AUGMENTED REALITY

Course 2024/2025

Contents

- **Introduction**
- Augmented Reality Displays
- AR interfaces
- Tracking technology
- Research directions

These slides are based on:

- Oliver Bimber, Ramesh Raskar, 2004. Spatial Augmented Reality: Merging Real and Virtual Worlds
- Doug Bowman, et al. 2004. 3D User Interfaces. Theory and Practice. Addison-Wesley.
- Mark Billinghurst AR couses (www.hitlabnz.org)

Introduction to AR

- Augmented Reality is a combination of a **real scene** viewed by a user and a synthetic **virtual scene** that augments the scene with additional information.
- AR environments differ from VEs in that we have access to both real and virtual objects at the same time.

Augmented vs Virtual Reality

Augmented Reality

- System augments the real world scene
- User maintains a sense of presence in real world
- Needs a mechanism to combine virtual and real worlds

Virtual Reality

- Totally immersive environment
- Visual senses are under control of system (sometimes aural and proprioceptive senses too)

Augmented vs Virtual Reality

Virtual Reality

- Scene generation: require realistic images.
- Display device: fully immersive, wide FOV.
- Tracking and Sensing: low accuracy is ok.

Augmented Reality

- Scene generation: mimimal rendering ok.
- Display device: non-immersive, small FOV.
- Tracking and Sensing: high accuracy needed.

Goal of AR

- Goal: enhance user **performance** and **perception** of the world.
- Challenge: keep users from **perceiving the difference** between the real world and the virtual augmentation of it.

AR applications

Most AR apps are focused on enhancing real-world activities:

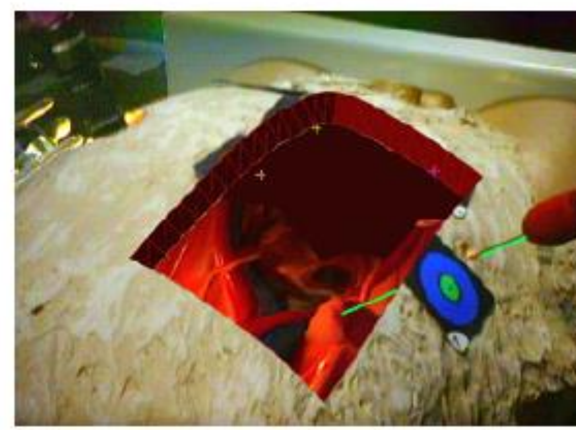
– Guidance to surgeons by displaying possible paths for needles

– Evaluate cockpit designs in relation to the real physical cockpit.

AR applications

Medical applications



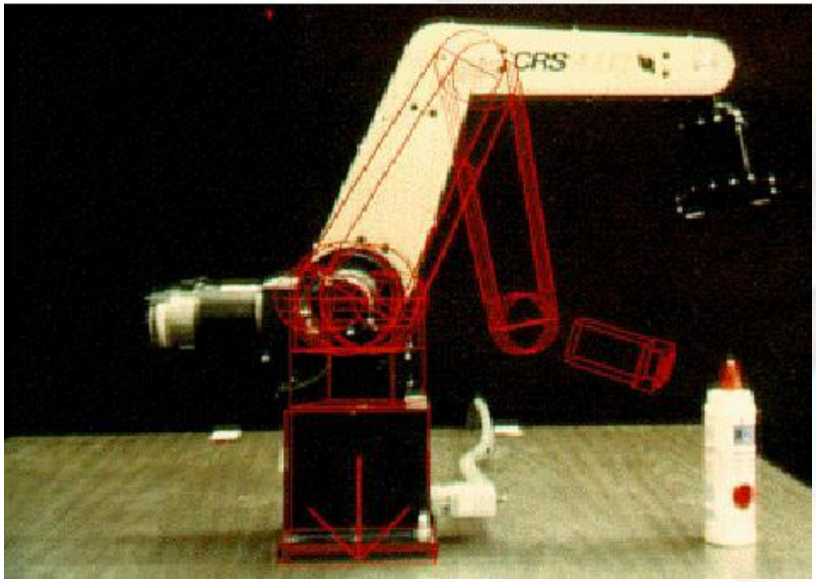



AR applications

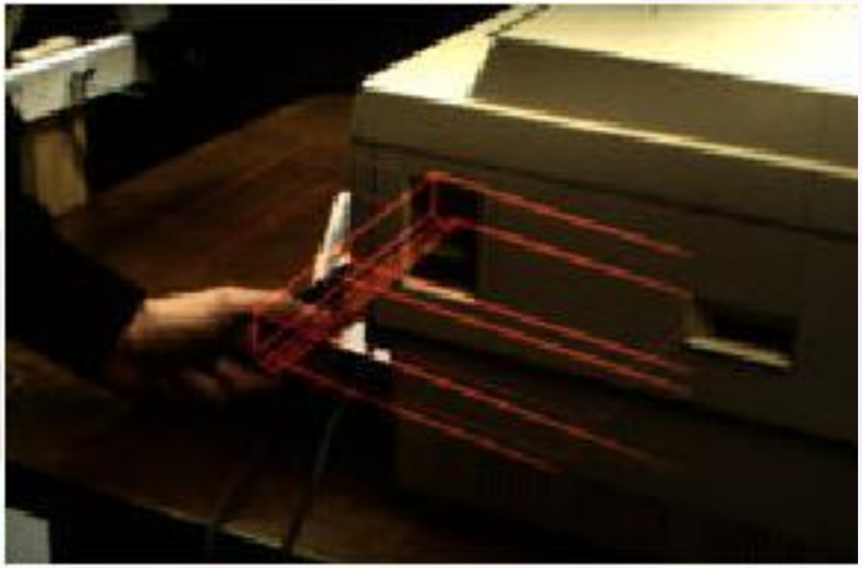
Collaborative applications

AR applications

- Robotics and Telerobotics

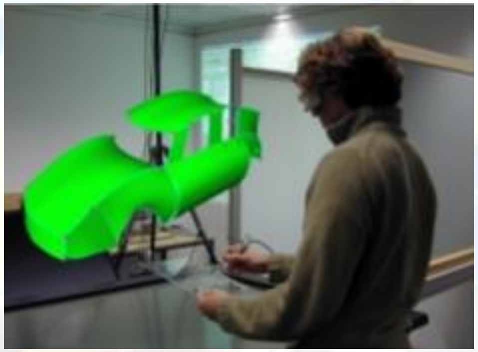


- Manufacturing, Maintenance, and Repair
- Hazard Detection



AR applications

- Archeology
- Entertainment
- Engineering design
- Consumer design



Mixed Realities

Mixed Reality view [Paul Milgram]:

P. Milgram and A. F. Kishino, Taxonomy of Mixed Reality Visual Displays IEICE Transactions on Information and Systems, E77-D(12), 1994.

Mixed Reality (MR)

Real Environment → Augmented Reality (AR)

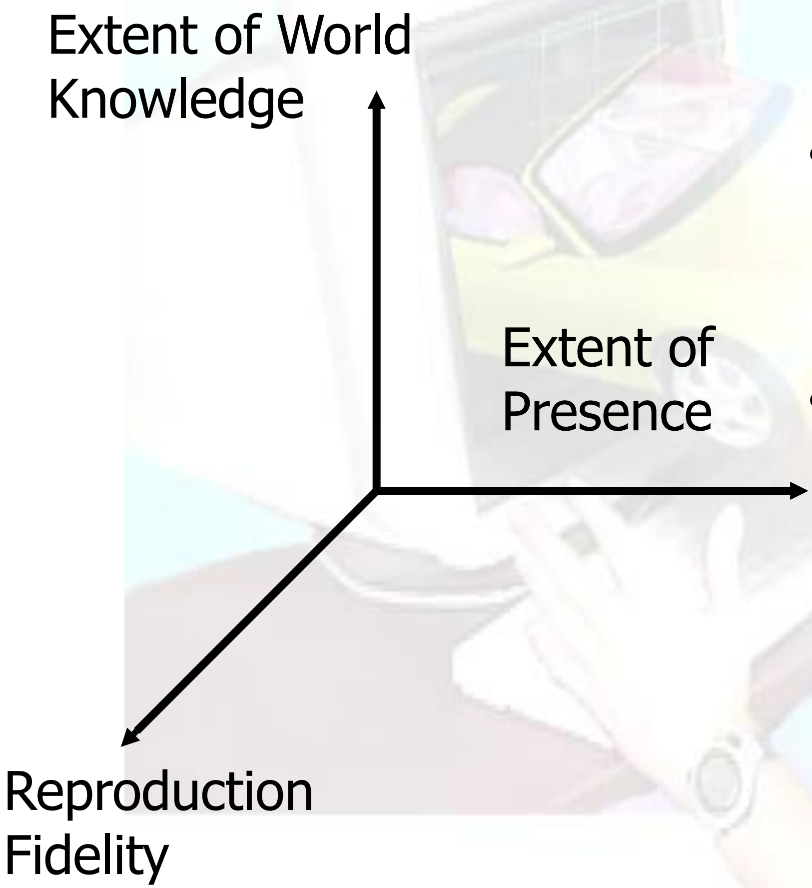
Augmented Virtuality (AV) ← Virtual Environment

AR characterization

Characterization of AR interfaces [Azuma 1997]:

- Superimpose virtual and real objects in the same interaction space (combines real and virtual images)
- Interactive in real-time (virtual content can be interacted with)
- Registered in 3D (virtual objects appear fixed in space)

Milgram’s Taxonomy for MR

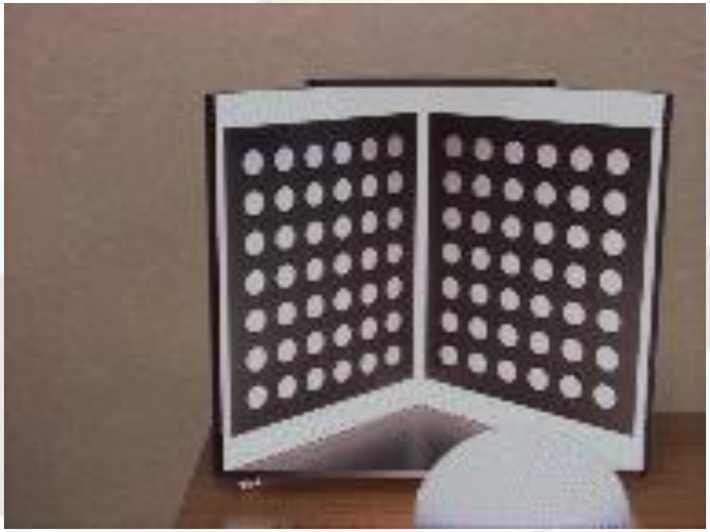


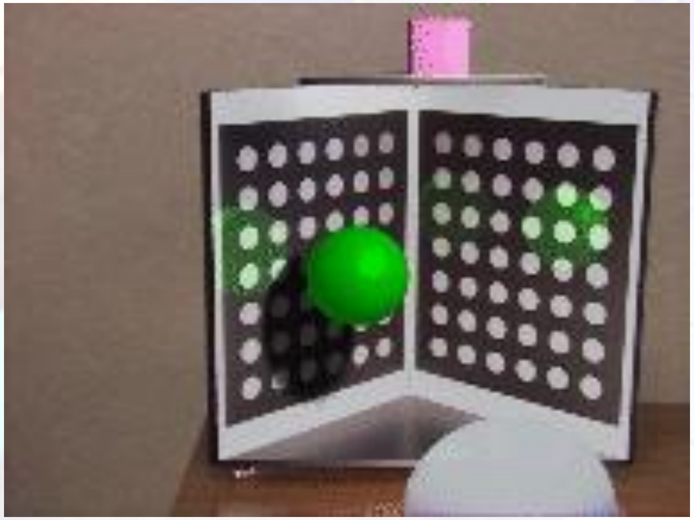
- **Reproduction Fidelity** – quality of computer generated imagery
- **Extent of Presence** – level of immersion of the user within the displayed scene
- **Extent of World Knowledge** – how much the computer knows about the real world, the camera viewing it, and the user

AR technologies

Basic technologies involved in AR

- Display technology → Combines Real and Virtual Images
- Interaction technology → Interactive in real-time
- Tracking technology → Registration in 3D



