ms	2	~ 1,5 m	0,25 cm; 0,75 degree	€ 3.500
LongRanger		4	~ 10 m		
InsideTrak	12 ms	2	~ 1,5 m	1,25 cm; 2 degrees	€ 1.200
UltraTrak Pro	6 ms	32	4,5 m	2,5 – 7,5 cm, 3 deg.	€ 75.000

DC

Model	Latency	Max. # of sensor	Range	Accuracy	Costs
Flock	20 ms	30	~ 1 m	0,25 cm; 0,5 degree	€ 3.000
Flock 10	20 ms	30	~ 3 m	3 cm; 1 degree	€ 8.000
MotionStar	20 ms	120	~ 3 m	3 cm; 1 degree	33.000
MotionStar Wireless	20 ms	18	~ 3 m	3 cm; 3 degree	€ 64.000

Magnetic: MotionStar

- Wireless Magnetic Tracking System
- Up to 20 sensors
- Motion Capturing
- More than one transmitter possible



Sensors: 20/suit

**100 updates/sec
3 meters range
from base unit
Resolution < 2 mm
and < .2 degrees**

**Electronic unit
(2 hours battery life)**



Razer Hydra & Sixense STEM System

- Low cost PC controller
- 6 DOF AC magnetic tracking
 - Very low field strength (1/40 earth magn. field)
 - Switching polarity with $\sim 8,000$ Hz
- 1 mm, 1 deg accuracy
- Hydra: Range ~ 1.5 m from base
- Low latency, no line of sight
- Distortions from metallic objects
- Sixense STEM System
 - Wireless, longer range (3.5m), 5 sensors..
 - Additional inertial sensor as distortion compensation!



Passive Magnetic

- Devices
 - Compass (magnetometer)
- Measure
 - Earth's magnetic field
 - Heading only (2DOF)
- Pros:
 - No infrastructure, Available outdoors, Absolute reference
- Cons
 - Affected by
 - Other ferrous and magnetic materials
 - Active magnetic sources
 - Limited information (2D orientation only)
 - Field is very tilted in significant parts of the world



Optical Tracking

Optical Flow Motion Estimation



- Easy to implement
- Watch pixel movement
- Compare movement vectors of pixels
- Indicates movement direction
- Not very accurate!

Vision Trackers

Types

– Fiducials

- ARToolkit
 - [Video1: Magic Book](#)
 - [Video2: Cockpit Layout](#)
 - [Video3: Invisible Train](#)

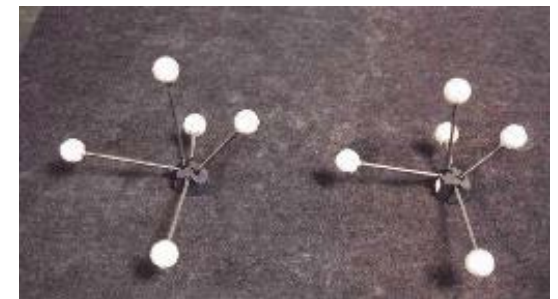
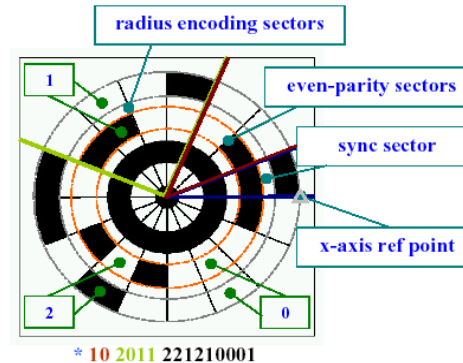
- ARTag
- Intersense

– Markers/LEDs

- Active / Passive

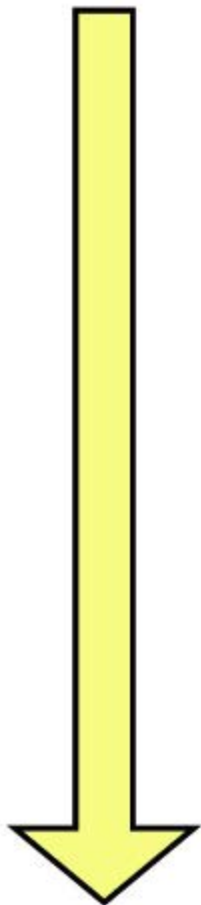
– Natural features

- Images
 - [Video: IKEA Katalog](#)
- 3D objects
- Various research projects



Vision-based Tracking

- Most devices have cameras



- Marker-based
 - Artoolkit port



- Natural Features
 - SIFT, Surf, FERNs
 - Active search tracking



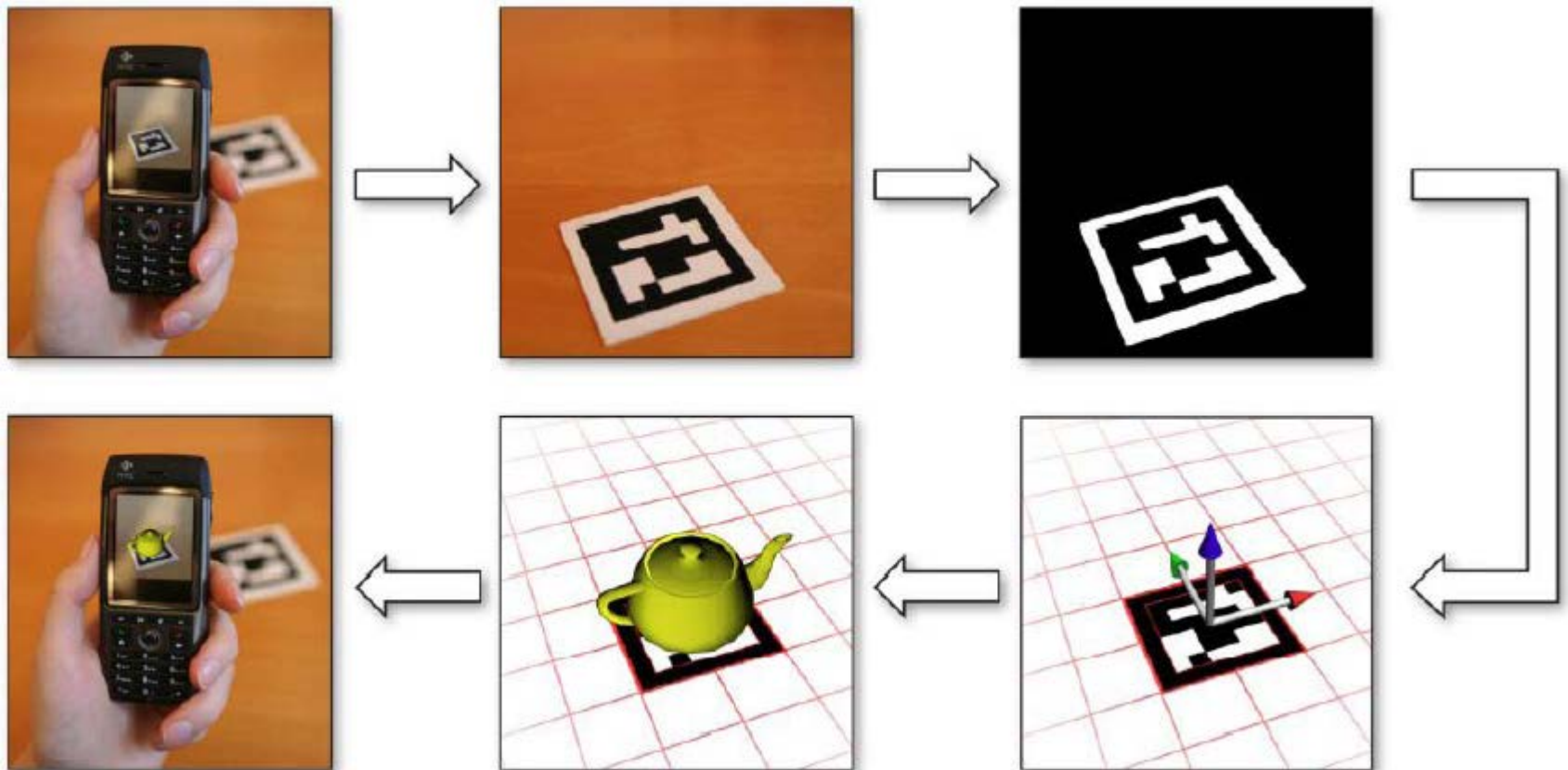
- Recognition, Initialization, Mapping and Tracking

Marker Tracking



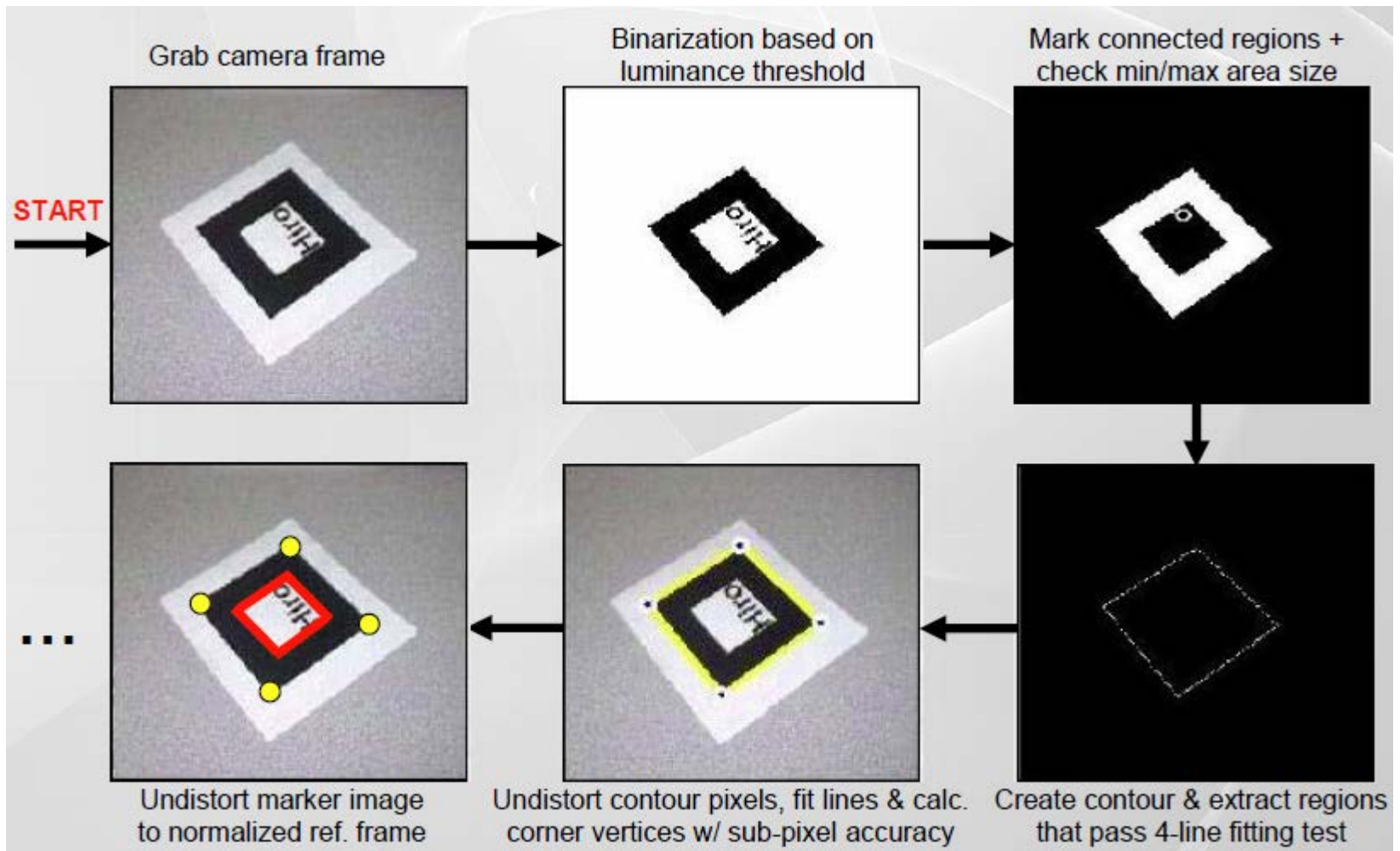
- Has been done for more than 15 years
 - Mobile phones today are faster than computers of that time
- Several open source solutions exist
- Fairly simple to implement
 - Standard computer vision methods
- A rectangular marker provides 4 corner points
-> enough for pose estimation!

Marker Tracking: Pipeline Overview

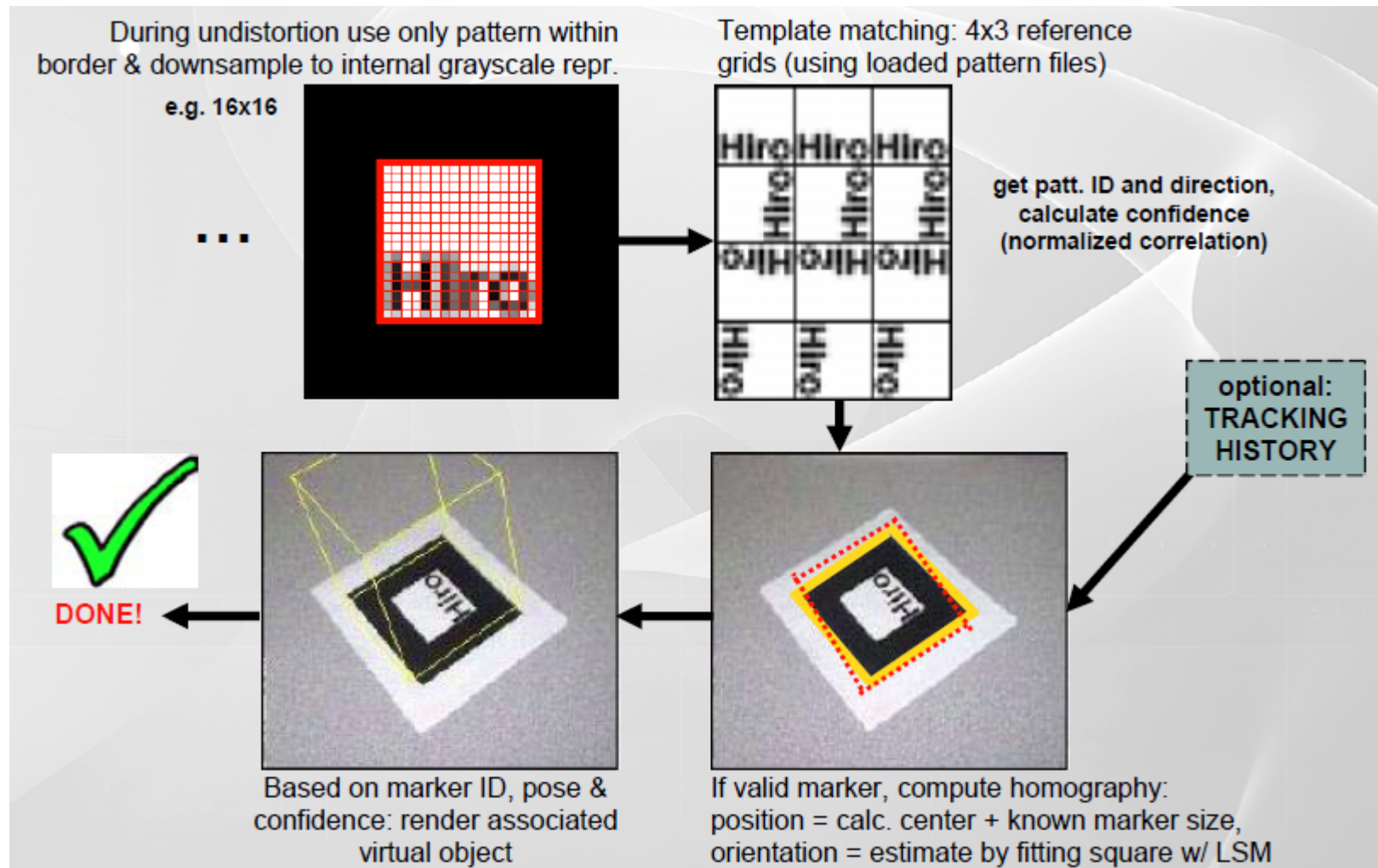


Goal: Do all this in less than 20 milliseconds on a mobile phone

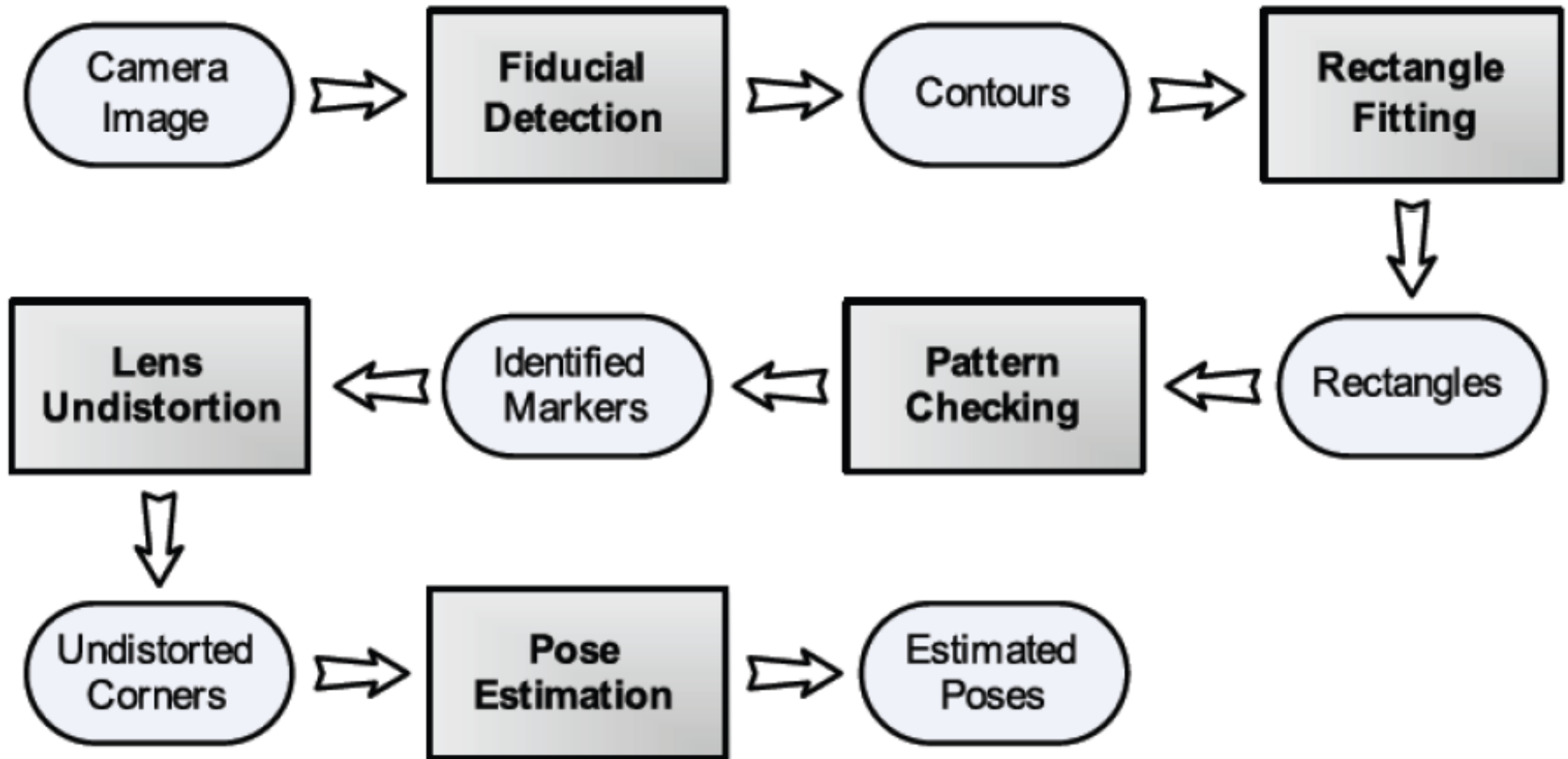
Marker Tracking: Overview 1/2



Marker Tracking: Overview 2/2

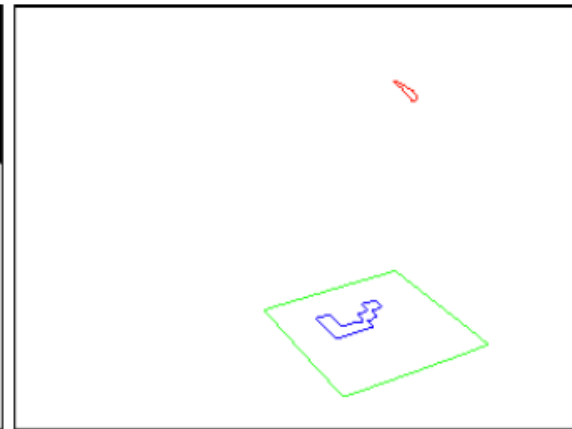
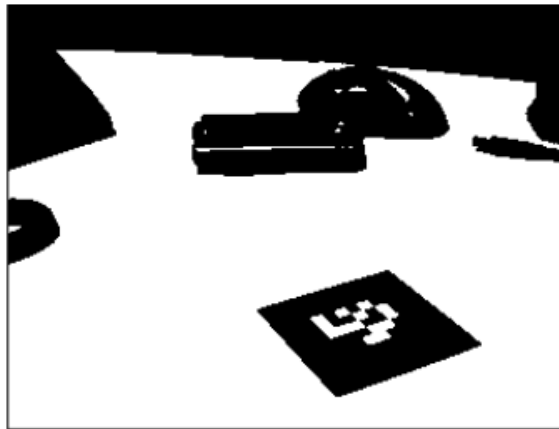


Marker Tracking: Overview Pipeline



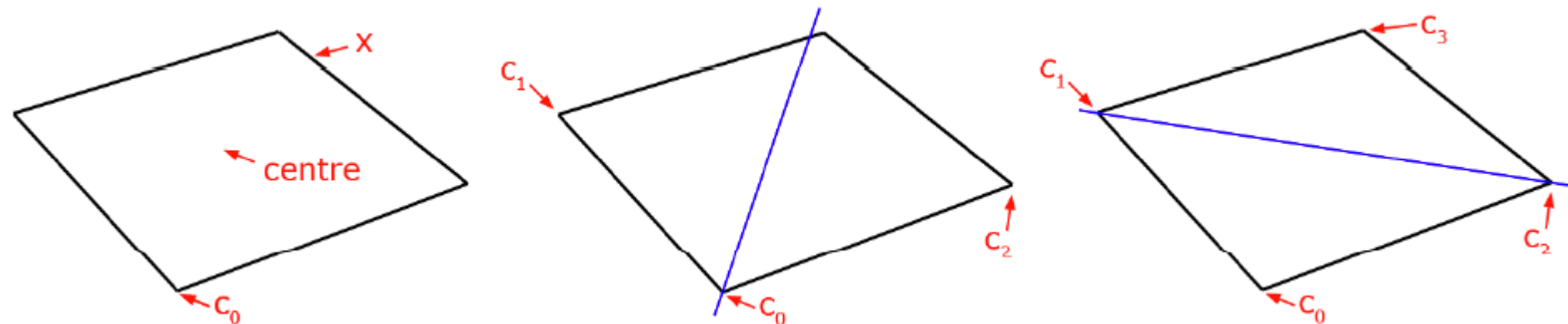
Marker Tracking: Fiducial Detection

- Threshold the whole image
- Search scan-line per scan-line for edges (white to black steps)
- Follow edge until either
 - Back to starting pixel
 - Image border
- Check for size
 - Reject fiducials early that are too small



