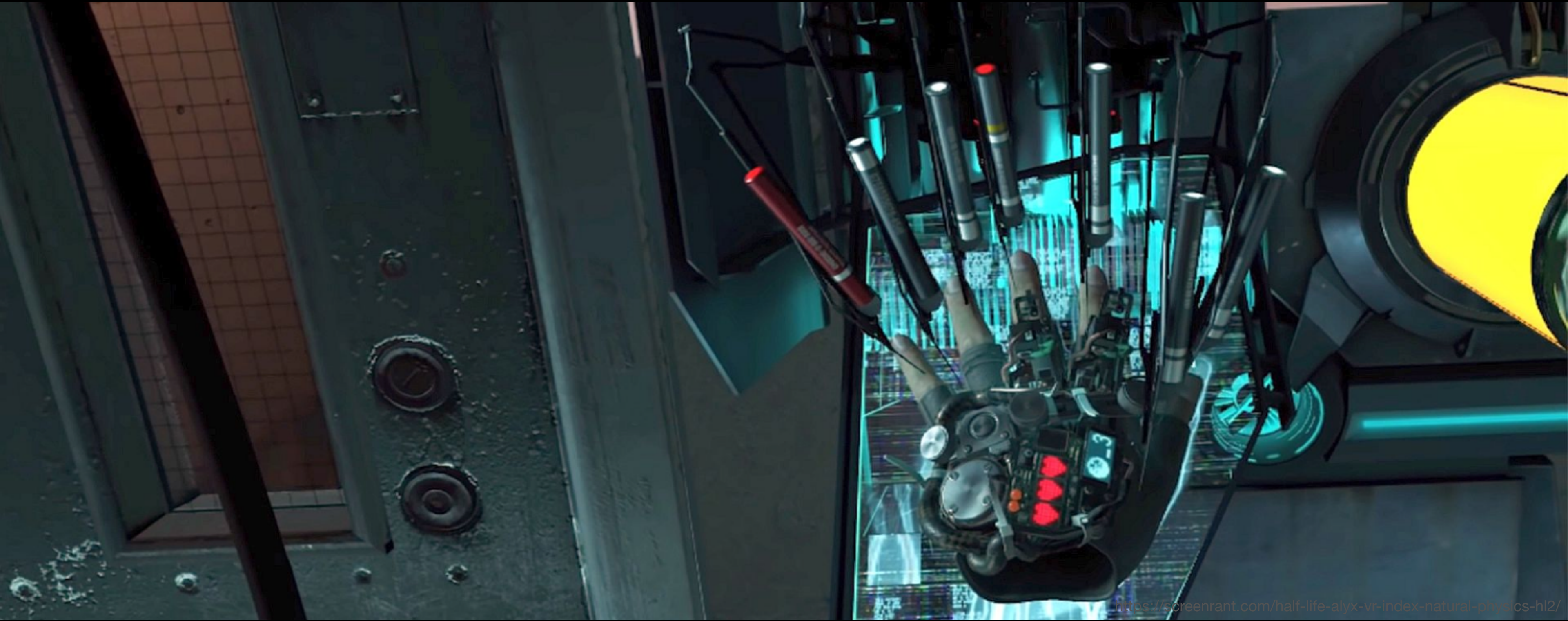
Virtual & Augmented Reality

WS 2025

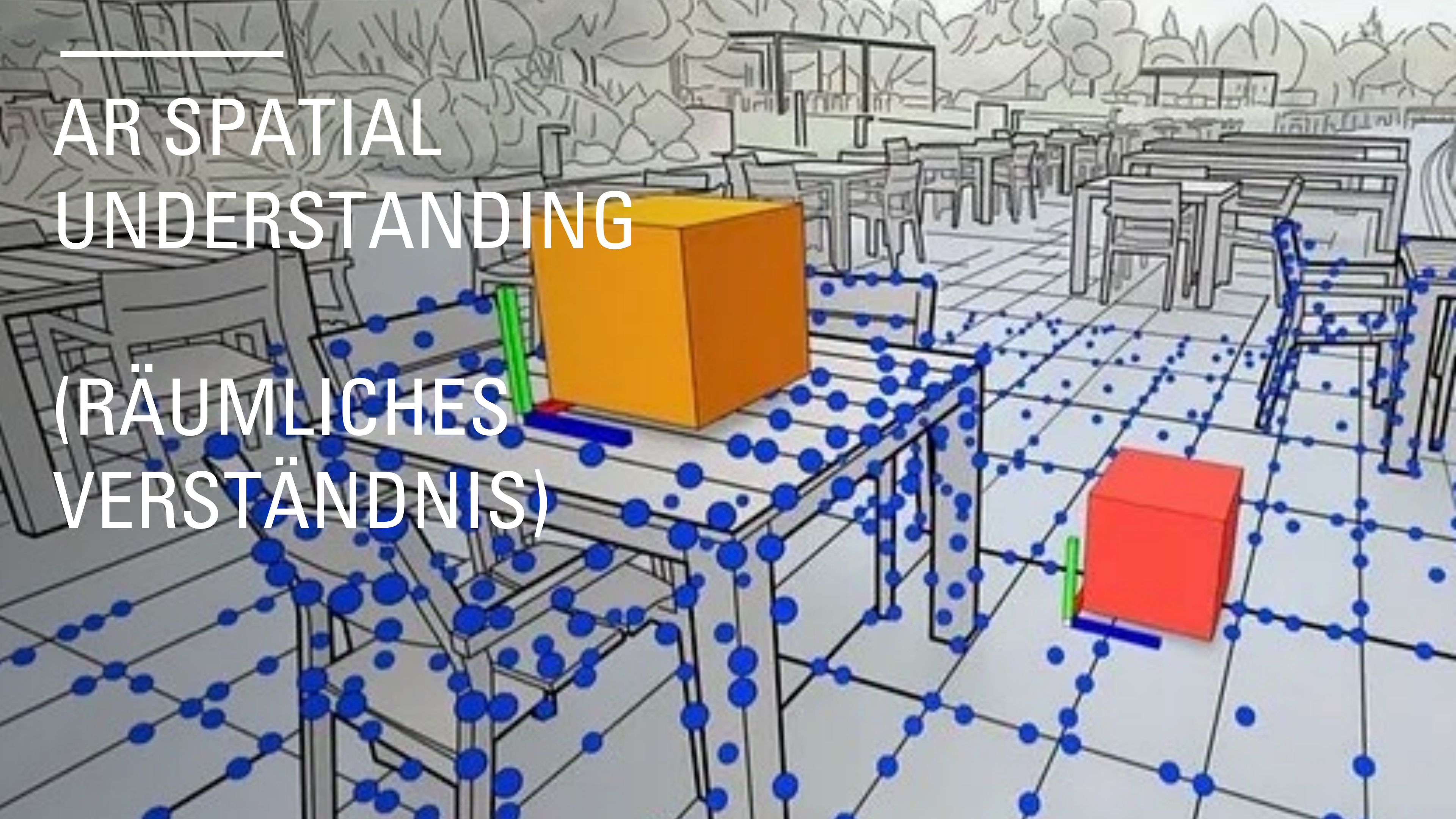

https://screenrant.com/half-life-alyx-vr-index-natural-physics-hl2/

AR Spatial Understanding

BHT

AR SPATIAL UNDERSTANDING

(RÄUMLICHES VERSTÄNDNIS)



AR SPATIAL UNDERSTANDING TOPICS

- Types of Scene Understanding
- Feature Extraction
- Marker Tracking
- Plane Detection

TYPES OF SCENE UNDERSTANDING

- Small overview:
  - Image Detection
  - Curved Images
  - Plane Detection
  - Face Detection
  - Model Tracking
  - Image Segmentation

FEATURE EXTRACTION

- Technique in Computer Vision
  - Looking at the most relevant areas in an image
  - try to find these areas in another image

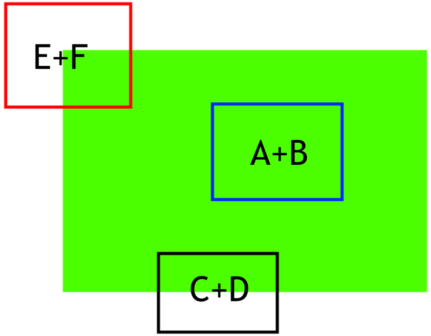
- For reference:
  - https://docs.opencv.org/3.4/db/d27/tutorial_py_table_of_contents_feature2d.html