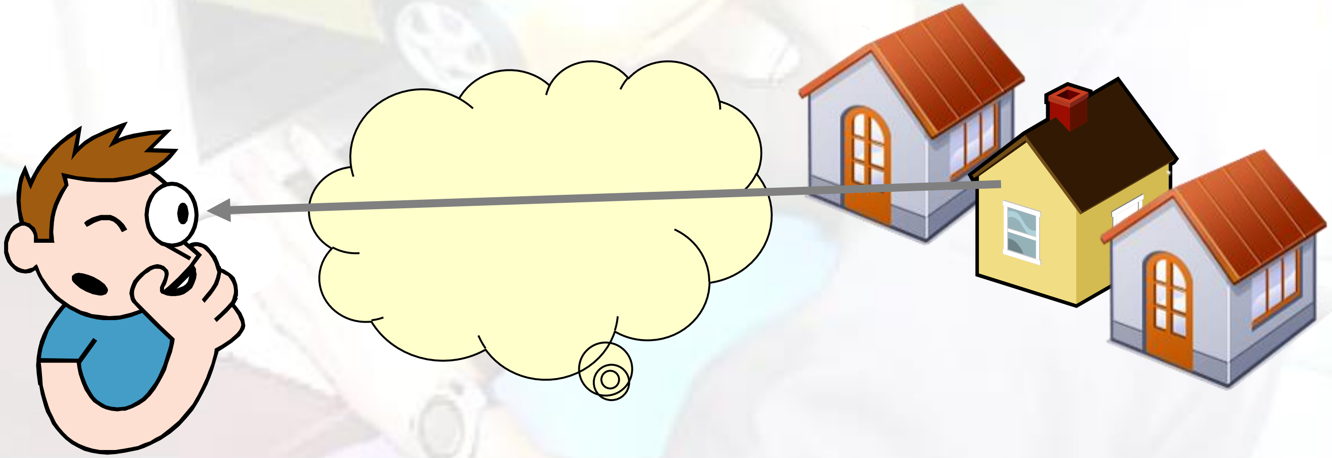


Contents

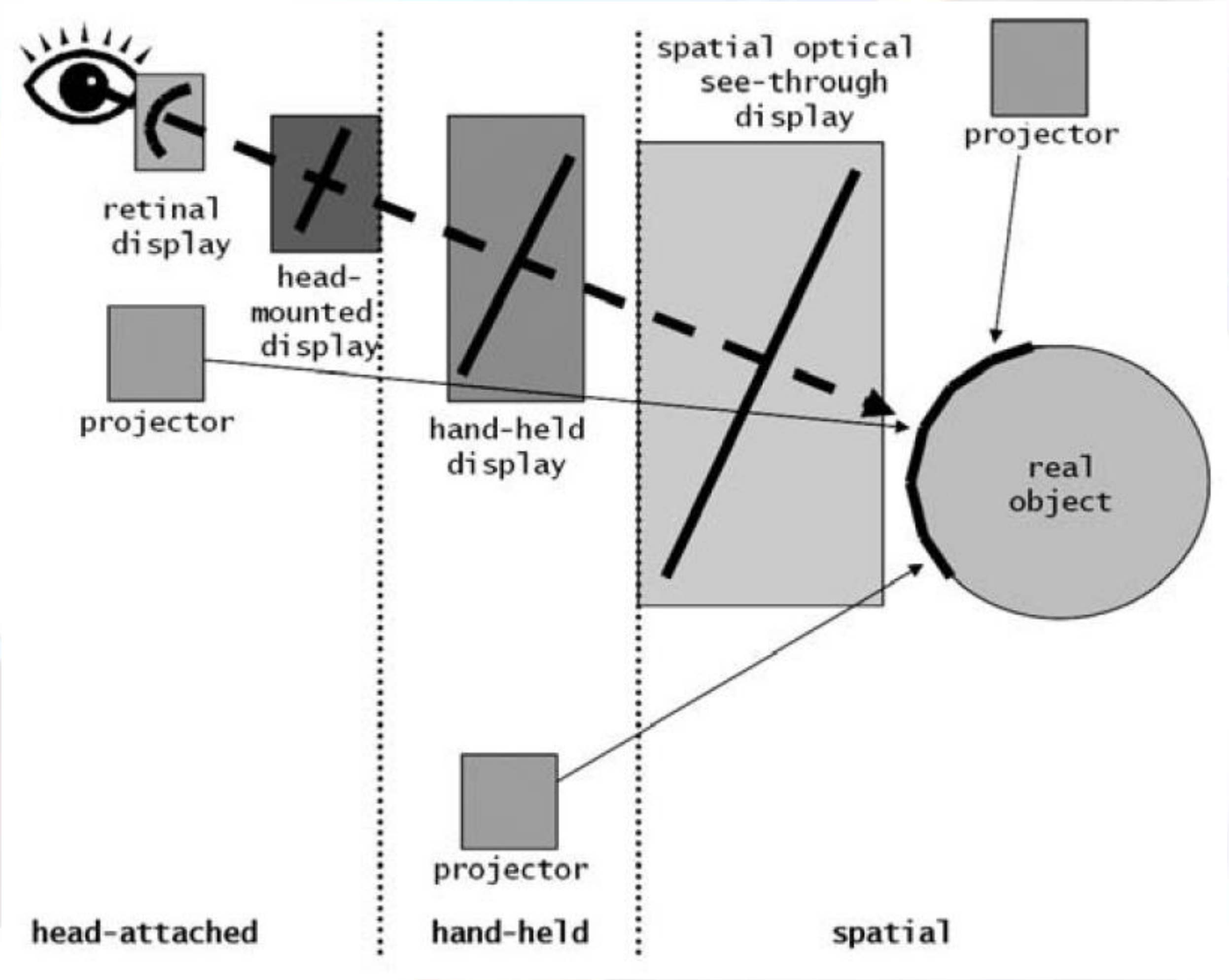
- Introduction
- **Augmented Reality Displays**
- AR interfaces
- Tracking technology
- Research directions

Augmented Reality Displays

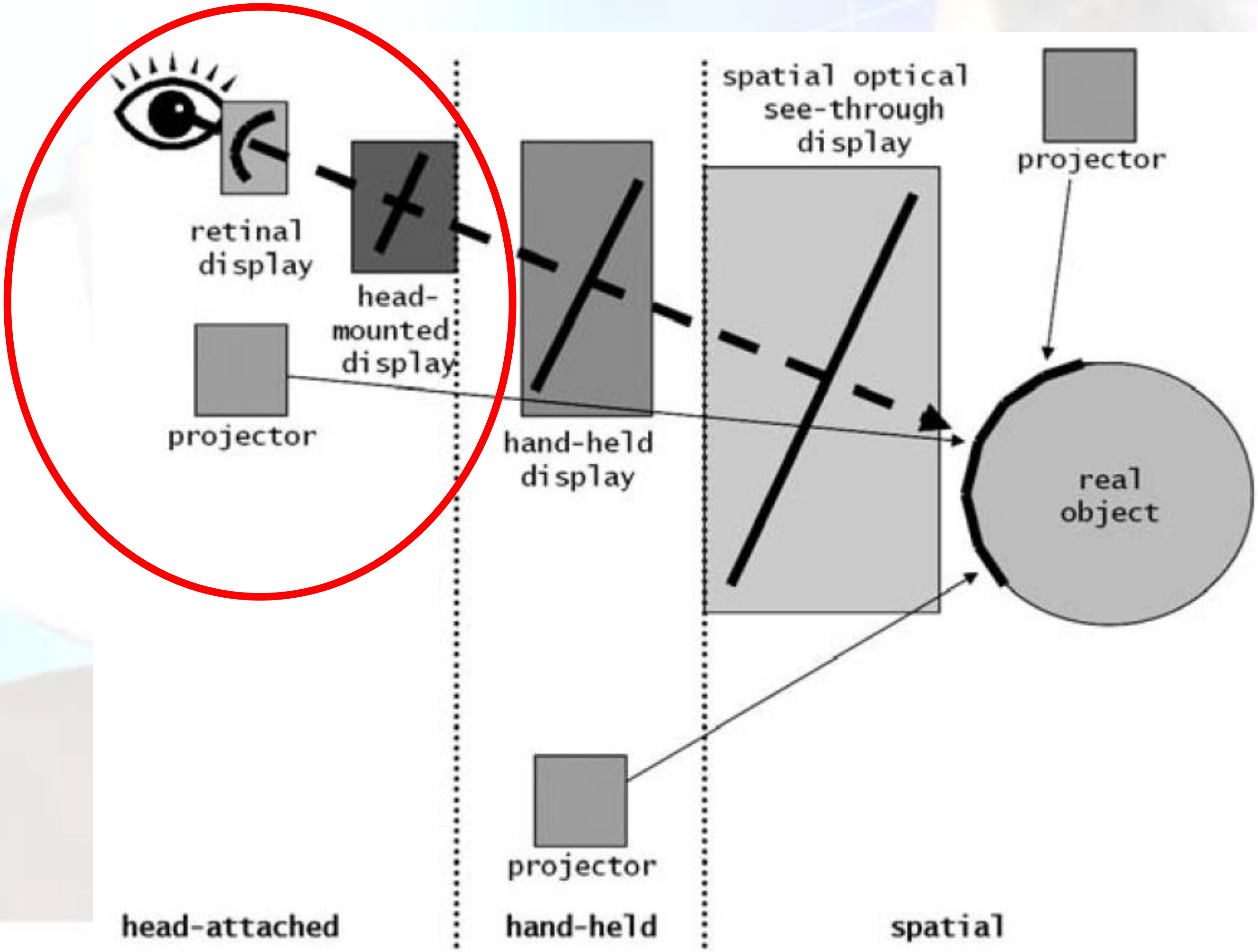
- An AR display uses optical, electronic, and mechanical components to generate images somewhere on the ***optical path*** in between the observer's eyes and the physical object to be augmented.



Choices for image generation



Head-attached Devices

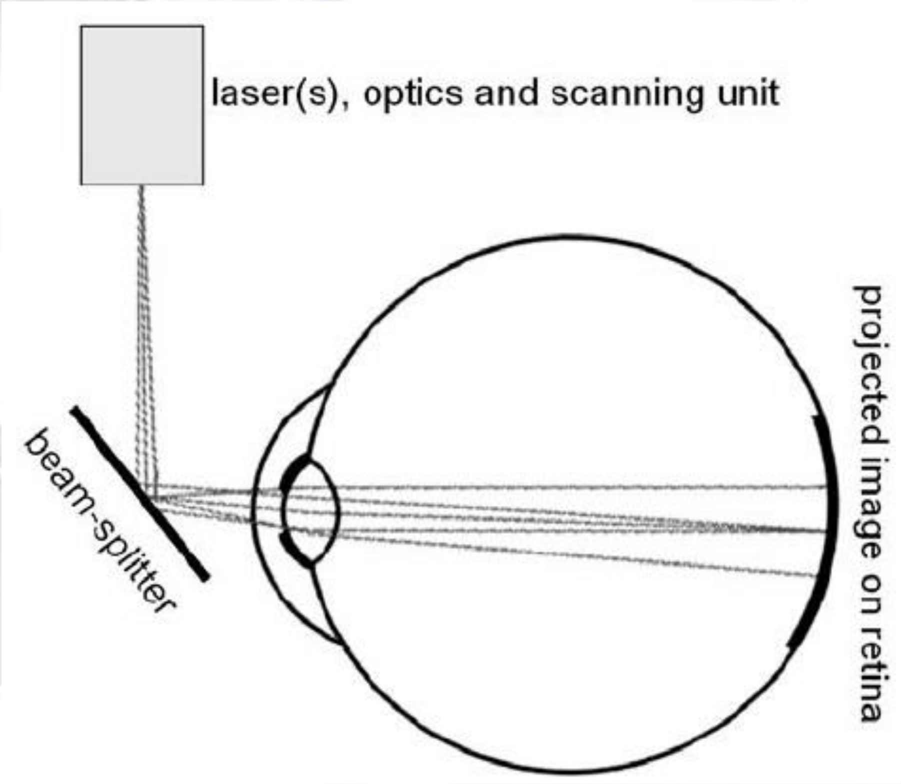


Head-attached Displays

- Head-attached displays require users to wear the display system on their head.
- Three main types:
  - Retinal displays
    - Make use of lasers to project images directly onto the retina.
  - Head-mounted displays
    - Make use of miniature displays in front of the eyes.
  - Head-mounted projectors
    - Make use of miniature projectors that project images on the surfaces of the real world.

Retinal Displays

- Use low-power semiconductor lasers to scan modulated light directly onto the retina.



Retinal Displays

Until 2018 ?

See later Magic Leap

- Pros:
  - Wide field of view
  - High resolution
  - High brightness and contrast
  - Low-power consumption - suitable for mobile outdoor AR
- Cons (existing versions)
  - Mostly monochromatic (no cheap blue and green lasers yet)
  - Stereoscopic versions are expensive.

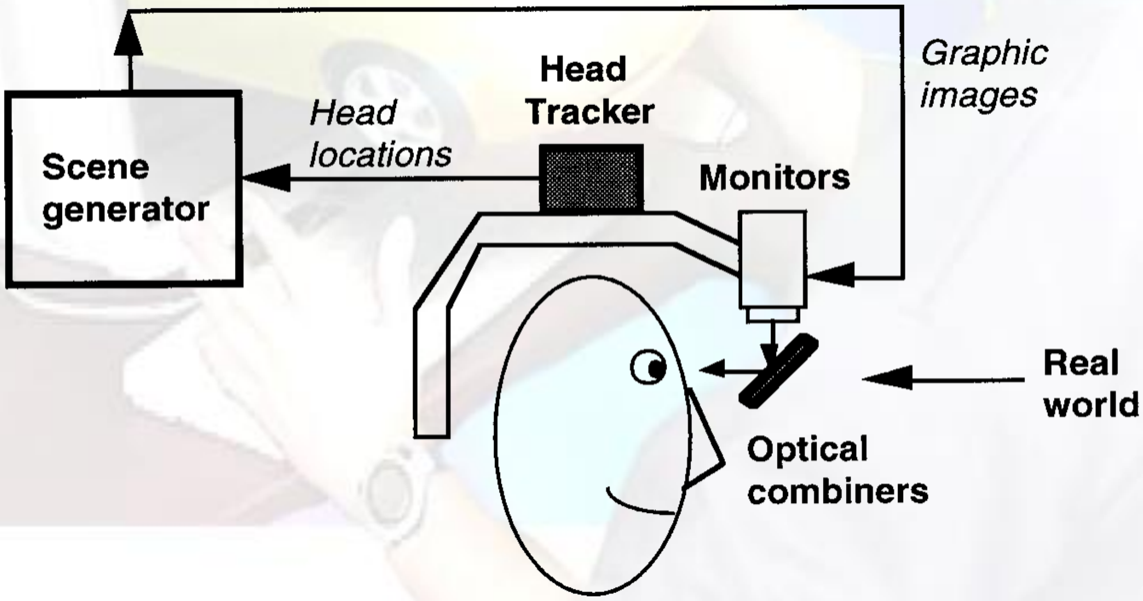
Head-Mounted Displays

Two different HMD technologies to superimpose graphics:

- Optical see-through
  - The user sees the real world **directly**
  - Use optical combiners
- Video see-through
  - The uses sees the real world through a **video camera**
  - Use closed-view HMDs

Optical see-through HMDs

- The user sees the real world **directly**
- Make use of optical combiners:
  - Half-silvered mirrors (partially transparent, partially reflective)
  - Transparent LCD



Optical see-through HMDs

NVIS nVisor ST

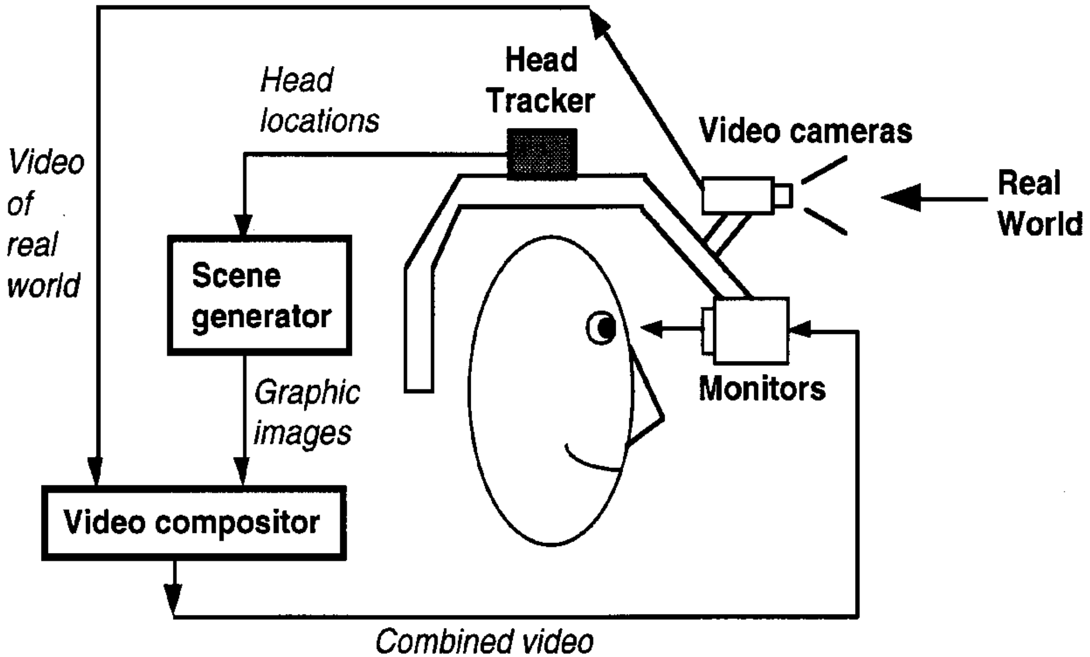
Rockwell Collins ProView XL40

Rockwell Collins Sim Eye XL100A

Optical see-through HMDs

Video see-through HMDs

- Video see-through
  - Use closed-view HMDs.
  - Combine real-time video from head-mounted cameras with virtual imagery.