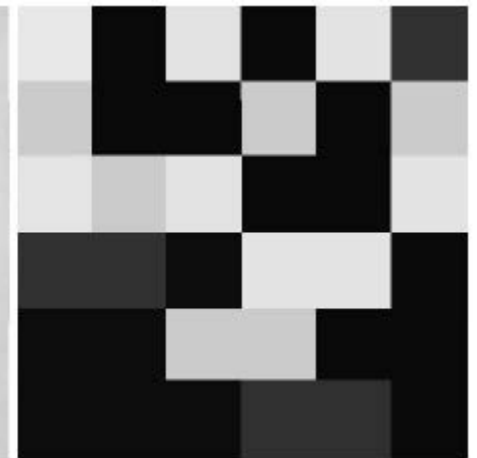
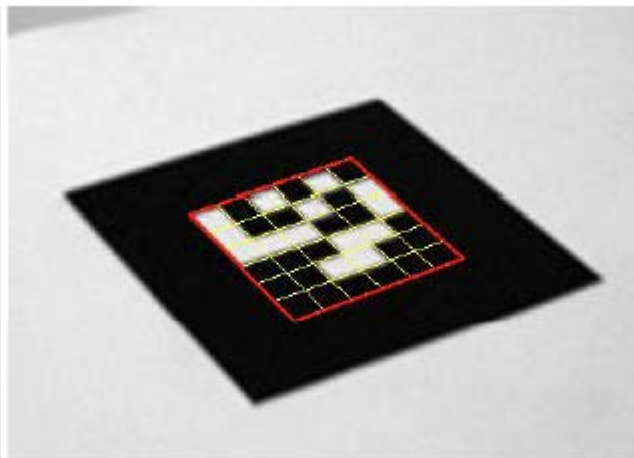
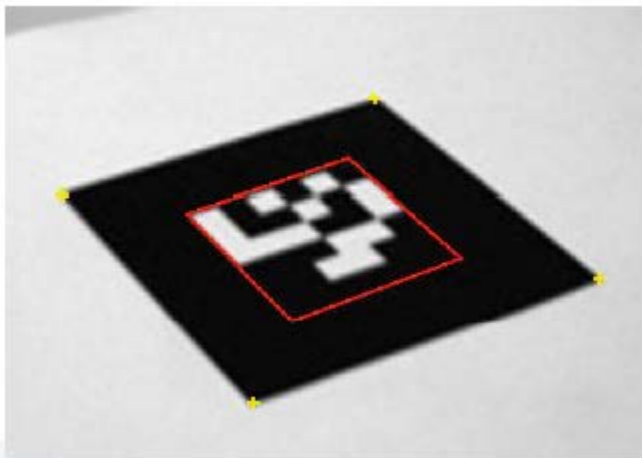
Marker Tracking: Rectangle Fitting

- Start with an arbitrary point “x”
- The point with maximum distance must be a corner c_0
- Create a diagonal through the center
- Find points c_1 & c_2 with maximum distance left and right of diagonal
- New diagonal from c_1 to c_2
- Find point c_3 right of diagonal with maximum distance
- Repeat to check if no more corners exist

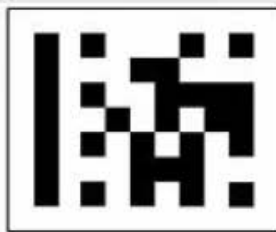


Marker Tracking: Pattern checking

- Calculate homography using the 4 corner points
 - “Direct Linear Transform” algorithm
 - Maps normalized coordinates to marker coordinates (simple perspective projection, no camera model)
- Extract pattern by sampling
- Check pattern
 - Id (implicit encoding)
 - Template (normalized cross correlation)
- Four 2D-3D correspondences - pose estimation



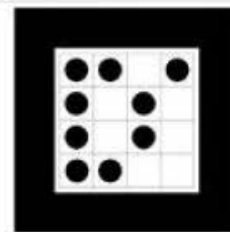
Other Popular Fiducial Marker Libraries for AR



CyberCode



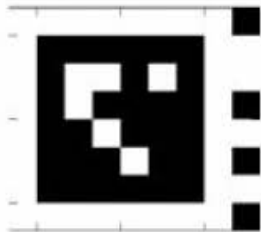
ARStudio



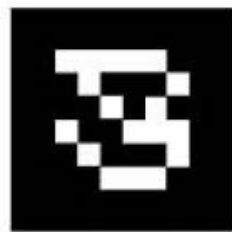
SCR marker



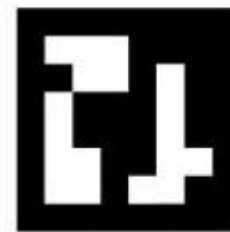
IGD marker



HOM marker



ARTag



ARToolKitPlus



Visual Code
from ETHZ



ReactIVision



USC's multi-ring
marker



Intersense
IS-1200 marker



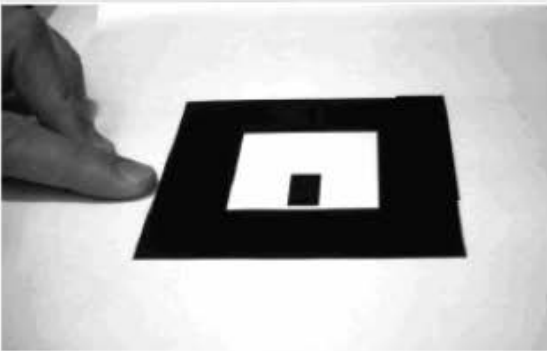
Shotcode

Improvements / diffs:

- Improved robustness (e.g. partial occlusion, error detection in pattern)
- Speed-ups and code optimizations
- Commercial libraries

Non-exhaustive list

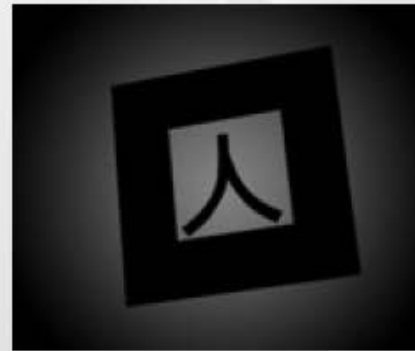
Tracking Challenges in ARToolKit



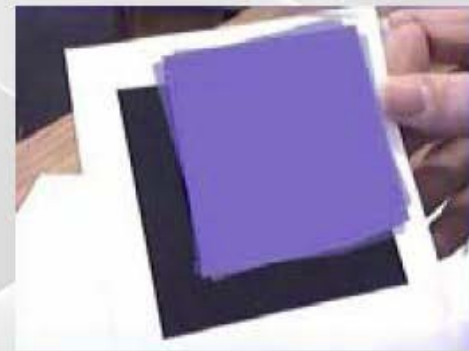
Occlusion
(image by M. Fiala)



Unfocused camera,
motion blur



Dark/unevenly lit
scene, vignetting



Jittering
(Photoshop illustration)



False positives and inter-marker confusion
(image by M. Fiala)

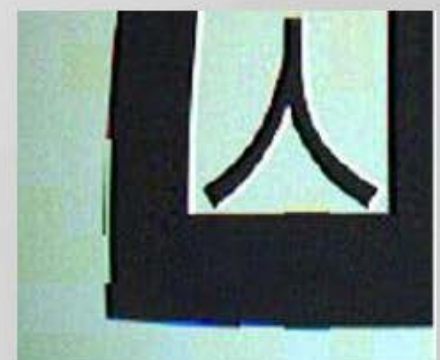


Image noise
(e.g. poor lens, block coding /
compression, neon tube)

Fiducials - Sources of Errors

- „Optical noise“ – low resolution/pixel flickering
- Too far: Marker too small; too close: marker does not fit into image
- Ambiguity during detection
- Changing lighting conditions

Solution: **Infrared light**

- No problems in dark environments
- **Active Markers**: Emit light (e.g. IR LEDs)
- **Passive Markers**: Reflect light

Inside-Out / Outside-In Tracking

- Outside-In:
 - Sensors (Cameras) are located at fixed reference point
- Inside-Out:
 - Sensors are located on mobile target
 - Higher resolution and accuracy in orientation than outside-in
 - Reason: produces larger motion in image -> larger displacement of pixels -> higher accuracy

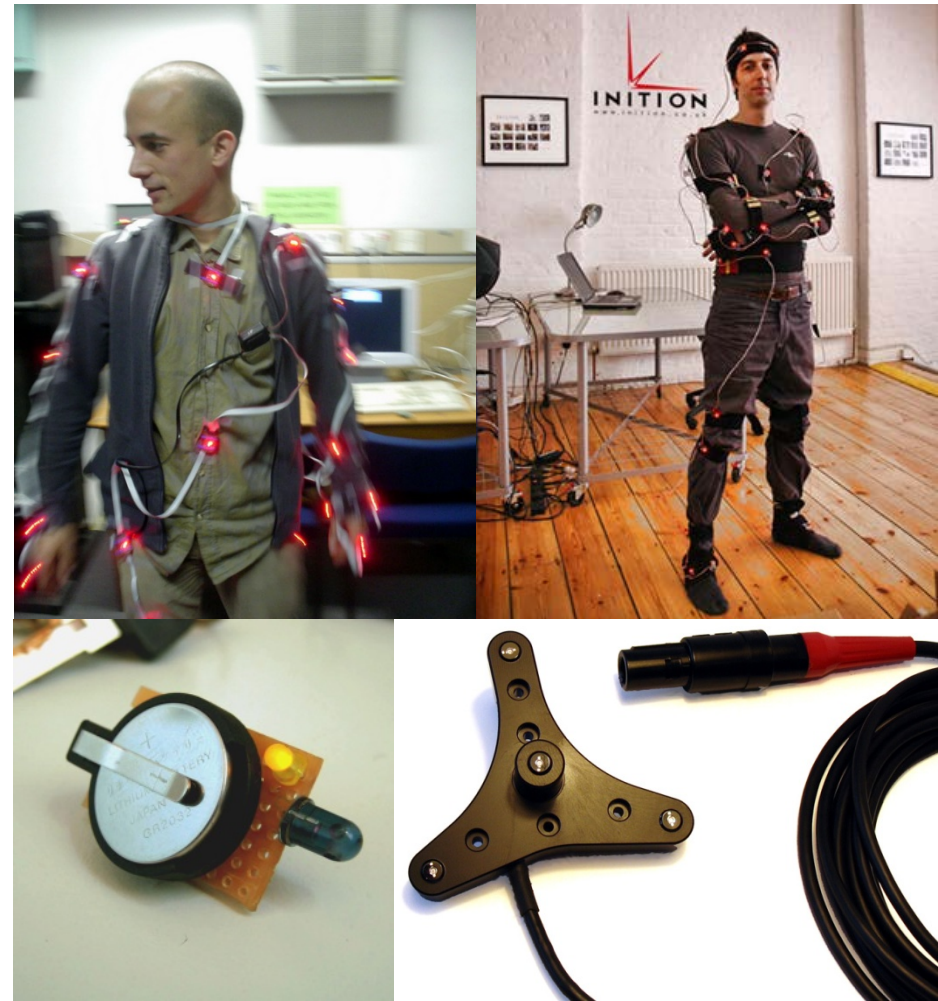
Inside-Out: HiBall

- HiBall-3100 wide area tracker
- HiBall: 6 lenses and photodiodes
- Active markers (IR LEDs)
- 6-DOF
- Hundreds of LEDs mounted on ceiling
- Very high accuracy
- High update rate: 2000 Hz



Overview: Tracking Active Markers

- Frequency-encoded IDs solve correspondence problem, but require high-speed cameras.
- Tethered or battery-powered markers restrict user movement.
- Higher cost of equipment, additional maintenance overhead.

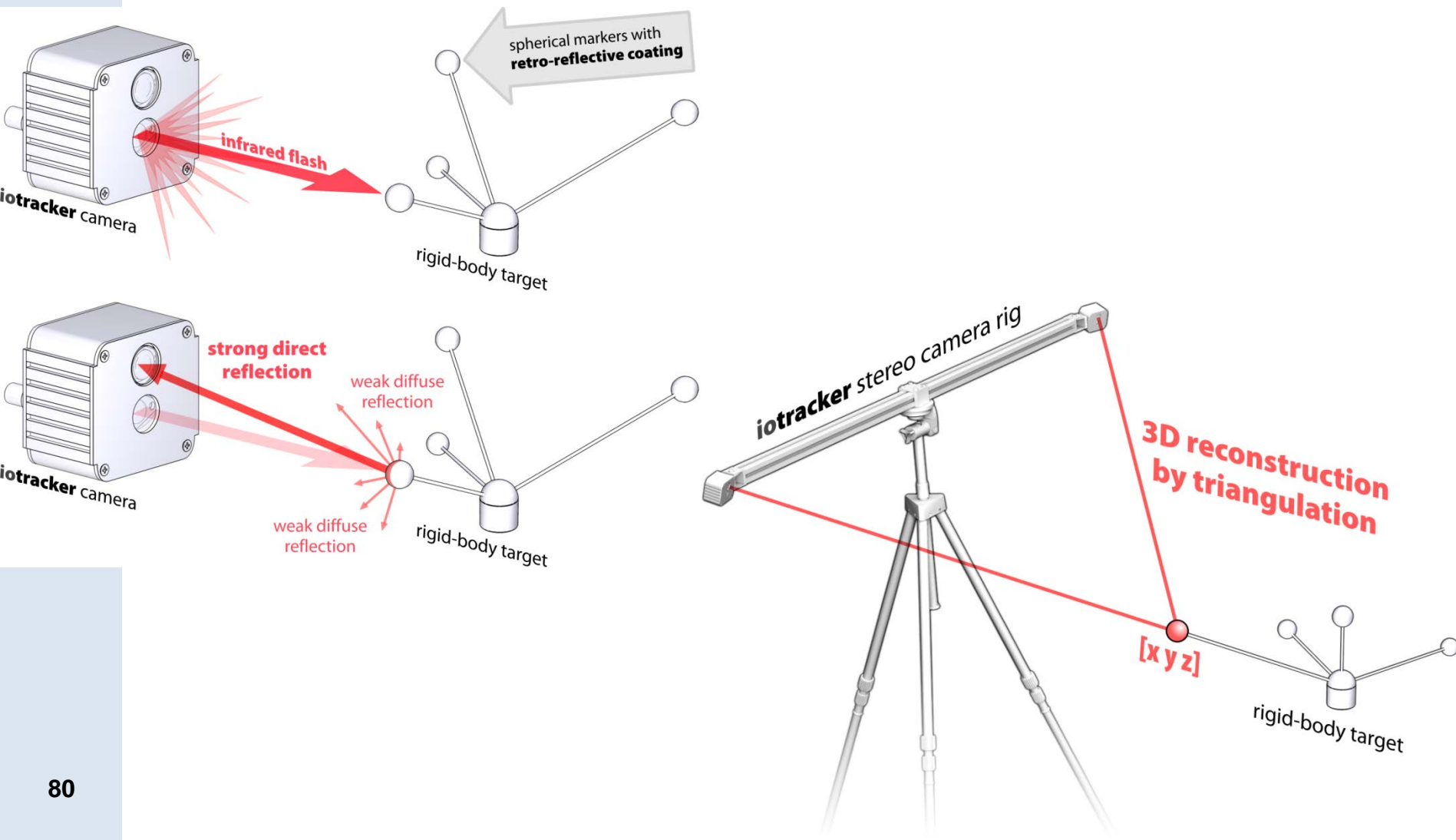


Optical Tracking - Polaris

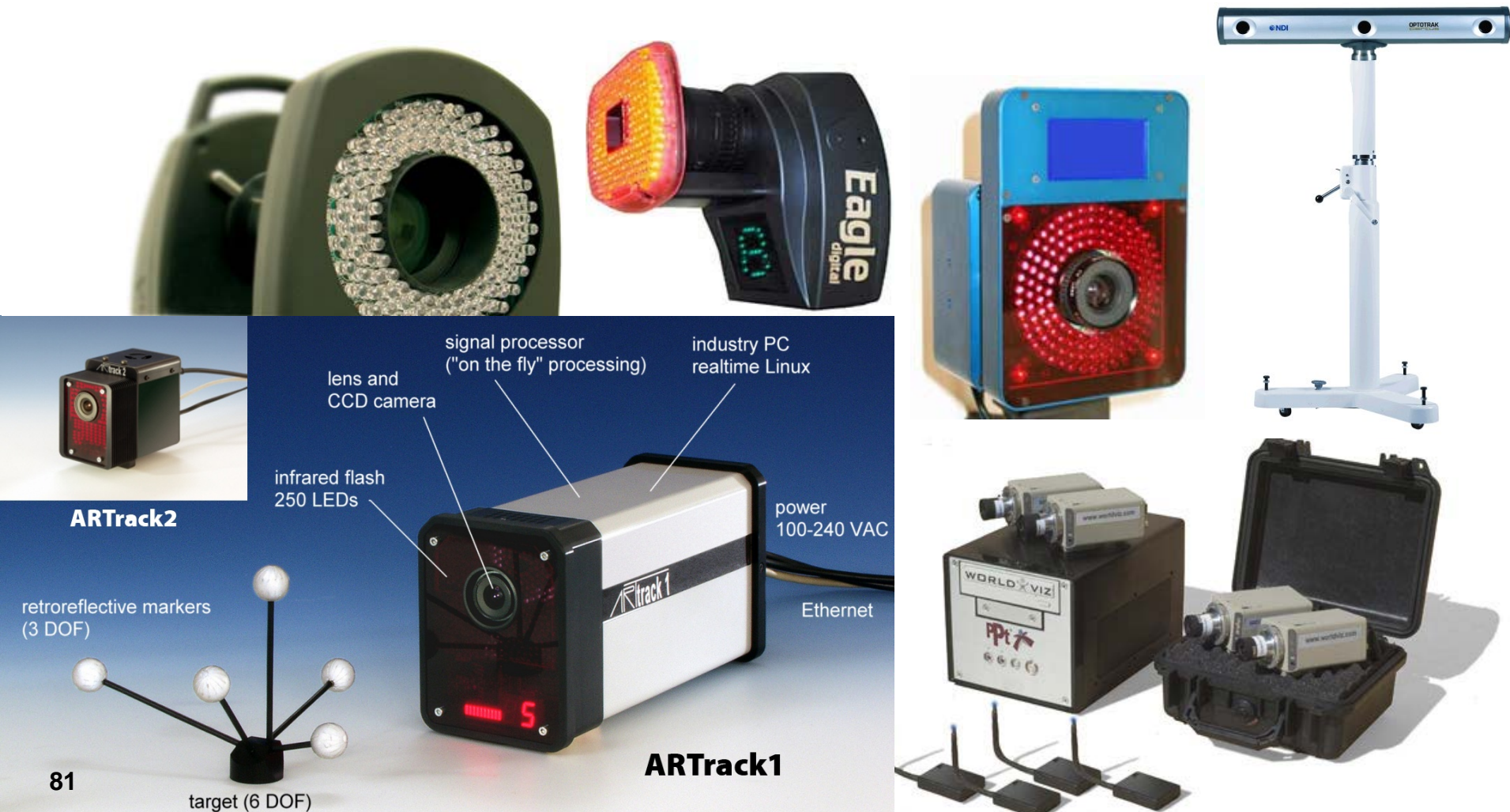
- Optical
- Passive and active targets



Marker-based Optical Tracking: Passive Targets



Overview: Passive Infrared-Optical Tracking Systems

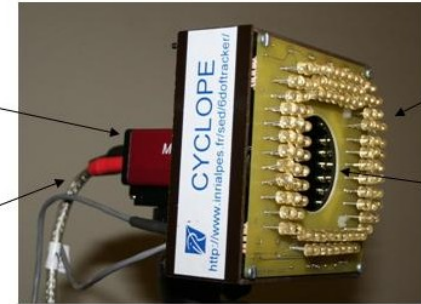


Overview: Passive Infrared-Optical Tracking Systems



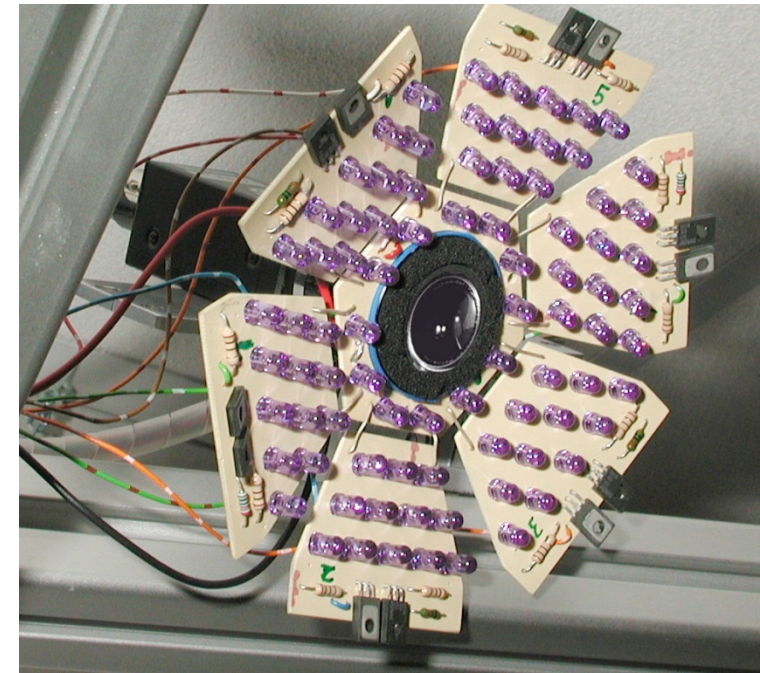
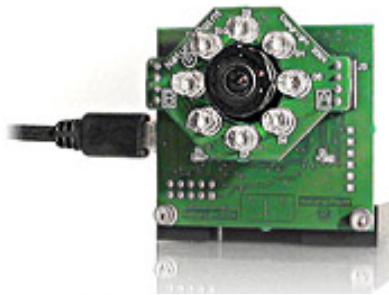
INDUSTRIAL
CAMERA

FireWire
Link



IR FLASH

IR FILTER



Optical Tracking - A.R.T.

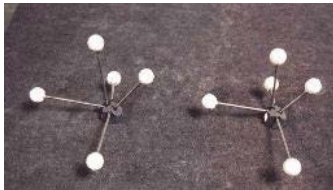


Tracking camera:

IR flash

IR camera

Embedded Linux System for image processing



Targets:

Retroreflective spheres

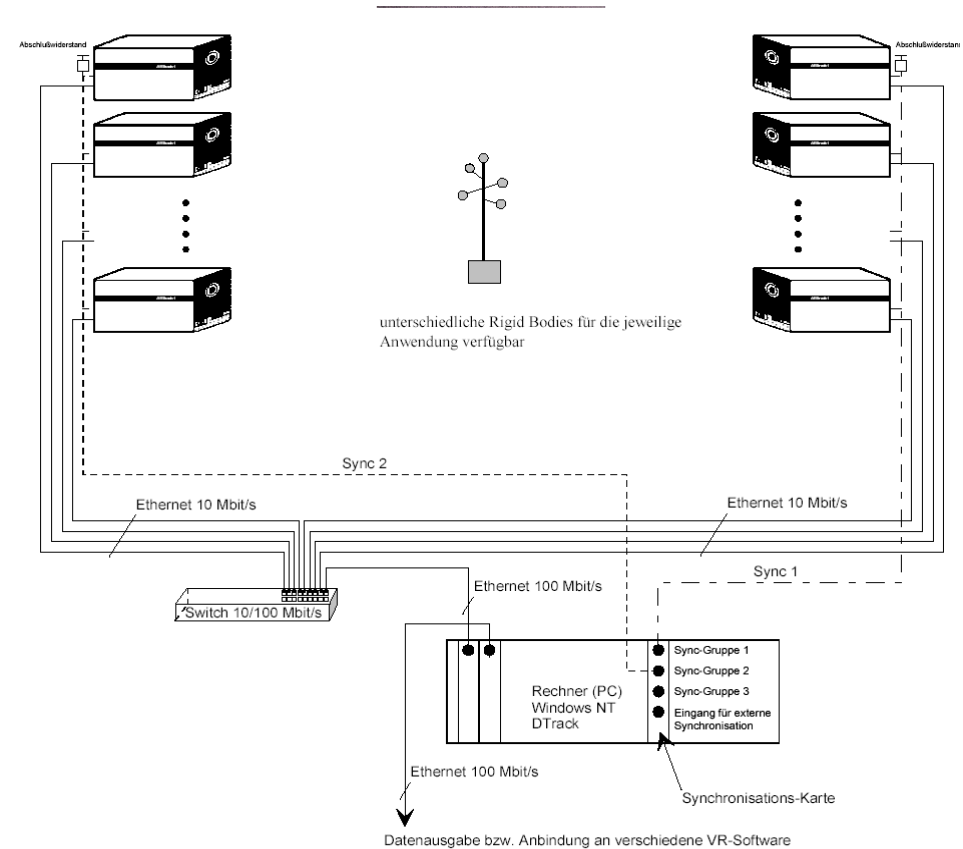
Unique distances/arrangement



Example: [Fingertracking](#) (Dorfmueller)

A.R.T. Tracking - Schema

- Multiple cameras
- Synchronization:
external PC
- Autonomous
image
pre-processing
- Central processing:
external PC



Optical Tracking - Specifications

Model	Latency	Max. # of sensor	Range	Accuracy	Costs
A.R.T. Dtrack	20 - 40 ms	10	~ 6 m	250 μ m; 0,4 degree	€ 30.000
ND Polaris (active)	20 ms	20	~ 3 m	350 μ m	?
ND Polaris (passive)	20 ms	6	~ 3 m	100 μ m	?
OptoTrak	20 ms	256	~ 3 m	100 μ m	€ 64.000

