Mobile Information Systems

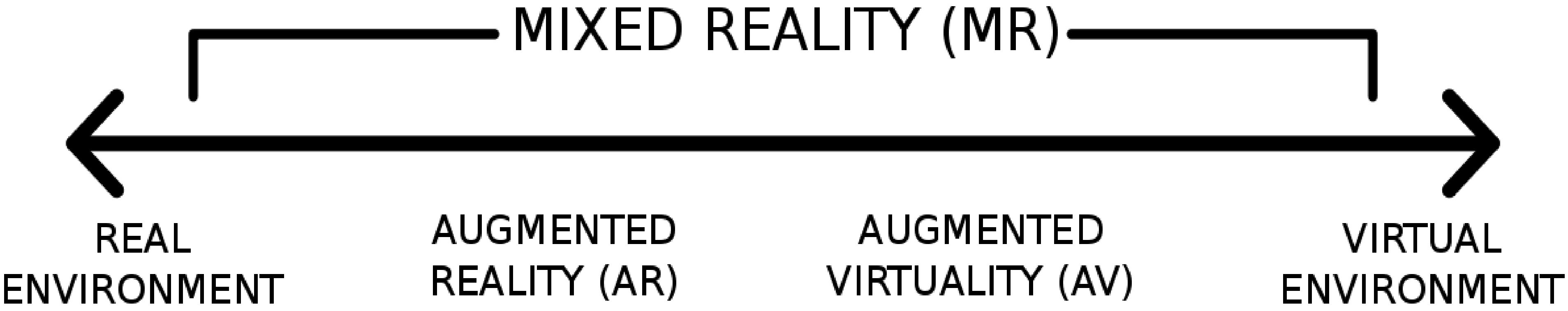
Lecture 09: Mixed Reality

© 2015-20 Dr. Florian Echtler
Bauhaus-Universität Weimar
<florian.echtler@uni-weimar.de>

Augmented Reality (AR)

Image source (CC): https://en.wikipedia.org/wiki/Mixed_reality#/media/File:Reality-Virtuality_Continuum.svg

- Part of the "Virtuality Continuum" (see below)
  - Milgram & Kishino, "A Taxonomy of Mixed Reality Visual Displays", 1994
- Subclass of *mixed reality*
  - Roughly categorized by ratio between real and virtual content



Augmented Reality (AR) (2)

- Definition/characteristics
  - R. Azuma, “A Survey of Augmented Reality”, 1997
    - Mix of *real-world* and *virtual* visual information
    - Information can be *interacted* with in *real-time*
    - Real & virtual elements *spatially aligned* in 3D

→ Requirements:

  - Visual input (real world) & output (virtual)
  - User input & fast graphics (to enable interaction)
  - 6D head tracking (for spatial alignment)

Is this AR?

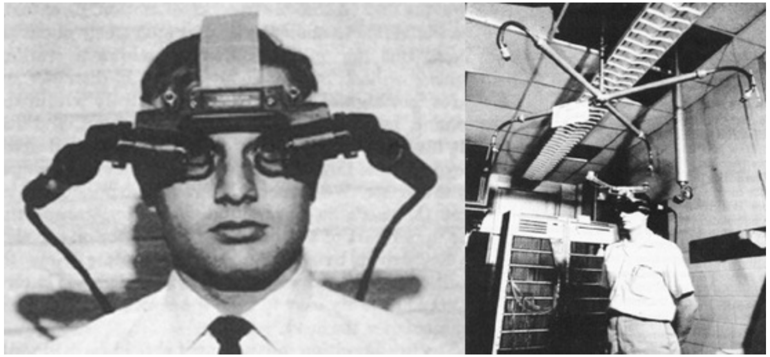
Image source (FU):
http://archive.wired.com/culture/culturereviews/magazine/16-01/found

(apart from the fact that there can be too much of a good thing)