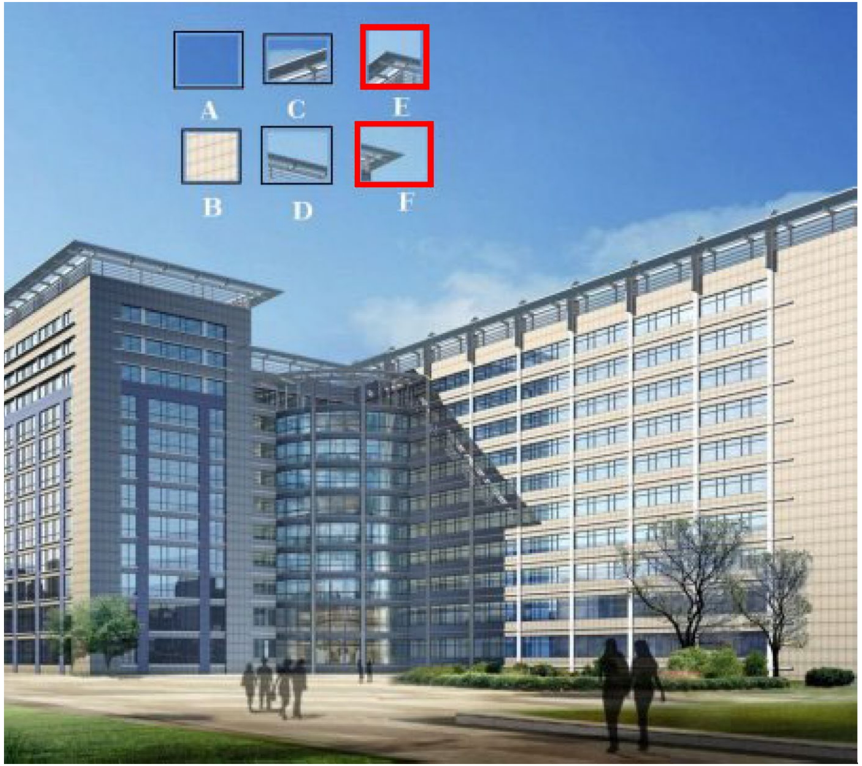
FEATURE EXTRACTION

- Which of these features is most relevant / remarkable?

- E + F

→ Corners



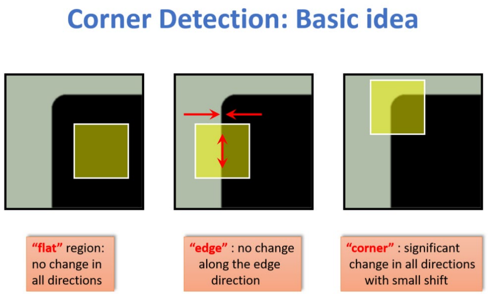


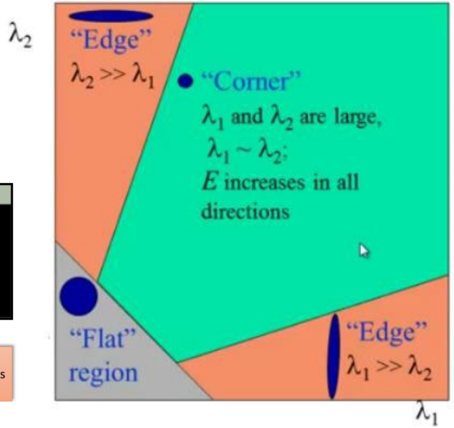
FEATURE DETECTION ALGORITHMS

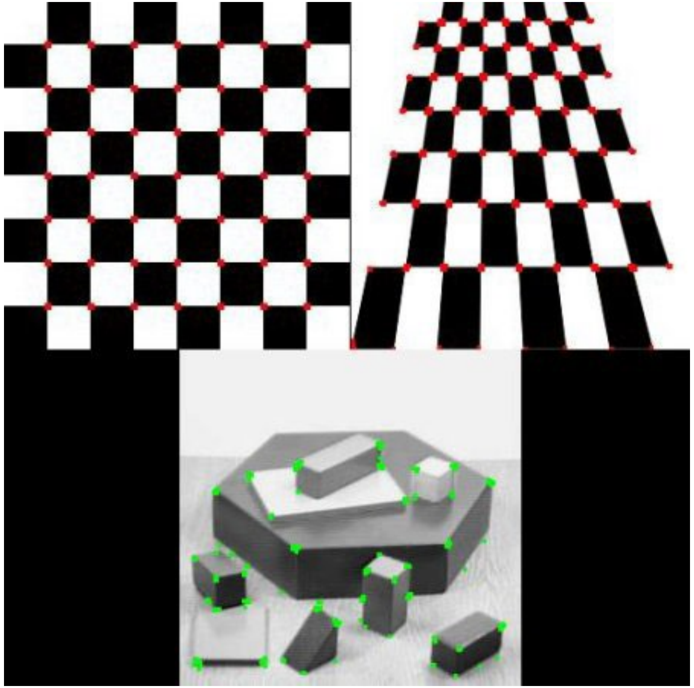
- Many algorithms that can find these features
  - Harris-Corner Detection
  - Shi-Tomasi & GoodFeaturesToTrack
  - SIFT & SURF (& FAST & BRIEF & ORB)

FEATURE DETECTION ALGORITHMS

- Example: Harris-Corner Detection
  - kernel (e.g. 3x3, 5x5)
  - traverses grayscale image
  - look for change of color change in both directions