All latency times without filtering

Prices: Optical Tracking Systems

- However, commercially available infrared-optical trackers are **not exactly affordable...**
 - **Vicon Peak** 6-camera system € 50,000 (approx.)
 - **A.R.T.** 4-camera *ARTrack2* package € 30,000 (approx.)
 - **PhaseSpace** 4-camera system € 25,280
- Constant prices for more than ~10 years
- Clients:
 - Movie industry (Motion Capture)
 - Medical research
 - Industrial AR/VR applications & research
 - Not affordable for many other areas



TU Wien IMS: iotracker

- Self-assembled tracking cameras by using commercial parts
- connected via Firewire
- Combined tracking + application on 1 PC
- Scalable: 60Hz-200Hz
- Price: 1/3 – 1/5 of other commercial systems





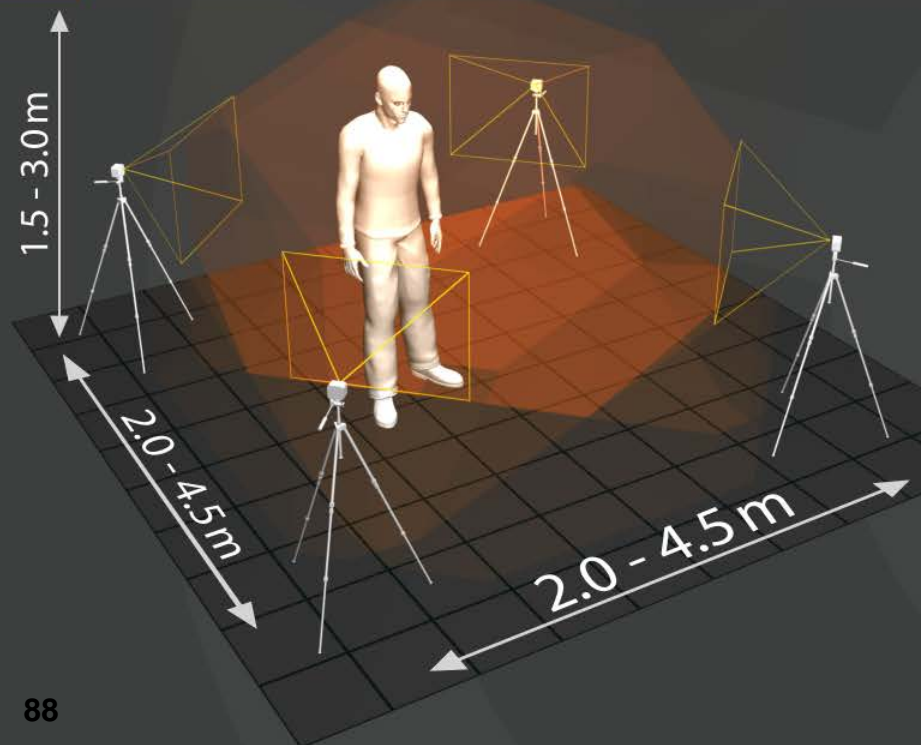
iotracker

affordable **infrared-optical** pose tracking



TECHNISCHE
UNIVERSITÄT
WIEN

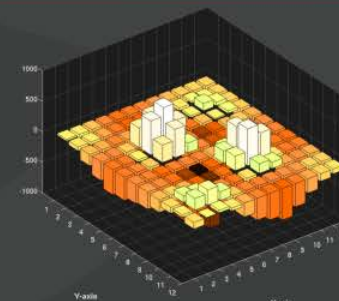
VIENNA
UNIVERSITY OF
TECHNOLOGY



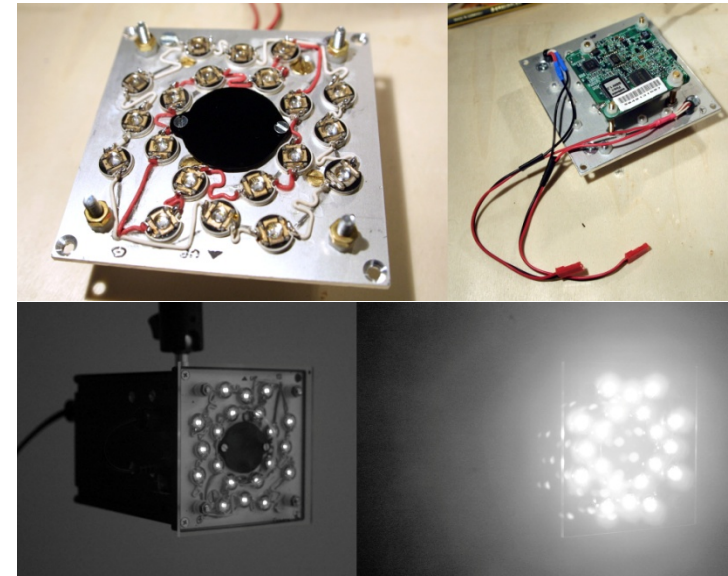
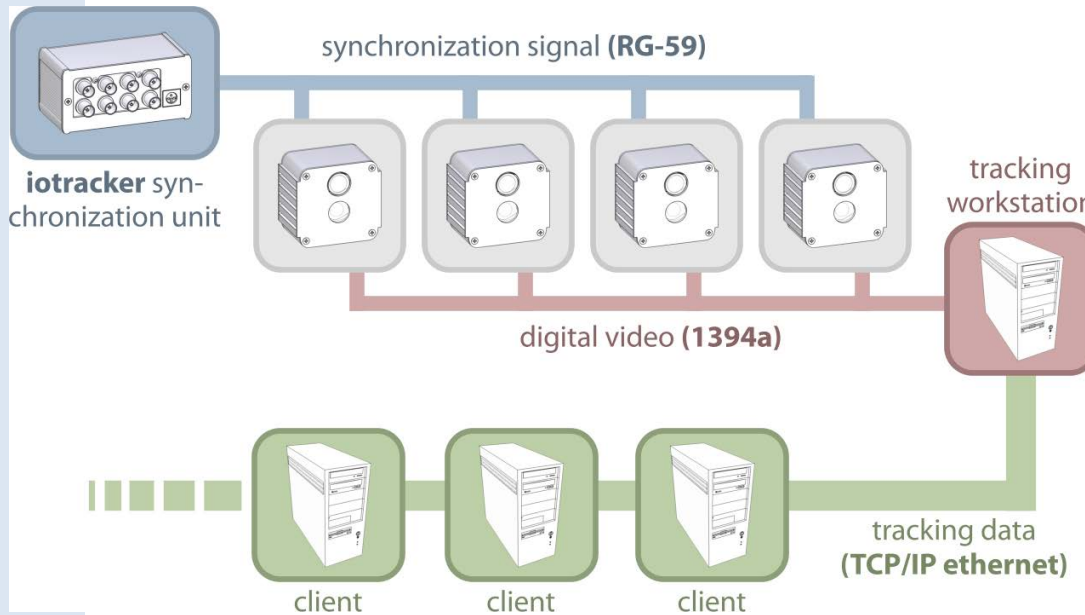
specifications



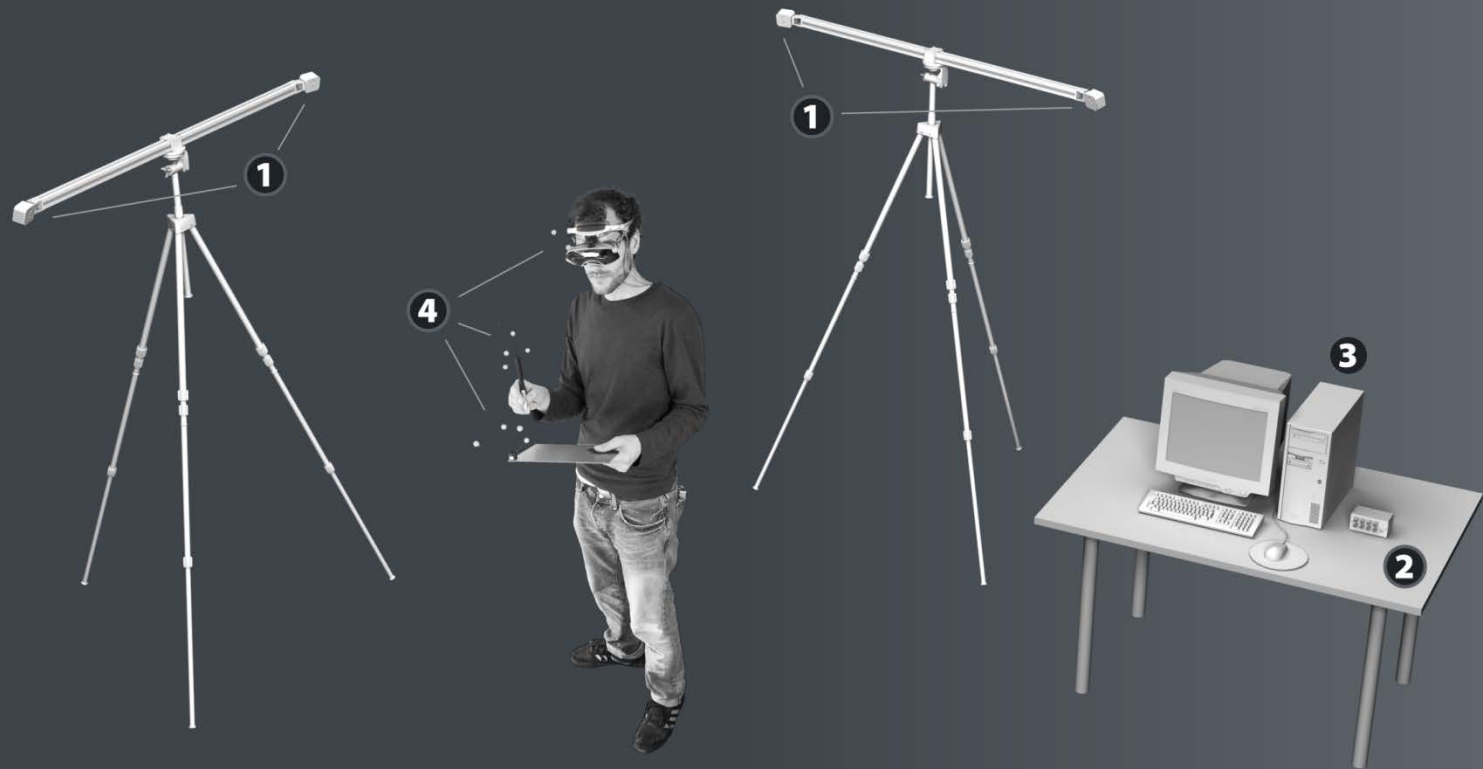
Update rate: **60 Hz**
Latency: **18 - 40 ms**
Jitter: **< 0.05 mm / 0.02°**
Accuracy: **± 0.5 cm**



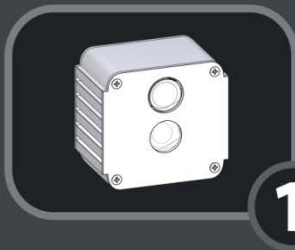
Camera Hardware



iotracker motion-tracking system

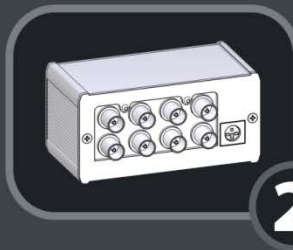


camera



1

synchronization unit



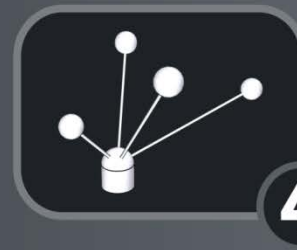
2

tracking workstation



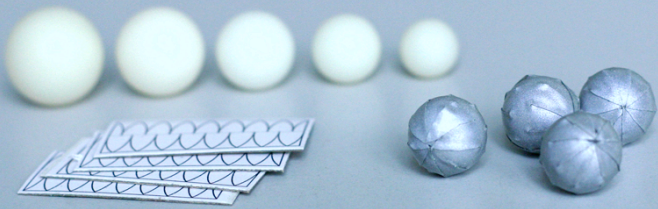
3

rigid-body target

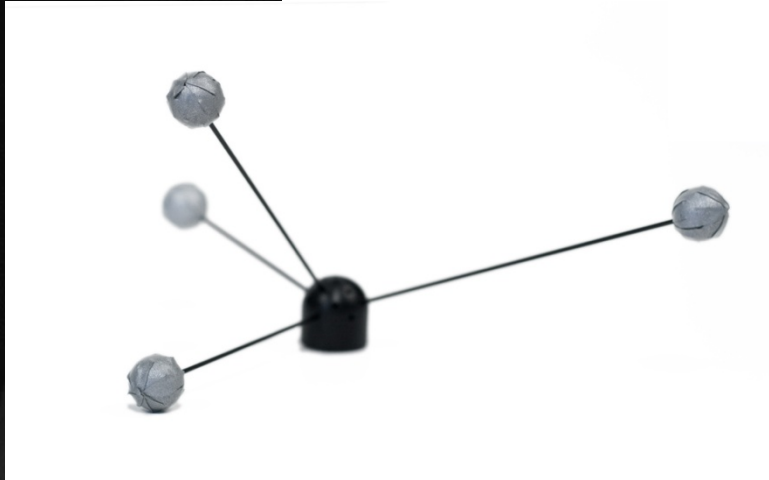
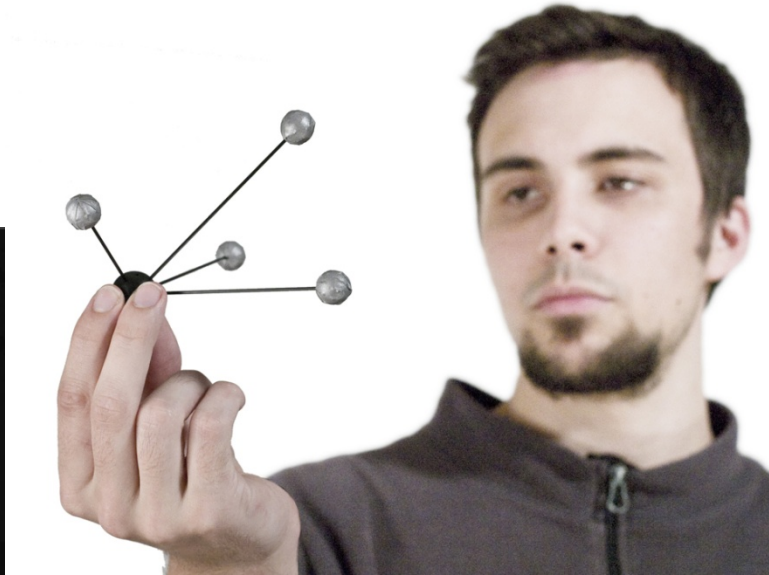


4

Retroreflective Markers



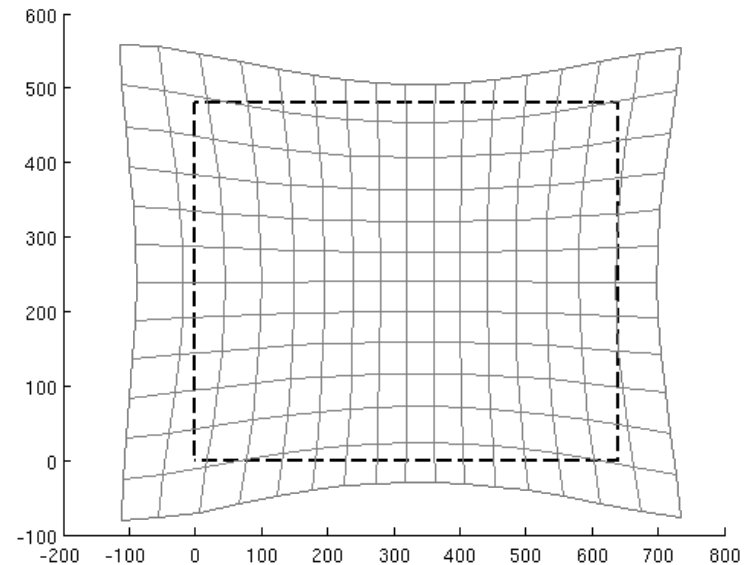
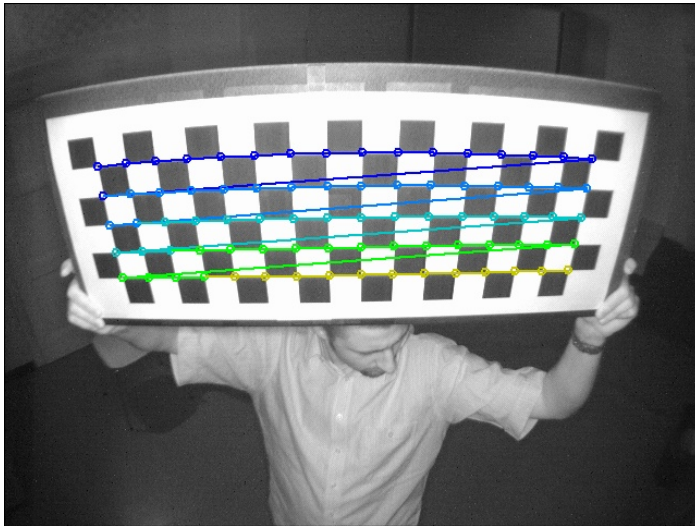
Ideal Camera Image



Optical Tracking System Design

1. **Camera calibration** (intrinsic, extrinsic parameters)
2. **Segmentation** & feature identification: detecting blobs
3. **Feature correlation**: finding multiple-view blob correspondences
4. **Projective Reconstruction**: acquire 3-DOF marker positions
5. **Model-fitting**: find pre-calibrated rigid constellations within the marker point cloud
6. **Pose estimation**: obtain 6-DOF pose (rotation/translation) for each rigid constellation

Camera Calibration (Intrinsic Parameters)



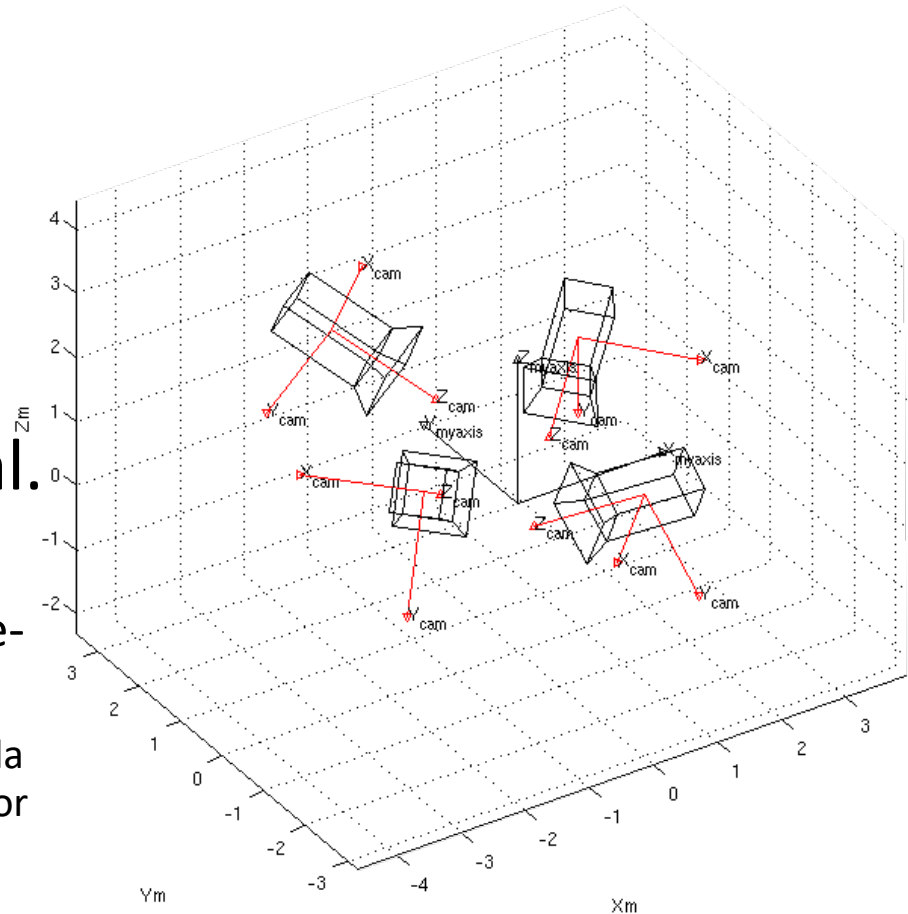
Non-linear lens distortion model (*Heikkilä 1997*):

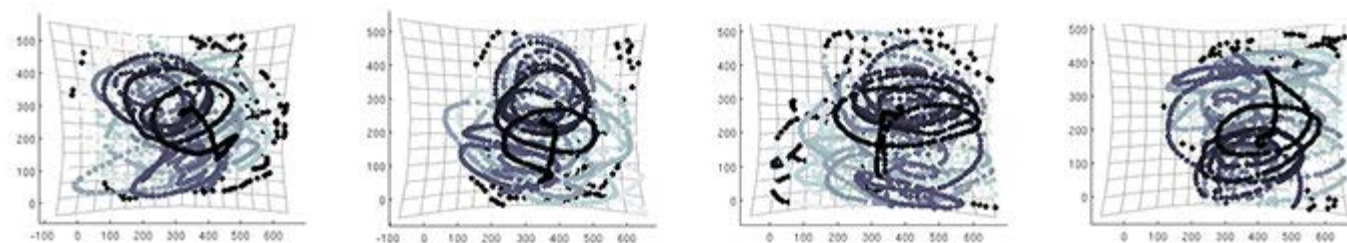
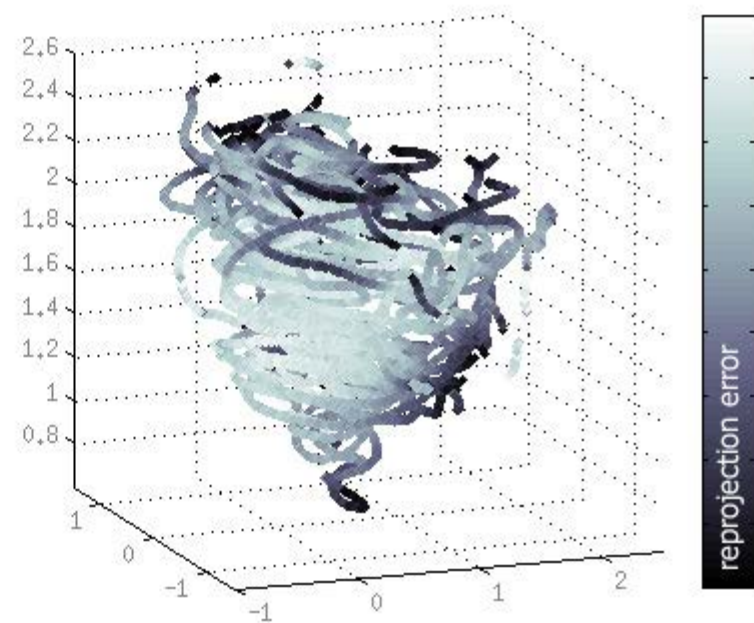
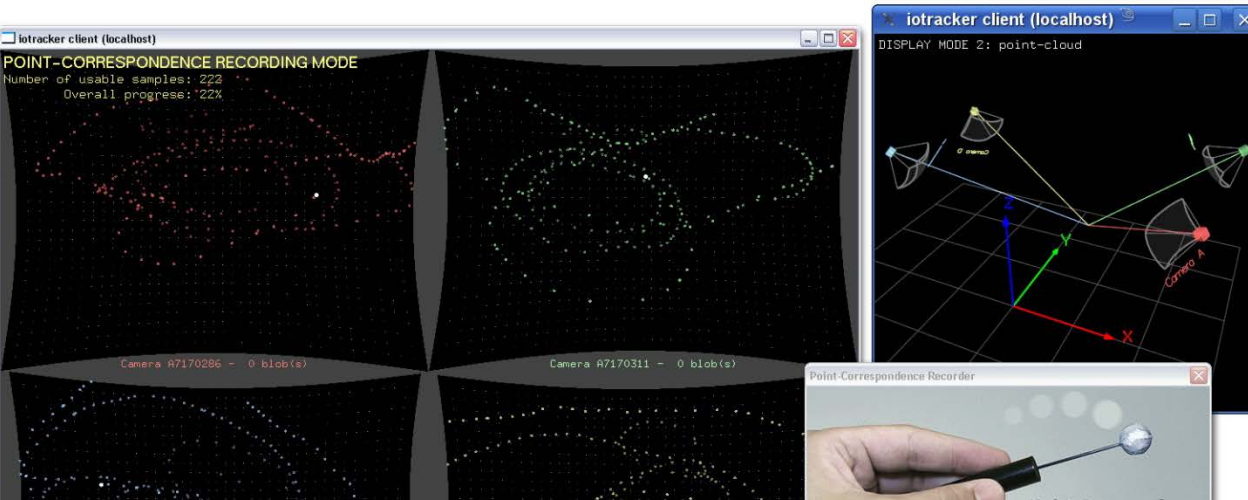
- Radial / tangential distortion
- Focal distance, ...