Video see-through HMDs

Trivisio AR-visio goggles

Head-Mounted Displays

Limitations (common to optical and video-based HMDs):

- Low resolution
  - Optical: real objects OK, synthetic objets low-res
  - Video: both real and synthetic objects low-res
- Limited field of view (limitations of the applied optics)
- Trade-off between ergonomy and image quality (heavy optics)
- Discomfort due to simulator sickness (especially during fast head movements).

Optical vs video see-through

Fixed focal length problem:

- **Video see-through**: real and virtual objects focused at the same distance.
- **Optical see-through:** real objects and virtual objects are sensed at different depths → the eyes are forced to either **continuously shift focus** between the different depth levels, or perceive one level as unsharp.

Calibration:

- **Video see-through**: graphics can be integrated on a pixel-precise basis.
- **Optical see-through:** require **difficult calibration** (user- and session-dependent) and precise head tracking to ensure a correct overlay.

Occlusion effects between real and virtual objects:

- **Video see-through:** well supported
- **Optical see-through**: **incapable of providing consistent occlusion effects**. To solve this problem, Kiyokawa et al. [79] use additional LCD panels to selectively block the incoming light from real objects.

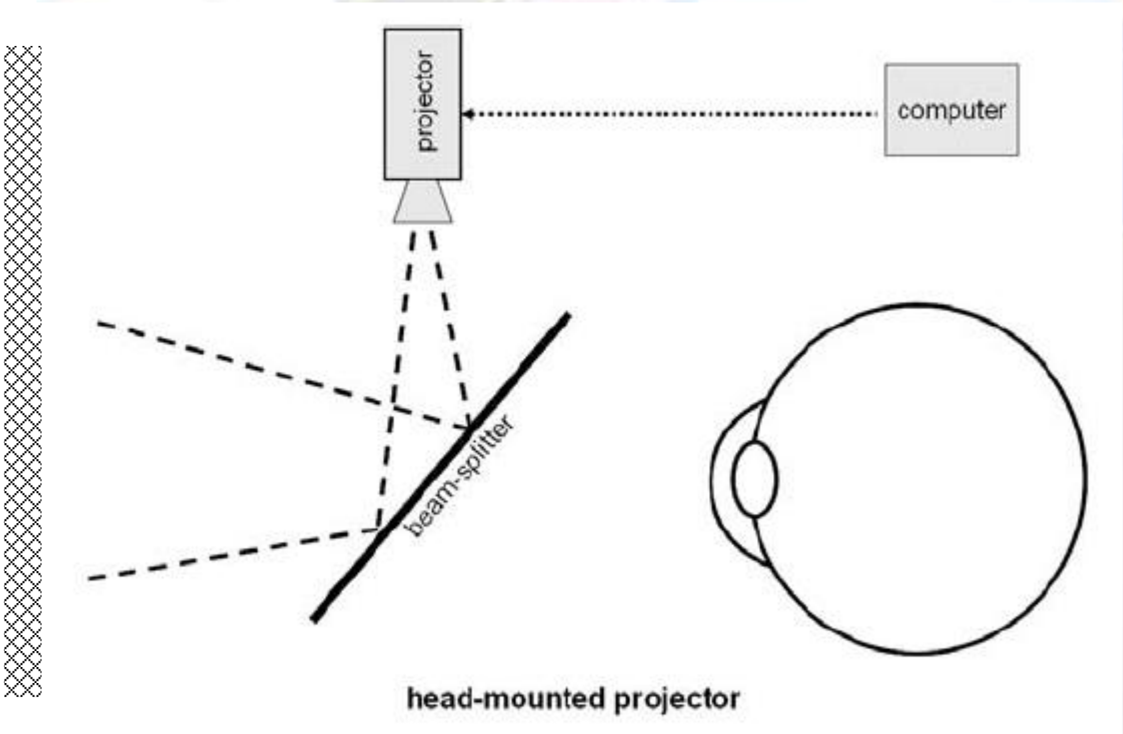
Head-Mounted Projectors

Two types of Head-Mounted Projectors:

- Head-Mounted Projective Displays (HMPDs)
  - Project onto **retro-reflective** surfaces in front of the viewer.
- Projective Head-Mounted Displays (PHMDs)
  - Project onto **diffuse** surface

Head-Mounted Projectors

- Both project images onto surfaces in front of the user.



Head-Mounted Projectors

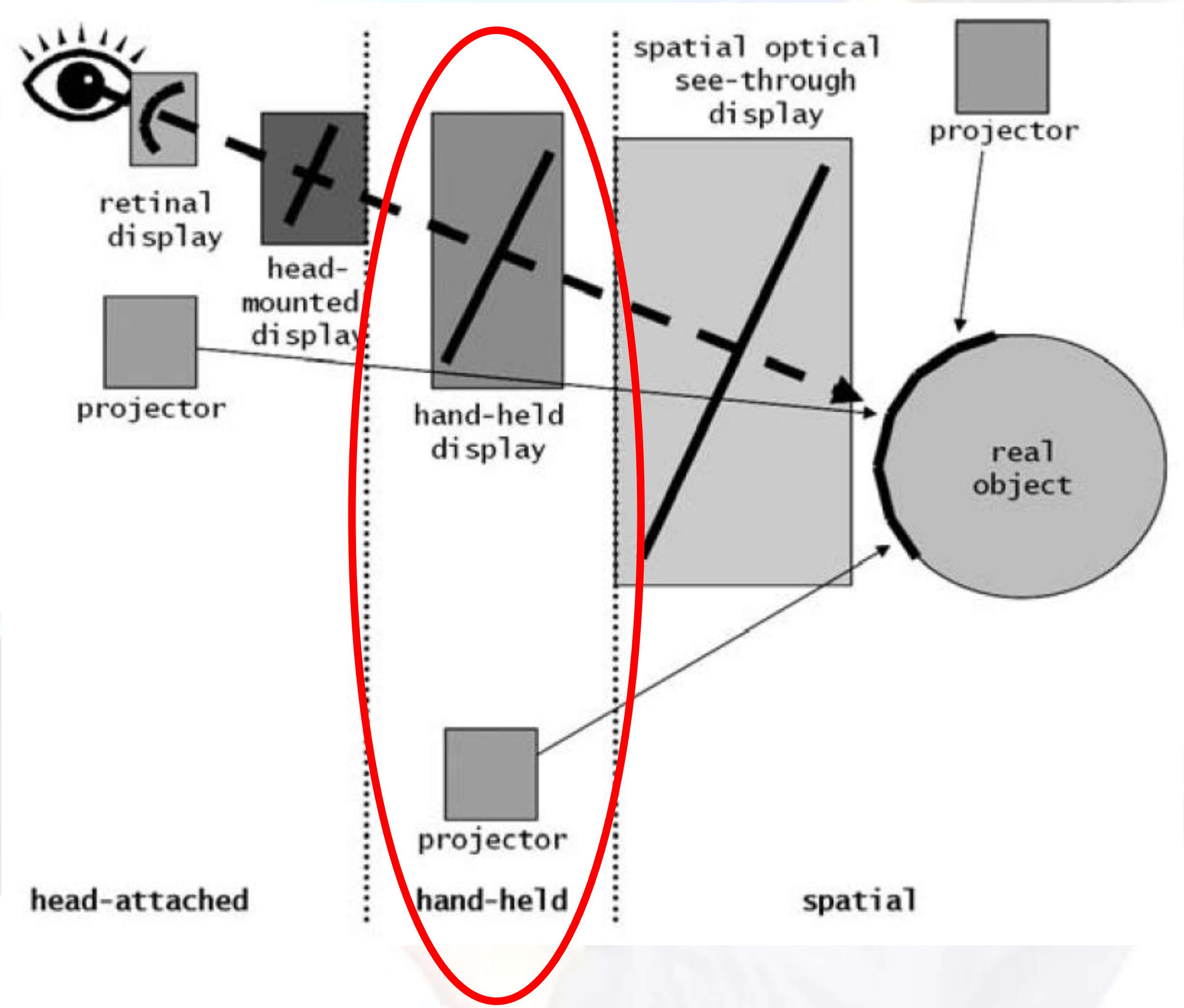
Pros

- Decrease the effect of inconsistency between accommodation and convergence that is related to HMDs
- **Larger field of view** without the additional lenses

Cons:

- Miniature projectors/LCDs offer limited resolution and brightness
- When **retro-reflective** surfaces are used:
  - Brighter images and stereo separation (for free)
- When **diffuse** surfaces are used:
  - Brightness depends strongly on the environmental light conditions

Hand-held devices



Hand-held devices

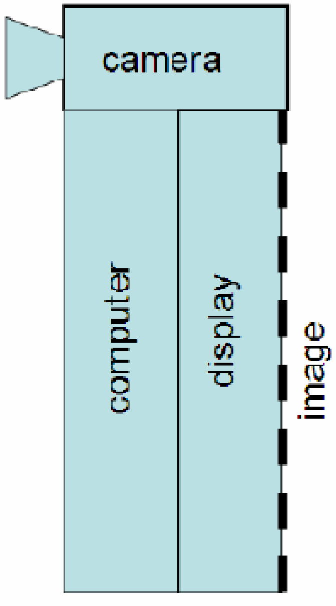
- Hand-held Display
  - Tablet PCs,
  - PDAs
  - Cell phones
- Hand-held Projector

Hand-held Displays

- Examples: Tablet PCs, PDAs, Mobile Phones...
- Suitable for wireless and unconstrained mobile handling.
- Types:
  - Video see-through (preferred approach)
  - Optical see-though

Hand-held Displays

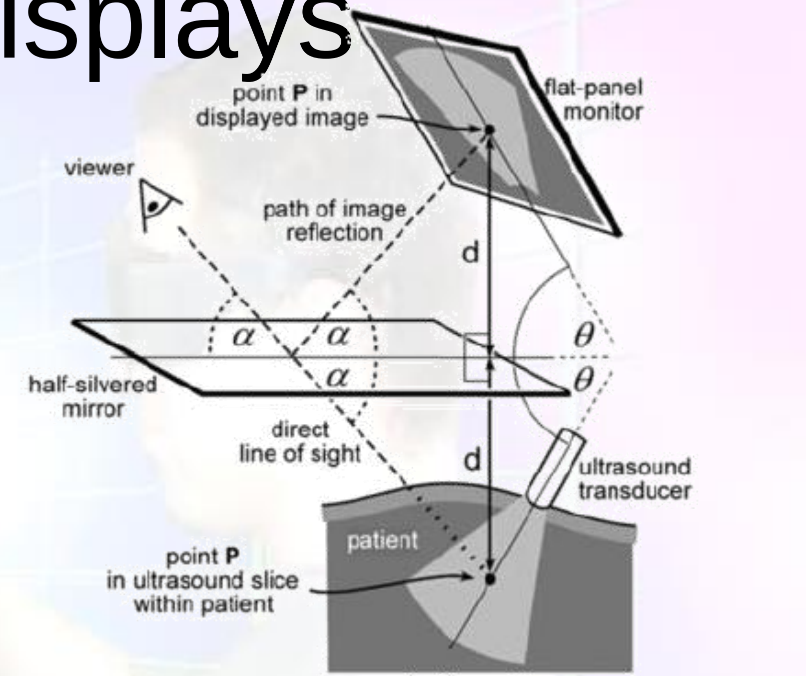
Video see-through

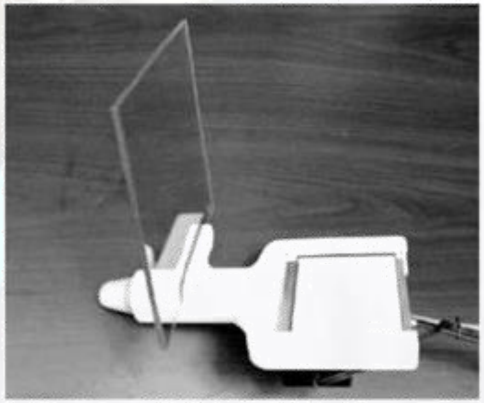


video see-through

Hand-held Displays

Optical see-through [Stetton et al].





Hand-held Displays

Pros:

- Alternative to head-attached devices for mobile applications.
- Consumer devices, such as PDAs and cell phones, have a large potential to bring AR to a mass market.