AR: History & Pioneers

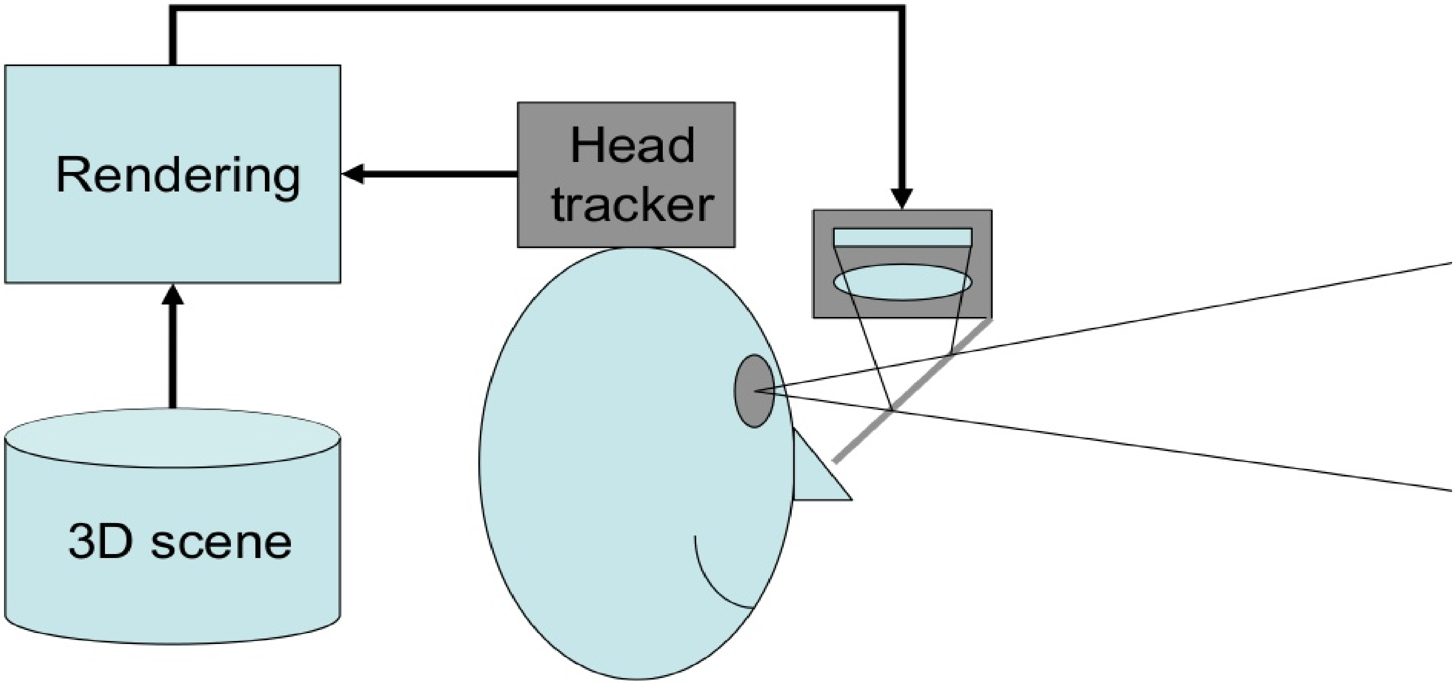
Image sources (FU): https://louislau91.[...]/sword_of_damocles.jpeg, http://eyetap.org/[...]/wearables.html

- 1901: first mention of idea in novel
  (“The Master Key” by L. Frank Baum)
- 1968: first head-mounted display (HMD)
  (“The Sword of Damocles” by Ian Sutherland)
- 1980: first wearable HMD by Steve Mann



HMD: optical see-through (1)