Camera Calibration (Extrinsic Parameters)

- Calculate position and orientation of cameras to each other and points in space
- We use the calibration algorithm by Svoboda et al. [*]
 - uses multiple-camera (>3) single-point correspondences

[*] Tomas Svoboda, Daniel Martinec, Tomas Pajdla
A Convenient Multi-Camera Self-Calibration for
Virtual Environments (2005)

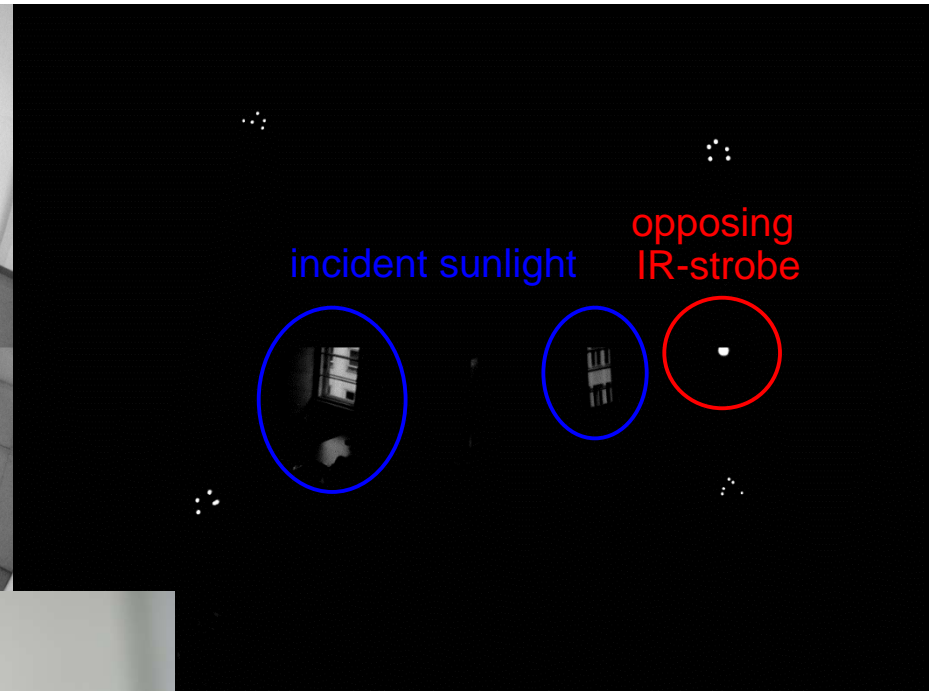




Blob Detection

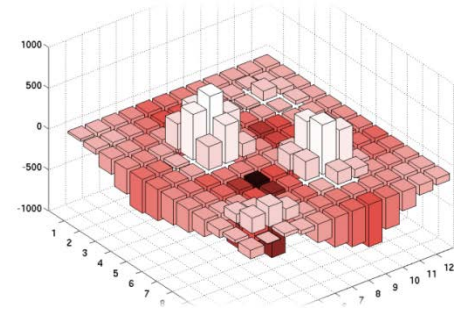
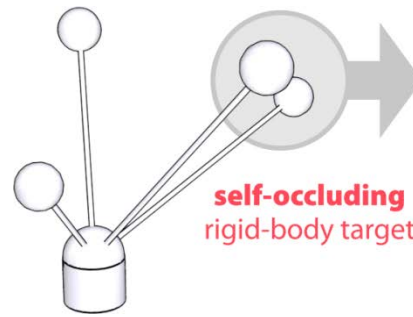


Shutter: 1/31s
Maximum Gain



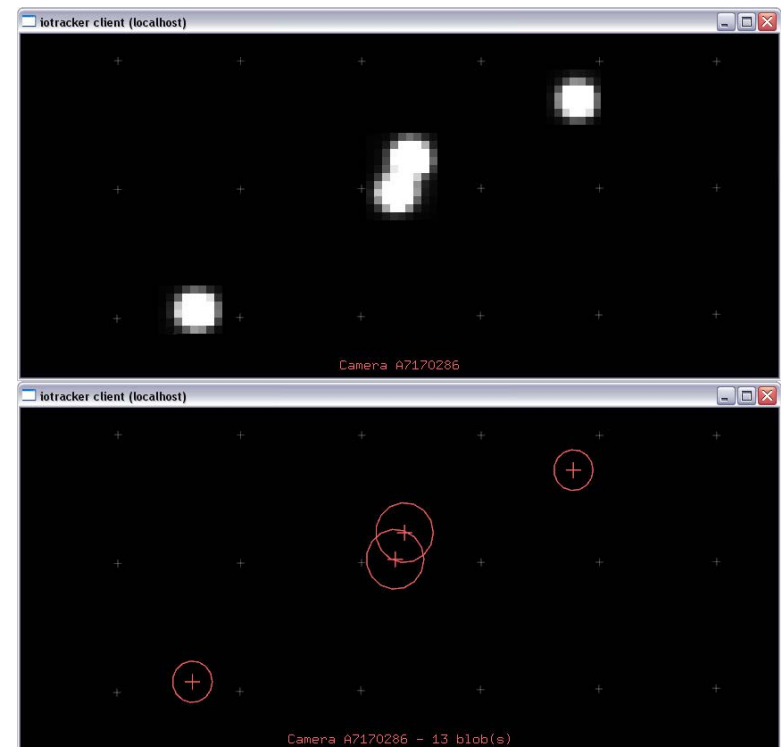
Shutter: 1/200s
Low Gain

Blob Detection

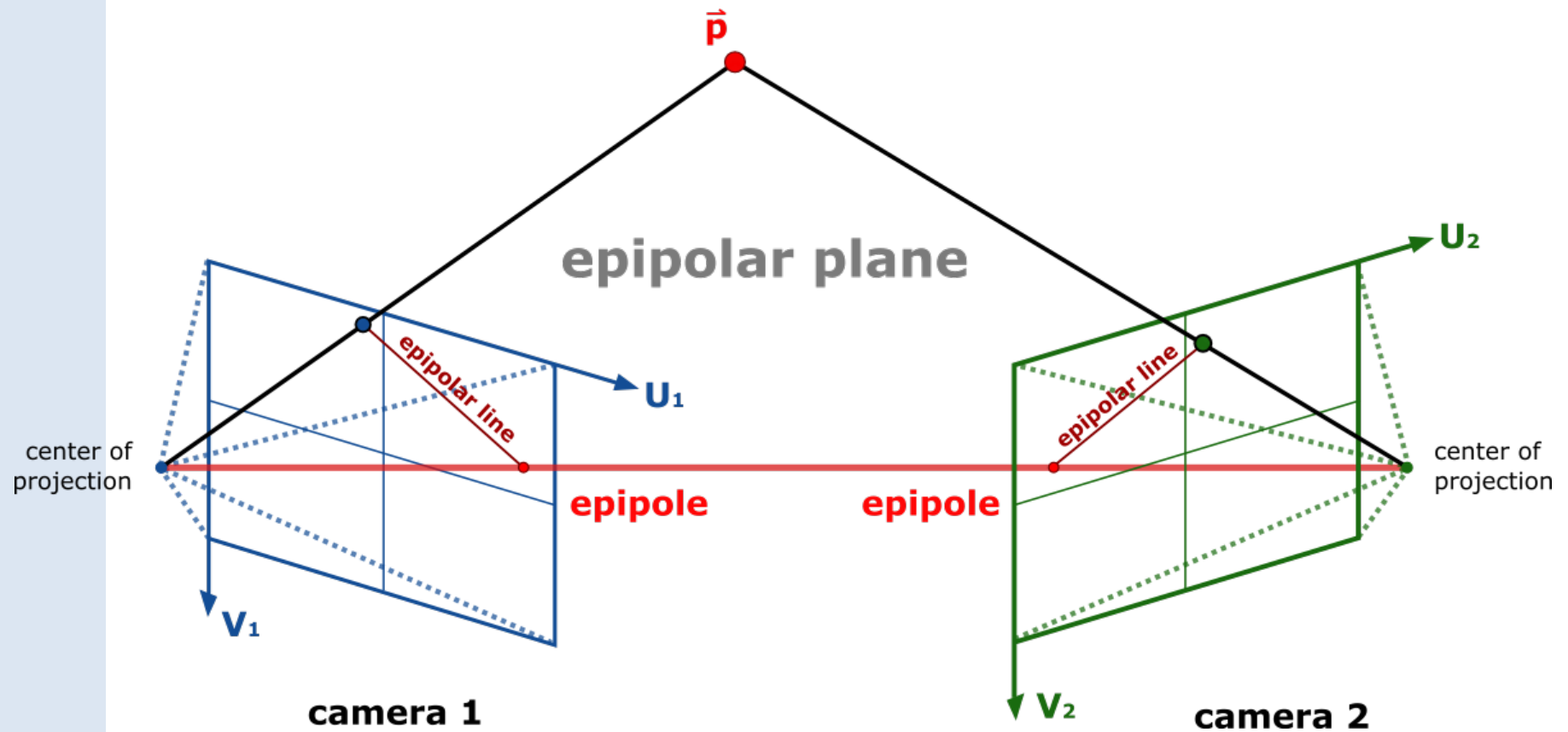


Two-step procedure:

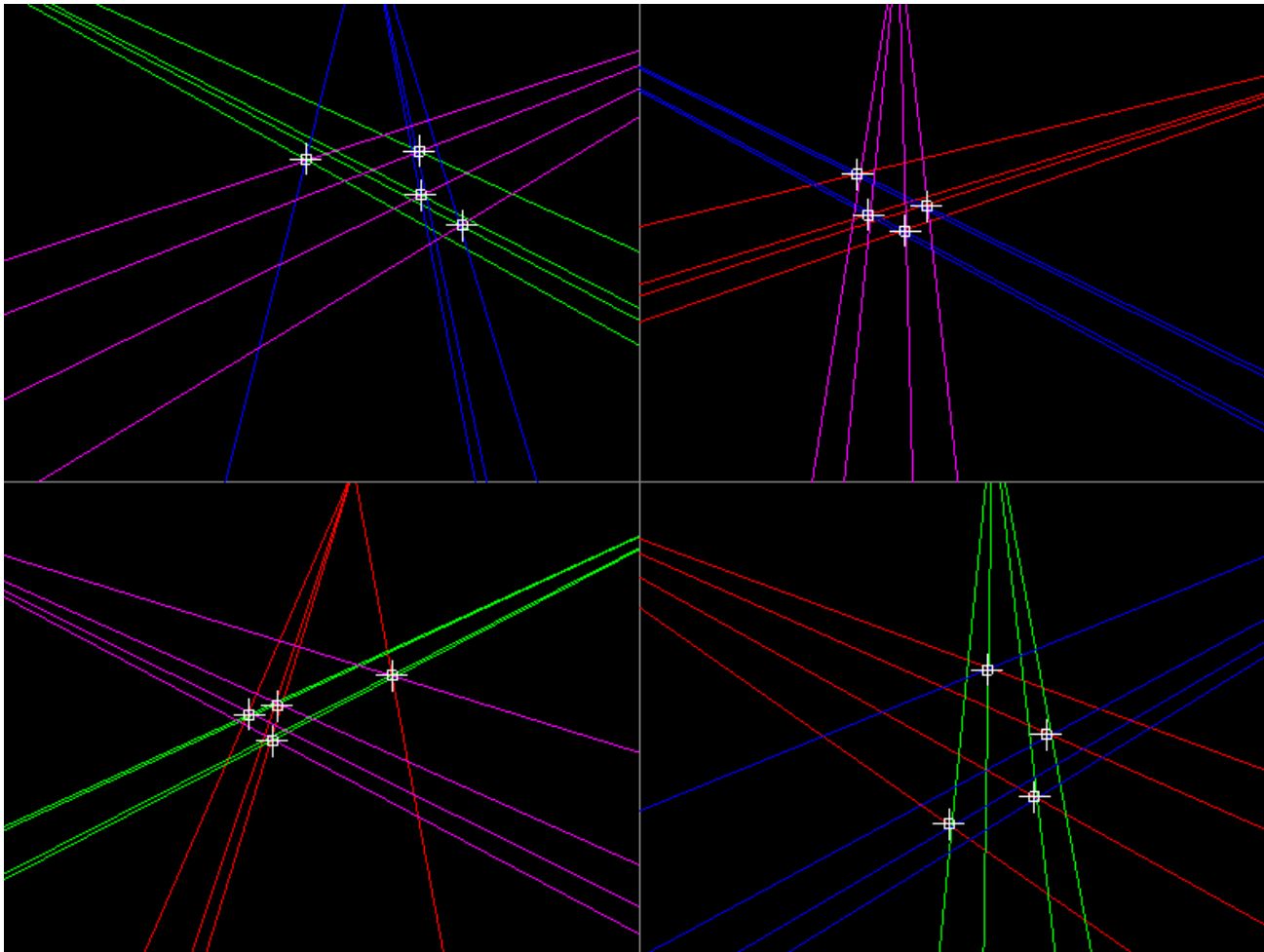
- Binary thresholding to discover blob candidates.
- Weighted-luminance centroid computation, using per-pixel difference matting
 - Background model is trained beforehand
- Problem: overlapping blobs
 - Solutions: contour-based segmentation or Hough circle finder



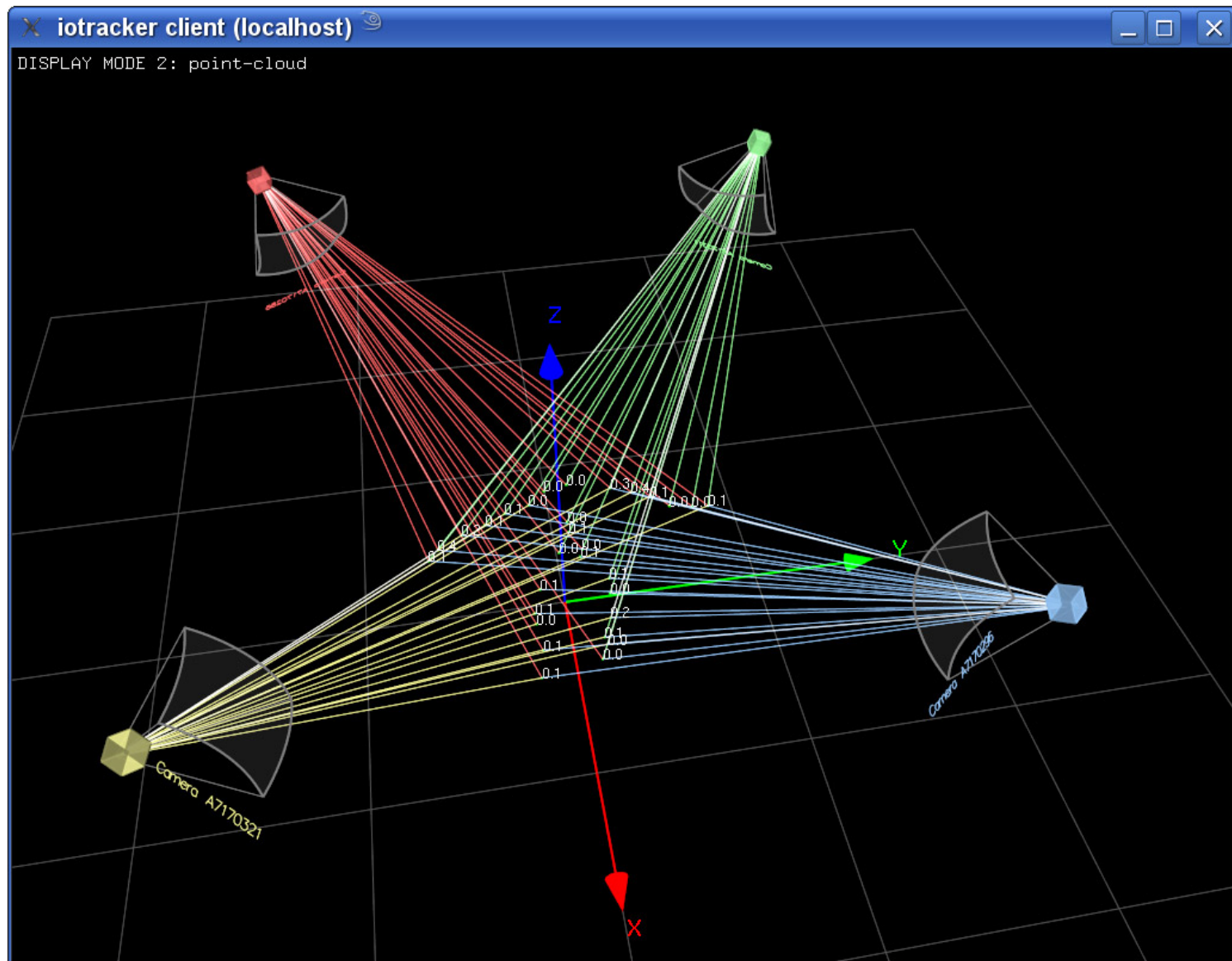
Epipolar Geometry



Epipolar Geometry: Feature Correlation

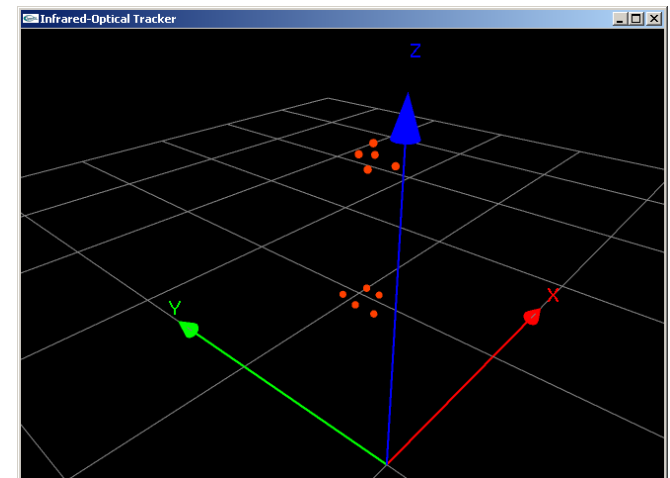
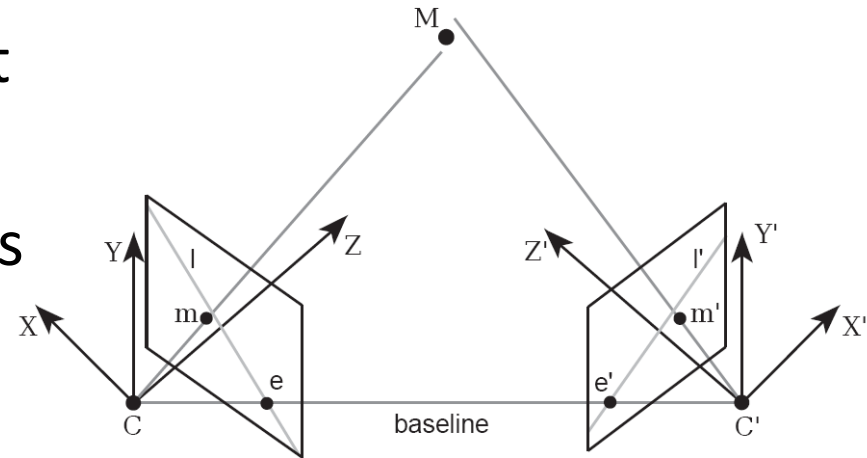


Epipolar Geometry:



Projective Marker Reconstruction

- Rays do not exactly intersect (in practice)
 - Currently using least-squares approach for projective reconstruction
 - Small number of markers (<40): Sparse bundle adjustment (Levenberg-Marquardt) to refine results (recursive)
- 3D Point cloud



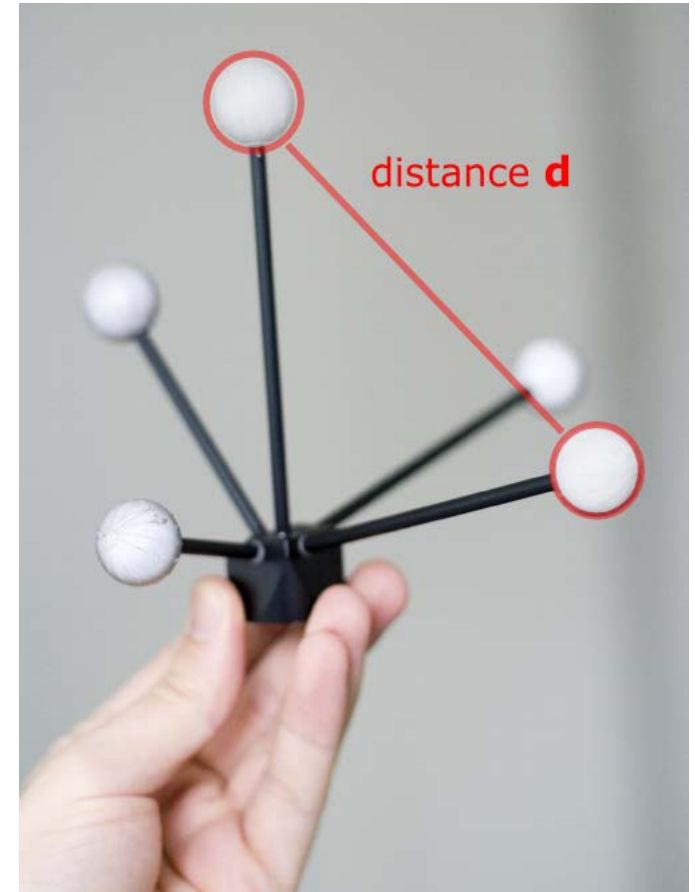
Model-Fitting

- pre-compute intra-constellation marker distances

$$n! / [2*(n-2)!]$$
 entries

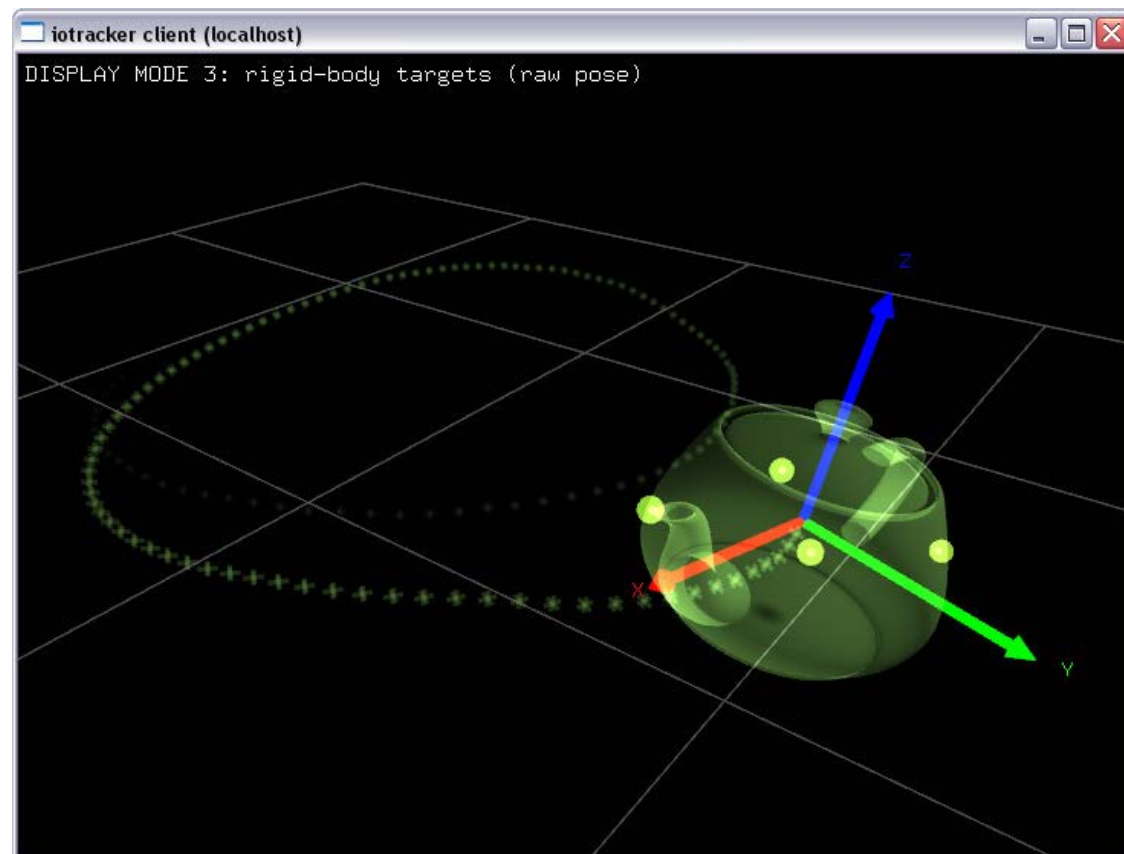
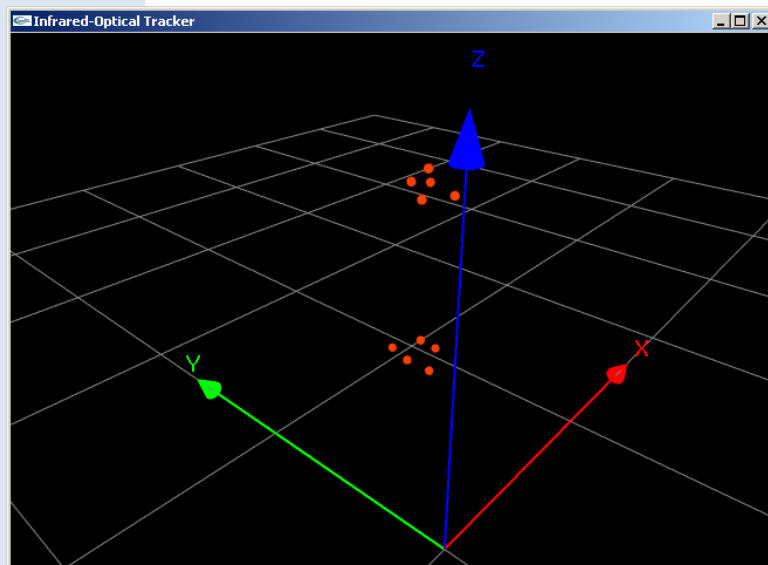
Once per frame/constellation:

- for every pair compute the Euclidean distance
 (about 5000 computations for 100 markers)
- Selection out of multiple hypotheses



6-DOF Pose Estimation

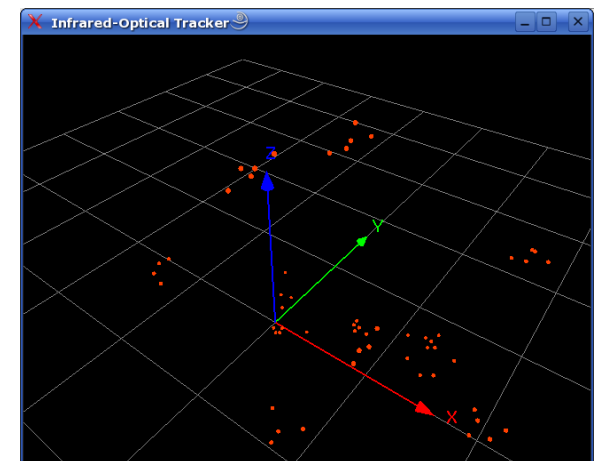
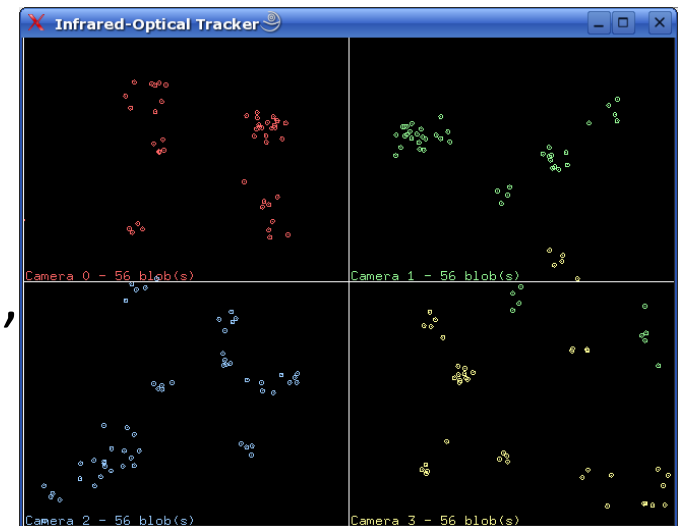
- Based on stored target model:
Perform a **least-squares estimation** of rotation and translation.



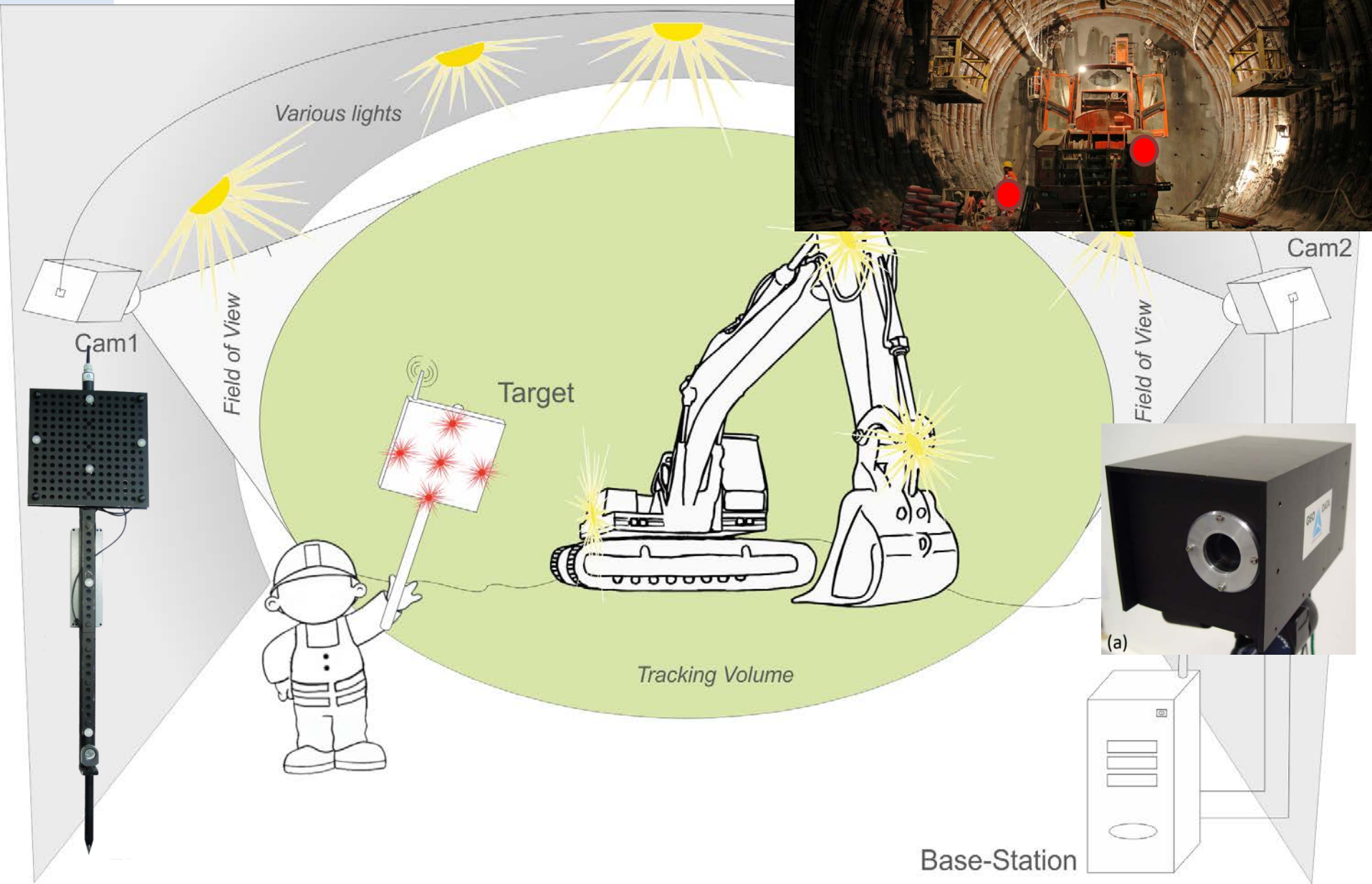
Results: Performance

Artificial benchmark:

- 4 cameras
- 224 blobs (56 per camera), with simulated 2D jitter (gaussian noise), no segmentation
- 56 markers
- 8 rigid constellations (consisting of 36 markers)
- all 8 constellations tracked at 60Hz with approx. 70-75% combined CPU load (2.16 GHz Intel T2600)



Measurements in a Tunnel



Consumer Devices

Usually tracking in small area (monitor):

- Dynasight (2500 USD)
- TrackIR (150 USD)
 - 6DOF
- OptiTrack
 - >100fps
 - On-board blob detection



Active/Passive Marker Tracking

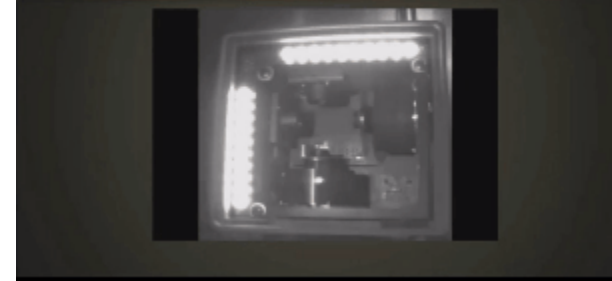
- Advantages
 - Very accurate (millimeter or sub-millimeter)
 - Supports tracking in a large volume
 - allow for untethered tracking
- Disadvantages
 - Requires active (LED) or passive markers
 - Occlusion problem

HTC Vive Lighthouse

- Inside-Out Trackingsystem
- 2 line lasers per base station + synch IR LEDs (every 8.3ms)
- Photo diodes on HMD and controller
- Time difference -> angle to base

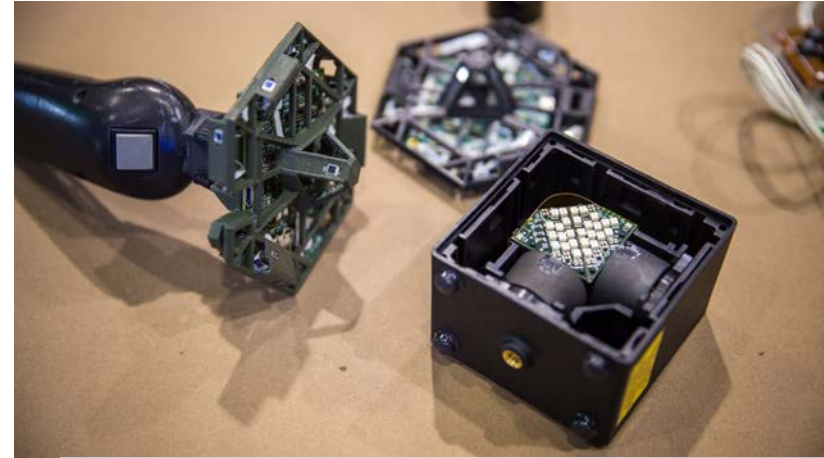


Lighthouse – How it works

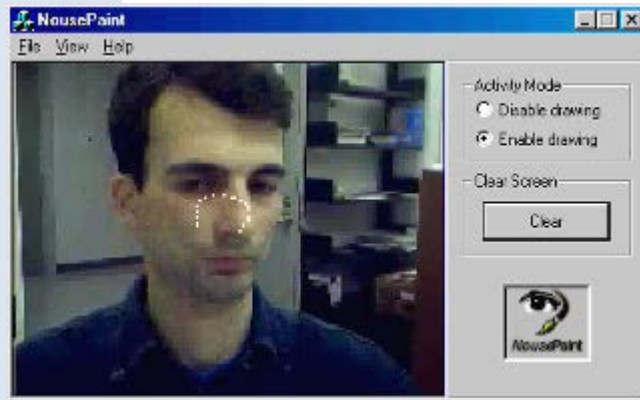


Steam VR (Valve) Lighthouse Tracking

- 1006Hz inertial sensor mainly used for tracking (hybrid tracking!)
- 60Hz optical tracking update rate
- 1ms worst case latency (head), 2.7ms (controllers)
- Very precise. Static jitter: 0.3mm lateral, 2.1mm in distance direction (single base)



Optical Tracking without Markers



Natural Feature Tracking (NFT)

- More difficult than marker tracking
 - Markers are designed for their purpose
 - The natural environment is not...
- Less well established methods
 - Every year new ideas are proposed
- Slower than marker tracking

Idea:

1. Detect reliable features in image (e.g. corners, lines, ...)
2. Compare with reference image/3D model

Many different methods and combinations

- Features
 - Points
 - Edges
 - Horizon
- Sensor fusion
 - Gyroscopes
 - Accelerometers
 - GPS, ...
- Requires Initialization!
- Not (very) robust
 - Sensor fusion
 - Recovery methods



Azuma, '99



Kretschmer et al., '02

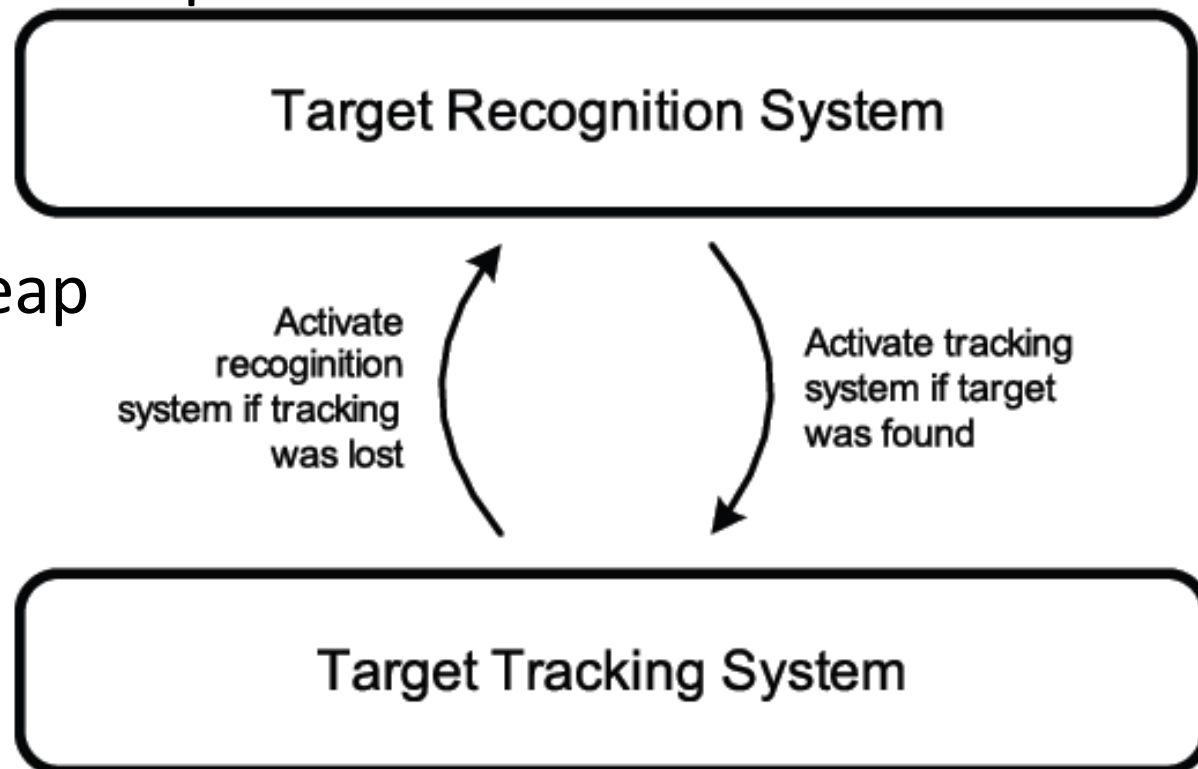


Ribo et.al., '02

NFT - Initialization vs. Tracking

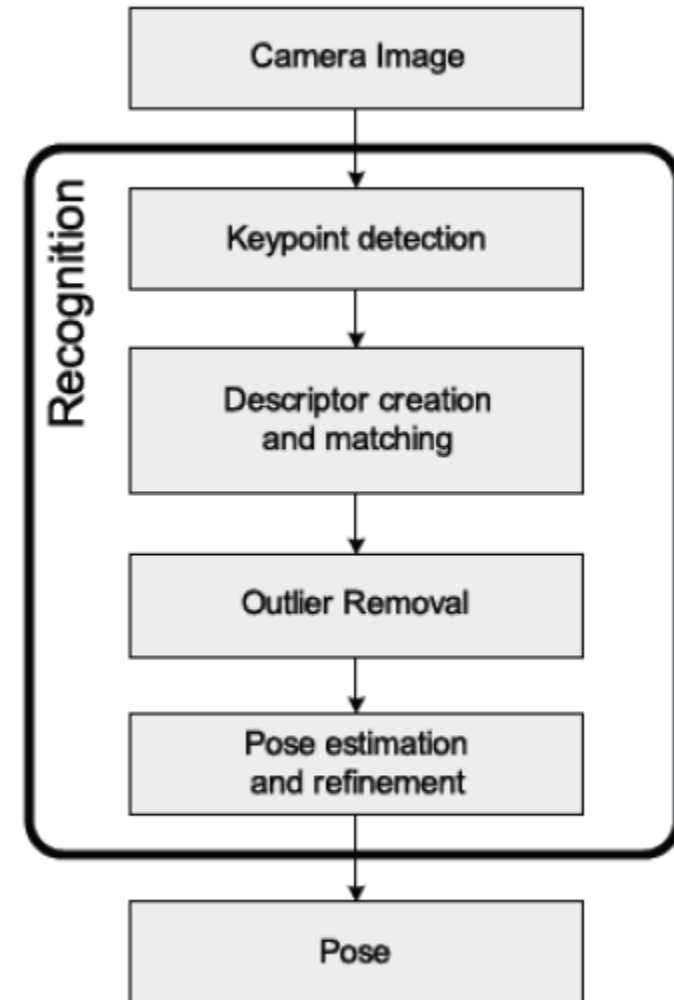
- Initialization = Recognition
 - Computationally expensive
 - Required at startup

- Tracking
 - Fast and cheap
 - Robust
 - Accurate



Detection in Every Frame

- This is what most „trackers“ do...
- Targets are detected every frame
- Popular because detection and pose estimation are solved simultaneously

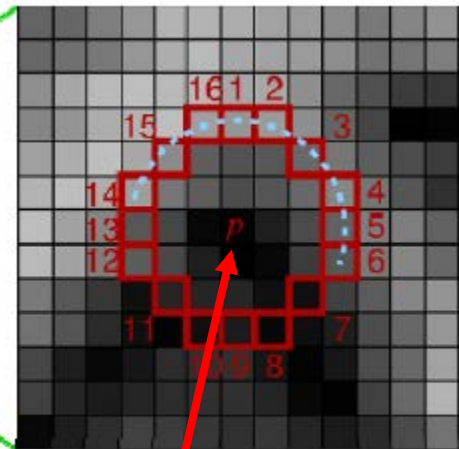
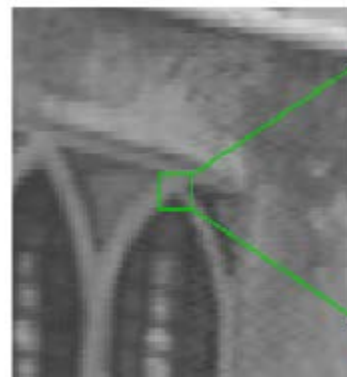


Natural Feature Tracking: Examples

- EPFL Book: [Video1](#), [Video2](#)
- [Stricker: ISMAR Demo](#)
- [Reitmayr: Cambridge Going-out Demo](#)
- [PTAM on iPhone](#)
- 13th Lab: [SLAM](#) and [Room measurements](#)
- InsiderNavigation: [Indoor Navigation at Airport](#)

FAST Corner Detector

- FAST corner detector frequently used for mobile apps
- Produces very stable features
- Simple and fast algorithm to compute a feature:
If n contiguous pixels are all brighter than the nucleus (center pixel) by at least threshold t or all darker than the nucleus by t , then the pixel under the nucleus is considered to be a feature. - Best results with Fast-9 ($n=9$)
- Robustness by using FAST on a database of different scales



* [Rosten, 2005]



PTAM: Algorithm Overview

- Georg Klein and David Murray. PTAM: Parallel Tracking and Mapping, 2007.
PTAM is a SLAM algorithm (Simultaneous Localization and Mapping)
- Tracking and mapping are separated and run in two parallel threads
- Mapping is based on keyframes, which are processed using batch techniques (Bundle Adjustment)
- The map is densely initialized from a stereo pair
- Large numbers (thousands) of points are mapped

PTAM in the Hofburg Festsaal



PTAM Tracking - Part 1

A new frame is acquired from the camera,
and a prior pose estimate is generated
from a motion model

Map points are projected into the image
according to the frame's prior pose
estimate

A small number (50) of the coarse-scale
features are searched for in the image

PTAM Tracking - Part 2

The camera pose is updated from these coarse matches

A larger number (1000) of points is re-projected and searched for in the image

A final pose estimate for the frame is computed from all the matches found

Natural Feature Tracking by Detection

SIFT: Scale-invariant feature transform [Lowe, 1999]

- State of the art for object recognition
- Known to be slow
- Typically used off-line
- Invariant to uniform scaling, orientation, affine distortion, and partially invariant to illumination changes

Ferns [Ozuysal, 2007]

- State of the art for fast pose tracking
- Memory intensive
 - requires ~10x too much memory for phones
- Long training phase

NFT with SIFT on a Mobile Phone



Natural Feature Tracking

- Highly desirable to get rid of markers
 - More beautiful setups
 - More acceptance from commercial side
- Very precise
- However, suffers from
 - Light conditions (spotlights)
 - Environmental changes
 - Camera's field of view
 - Large occlusions
 - Performance issues
 - Low CPU performance
 - Memory requirements
 - Bad or slow cameras (blur)

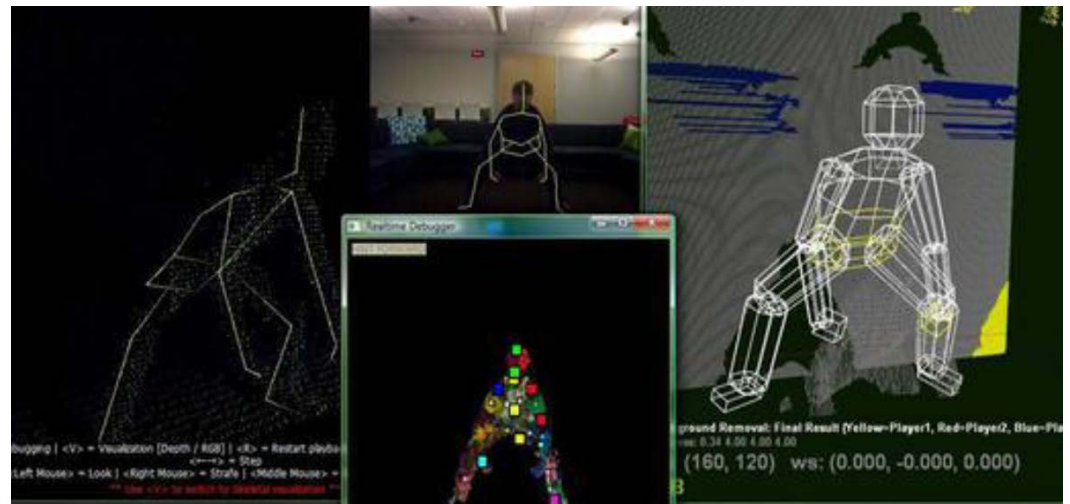
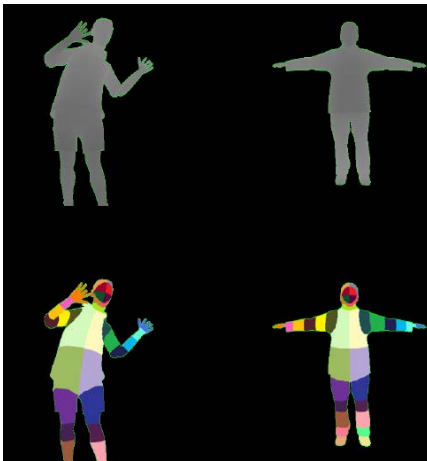
Depth Cameras

- 3 different technologies to capture an RGB + depth (RGBD) image with a single camera
 - Structured Light
 - Time of Flight (ToF)
 - Light Field Camera