FEATURE DETECTION ALGORITHMS

scale invariant feature transform features from accelerated segment test

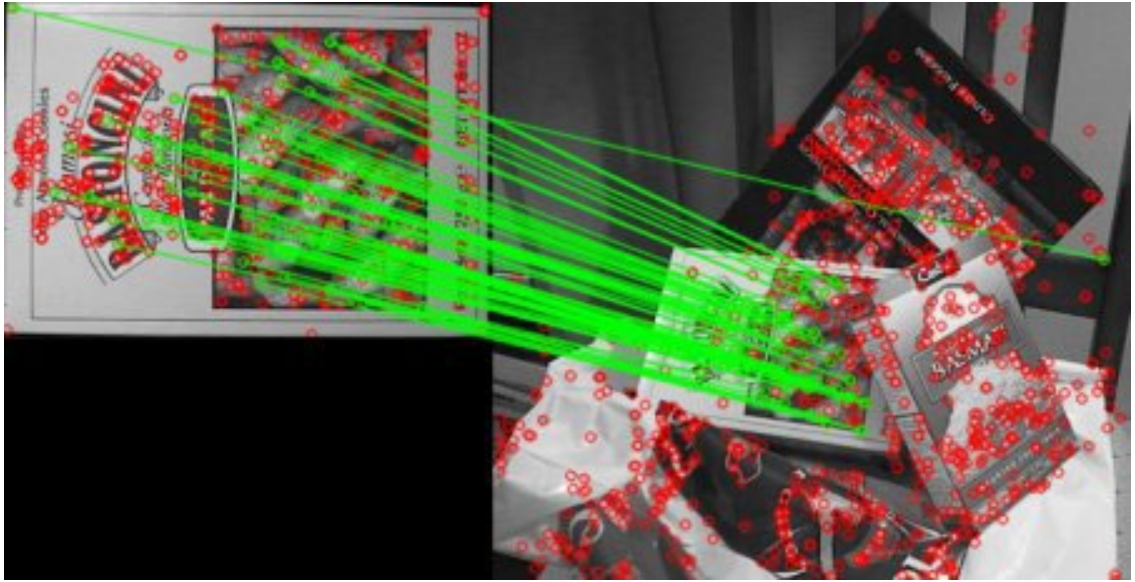
- Example: SIFT & FAST
  - scale independent
  - different levels of confidence

FAST SIFT

FEATURE MATCHING

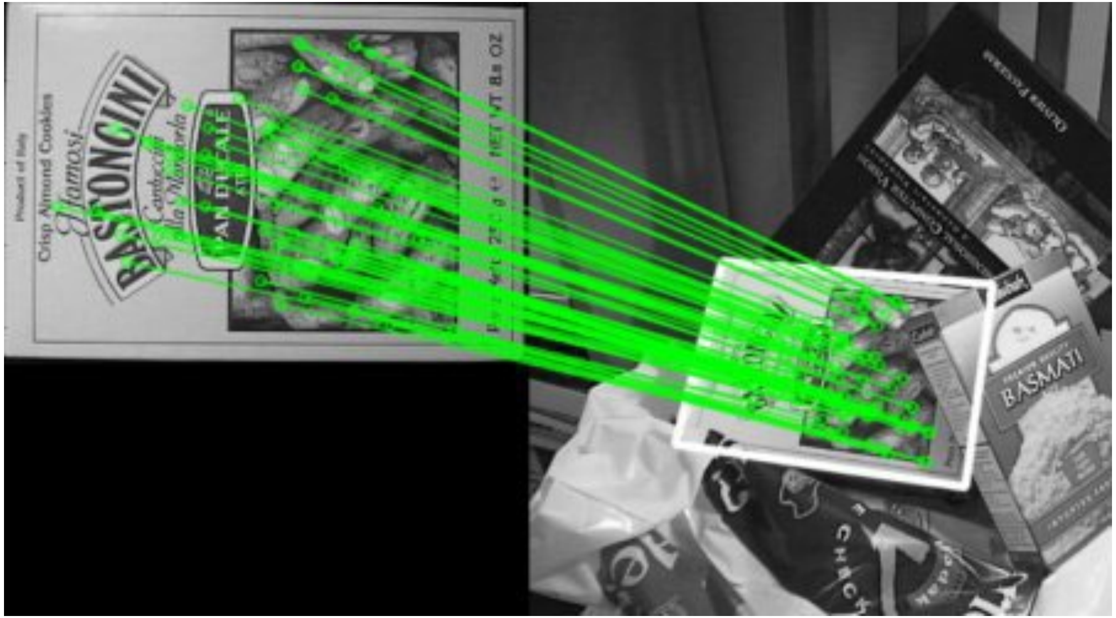
- Compare features in one image to another
  - Brute Force
  - FLANN matchingFast