Image source (FU): LMU Lecture by J. Wagner

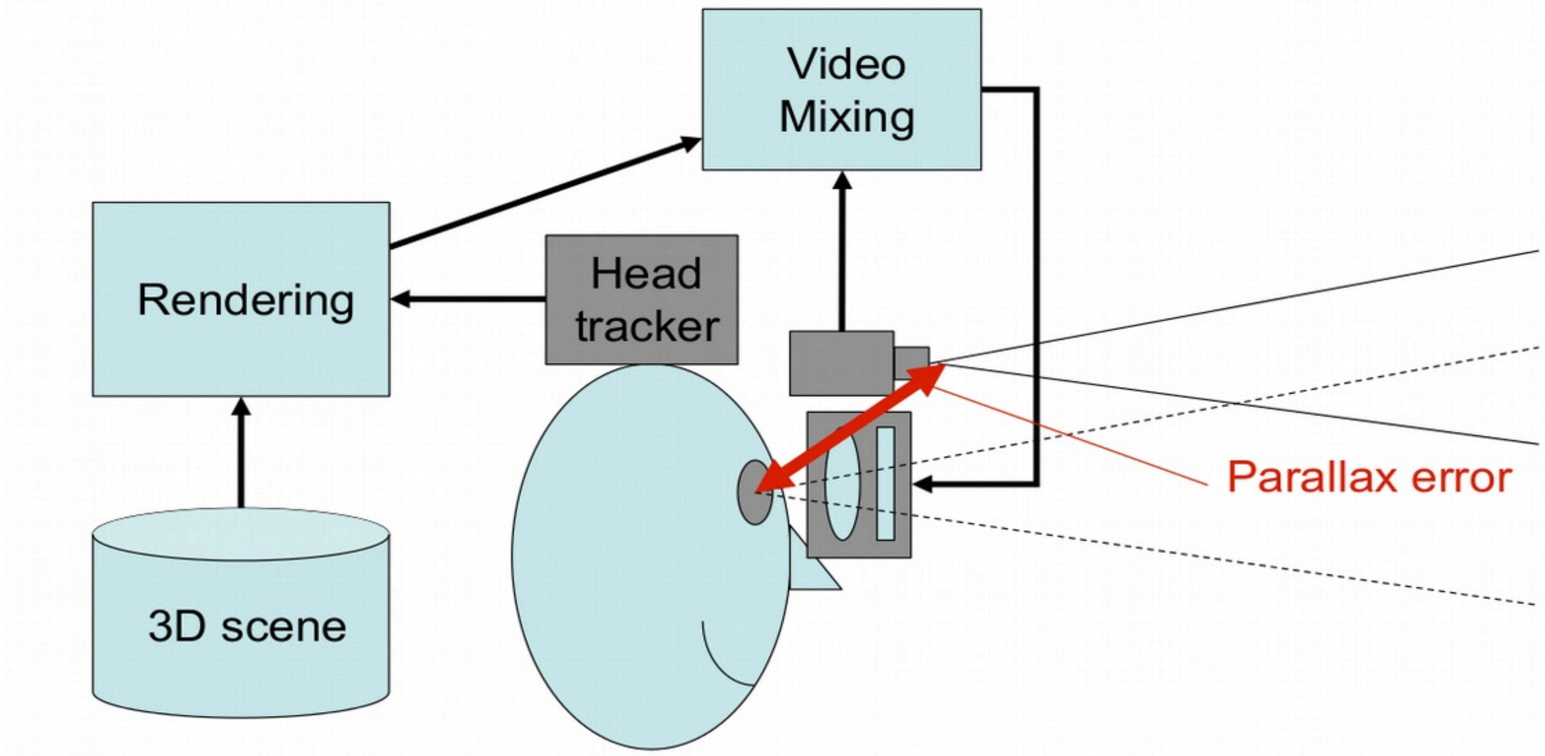


HMD: optical see-through (2)

- Use “beam combiner” + focus optics to overlay display image over real-world view
- Pro:
  - Direct view of real world
  - Eye can focus at different distances
- Contra:
  - (Mostly) not able to “cover” objects (problem with 3D occlusion, i.e. real object behind virtual one)
  - Lag/alignment issues reality ↔ virtual content more obvious

HMD: video see-through (1)

Image source (FU): LMU Lecture by J. Wagner



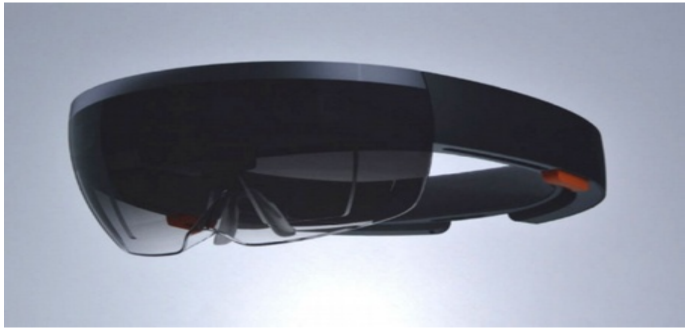
HMD: video see-through (2)