Structured Light: e.g. Microsoft Kinect

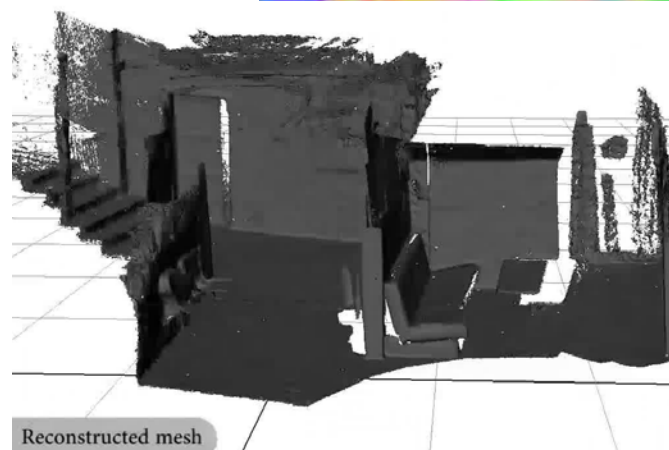
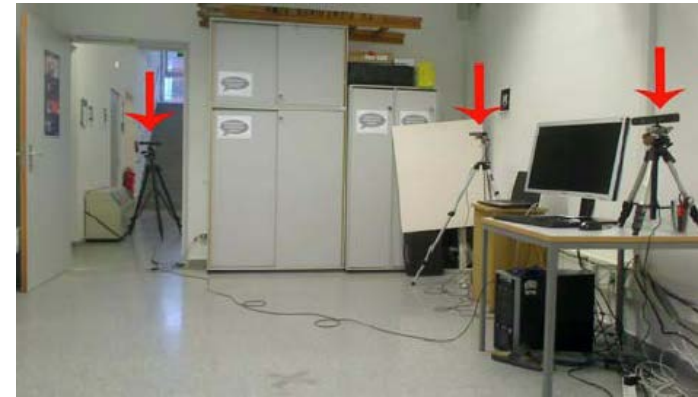


- 3D Sensor“ developed by Primesense
 - projects infrared pattern onto scene
 - [detects pattern shift in camera image](#)
 - 9 sub-regions with „random“ pattern
- Heavy machine learning to „guess“ pose
 - Mocap studios generated images, filmed people



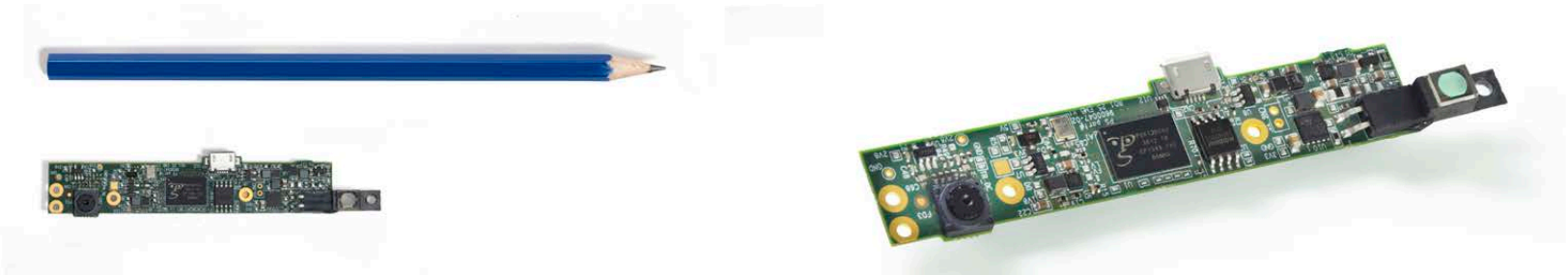
Some Kinect Projects

- Wide Area Tracking with Multiple Kinects
- Reconstruction – Kinectfusion
- Profitex – Large Scale 3D Reconstruction for Firefighters



Mobile Kinect-Like Sensors

- Primesense Capri (bought by Apple)



- Structure Sensor





Time of Flight Cameras

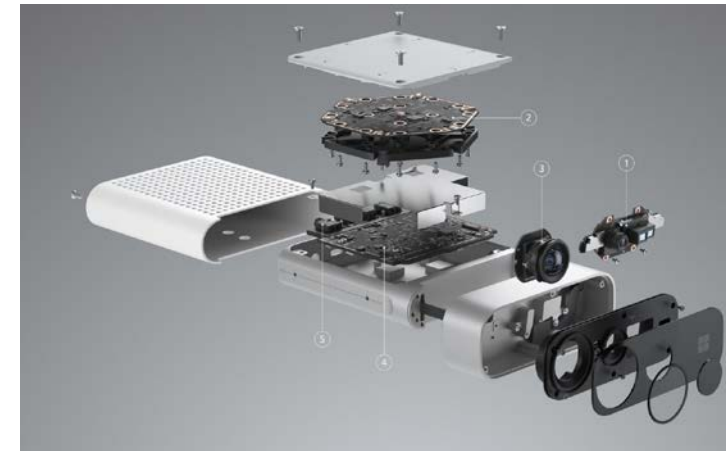


- Compute time of flight of IR pulse for each pixel
- Output depth image in addition to RGB
- Usually 0.5 - 7m range (or 7-14m etc.)
- Usually 8-bit depth resolution
-> range split into 256 parts
- Currently max. 1MP resolution for depth image
- New possibilities for new applications



MS Azure Kinect

- Time-of-Flight Sensor
 - 1024x1024 px depth image, 120° FOV
 - 3840x2160 (RGB camera), 90°x59° FOV
 - 30 Hz
 - Range: 0.5–5 meters
 - IMU, 7 microphones (array)



Infrared in V2



Depth sensing in V2



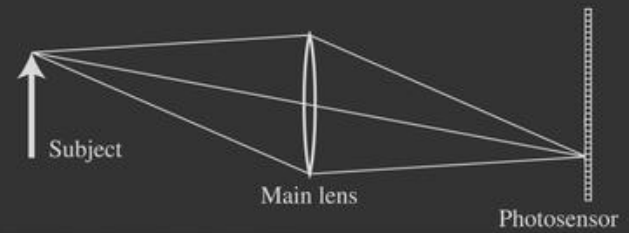
1080p color camera in V2

Light Field Camera (Plenoptic camera)

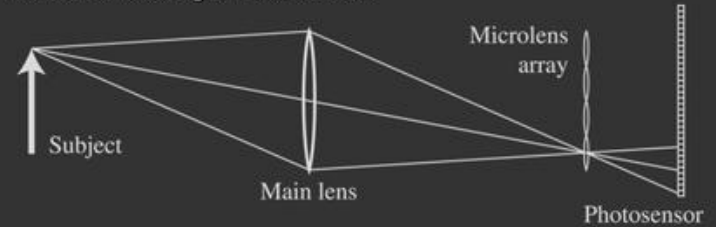
- is a camera that uses a microlens array to capture 4D light field information about a scene
- Allows reconstruction of the whole light field situation



Light Field Inside a Camera



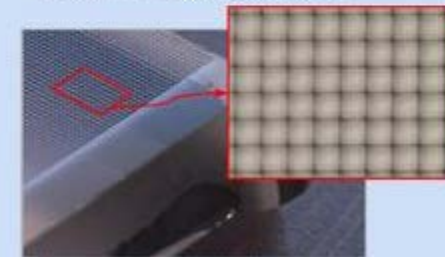
Lenslet-based Light Field camera



[Adelson and Wang, 1992, Ng et al. 2005]



Kodak 16-megapixel sensor



125 μ square-sided microlenses

4000x4000 pixels; 292x292 lenses =
14x14 pixels per lens

Light Field Cameras



- Cameras: Raytrix, Lytro
e.g. Raytrix R5, 1000x1000 resolution
(2Kx2K sensor, grayscale), 60Hz, dual GigE, >100 depth layers
- Interactive Examples: <http://lightfield.stanford.edu/lfs.html>
- Enables applications such as
 - Changing focus
 - Deblurring an image
 - Depth image computation



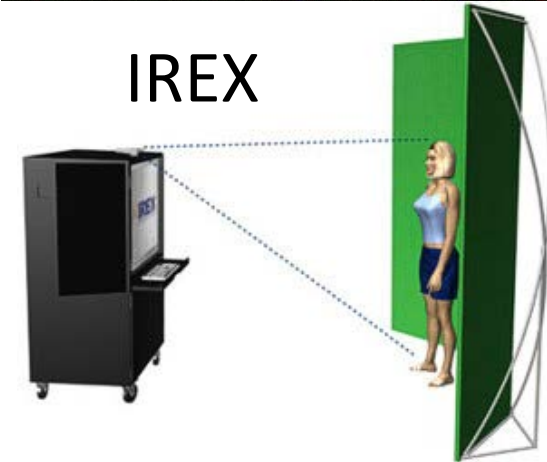
- Slides about light fields:

http://graphics.stanford.edu/courses/cs478/lectures/02272012_lightfields.pdf



Low-cost Gesture Recognition

Low cost for various applications:
Kinect / SONY Playstation Eye Toy
for Rehabilitation





Motion Capture

Motion Capture Technologies:

- Exoskeleton (Mechanical)
- Wireless Magnetic Sensors
- Wireless Inertial Sensors
- Marker based (Optical)
- Pure vision based (no markers):
 - First systems exist
 - High computational power needed (cluster)

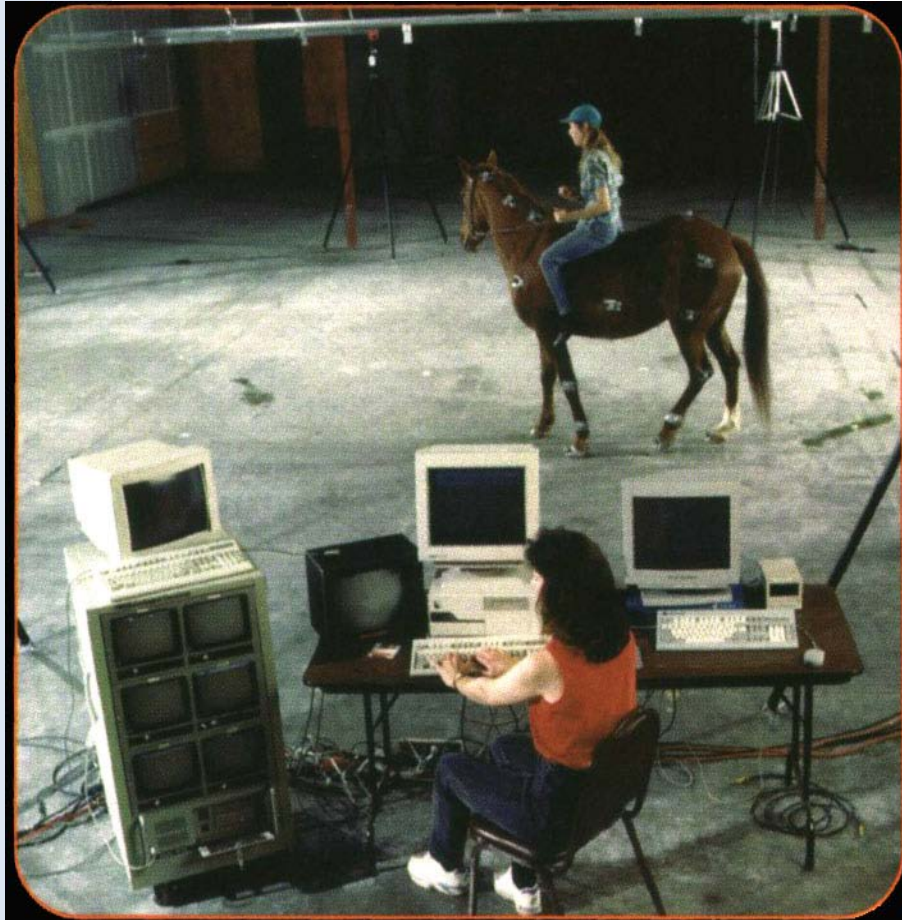


Optical: Vicon

- 3 different camera models (0.3 – 4M pixels; 240-1000fps)
- Very high quality
- Price starts at 70K



Tracking Animals

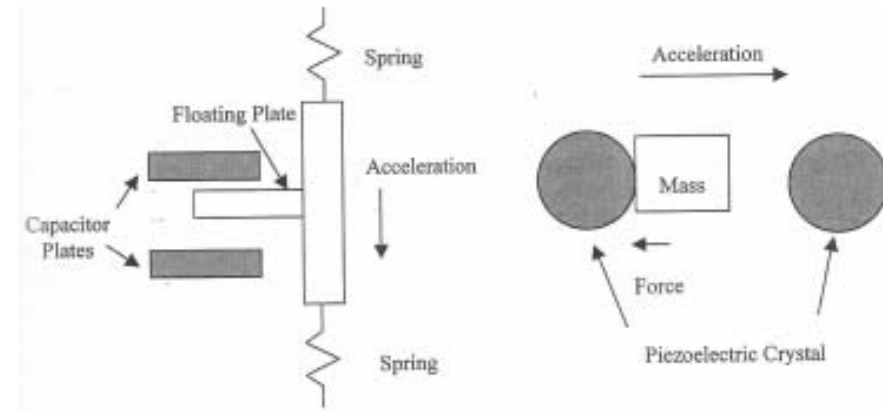


Inertial Sensing



Inertial or Gravimetric

- Accelerometers
 - Linear accelerations
 - Mass on spring in tube; Piezoelectric
- Rate sensors (Gyroscope)
 - Angular velocity
 - Excentric mass....
- Gravity sensors
 - Orientation in gravity field



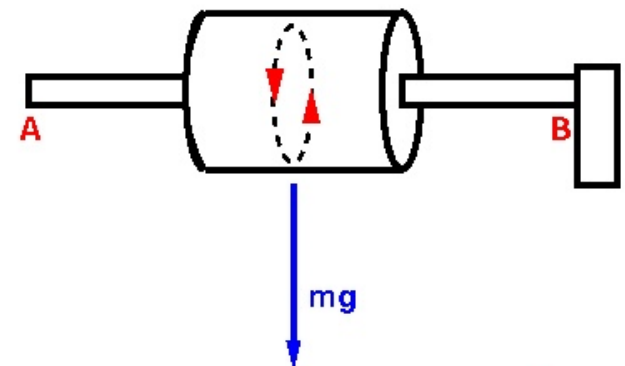
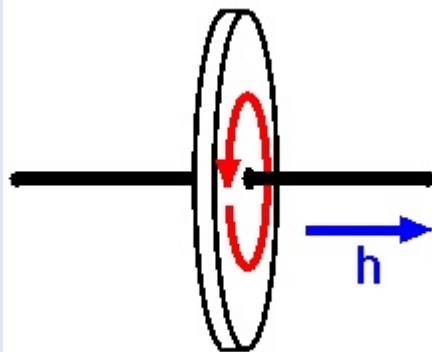
(Explanation:
Piezoelectricity is the
ability of certain
crystals to generate a
voltage in response to
applied mechanical
stress.)

Inertial Navigation

- **Relative** navigation from known starting point
- Sense rate, integrate once
- Sense Acceleration, integrate twice



How to Measure Angles with a Gyroscope?



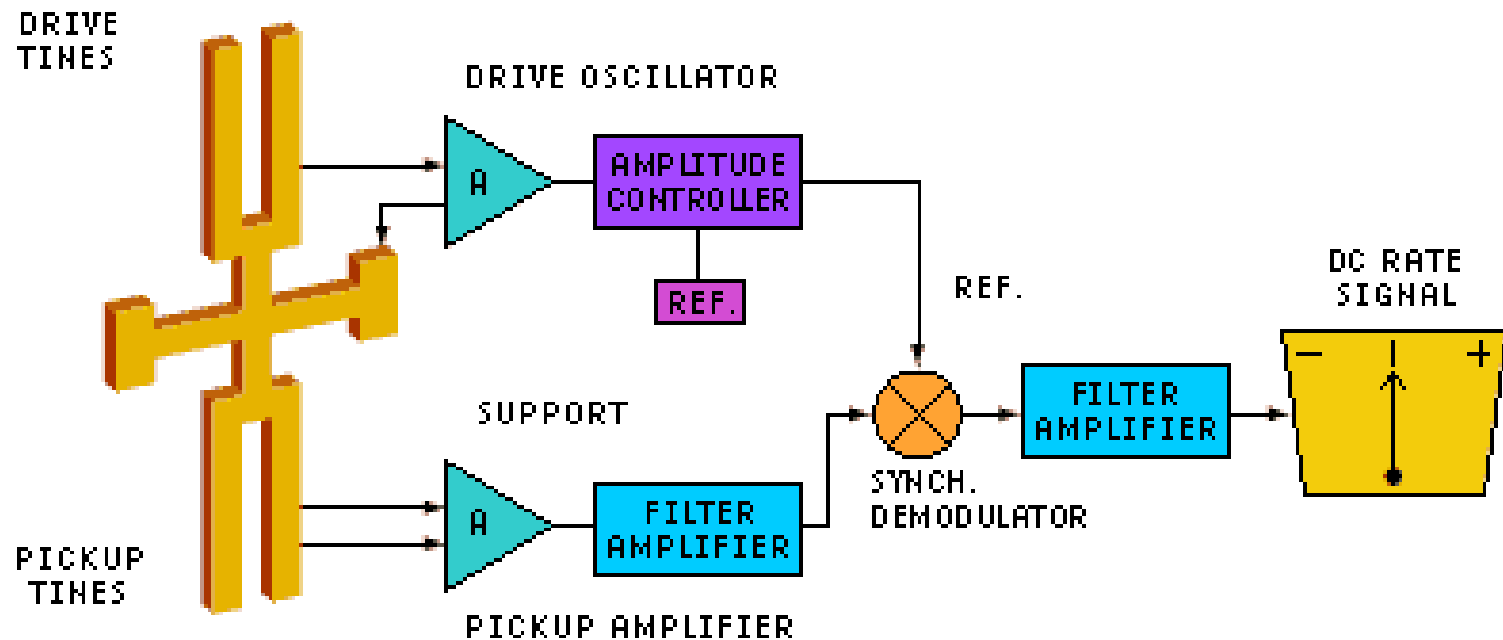
„Drehimpulserhaltungssatz“

For any collection of spinning objects, the total angular momentum must stay constant.

Vibrating Gyroscope

- When rotated a **vibrating element** (vibrating resonator) is subjected to Coriolis's effect.
- Coriolis effect causes secondary vibration orthogonal to the original vibrating direction
- Sensing the secondary vibration -> the rate of turn can be detected
- For vibration exert and detection the piezo-electric effect is often used -> often called "piezo", "ceramic", or "quartz" gyro

The GyroChip™



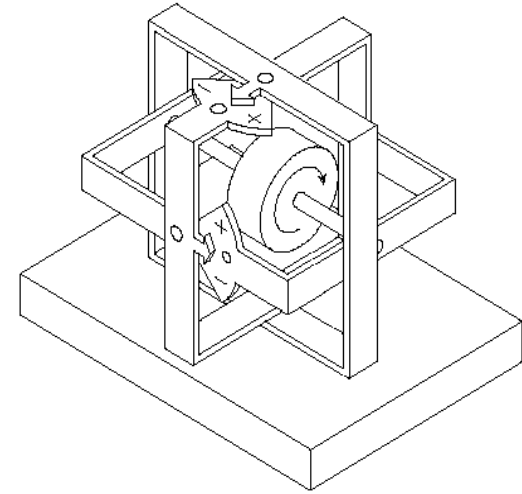
Quartz tuning fork -> rotation -> coriolis force-> voltage

Vibrating Gyro

- suitable for mass production and almost free of maintenance
- Drawback
 - If used under external vibration, it cannot distinguish between secondary vibration and external vibration
 - can be partly overcome by using other vibrating elements (rings) instead of tuning fork

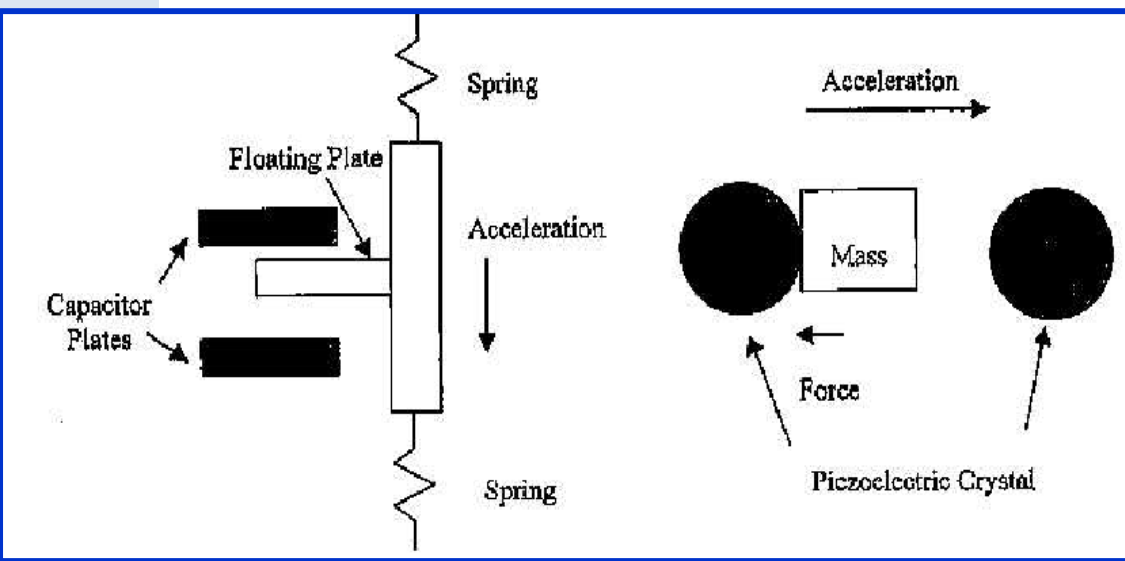
Inertial Trackers

- Intersense IS-300
- Less noise, lag
- Mostly only 3 DOFs (orientation)
- Relative measurements
--> integration errors --> drift



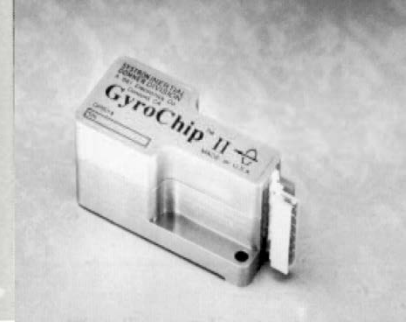
Types:

Gyroscopes
Accelerometers
Inclinometer
Electronic compass



Application Areas

- Stabilization
 - Satellite Communication Antennas
 - Optical Line-of-Sight Systems
 - Missile Seekers
 - Electro-Optical Infrared Radar Systems
- Controls
 - Aircraft & Missile Flight Control
 - Attitude Control
 - Yaw Dampers
- Guidance
 - Missile Mid-Course Guidance
 - Inertial/GPS Navigation Systems
- Instrumentation
 - Sounding Rockets & Missiles
 - Simulation & Training Aids



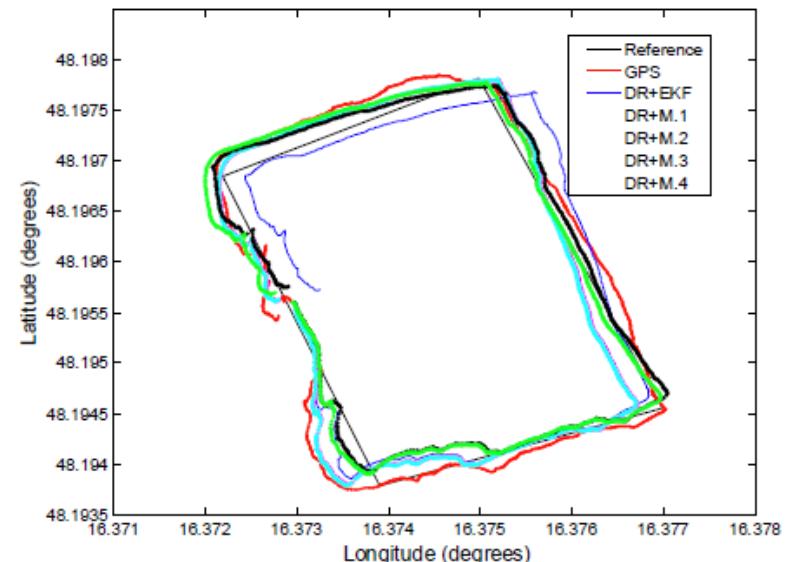
MotionPak 2™
6 DOF
Size ~ 6x6 cm



X-UFO

Navigation of Humans/Machines

- Traditionally used in robotics (inertial and/or optical)
- Outdoors additionally GPS



Motion Capture with Inertial Sensors

X-sense MVN



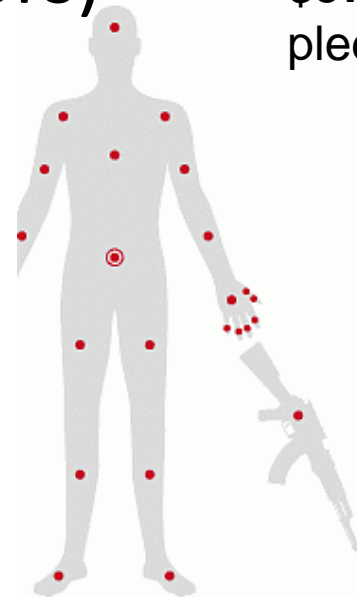
PrioVR

Inertial MoCap: Perception Neuron

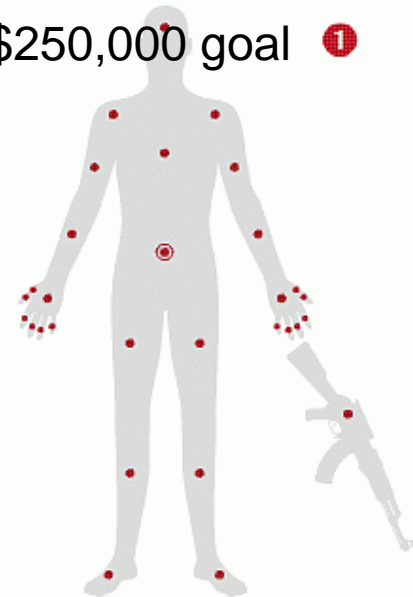
- Low cost: 1.500 EUR
- 32 sensors – flexible use
- USB or WiFi transfer
- 60 fps (with 32 sensors)
- Finger Tracking