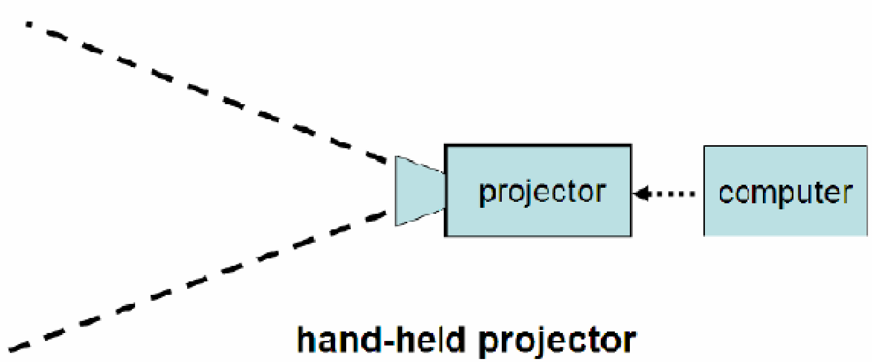
Hand-held Displays

Cons:

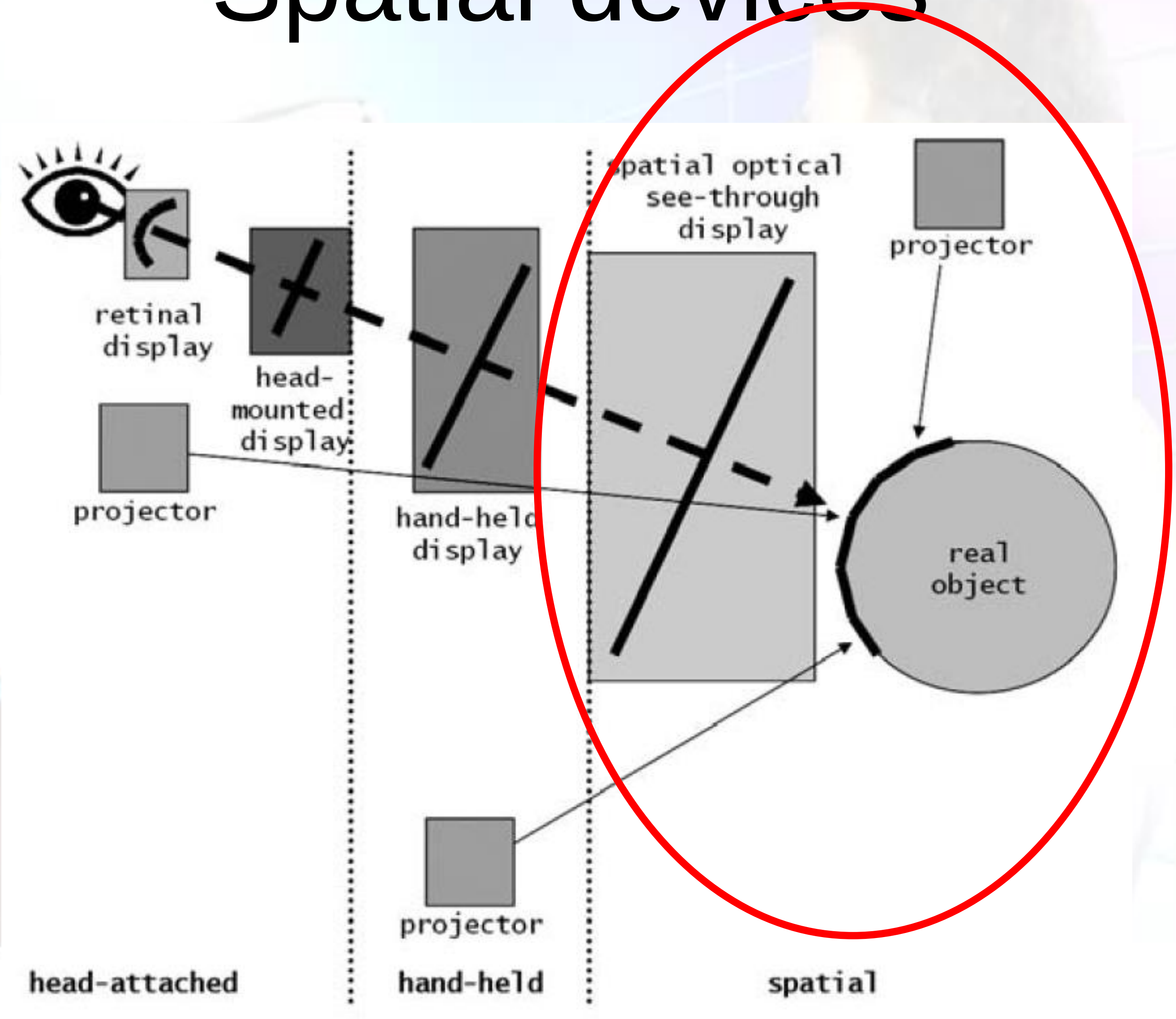
- Limited processing power: low frame rate, high delays.
- Limited screen size: limited field-of-view (but Parks Effect!)
- Parks Effect: moving a scene on a fixed display is not the same as moving a display over a stationary scene (persistence of the image on the retina)
- → If the display can be moved, a larger image of the scene can be left on the retina.
- Integrated cameras very limited (resolution, fixed focus, distorsion)
- Do not provide a completely hands-free working environment.

Hand-held Projectors

- Hand-held projectors can be used to augment the real environment with context sensitive content.



Spatial devices



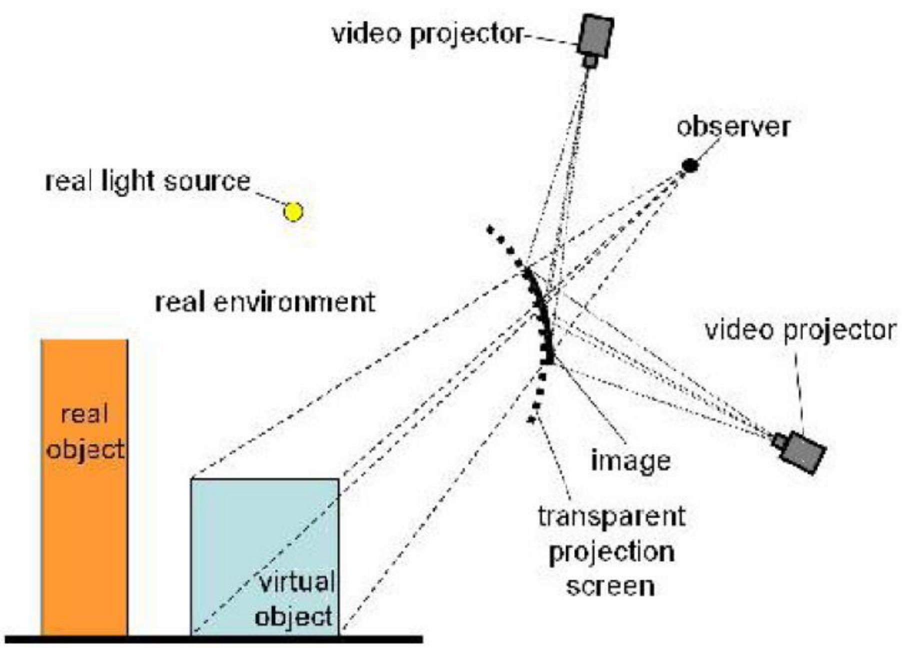
Spatial displays

- In contrast to body-attached displays (head-attached or hand-held), **spatial displays** detach most of the technology from the user and integrate it into the environment.
- Three different approaches:
  – Video see-through
  – Optical see-through
  – Direct augmentation

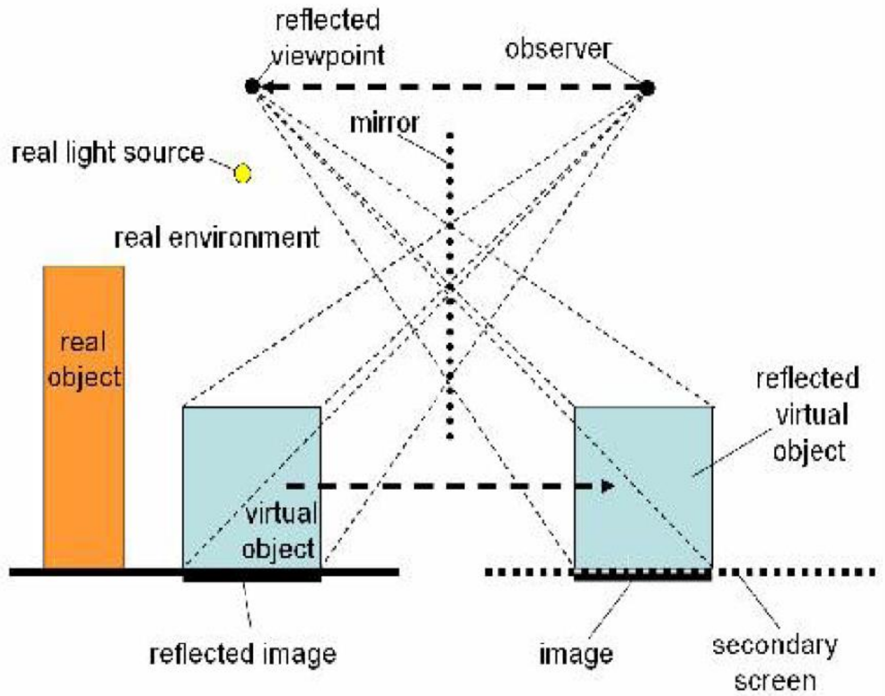
Screen-based video see-through

- Make use of video see-through on a regular monitor.
- Cons:
  - Small field of view (due to relatively small monitor) – but screen-size is scalable if projection screens are applied;
  - Limited resolution of the real environment
  - Mostly provides a remote viewing, rather than supporting a see-through metaphor

Spatial optical see-through



Transparent Screens



Mirror-Beam Combiners

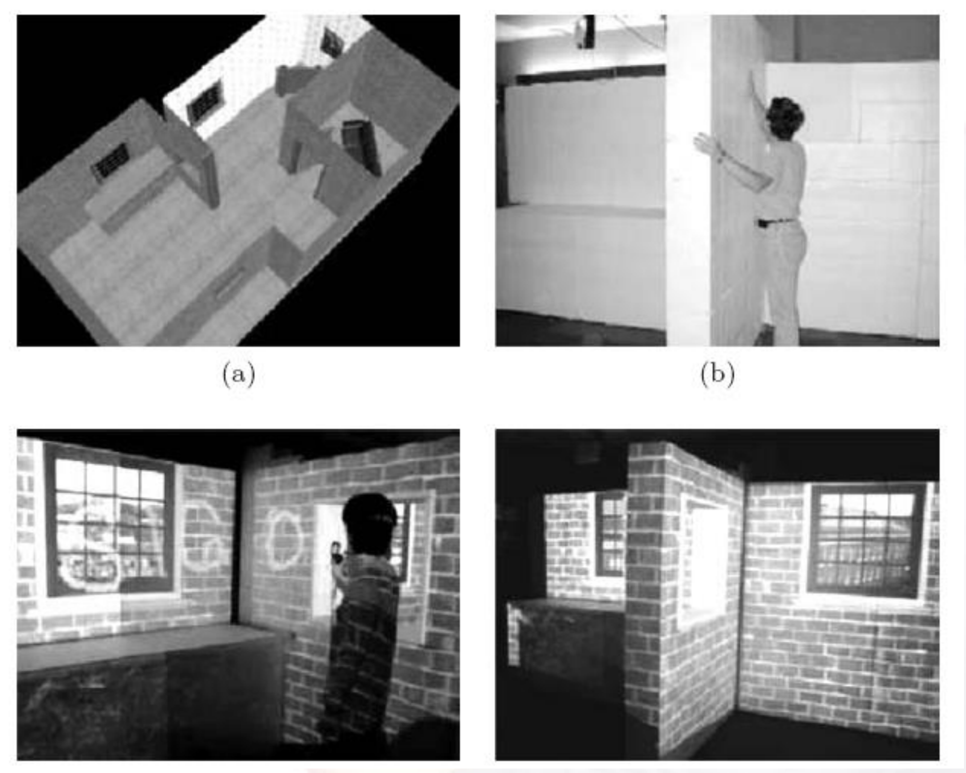
Spatial optical see-through

Cons:

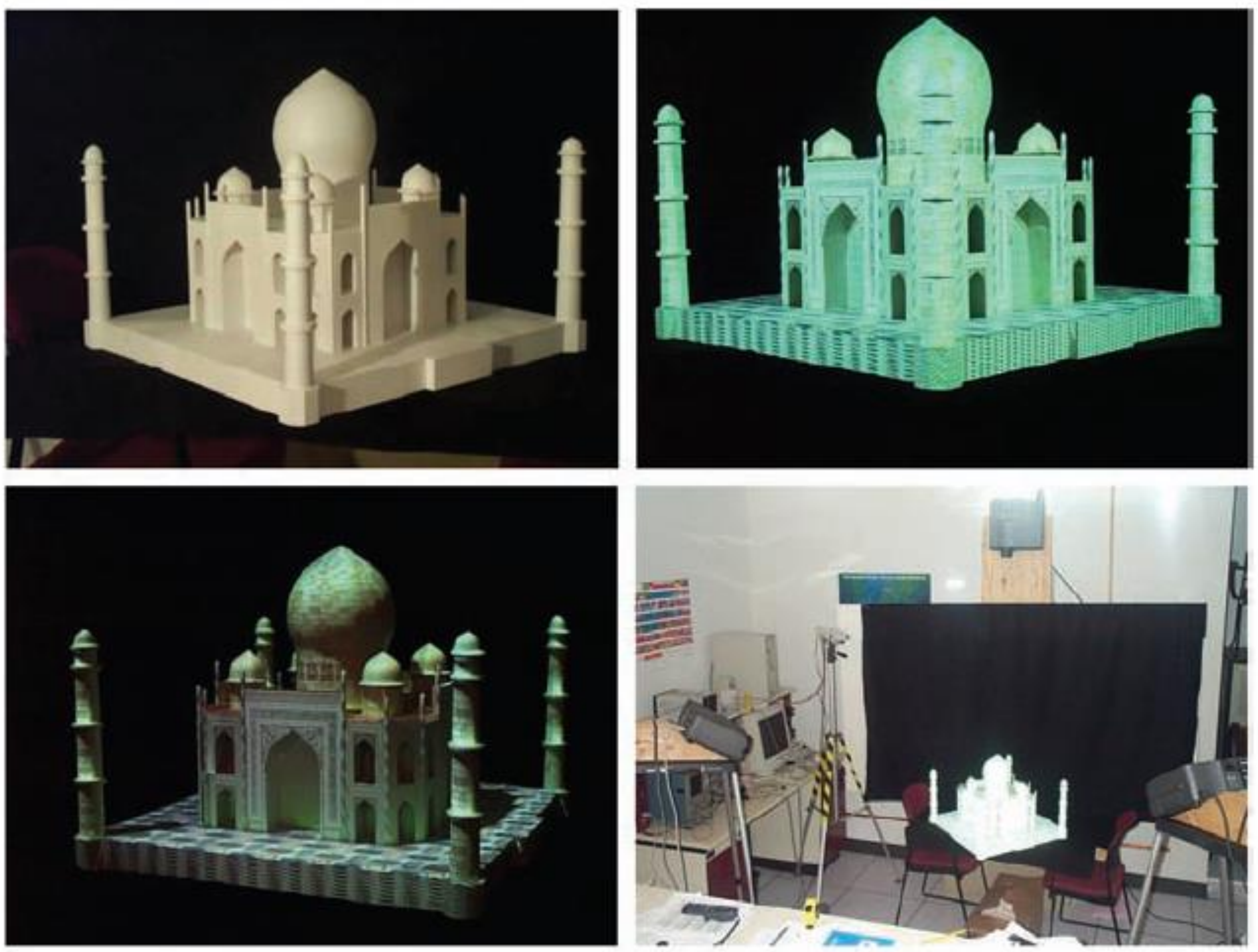
- Do not support mobile applications (spatially aligned optics and display technology).
- Only mirror-beam splitters allow for direct manipulation.
- As in optical see-through HMD, occlusion between real and virtual objects is not supported.