    Library for Approximate Nearest Neighbors



FEATURE MATCHING + HOMOGRAPHY

- Find how the object is perspectively transformed
    - calculate Homography



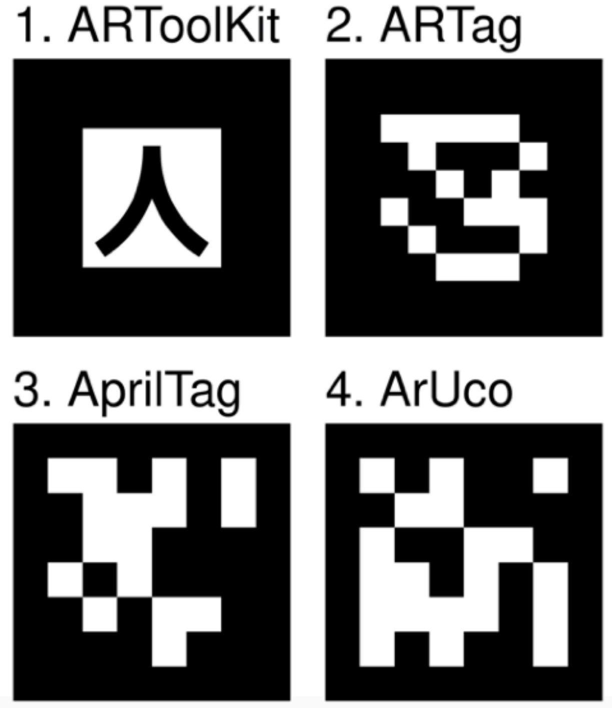
Homography = „similar drawing“
a bijection that maps lines to lines of the same plane object in two different projections

FEATURE MATCHING (WITH MARKERS)

- Some features are easier to find than others
- We can use optimized images

→ Fiducial Markers (high contrast, clear features)
(Referenzmarker)

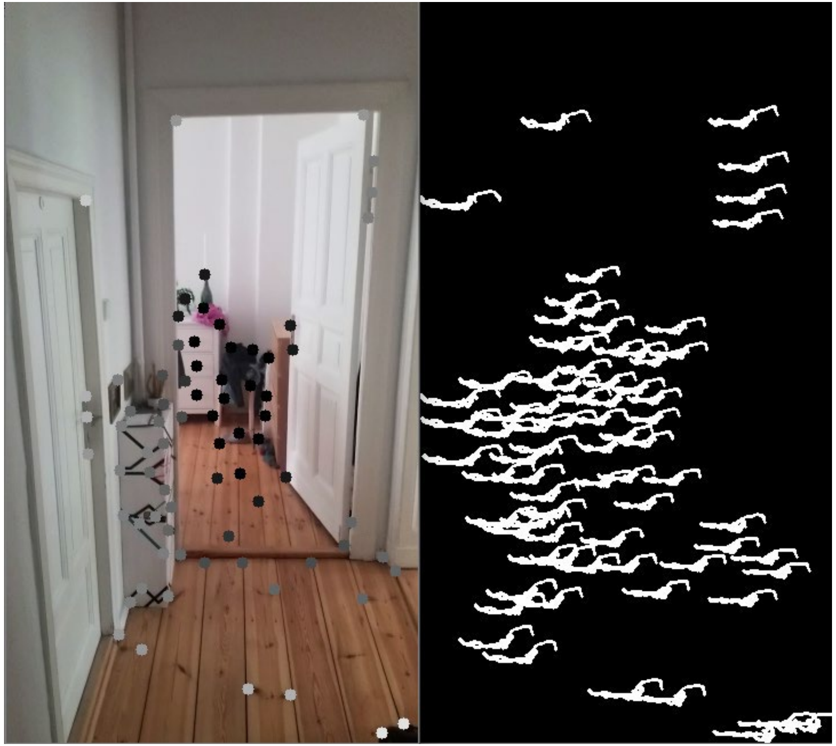
very similar to QR Codes



https://en.wikipedia.org/wiki/ARTag

PLANE DETECTION

- We can also track features in 3D space
- If we track those over multiple images we get more information about them
  - how do they move
    (if its far away it should move less)
  - path of points
    (extension: is it in line with device IMU)



PLANE DETECTION

- group of 3D points = point cloud
  - managed by SDK, e.g. ARCore
  - https://www.youtube.com/shorts/1-GDF18ig3M

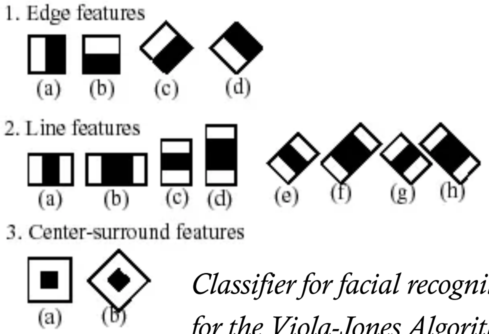
- if point clusters are horizontal or vertically aligned →its probably a plane
- anchors can be placed as self-created feature point
- serves as positioning for digital content