- Similar to VR displays (cf. lecture 5) + camera
- Pro:
  - Possible to use camera image for tracking
    → less lag, better alignment real ↔ virtual
  - Real objects can be hidden/substituted
- Contra:
  - Parallax error camera ↔ display
  - Similar problems to VR displays (motion sickness)
  - Single focus distance

Commercial HMDs for AR

Image source (CC): https://www.flickr.com/photos/jiff01/15867880743/
Image source (CC): http://photozou.jp/photo/show/124201/256868351

- Optical see-through:
  - Google Glass, North Focals (*not* AR!)
  - Microsoft Hololens →
  - Magic Leap
- Video see-through:
  - Smartphone goggles (Google Cardboard, Samsung Gear VR)
  - Lenovo Mirage Solo →
  - VR Headset + add-on cameras



Issues with HMDs

- Brightness (esp. for outdoor/sunlit context)
- Field of view (FOV)
  - Human FOV nearly 180°, many HMDs only ~ 20°
- Weight/comfort (can induce headaches)
- Lag between …
  - real and virtual content
  - head and image movement
- Safety – what if video see-through fails while you're crossing the street?

Issues with Cardboard & Co.

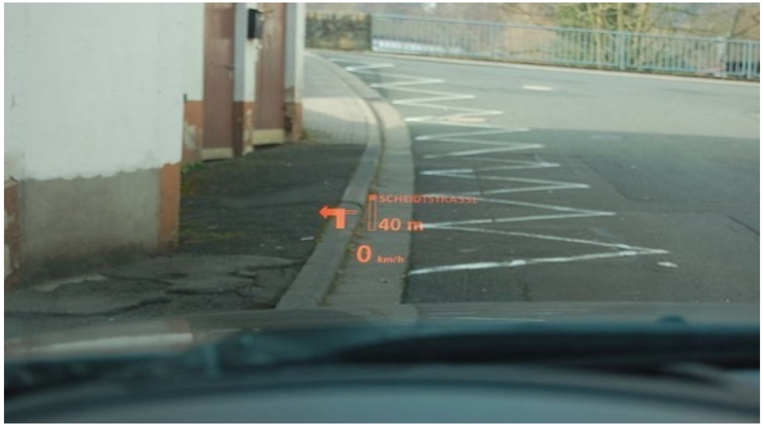
Image source (CC): https://www.flickr.com/photos/othree/14519574116

- Only single camera on almost all devices
  → no 3D view for AR possible, “flat screen effect”
- (Mostly) no positional tracking, only rotation
  → higher probability of motion sickness
- Unpredictable device capabilities
- Uncomfortable weight distribution
- Touchscreen inaccessible

Heads-up displays (HUDs)

Image source (CC): https://en.wikipedia.org/wiki/Head-up_display#/media/File:E60hud.JPG
Image source (CC): https://en.wikipedia.org/.../File:C-130J_Co_Pilot%27s_Head-up_display.jpg

- Mostly used in vehicles (car, airplane)
  - Navigational information, speed, horizon etc.
  - Not always spatially aligned (cf. left image)
  - Uses beam combiner (windshield) + 2D display



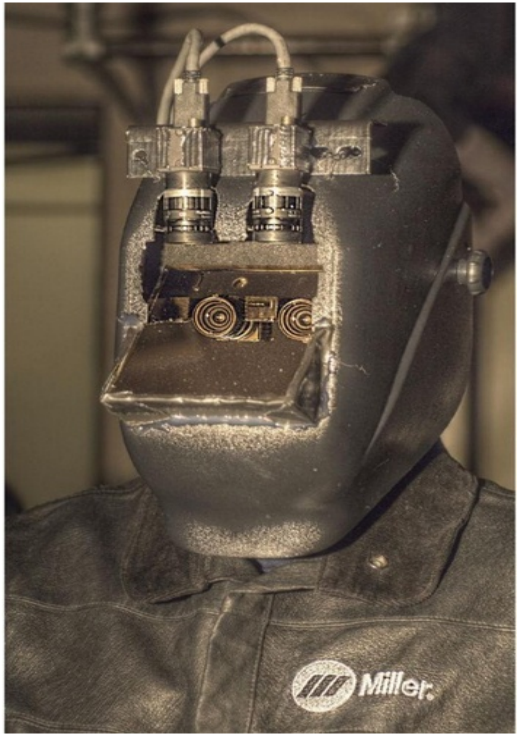
~~Augmented~~ Diminished Reality?