Projection-based spatial displays

- Images are projected directly into physical objects.
- Single static, single steerable or multiple projectors.



Projection-based spatial displays



Projection-based spatial displays

Projection-based spatial displays

- A stereoscopic projection and consequently the technology to separate stereo images is not necessarily required if only the surface properties (e.g., color, texture) of the real objects are changed by overlaying images.

- However, if 3D graphics are displayed in front of the object’s surfaces, a view-dependent, stereoscopic projection is required.

Projection-based spatial displays

Cons:

- Shadow-casting of user’s hands (due to front-projection)
- Display constrained to size, shape, and color of the physical objects
- Conventional projectors focus on a single plane located at a fixed distance. Projecting onto non-planar surfaces causes blur (laser-projectors are OK)
- Complexity of geometric and color calibration increases with the number of projectors

Pros:

- Ergonomics
- Unlimited field of view
- Scalable resolution
- Accommodation (virtual objects are rendered near real location)

Contents

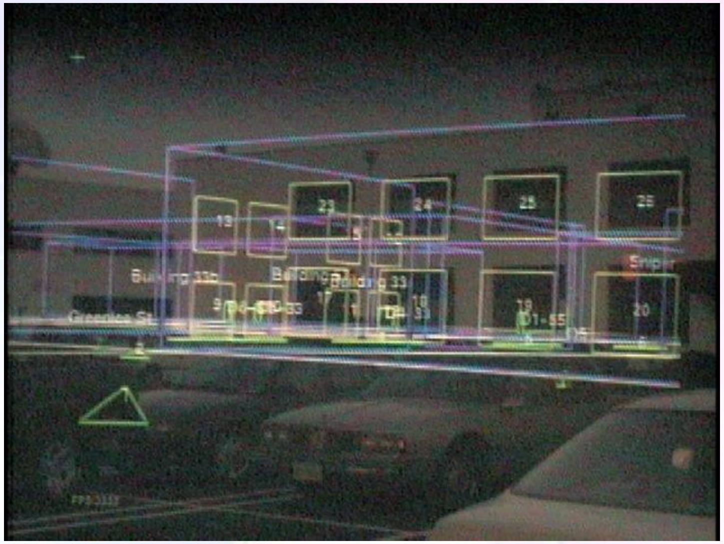
- Introduction
- Augmented Reality Displays
- **AR interfaces**
- Tracking technology
- Research directions

AR interfaces

Classification according to interaction with virtual objects:

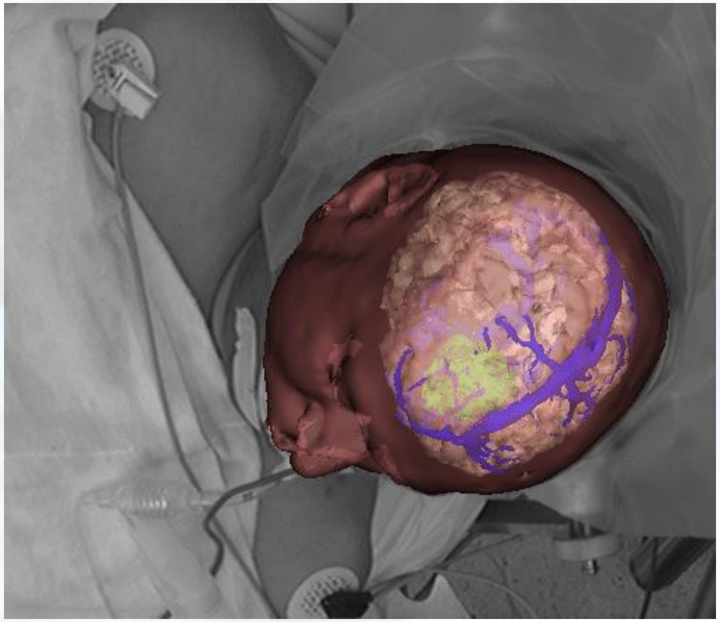
- Little or no interaction: 3D data browsers
- Interaction through 6-DOF devices
- Tangible User Interfaces (TUI)

3D data browsers



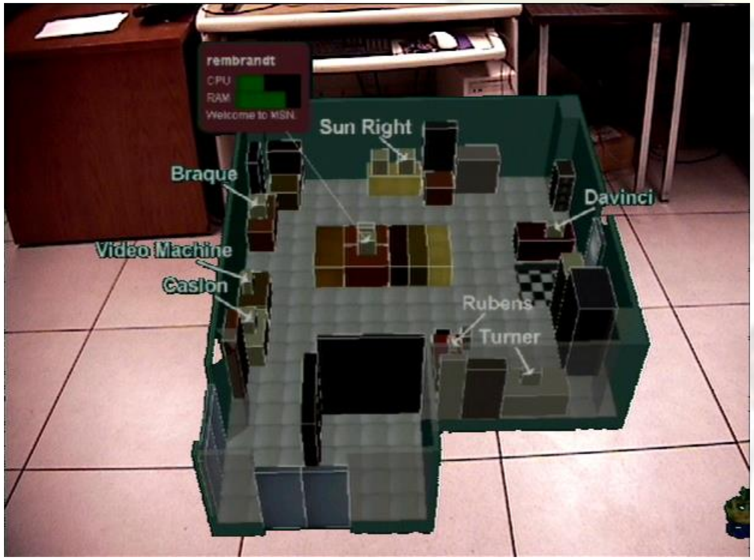
AR interfaces as 3D data browsers:

- One of the early applications.
- User can observe superimposed virtual objects (text, drawings…) but cannot manipulate them (little interaction).
- The main challenge is to correctly register virtual objects with the real-world.



3D data browsers

- Combination with WIM.
- WIM rotates according to the user’s orientation in the real world.

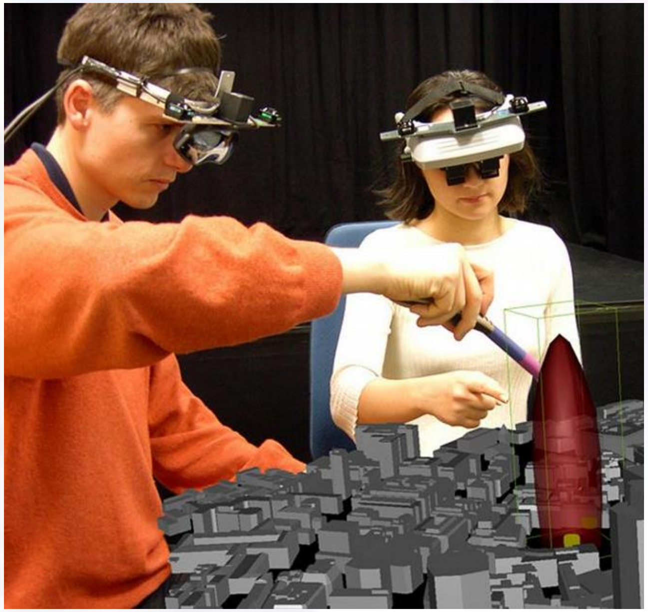


Interaction through 6-DOF devices

- Users can interact with virtual objects using 6-DOF devices.
- Typical interaction tasks: selection, manipulation, system control.

Typical problems:

- Different input modalities for virtual and real objects.
- Tactile feedback



Tangible UI

Tangible user interfaces (TUI):

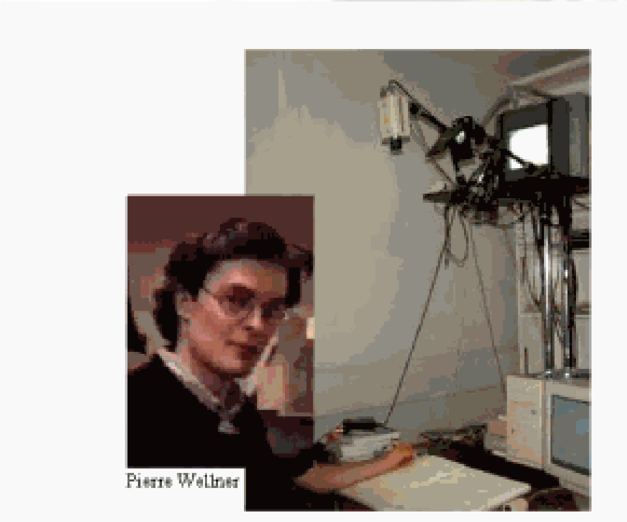
- Use of physical objects to interact with the application.
- Physical objects can be tracked by attaching 2D markers.
- Physical objects used as icons (phicons).

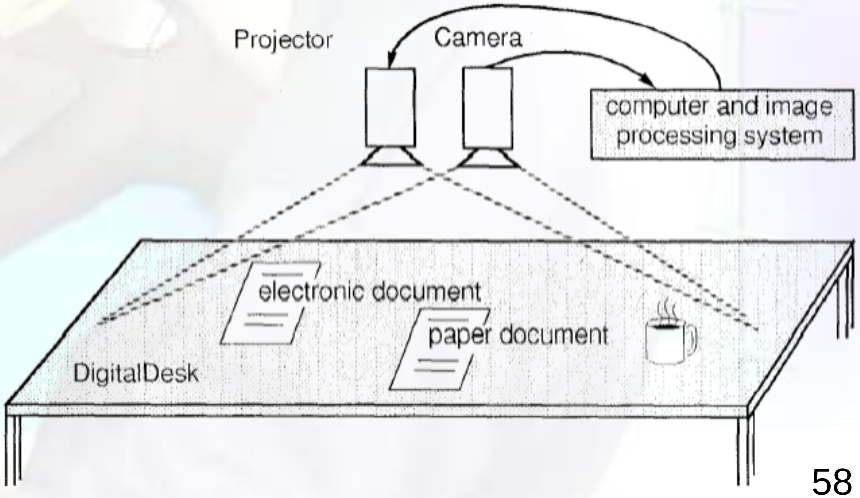
Comparison:

- Interaction through 6-DOF: different devices for physical and virtual interaction.
- TUI: both virtual and real objects are manipulated with the hand.

TUI: Digital Desk

- Digital Desk [Wellner 1993]
- Based on registering virtual objects only to a work surface.
- Overhead projection is used to display virtual objects.