FACE TRACKING

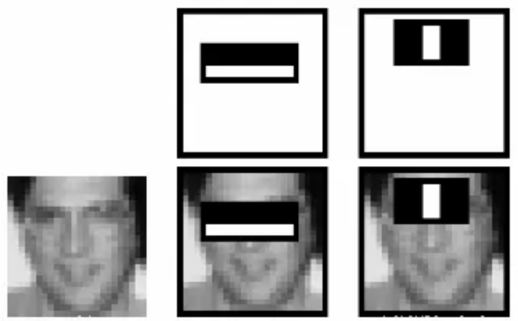
https://de.wikipedia.org/wiki/Viola-Jones-Methode

- Face detection – Find Face in image
- Face recognition – Finding Facial Features and comparison
- Face Tracking – Position, Orientation of the Face

- Algorithms
  - Viola-Jones
  - Single Shot Detector (SSD)
  - You Only Look Once (YOLO)

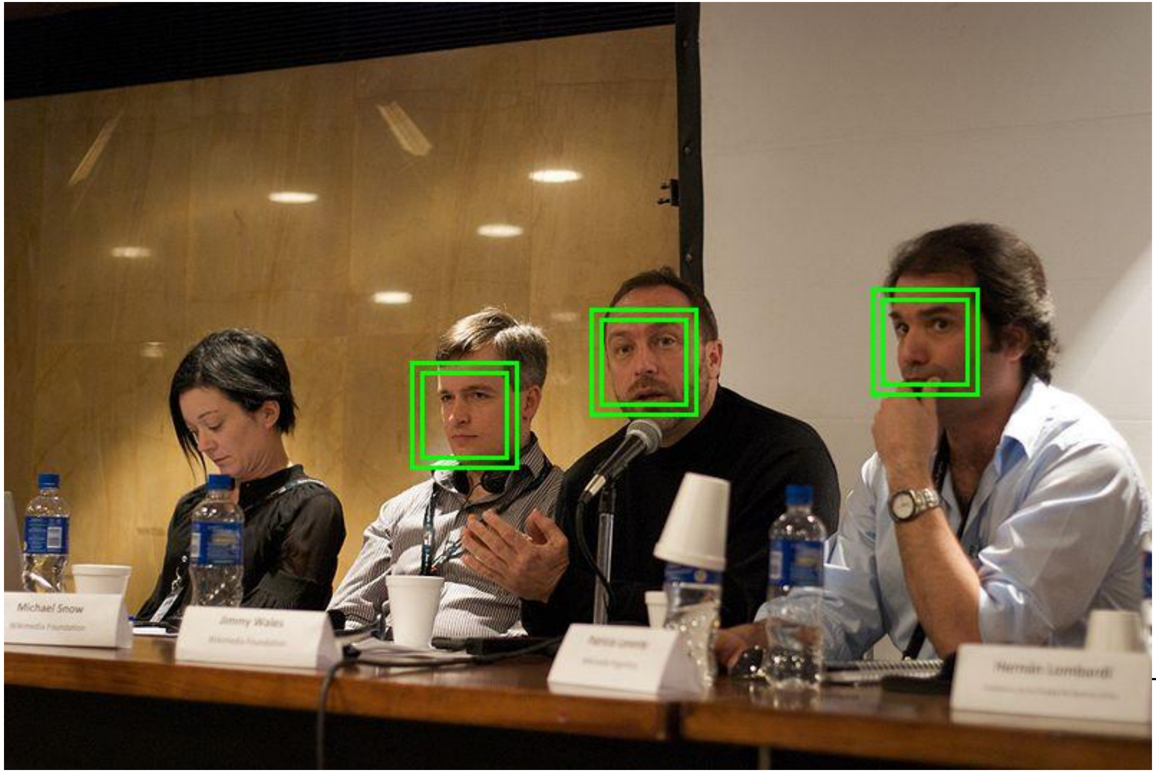


Classifier for facial recognition for the Viola-Jones Algorithm



FACE TRACKING

Von Jimmy answering questions.jpg: Wikimania2009 Beatrice Murchderivative work: Sylenius (Diskussion) - Jimmy answering questions.jpg,
CC BY 3.0, https://commons.wikimedia.org/w/index.php?curid=11309460



Viola-Jones-Algorithmus:

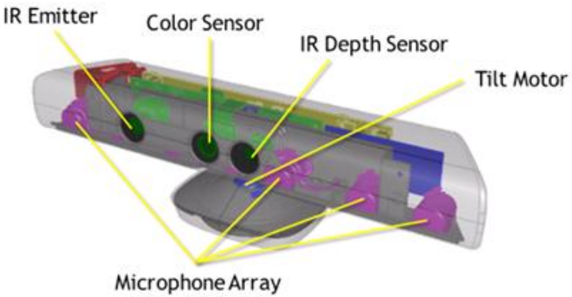
- effizient (Erkennung Echtzeit)