Image source (FU): http://wearcam.org/mannventions-password-stefanosmannaz13/mannglas.htm

- Not adding, but removing real-world content?
  - Any plausible scenario?

→ HDR welding mask/goggles

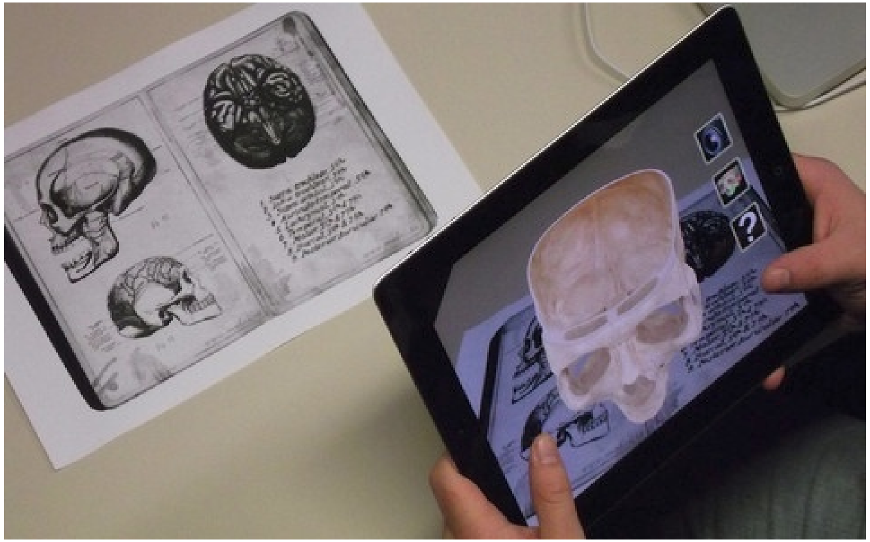
  - Filters out welding arc
  - Other details remain visible



Mobile devices as AR displays

Image source (CC): https://en.wikipedia.org/.../File:App_iSkull,_an_augmented_human_skull.jpg

- Uses “window into virtual world” metaphor
- Usually no 3D display
  - Except with tricks like Google Cardboard
  - Parallax issues
- Primary problem: location/tracking



AR: tracking & localization

- Fundamental requirement for AR: stable & reliable head/view tracking
- Multiple different approaches:
  - GPS + IMU
  - Vision-based
    - Stereo/depth cameras
    - SLAM algorithms

Tracking: GPS + IMU

Image source (CC): https://www.flickr.com/[...]/, https://www.flickr.com/[...]/27972275293

- Sensor fusion for 6D pose (geo coordinates)
  - 3D position from GPS, 3D orientation from IMU
  - (Relatively) low accuracy & position
  - Requires additional environment information
- Best suited for large-scale applications, e.g. buildings
- Examples: Pokemon Go, Layar, Wikitude Browser

Tracking: vision-based

- Primary goal (again): retrieve 6D pose
  - Secondary goal: create environment map
  - Enables interaction *with* environment
- … using specialized hardware:
  - Stereo cameras/depth cameras
  - Google Project Tango (tablet with depth cam)
- … for common devices (single RGB cam):
  - Marker-based (fiducial/image)
  - SLAM algorithms