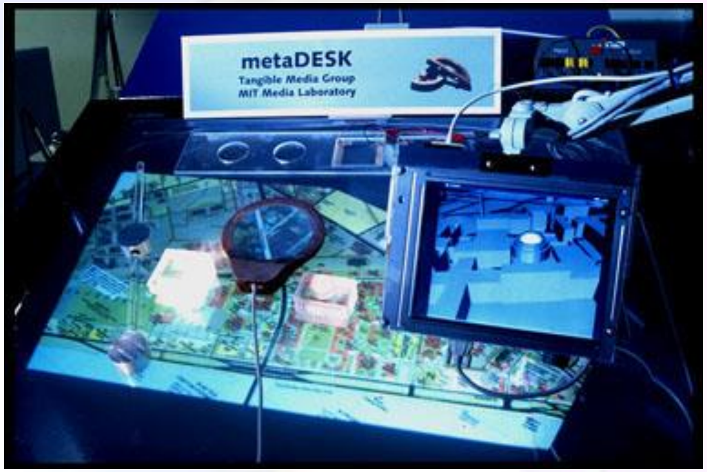




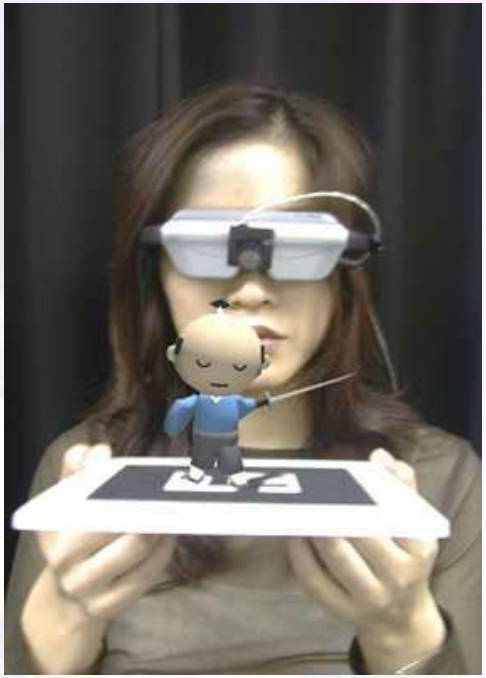
TUI: metaDesk

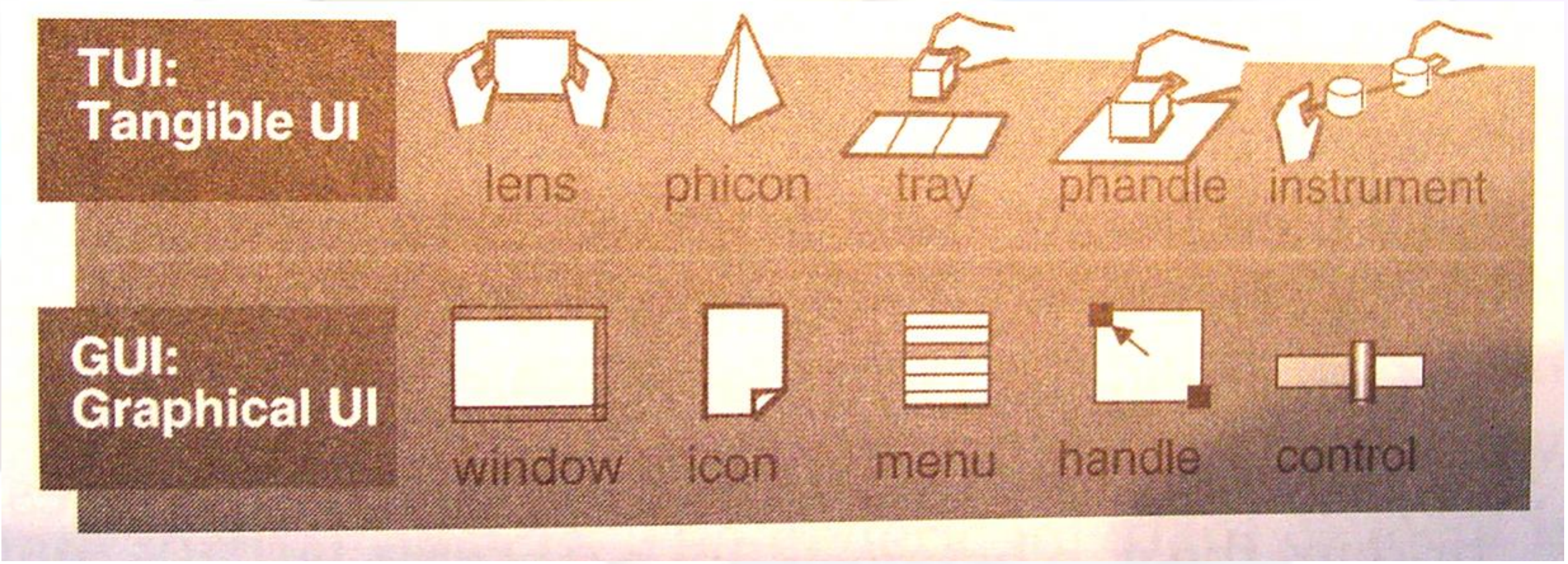
metaDesk [Ullmer and Ishii 1997]

- Images are backprojected on a screen.
- Uses back-projected infrared light for tracking.
- Physical objects reflect infrared light, captured by a camera under the table.



AR interfaces





Contents

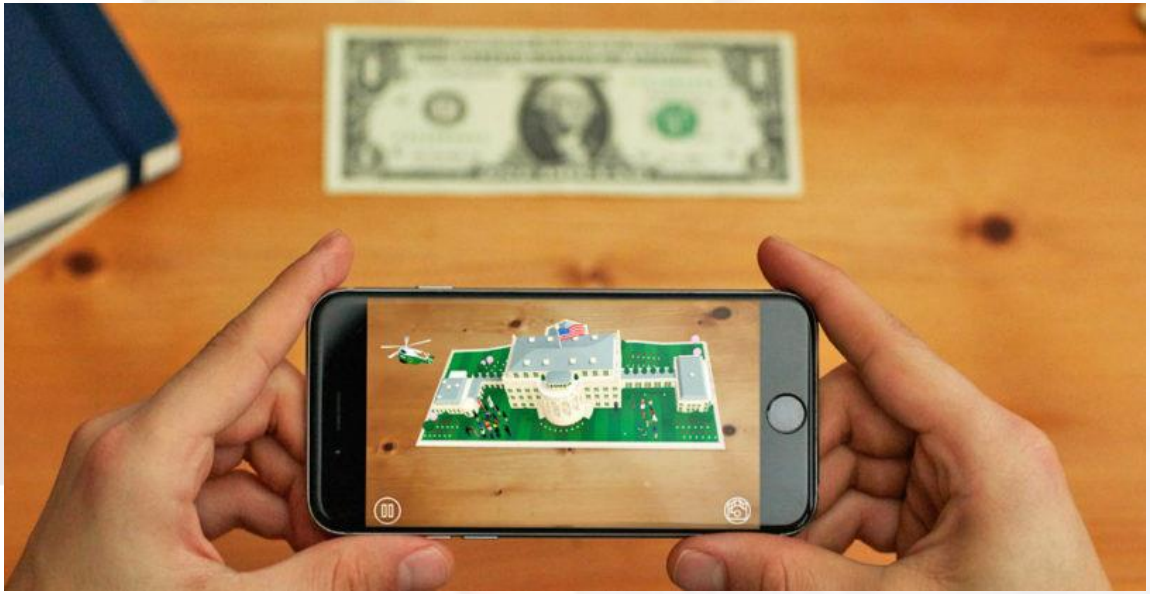
- Introduction
- Augmented Reality Displays
- AR interfaces
- **Tracking technology**
- Research directions

Objects registered in 3D

- Registration
  - Positioning virtual object wrt real world

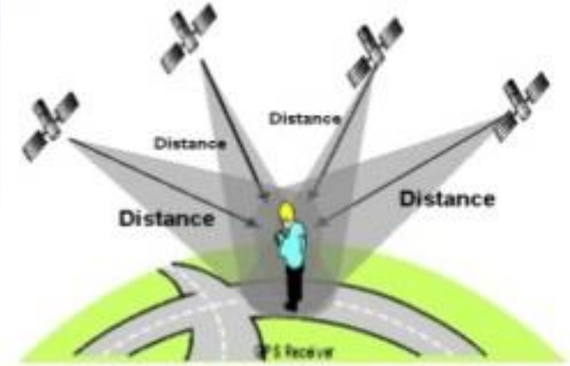
Objects registered in 3D

- Registration
  - Positioning virtual object wrt real world
- Tracking
  - Continually locating the user’s viewpoint
    - Position (x,y,z), Orientation (r,p,y)



Tracking Technologies

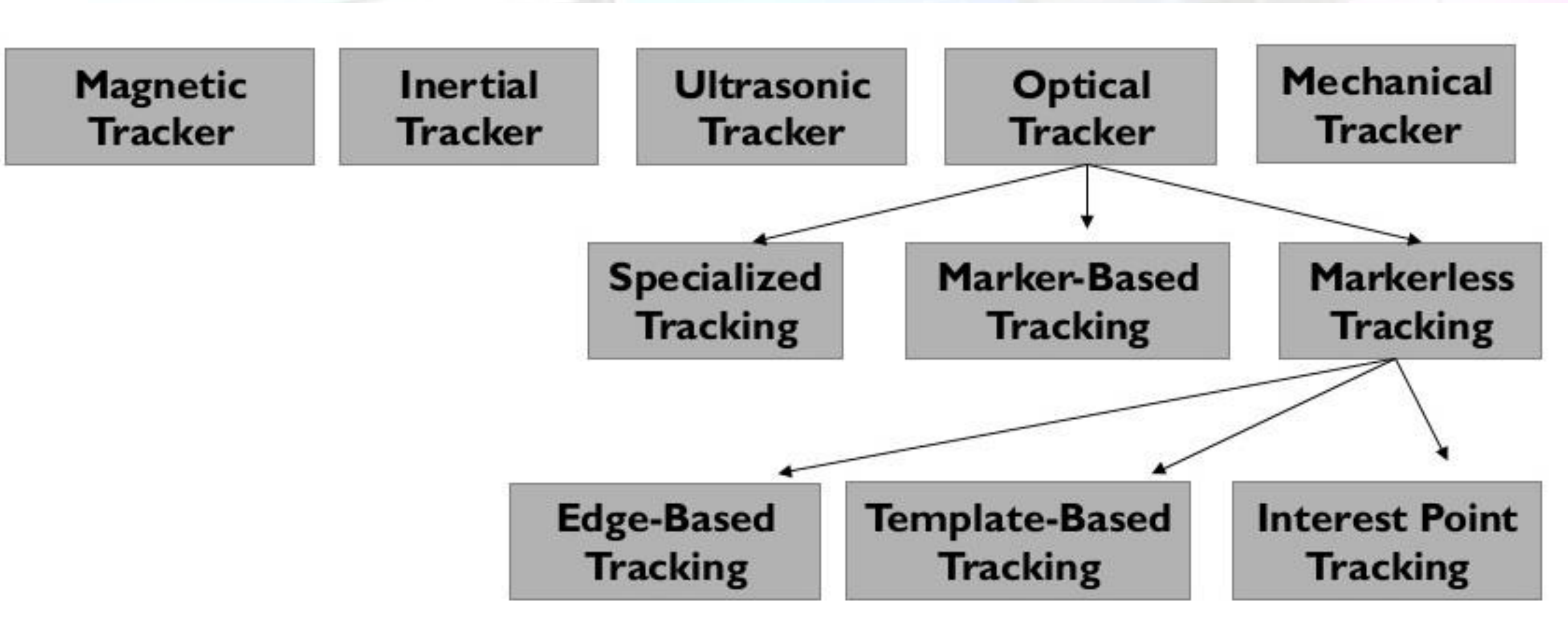
- Active
  - Mechanical, magnetic, ultrasonic
  - GPS, Wifi, cell location



- Passive
  - Inertial sensors (compass, accelerometer, gyro)
  - Computer vision
    - Marker based, Natural feature tracking

- Hybrid tracking
  - Combined sensors (eg Vision + Inertial)