- Gesichter müssen aber frontal zu sehen sein

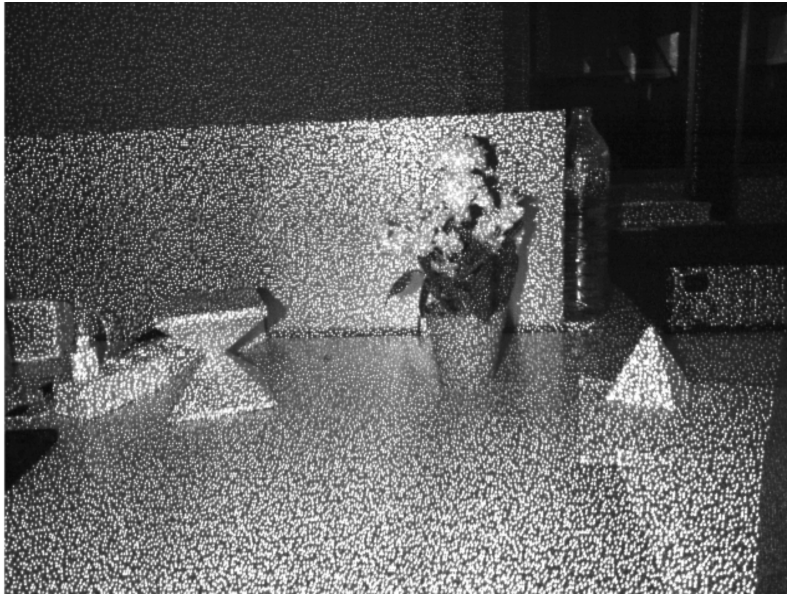
- müssen gut beleuchtet sein

“REAL” SPATIAL UNDERSTANDING

- using additional sensors helps in finding “real” surfaces
- usually using some form of InfraRed technology
  - ToF (Time of Flight)
  - LiDAR (Light Detection And Ranging)



Pattern →

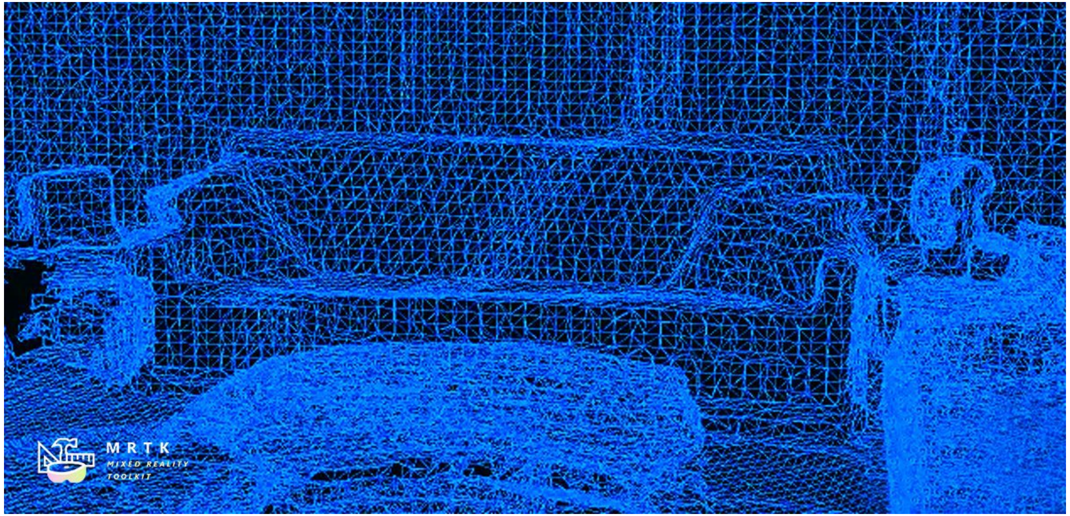


https://learn.microsoft.com/en-us/archive/msdn-magazine/2012/november/kinect-3d-sight-with-kinect

“REAL” SPATIAL UNDERSTANDING

- using additional sensors helps in finding “real” surfaces
- depends on
  - sensor resolution
  - time required for fusing data
  - coverage of the area