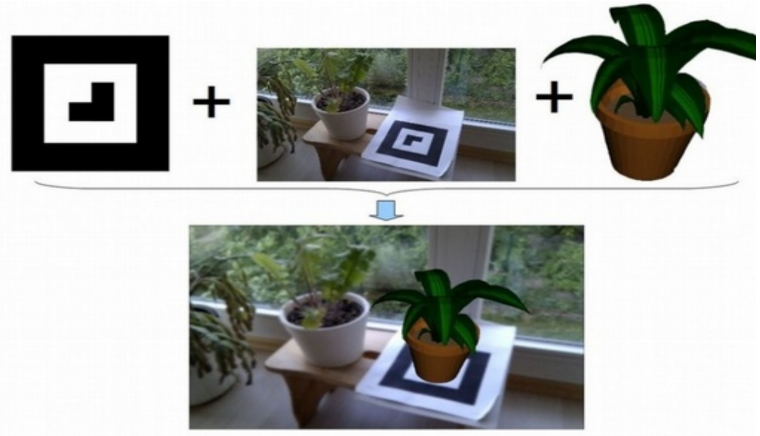
Tracking: vision – stereo/depth

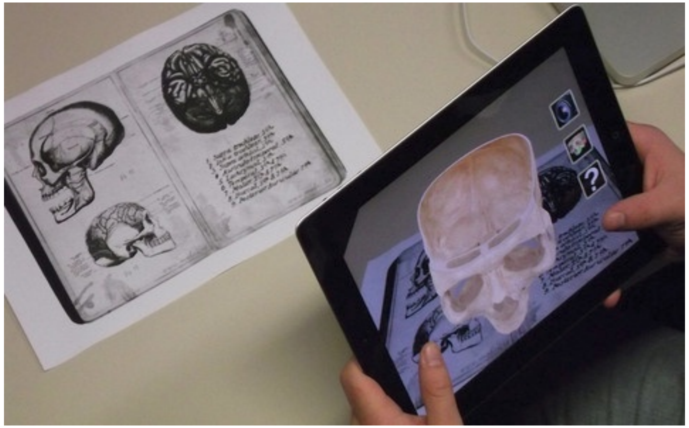
- Core idea: each pixel provides color + *distance*
- Stereo cameras: 2 cams mounted in parallel
  - Replicate (part of) human depth perception
  - Find *stereo correspondences* between 2 images
  - Use camera/lens geometry to calculate distance
- Depth cameras: single camera + emitter
  - ToF – Time-of-flight: measure travel time of IR flash
  - Structured light: shift in projected light pattern
- Both methods require extra hardware!

Tracking: vision – marker-based

Image source (FU): https://code.google.com/p/andar/
Image source (CC): https://en.wikipedia.org/.../File:App_iSkull,_an_augmented_human_skull.jpg

- Core idea: use pre-defined markers/targets
  - “Fiducials”: printed black-and-white patterns
  - Image targets: use (flat) real-world objects
- System detects pose relative to target

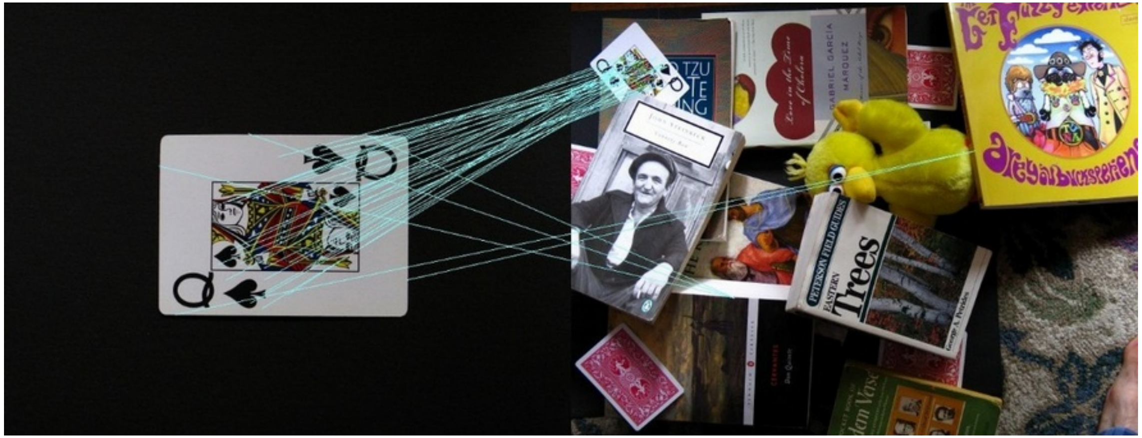




Marker-based tracking: SIFT & Co.