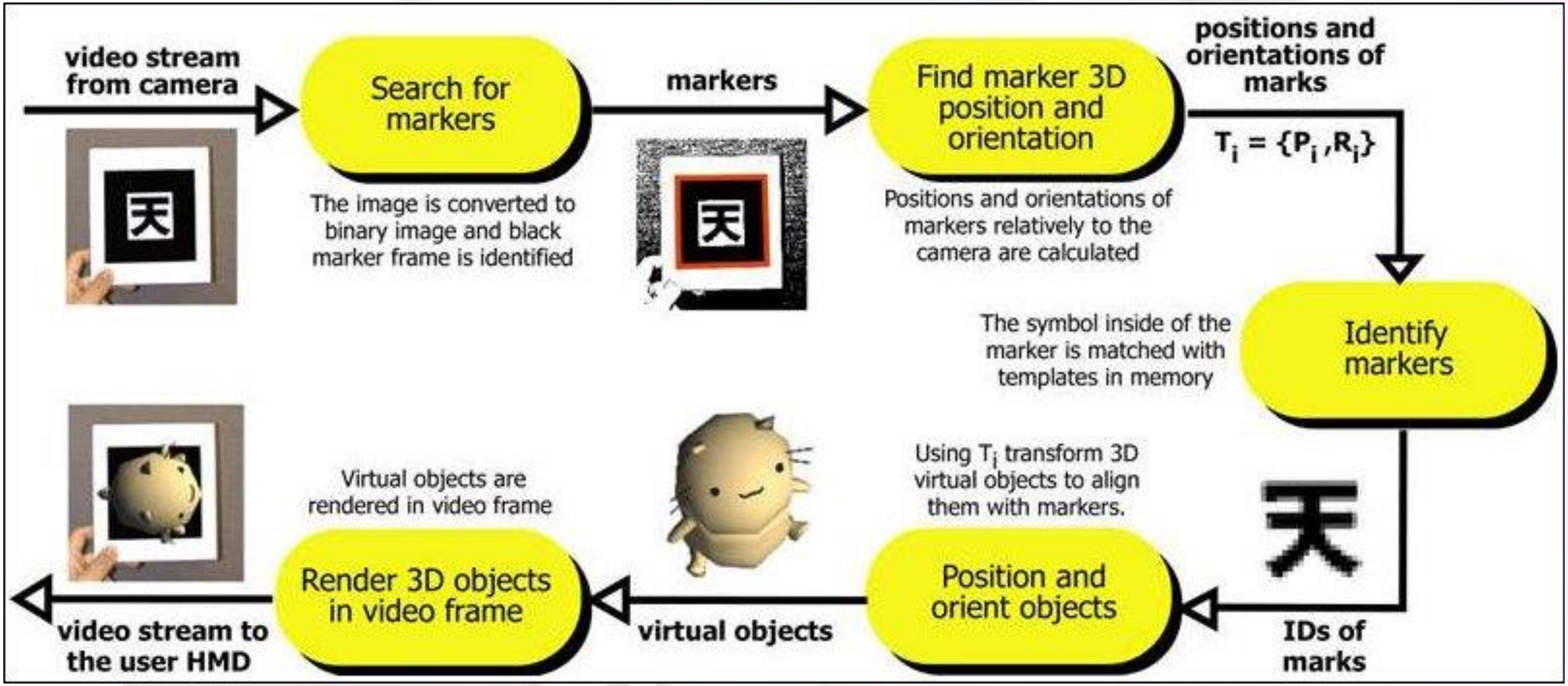
Tracking Types



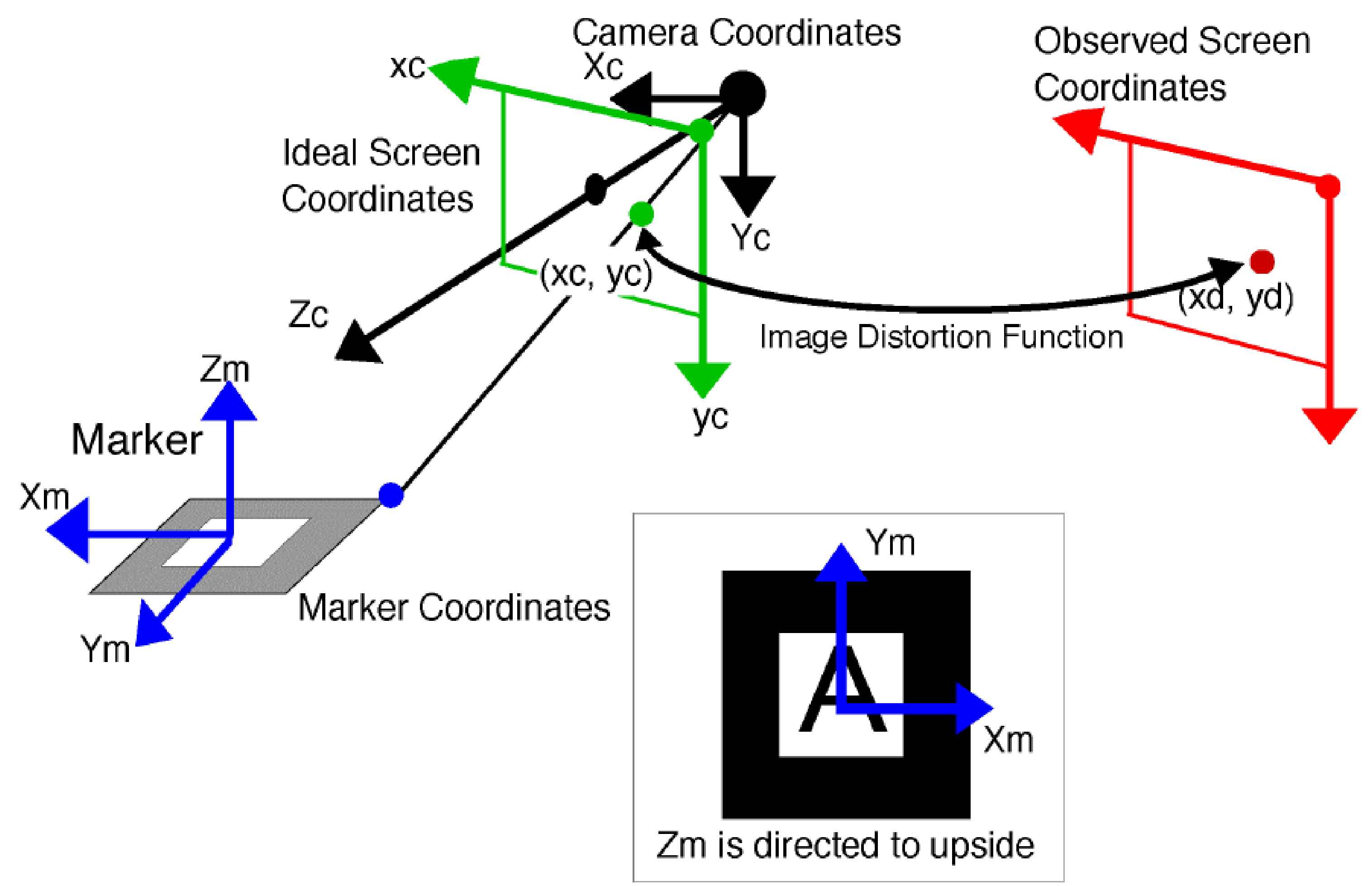
Example: Marker tracking

- Available for more than 10 years
- Several open source solutions exist
  - ARTolkit, ARTag, ATK+, etc
- Fairly simple to implement
  - Standard computer vision methods
- A rectangle provides 4 corner points
  - Enough for pose estimation!

Marker-based tracking: AR-Toolkit

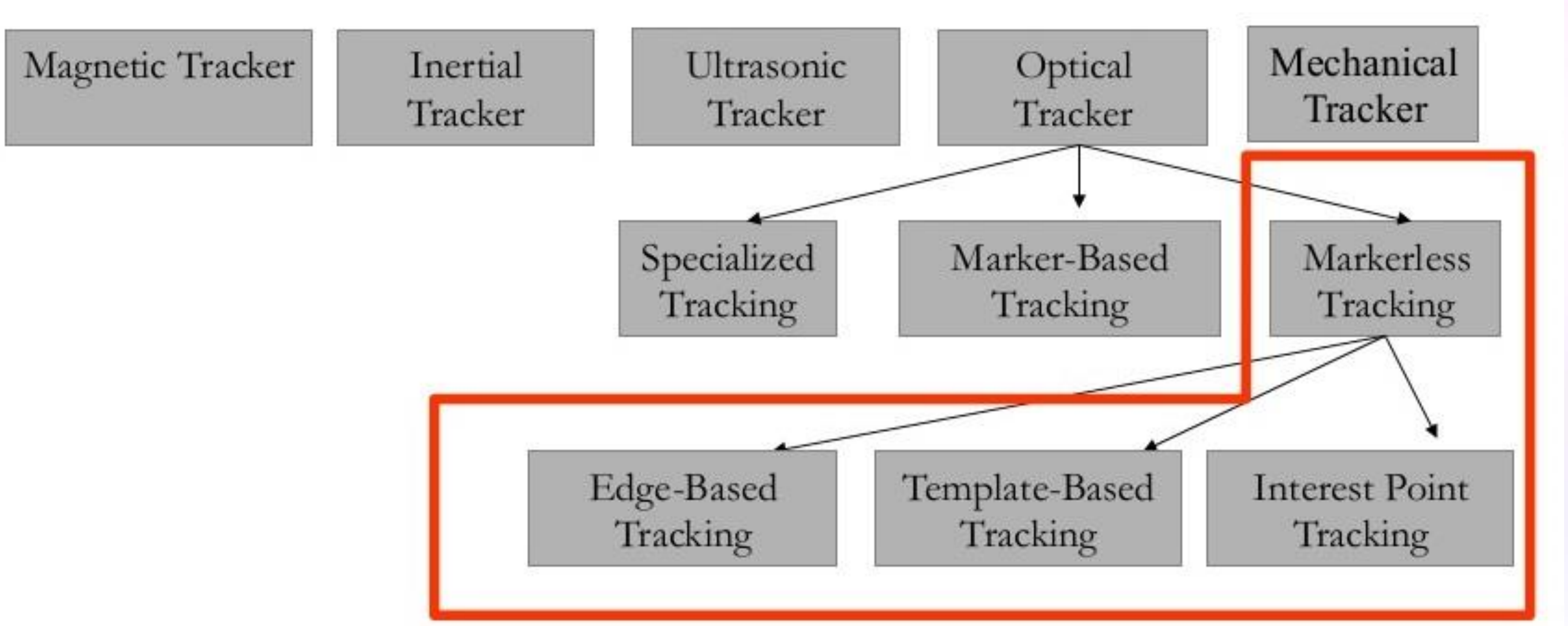


Different coordinate systems



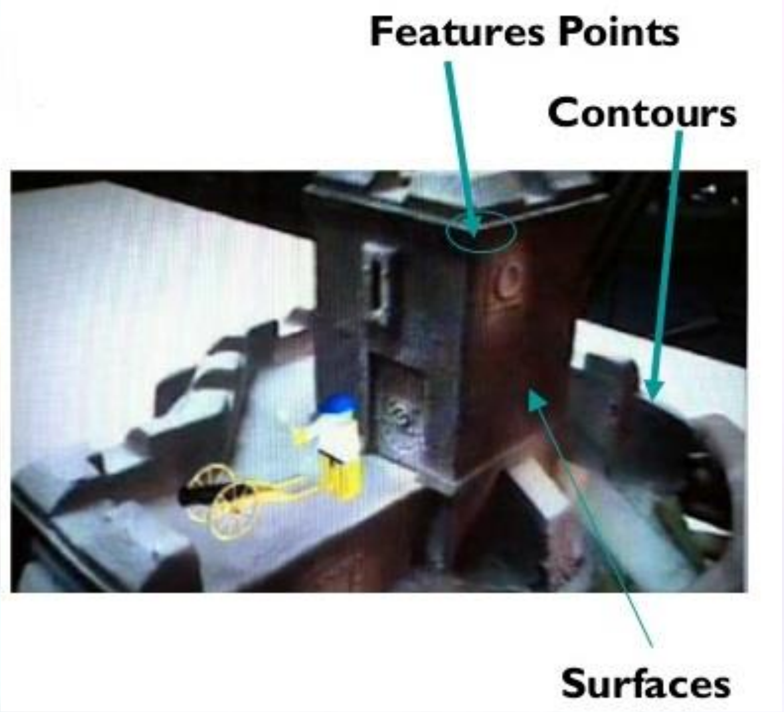
Markerless Tracking

- No more markers! ➔ Markerless Tracking



Natural feature tracking

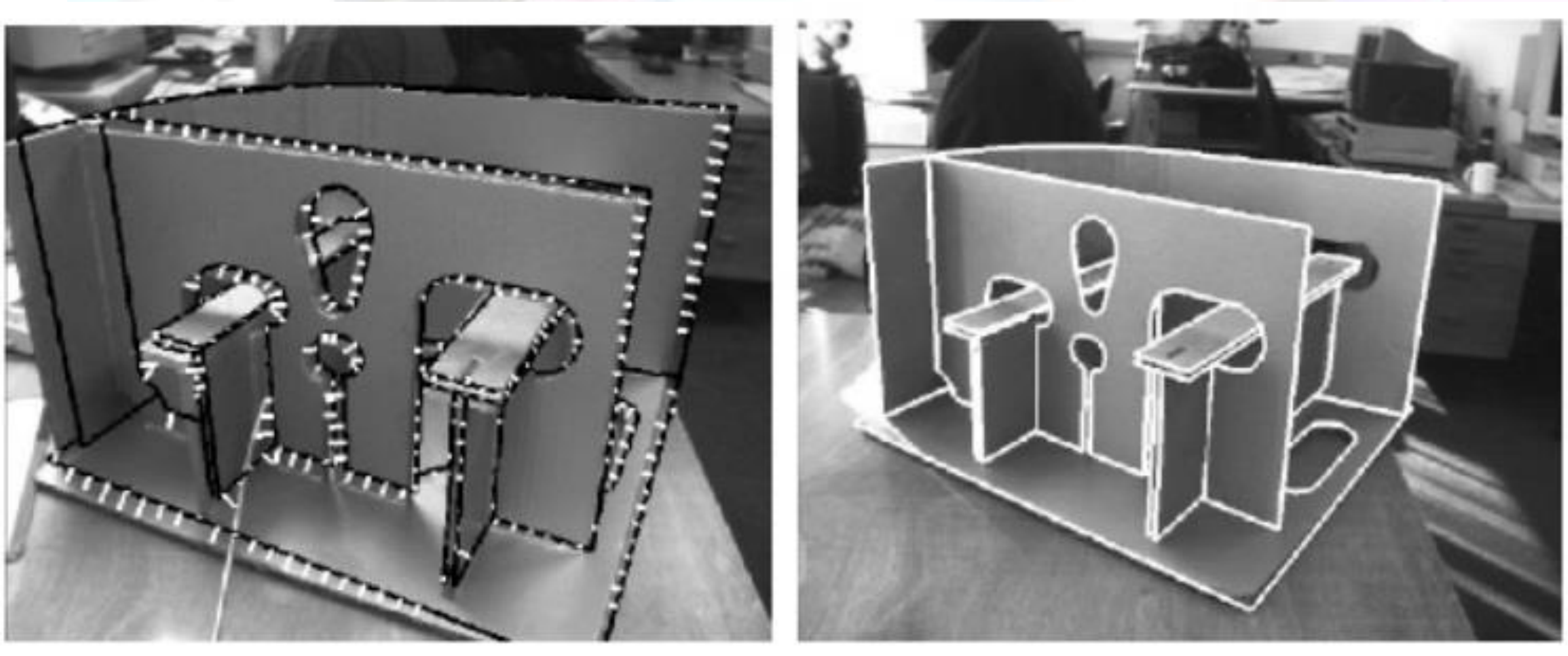
- Use natural cues of real elements
  - Edges
  - Surface texture
  - Interest points
- Model or Model-free
- No visual pollution



Texture tracking

Edge Based tracking

- RAPiD [Drummond et al. 02]
  - Initialization, Control points, Pose prediction (Global method)

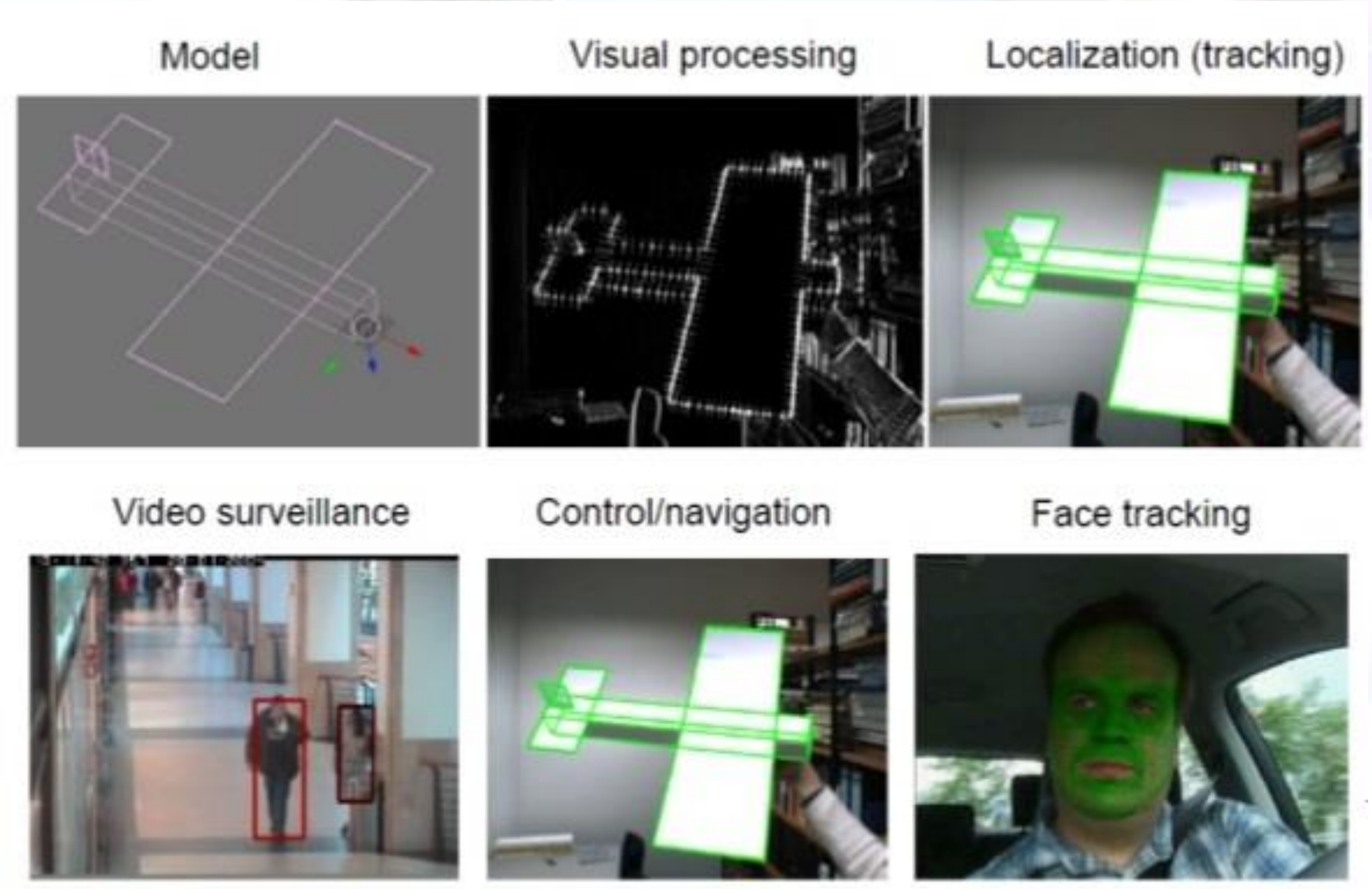


Line Based tracking

- Visual Servoing [Comport et al. 2006]
  - Used in the automated vision for robot's control.