Kickstarter 2015:
\$571,908
pledged of \$250,000 goal ¹



Body with two hands and a tool



Full body with two hands and a tool

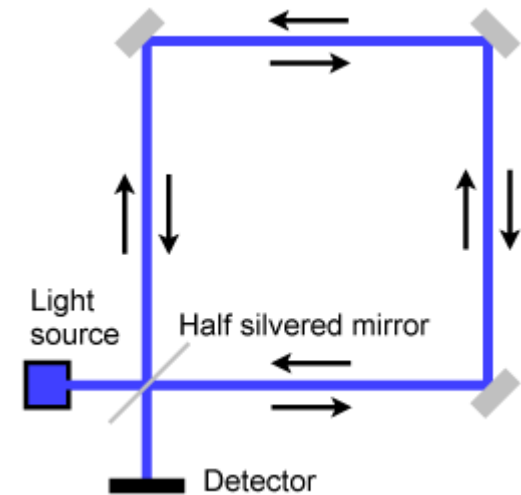
Inertial Tracking - Sensors

- Pros
 - Truly Sourceless
 - Very fast
 - Robust
- Cons
 - Bias, scale and alignment errors
 - Bias integration results in drift
 - Wireless versions require “bigger” sender



Other Gyroscopes & Additional Sensors

- Sagnac Effect
 - Ring Laser Gyroscope
 - Fiber Optic Gyroscope
- Quantum Gyroscope
- Electronic compass:
Measuring the earth's magnetic field
- Inclinometer – measures slope (viscous fluid)



Time-of-Flight and Frequency Measuring

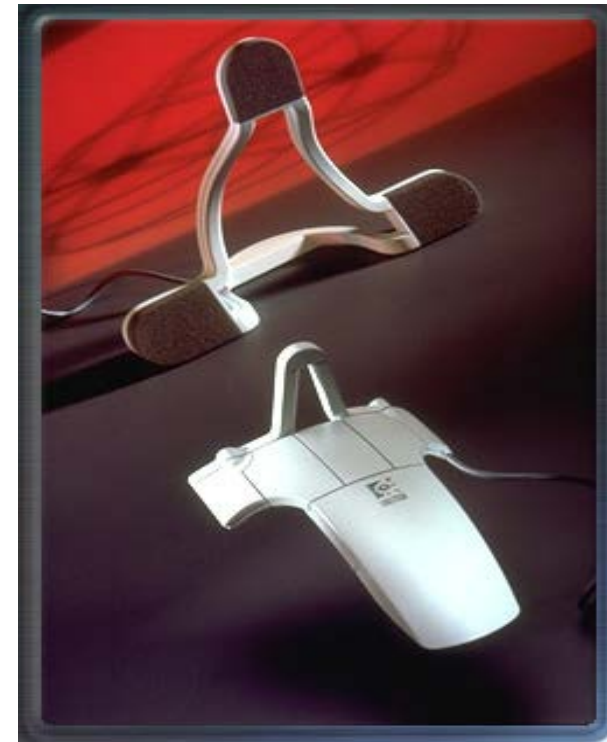
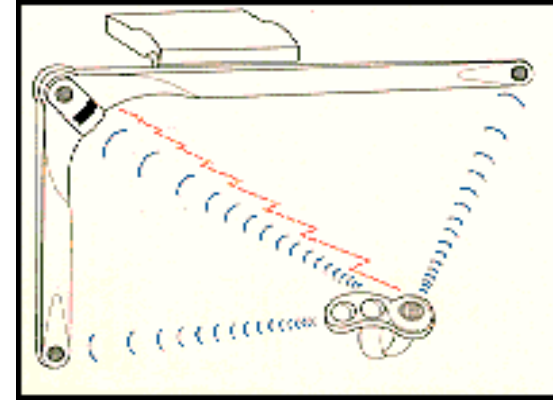
- Acoustic Trackers (Ultrasound)
- GPS / DGPS
- Radio Frequency
 - WLAN
 - RFID
 - UWB
 - ...

Time-of-Flight and Frequency Measuring

- Measure the time of propagation of a signal
- Compare phase difference of waves
- Measure frequency of waves to indirectly estimate time difference

Acoustic Tracking

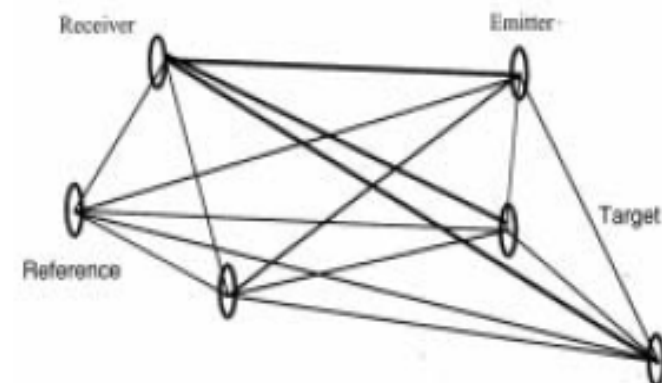
- Measure distance between reference features and moving target
- Determine distance: Measure **time of propagation** of pulsed signals (travel time of sound)
 - Speed of propagation must be constant (!)
- Pulsed signal: Ultrasound ($\sim 40\text{Khz}$)
- Use 3 or more emitters and 3 or more receivers (6 DOF)
- Typical setup for 3 DOF:
 - 3 microphones, 1 speaker



Logitech Fly Mouse

Ultrasonic Time-of-Flight

- Each emitter sends out signals sequentially
- Receivers receive signals simultaneously
- Compute position of target emitters with respect to the fixed reference via simple triangulation
- Repeat for (at least) 3 target emitters
 - position and orientation of target with respect to fixed reference



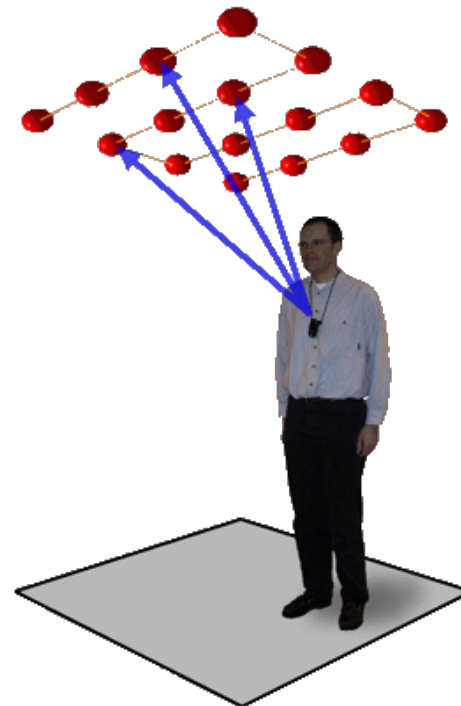
Acoustic Tracking: Pros & Cons

- Advantages
 - Small and light weight targets
 - Relatively inexpensive
- Disadvantages
 - Line of sight issues
 - Accuracy depends on constancy of velocity of sound
 - Speed of ultrasonic waves varies
 - Temperature, Pressure, Humidity, Turbulence
 - Ultrasonic noise
 - CRT tubes of monitors, disk drives, etc.
 - Reflection of signal

Acoustic Trackers: Wide Area

- Wireless targets are possible
- Example: AT&T Cambridge BAT, IS-900

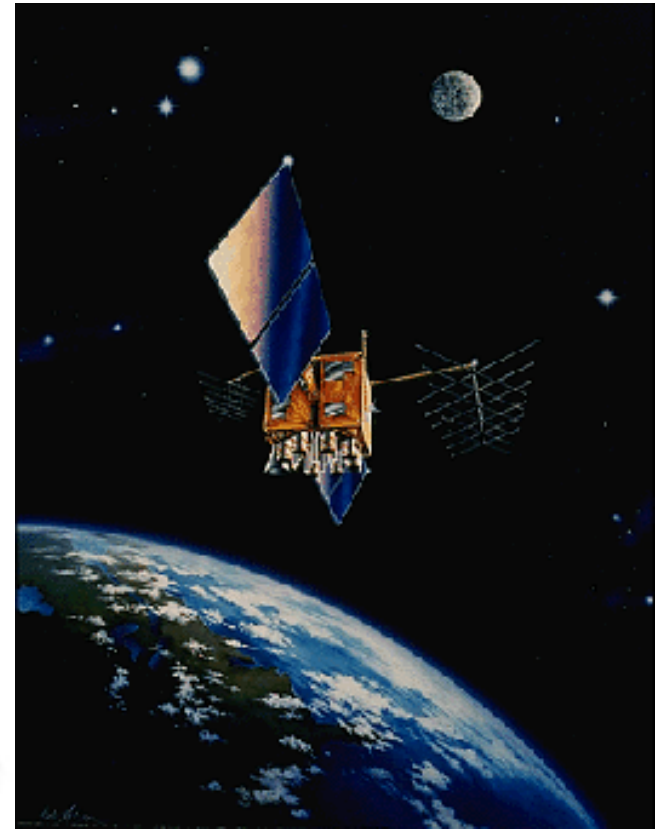
AT&T Bat



Global Positioning Systems (GPS)

Global Navigation Satellite Systems (GNSS)

- GPS (USA; Global Positioning System)
- GLONASS (Russia)
 - “Globalnaja Nawigazionnaja
Sputnikowaja Sistema”
- Galileo (Europe)



GPS

- Large scale
- Based on radio signals
- Uses 24 satellites in orbit -> 4 of them seen simultaneously from any point on earth
- + 6 monitoring stations, 4 ground antennas, 1 master control station

GPS

- Two levels of service:
 - Standard positioning service (SPS)
 - Precise positioning service (PPS) -> only for U.S. Government
- Each satellite has atomic clock (accuracy 340ns for SPS)
 - 1ms clock error produces horizontal measurement error of 300km
- Master control station controls orbit of satellites and corrects clocks

GPS Receiver

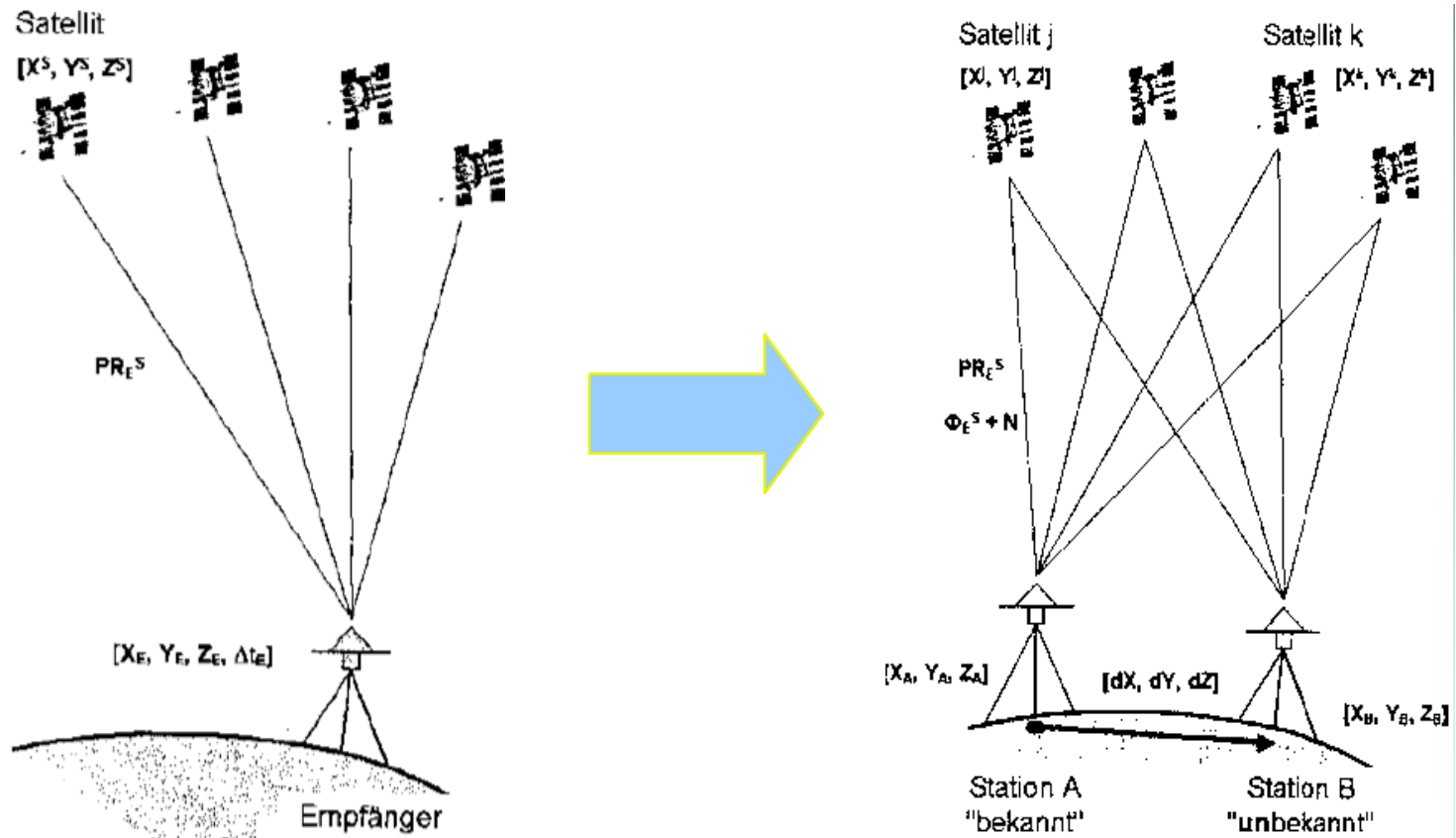
- Receiver needs to receive signals from at least 3 different satellites to determine position
- BUT receives clock as unknown time bias! (drift)
- Satellites more accurate (2ns/year) than GPS receivers (10ns/day)
- Therefore use 4 satellite signals to calculate position and time bias; in practice: 5 satellites!!
- Accuracy of SPS: 100m
- Accuracy of PPS: at least 10x better than SPS

GPS Pros & Cons

- Advantages:
 - Uniform global accuracy ($\sim 10\text{-}100\text{m}$) without other infrastructure
 - Cheap and easy to use
- Disadvantages:
 - Direct line of sight needed - no GPS data indoors
 - Problems in narrow street passages, wood,...
 - Resolution and precision

Differential GPS (DGPS)

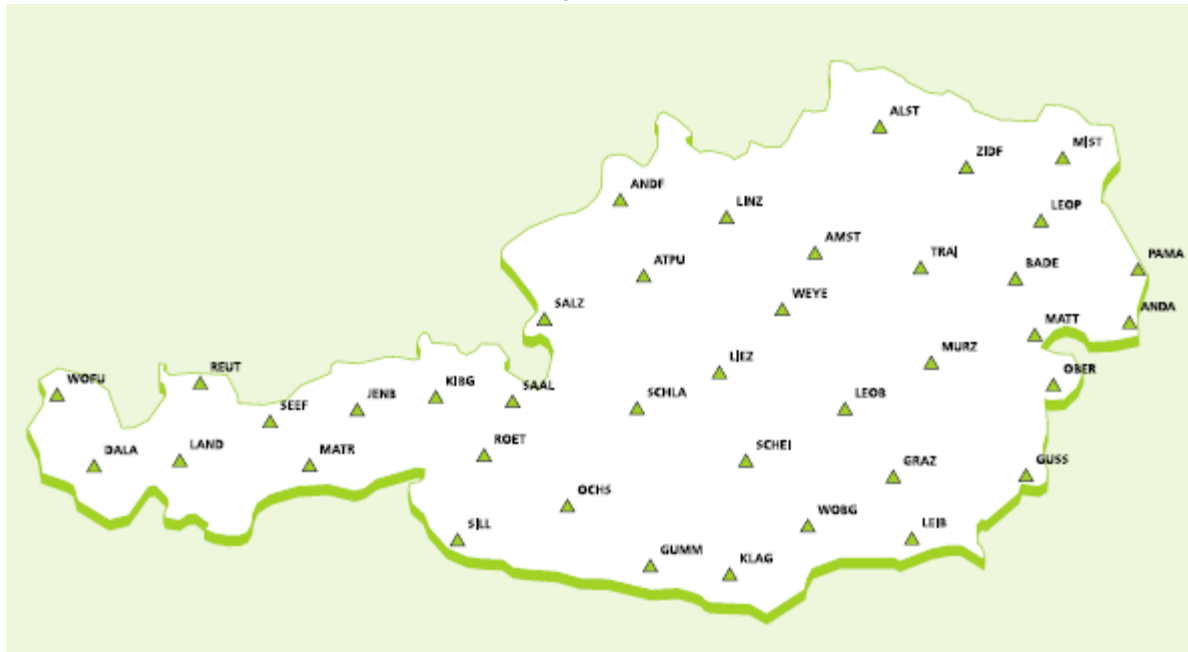
- Uses additional fixed ground station to refine resolution
 - Theoretically 0.1m, practically 3-5m



Austrian EPOSA System

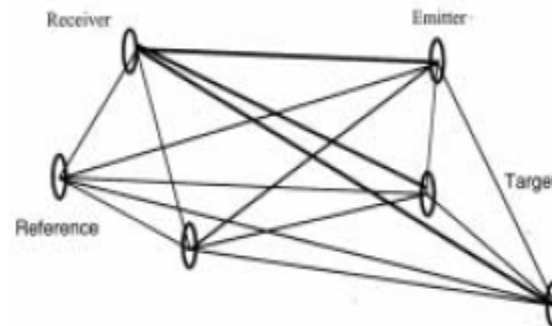
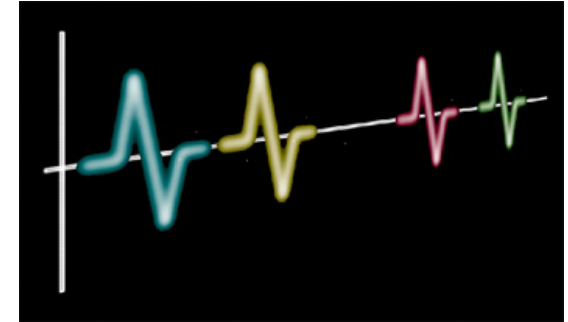


- Differential correction data sent over GMS, Internet (by RTCM or NTRIP standard)
- Supports GPS and GLONASS
- Cooperation between BEWAG, ÖBB-IKT and Wien Energie
- 1-3 cm Precision



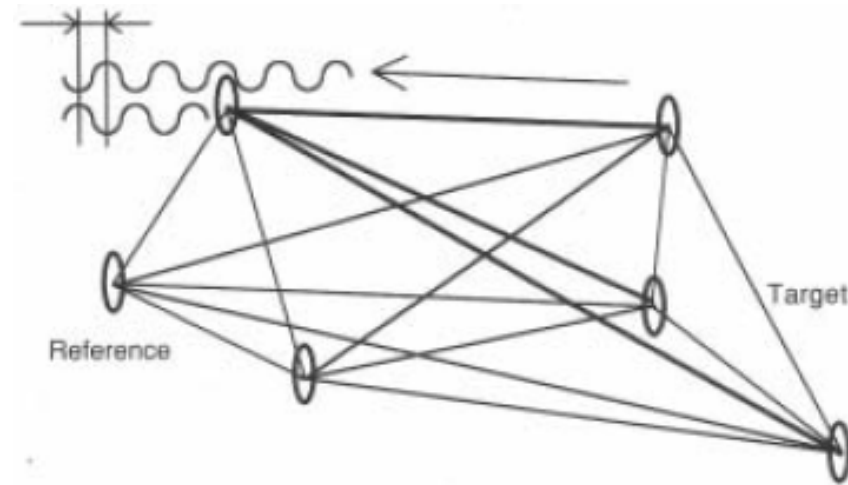
Radio Frequency Trackers

- Triangulation similar to acoustic tracking
 - Proximity
 - 802.11a/b = WLAN
 - RFID
 - GSM
 - Ultra Wideband
 - Phase difference



Phase Difference Time of Flight

- Similar to acoustic tracking
- Target sends signal
- Receiver has reference signal with the same frequency
- Receiver compares phase shift of target signal with reference
- Modification of phase indicates relative movement of target
- Example: laser measuring



Phase Shift Pro & Cons

- Advantages
 - High data rates possible (phase shift can be measured continuously)
- Disadvantages
 - Limited to environmental conditions
 - Temperature, pressure,
 - GPS: Ionospheric influences (sun winds)
 - **Reflections** in the environment are problematic
 - Multipath effect
 - Accuracy
 - Cumulative errors in measurement process

Ubisense Ultrawide-band (UWB)

- UWB transmits signal over multiple bands of frequencies simultaneously: 3.1 - 10.6 GHz (RFID systems operate on single bands of radio spectrum)
- Active tags used; Ubisensors 5.8 to 7.2 GHz
- Tags have unique 32-bit (4-byte) identifier and send location max. 40 Hz

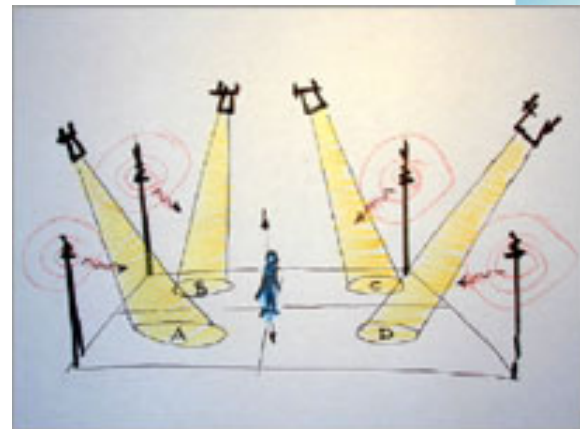


Inmotio: Local Position Measurement (LPM)

- Austrian company: Inmotiotec
- Large area tracking
 - Red Bull Air Race
 - Zactrack
- 500x500m area/sect.
- Max. 30 sections
- Like UWB: 5.8Ghz
- Update rate: 1000Hz
- Accuracy: +/- 5cm (outdoors)



Base station



Transponder

Kinexion – Decawave

- Decawave IEEE 802.15.4 standard
- UWB: over 15 diff. frequency bands
- +/- 10cm accuracy
- Inertial sensor integrated
 - $<0.1^\circ$ orientation data & acceleration
- 40m indoors, 300m outdoors
- 15g sensors
- Industrial & sports applications



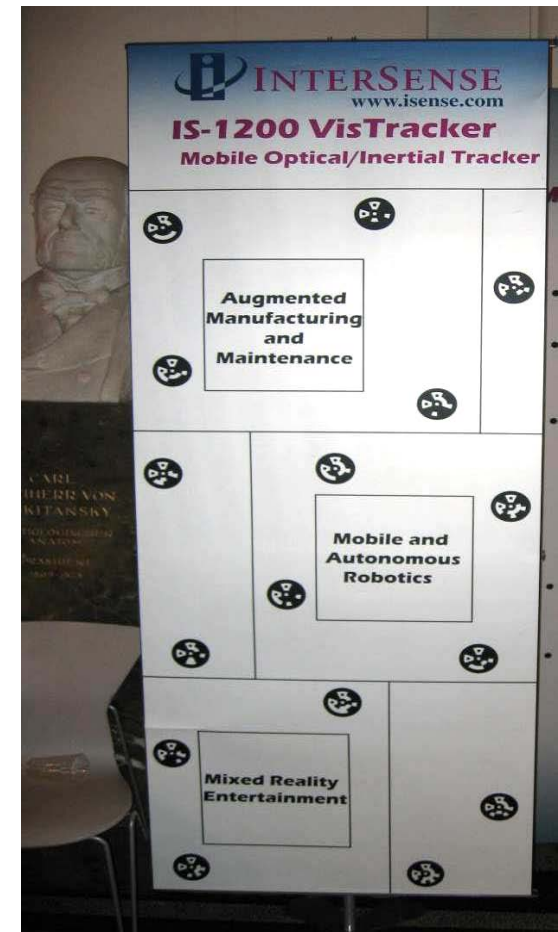
Hybrid Trackers

- Idea: one tracker's strength is other tracker's weakness
- For example: Intersense IS-600 / 900
 - inertial: +fast, -drift
 - acoustic: -slow, +accurate