Image source (FU): https://robwhess.github.io/opensift/

- SIFT = Scale Invariant Feature Transform
  - Detects key points in images
  - Describes them independent of size, rotation, …
  - Descriptors can be matches across images
    → calculate geometric transformation



Tracking: vision – SLAM (1)

Image source (FU): https://www.youtube.com/watch?v=F3s3M0mokNc

- Simultaneous Localisation & Mapping
  - Creates 3D environment map + 6D camera pose
  - Requires only single camera
  - Works on almost any device
  - Creates stereo corres-pondences from motion
  - (Optionally) uses IMU data to determine motion magnitude

Tracking: vision – SLAM (2)

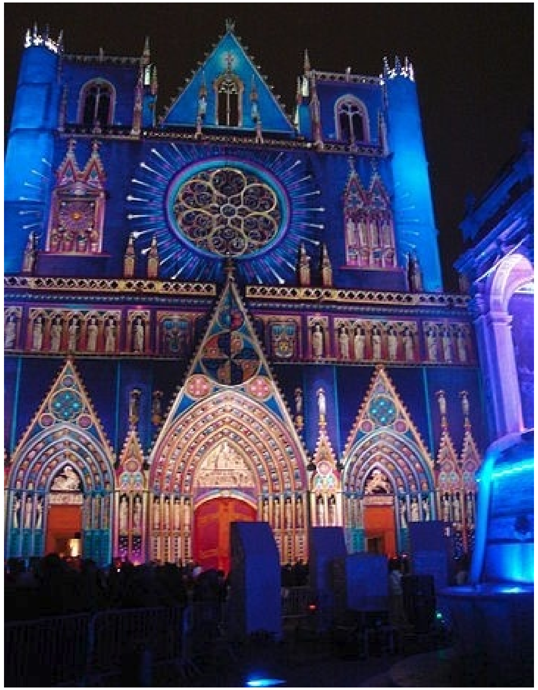
Image source (CC): http://photozou.jp/photo/show/124201/251690045

- Map generation requires significant processing power
  - Possible on many recent smartphones with ARKit/ARCore frameworks
  - Needs calibration data for camera (distortion), IMU (alignment to camera)
  - Includes plane detection (floor, tables, walls) for object placement + shadows
  - Touchscreen-based interaction

Spatial AR/Projection Mapping (1)

Image source (CC): https://en.wikipedia.org/.../File:Cath._St_Jean_%282%29.JPG
Video source (FU): https://www.youtube.com/watch?v=KBx4RGoB6XI

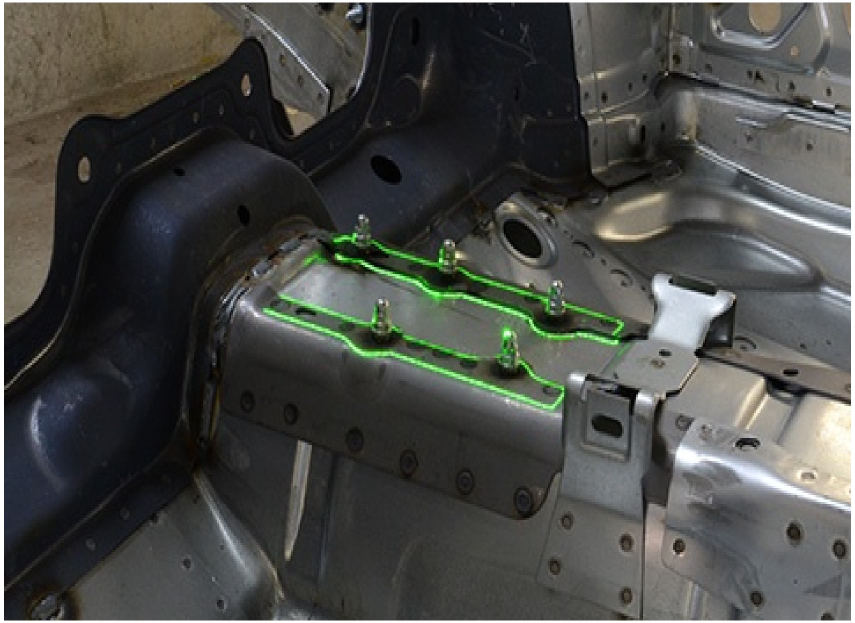
- Project augmentations directly onto real world
  - Mostly static projector/camera setups
    - E.g. for buildings, cars, other large objects
  - Pico projectors → mobile use?
    - Big problem (again): brightness vs. power supply
    - Secondary issue: focus distance → laser projectors?
    - Requires …
      - accurate environment maps
      - object tracking



Spatial AR/Projection Mapping (2)