Model Based tracking

- OpenTL – www.opentl.org
  - General purpose library for model based visual tracking

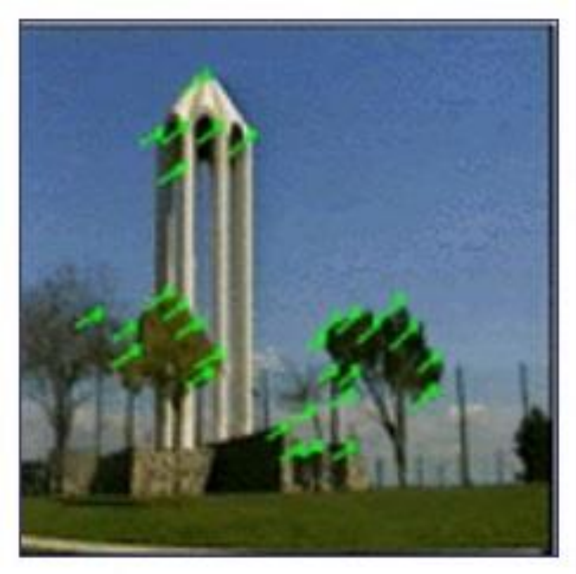


Marker vs. natural feature tracking

- Marker tracking
  - ++ Markers can be an eye catcher
  - ++ Tracking is less demanding
  - - - Environment must be instrumented by markers
  - - - Markers usually work only when fully in view
- Natural feature tracking
  - - - A database of keypoints must be stored/downloaded
  - ++ Natural feature targets might catch the attention less
  - ++ Natural feature targets are potentially everywhere
  - ++ Natural feature targets work also if partially in view

Example: Outdoor hybrid tracking

- Combines
  - Computer vision
    - Natural feature tracking
  - Inertial gyroscope sensors



- Both correct for each other
  - Inertial gyro – provides frame to frame prediction of camera orientation
  - Computer vision – correct for gyro drift

Robust Outdoor tracking

- Hybrid tracking
  - Computer vision, GPS, inertial

Contents

- Introduction
- Augmented Reality Displays
- AR interfaces
- Tracking technology
- **Research directions**

What makes a good AR experience?

- Compelling
  - Engaging, “Magic” moment
- Intuitive ease of use
  - Uses existing skills
- Anchored in physical world
  - Seamless combination of real and digital
    - ➔ **Mixed Reality**

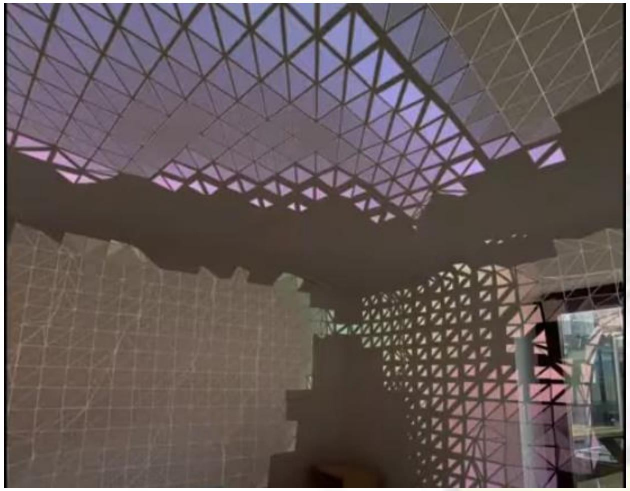
Possible research directions

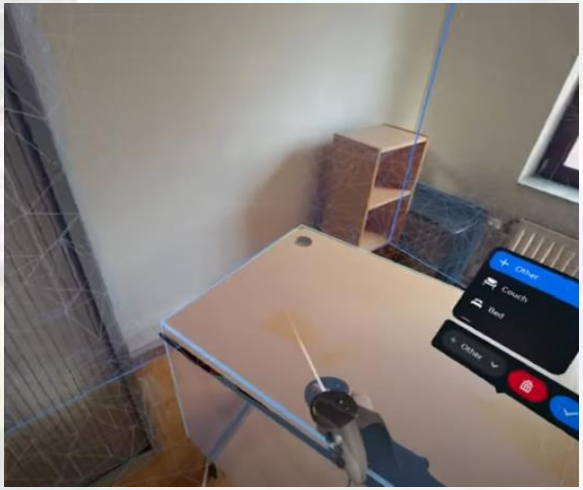
- Tracking
  - Wide area, Reliable indoor, Ubiquitous tracking
- Interaction
  - Intelligent systems, Gesture, Collaborative systems
- Displays
  - Wide FOV, Retinal scanning, Contact lens
- Social acceptance
  - Wearable AR, Handheld AR, Social interactions

Wide area tracking

- Process [Ventura & Hollener 2012]
  - Combine panorama's into point cloud model (offline)
  - Initialize camera tracking from point cloud
  - Update pose by aligning camera image to point cloud
  - Accurate to 25 cm, 0.5 degree over wide area

Recognition of the real world





- Computer vision to recognize reality
- Interaction between real and virtual objects

Gesture based interaction

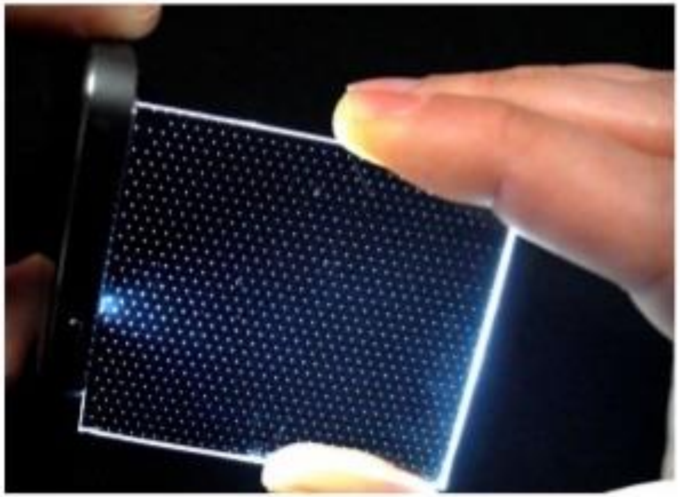
HIT Lab NZ

Microsoft Hololens

- Use free hand gestures to interact
  - Depth camera, scene capture
- Multimodal input
  - Combining speech and gesture

Meta SpaceGlasses

Wide FOV Displays



- Wide FOV see-through display for AR
  - LCD panel + edge light point light sources
  - 110 degree FOV

Social acceptance

- People don’t want to look silly
  - Only 12% of 4.600 adults would be willing to wear AR glasses
  - 20% of mobile AR browser users experience social issues
- Acceptance more due to Social than Technical issues
  - Needs further study (ethnographic, field test, longitudinal)

Google glasses

Google glasses

Objective:

https://www.youtube.com/watch?v=5R1snVxGNVs

By now:

http://www.youtube.com/watch?v=jK3WLILYhQs

Other glasses similar

Epson Moverio

Recon Jet

ODG R-7

Snapchat Spectacle

Power Wolf G1

Vuzix M100

Microsoft Hololens

Mixed Reality:

- Augments the real world with helpful information (like AR)
- Can transport you to a virtual world (like VR)
- Blends holograms with your real world

Microsoft Hololens

Abandoned!

Magic Leap

- Scans the surrounding to know the real environment
- Retinal display using photonics chip to send photons to the retina.
- Includes an eye tracker

Magic Leap

- Scans the surrounding to know the real environment
- Retinal display using photonics chip to send photons to the retina.
- Includes an eye tracker

https://www.youtube.com/watch?v=Opc0F5mkKkg

Meta Quest 3

- Scans the surrounding to know the real environment
- Allows hand tracking

https://www.youtube.com/watch?v=Exu7r2vZpcw

Some projects

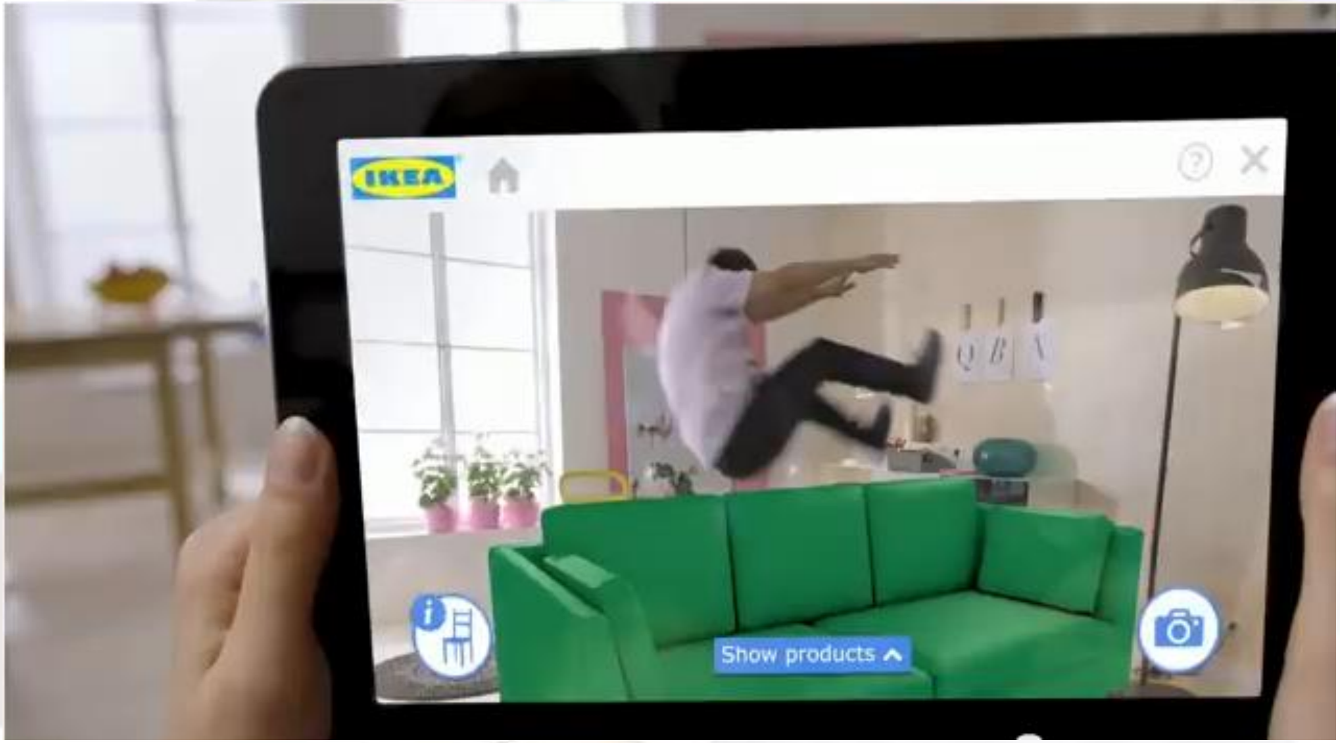
Virtuoso - Studierstube.icg.tu-graz.ac.at

Video: ./HandheldAR_Virtuoso_low.avi

Some projects

IKEA AR

Video: http://www.youtube.com/watch?v=vDNzTasuYEw



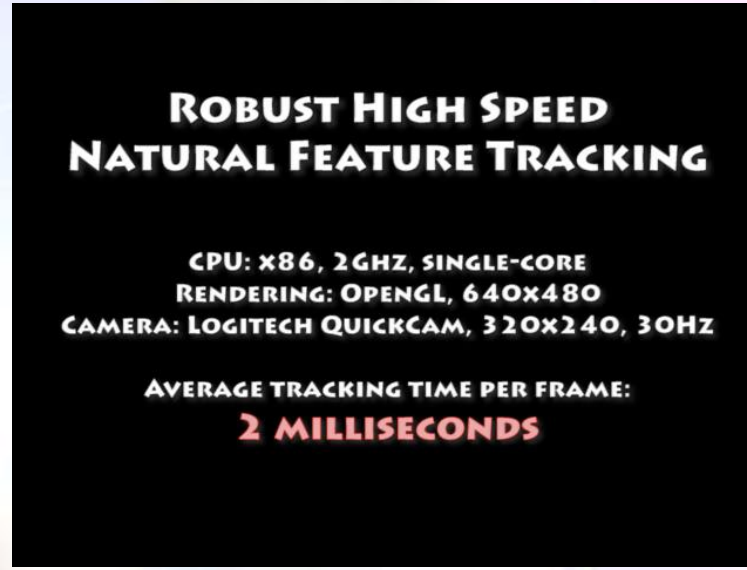
Some projects

Quiver http://quivervision.com/

Other videos

- Robust high speed feature tracking:

  ./RobustHighSpeedTracking_PC_v2.mp4



- https://www.youtube.com/watch?v=HxsN4i2ES6s
- https://www.youtube.com/watch?v=GB_qT6rAPyY
- http://www.youtube.com/watch?v=oH_LfXnklRw

Other references

- [Comport et al 2006]
  - Comport, A.I., Pressigout, M. and Chaumette F. *Real-time Markerless Tracking for Augmented Reality: The Virtual Visual Servoing Framework.* IEEE Trans. On Visualization and Computer Graphics, 12 (4), Aug. 2006.
- [Ventura & Hollener 2012]
  - Ventura, J. and Hollener, T. *Wide-area scene mapping for mobile visual tracking.* In Mixed and Augmented Reality (ISMAR). 2012.
- [Drummond et al. 2002]
  - Drummond, T. and Cipolla, R. *Real-time visual tracking of complex structures.* IEEE Transactions on Pattern Analysis and Machine Intelligence. 24 (7). 2012.

AUGMENTED REALITY

Course 2024/2025