Spatial Understanding of HoloLens →



CHALLENGES OF COMPUTER VISION FOR AR

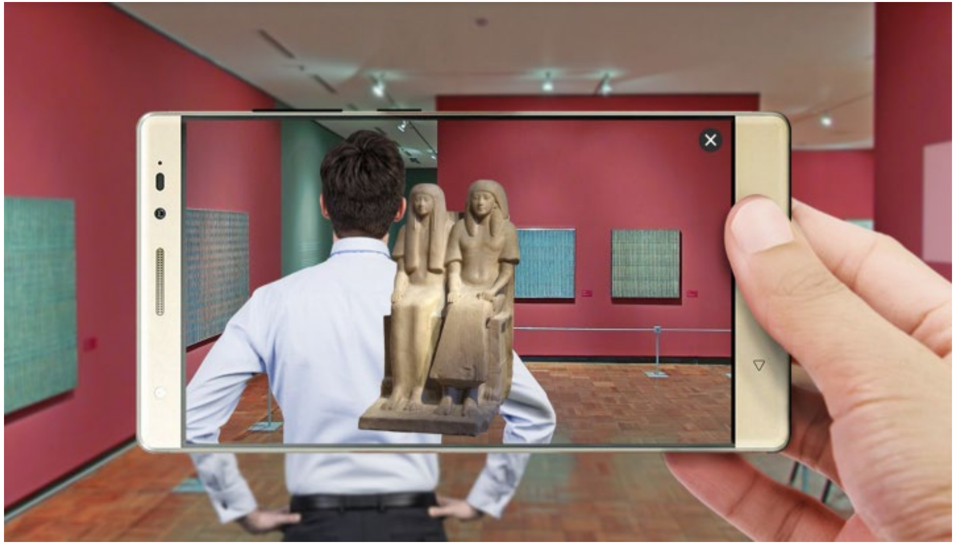
- scene understanding is imperfect
  - dependent on scene
  - on lighting
  - on device movement
- and is not instant (not as instant as video-stream)

AR RENDERING CHALLENGES

- Based on the platform
  - HMD with video see through→ video stream and rendering for each eye
  - HMD with optical see through → rendering for each eye
  - Mobile → video stream and „normal“ rendering
  - Spatial Projection → „normal“ rendering
- Video Stream
  - Slower than optical see through and „lower resolution“
- Occlussion behind real world objects
- Object Lighting

OCCLUSION IN AR

- ongoing research on this topic
- Problem:
  - by default virtual content is always in front
  - e.g. when using plane detection, we have no other knowledge about what could be in front



OCCLUSION HANDLING APPROACHES

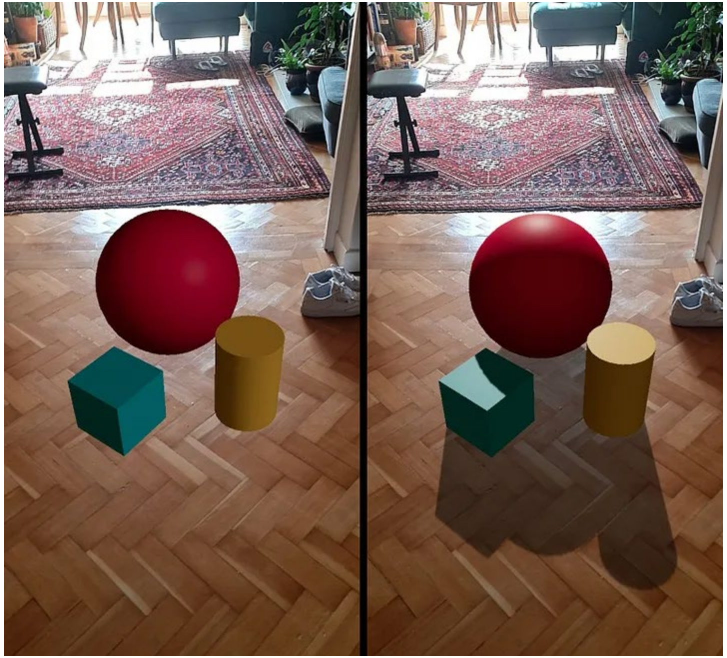
https://www.researchgate.net/figure/Phantom-Rendering-in-AR-The-invisible-phantom-object-prevents-occluded-virtual-fragments_fig6_227088766

Phantom objects (which represent the virtual counterparts of real world objects) are rendered invisible (only to the z-buffer), before purely virtual objects are being rendered.

- estimated/calculated occlusion
  - takes some time to gather information about scene
  - often times not very precise
- pre-known occlusion (phantom objects)
  - precise if tracking is good
  - unflexible, must be known beforehand
- sensor-based occlusion
  - sensor resolution often times too low
  - in theory best solution

LIGHTING FOR AR OBJECTS

- Objects placed in a real environment feel out of place if the lighting doesn't match
  - Shadows
  - Light Color, Intensity, Direction etc.
  - Ambient Lighting



FUTURE RESEARCH

- Different focus depths
- Form factor, comfort
- Higher FOV
- Better visuals
- Battery time
- …

EXAMPLES

Two Marker Tracker Example
https://www.youtube.com/watch?v=dIEZwmjuaUA

Plane Detection with ARKit
https://www.youtube.com/watch?v=-aSIfHl0Me8

Creal focus depth
https://www.youtube.com/watch?v=SVJDav1tE2g