Image source (FU): https://www.extend3d.de/

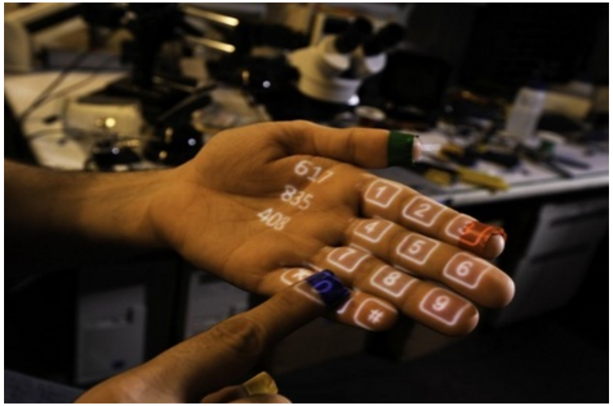
- Industrial usage scenario: project assembly instructions directly onto workpiece (“Werklicht” by Extend3D)
  - Needs high brightness & focal depth (green outline) → laser projector required
  - Workpiece geometry needs to be known in advance → target for visual tracking (similar to image targets)



Interaction with AR

Image source (FU): http://www.pranavmistry.com/projects/sixthsense/
Image source (FU): http://ilab.cs.ucsb.edu/projects/taehee/HandyAR/

- Touch screen, e.g. “Touch Projector”, Wikitude
  - Not feasible with HMDs, only for handheld use
- Individually tracked tools (e.g. Wiimote, gloves)
- Optical recognition/tracking of hands



AR applications

Image source (FU): http://www.trylive.com/
Image source (FU): http://campar.in.tum.de/Chair/ResearchIssueMedAR

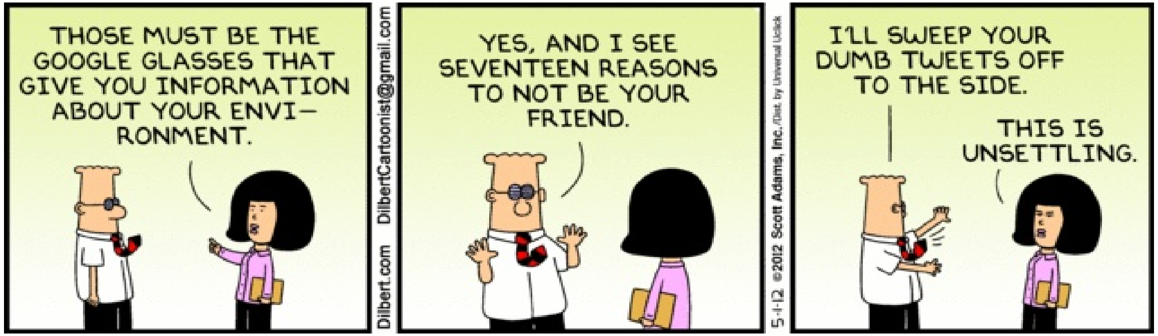
- AR: still waiting for the “killer app”
  - Advertisements (Ikea, fashion) →
  - Military (sadly): cockpits, weapon sights
  - Education (e.g. iSkull, slide 14)
  - Repair & Maintenance (e.g. see video)
  - Gaming (e.g. Ingress, Pokemon Go)
  - Medical applications (X-ray vision) →
  - Navigation (only prototypes)

Outlook: “mobile” VR?

Image source (CC): https://de.m.wikipedia.org/wiki/Datei:Oculus_Quest.jpeg
Image source (FU): https://de.m.wikipedia.org/wiki/Datei:Oculus_Touch_Hand_Controller_2.jpeg

- Oculus Quest: no cables (mobile HW)
- 4 low-res high-speed IR cameras
  - Tracking of hand-held controllers
  - “Room-scale” 6DoF tracking
- No obstacle detection, only pre-defined safety boundaries → needs large space
- Possible workaround: “redirected walking”
- General research topic in VR: locomotion

The End


THOSE MUST BE THE GOOGLE GLASSES THAT GIVE YOU INFORMATION ABOUT YOUR ENVI-RONMENT.
Dilbert.com DilbertCartoonist@gmail.com
YES, AND I SEE SEVENTEEN REASONS TO NOT BE YOUR FRIEND.
5-1-12 ©2012 Scott Adams, Inc./Dist. by Universal Uclick
I'LL SWEEP YOUR DUMB TWEETS OFF TO THE SIDE.
THIS IS UNSETTLING.