Intersense IS-1200

- Combination of inertial + optical tracking
- Advantage: Very accurate
- Disadv.: Markers needed



Miscellaneous Devices

- Gloves
- Pens / Wands
- Hybrid Devices
- Game Controllers

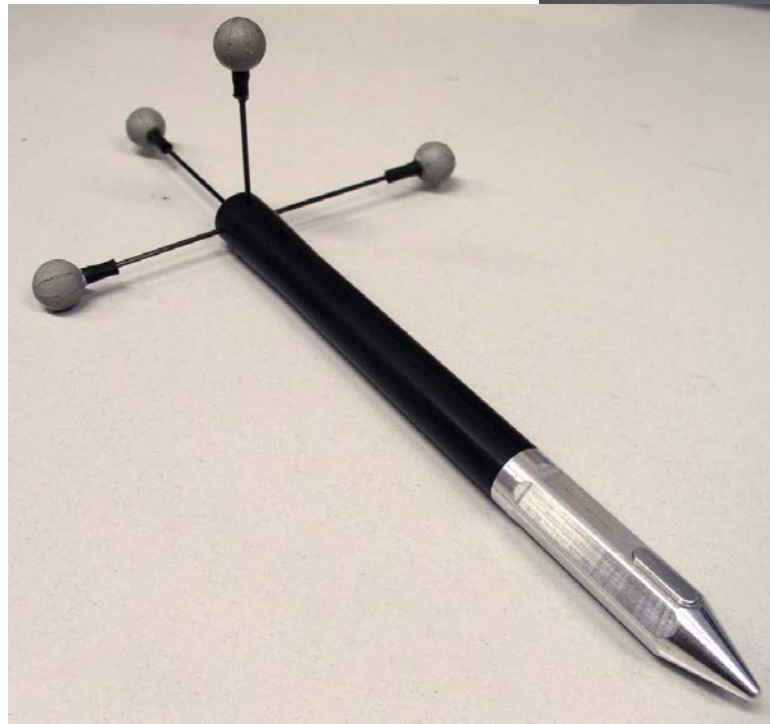
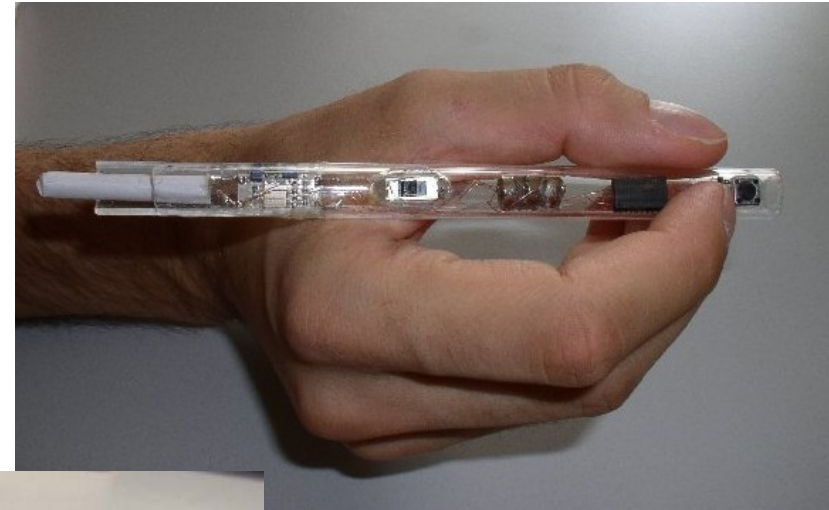
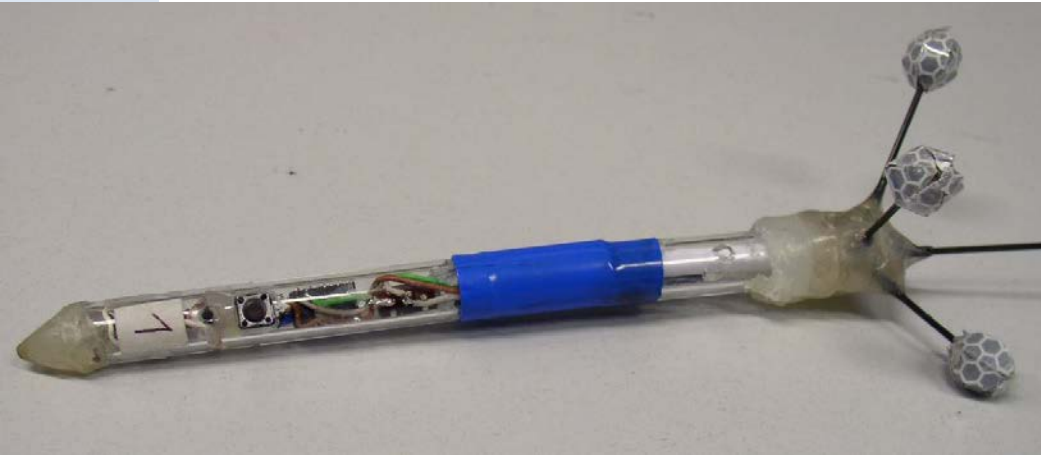
- Haptics
- Locomotion

Data Gloves

- Used to track the user's finger movements
 - for gesture and posture communication
- Almost always used with a tracker sensor mounted on the wrist
- Common types
 - CyberGlove
 - 18 sensors, 22 sensors
 - 5DT Glove
 - 5 sensors, 16 sensors
 - P5 Glove
 - bending sensors



Tracked Wireless Pens (IMS)



Hybrid devices

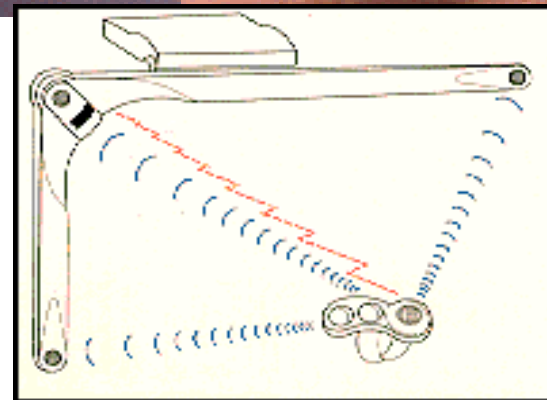
Continuous and discrete input Examples

- Button device + tracker
- Flex & Pinch
- Ring mouse
- LCD tablet
- Shape Tape
- Cubic Mouse
- Spaceball



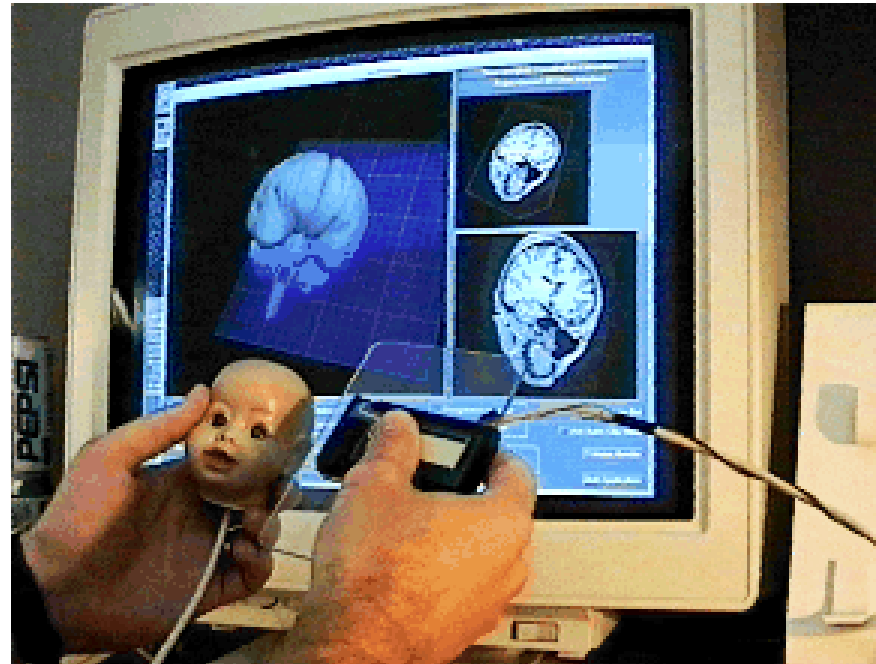
Mouse Type Devices

- Space Mouse
- Ring Mouse (pictured)
- Fly Mouse
- Gyro Mouse



Various Props – Real Objects

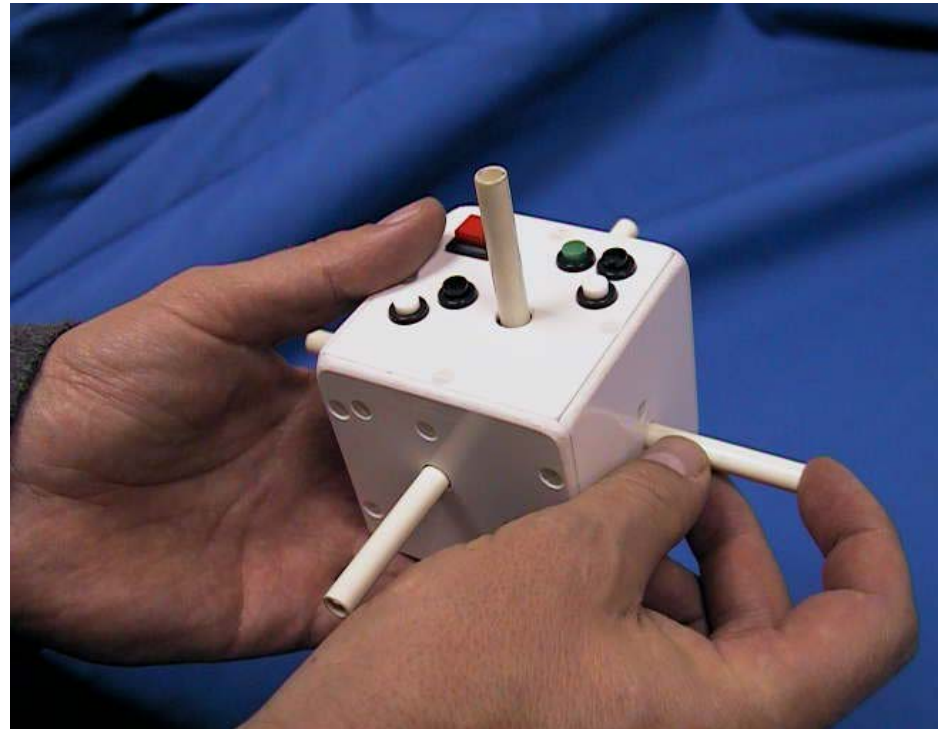
- Head prop
- Car prop
- ...



Brain segmentation

Cubic Mouse

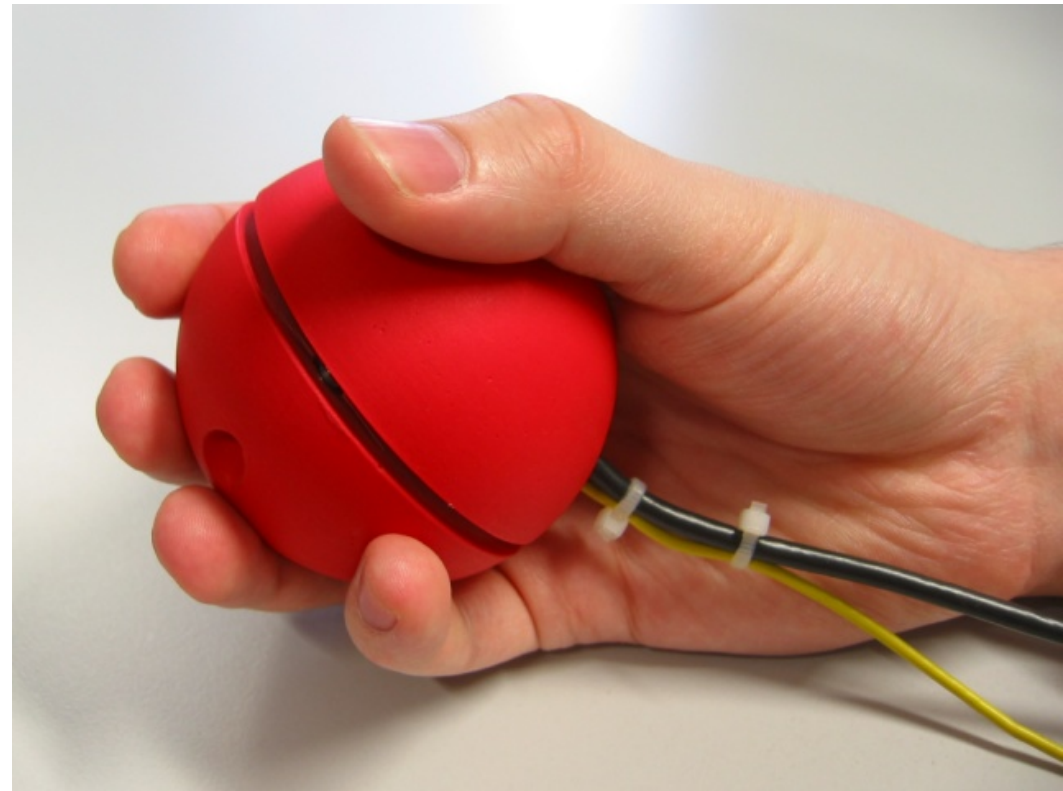
- First 12 DOF input device
- Tracks position and rotation of rods using potentiometers
- Other shapes and implementations possible
 - Mini Cubic Mouse
 - ...



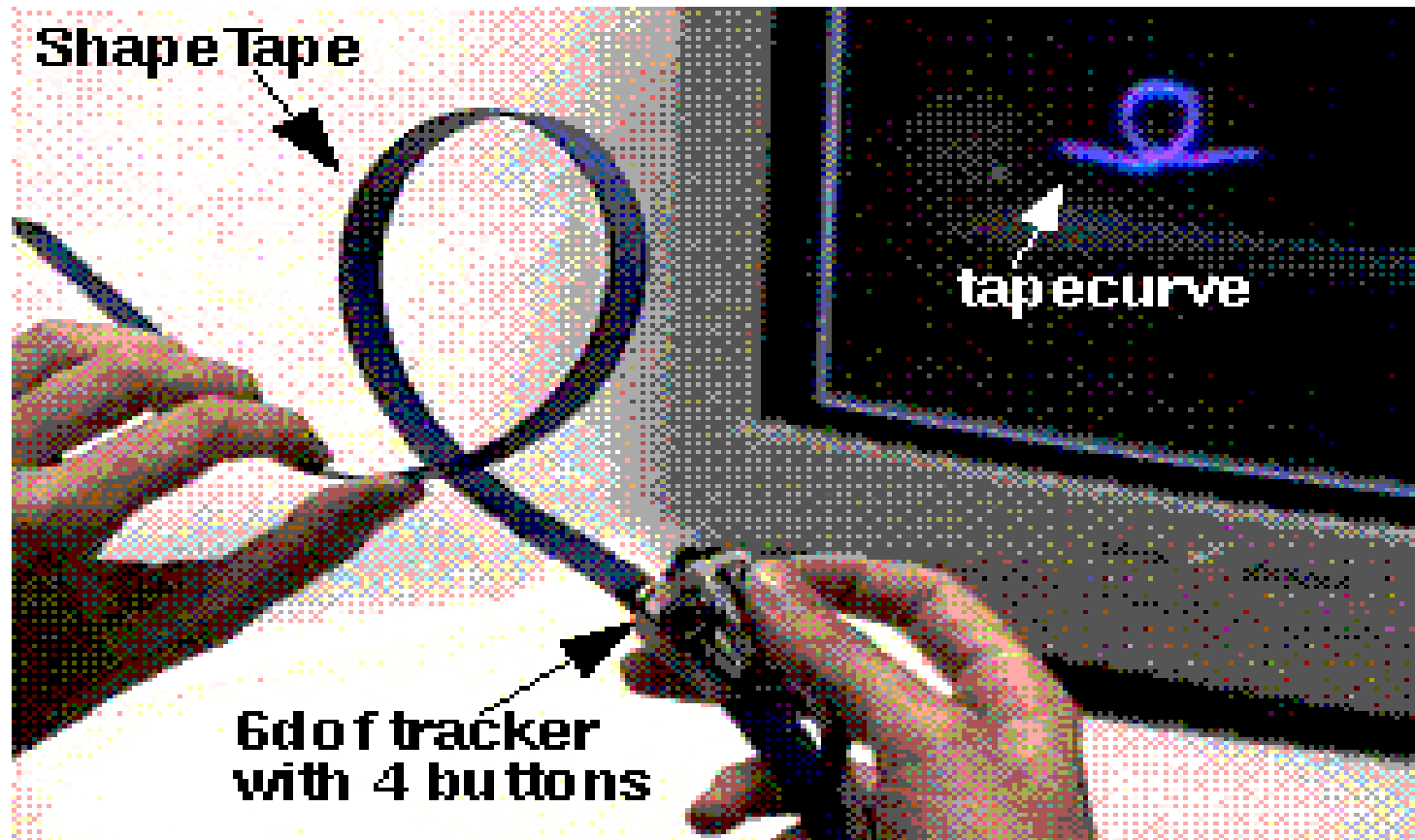
iOrb

- Inertial Tracker inside
- 1 Button
- Freely rotatable

Reitmayr et al., 2004



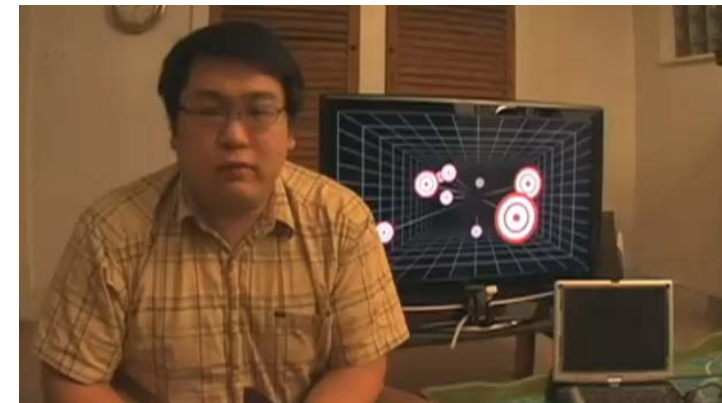
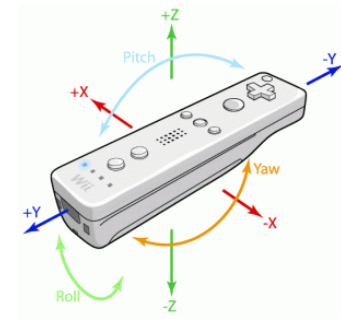
Shape Tape



Courtesy Balakrishnan et al

Wiimote

- 3-axis accelerometer
- Optical (IR) sensor
 - IR LEDs on a sensor bar used for recalibration
 - Calc. distance to sensor bar
 - max. 5 points simultaneously
- Rumble functionality



Sony MOVE Motion Controller

- Inertial sensor (gyro, accel., magnetom.) – measures orientation
- 60 Hz camera used for optical tracking of colored sphere
 - High accuracy (cm/mm)
 - Controller can change colors (eases segmentation)



Haptic Devices



Tactile Technologies – Feeling Pressure

- Tactile information is produced by perturbing the skin
 - **Pins** -either alone or in an array, as in devices for Braille display
 - typically used for fingertip stimulation
 - **Air jets blow** to produce a disturbance
 - **Cushions of air** can be inflated or deflated to vary pressure on skin
 - **Electrical stimulation** -low levels of current provide a localised tingling sensation
- Typically used in gloves, or for larger body areas
- These technologies all share the same lack of realism

Haptics

- Human tactile sensing sensitive to vibration 10-1000Hz -> 1000Hz required for stable force feedback
- Physical accuracy difficult to achieve in real time
- Simplified/approximate (deformation) models
- GPU based haptics modelling approaches
- Haptics devices & photos:
<http://www.personal.rdg.ac.uk/~shshawin/LN/L8hapticdesigns.html>
- Further literature: Good introduction to haptics:
<http://www.dcs.gla.ac.uk/~stephen/lectures/HCI4/HCI4%20kinaesthetic%20haptics.pdf>

Haptics – Special Devices

- [Haptic Smarttool](#)
 - Could be useful for medical applications
- Medical Training Devices (e.g. Endoscopy)



Hysteroscopy AccuTouch® System
Powered by TouchSense® Technology



Locomotion devices

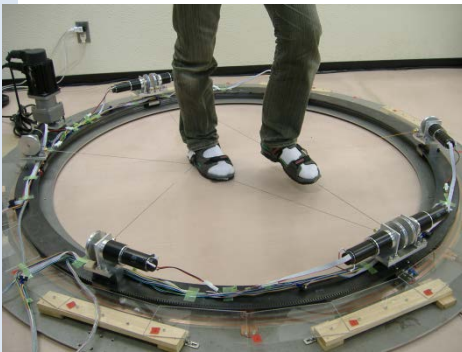
Locomotion = Active Movement

- Treadmills
- Stationary bicycles
- Walking/flying simulations (use trackers)



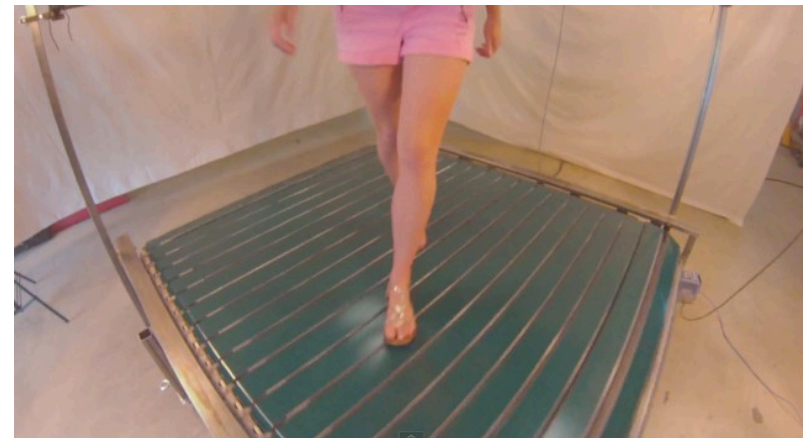
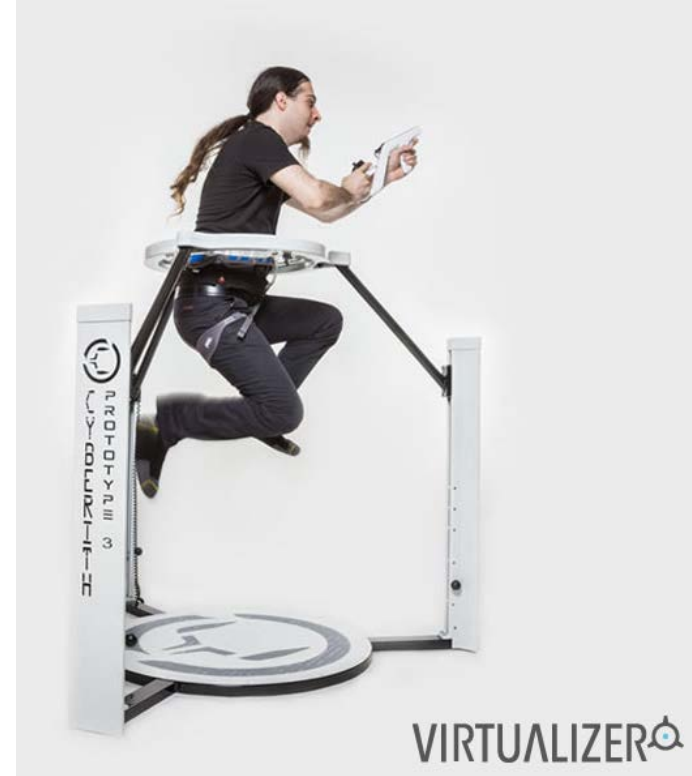
Locomotion - Prof. Iwata

- [Virtual Preampulator](#) (1996)
- [TorusTreadmill](#) (1999)
- GaitMaster 2 (2000)
- CirculaFloor (2005)
- Powered Shoes (2006)
- StringWalker (2007)



Current Locomotion Devices (2014 -)

- Cyberith Virtualizer
 - Developed by
Tuncay Cakmak (TU Vienna)
- Infinadeck
Omni-Directional Treadmill (ODT)
- Virtuix Omni



Literature

- 3D User Interfaces – Theory and Practice (2nd ed.)
J. LaViola, Ernst Kruijff, Ryan P. McMahan, Doug Bowman, Ivan Poupyrev; Addison Wesley, 2017.

