HoloLens, Mixed Reality and Spatial Computing

Marc Pollefeys | ETH Zurich and Microsoft

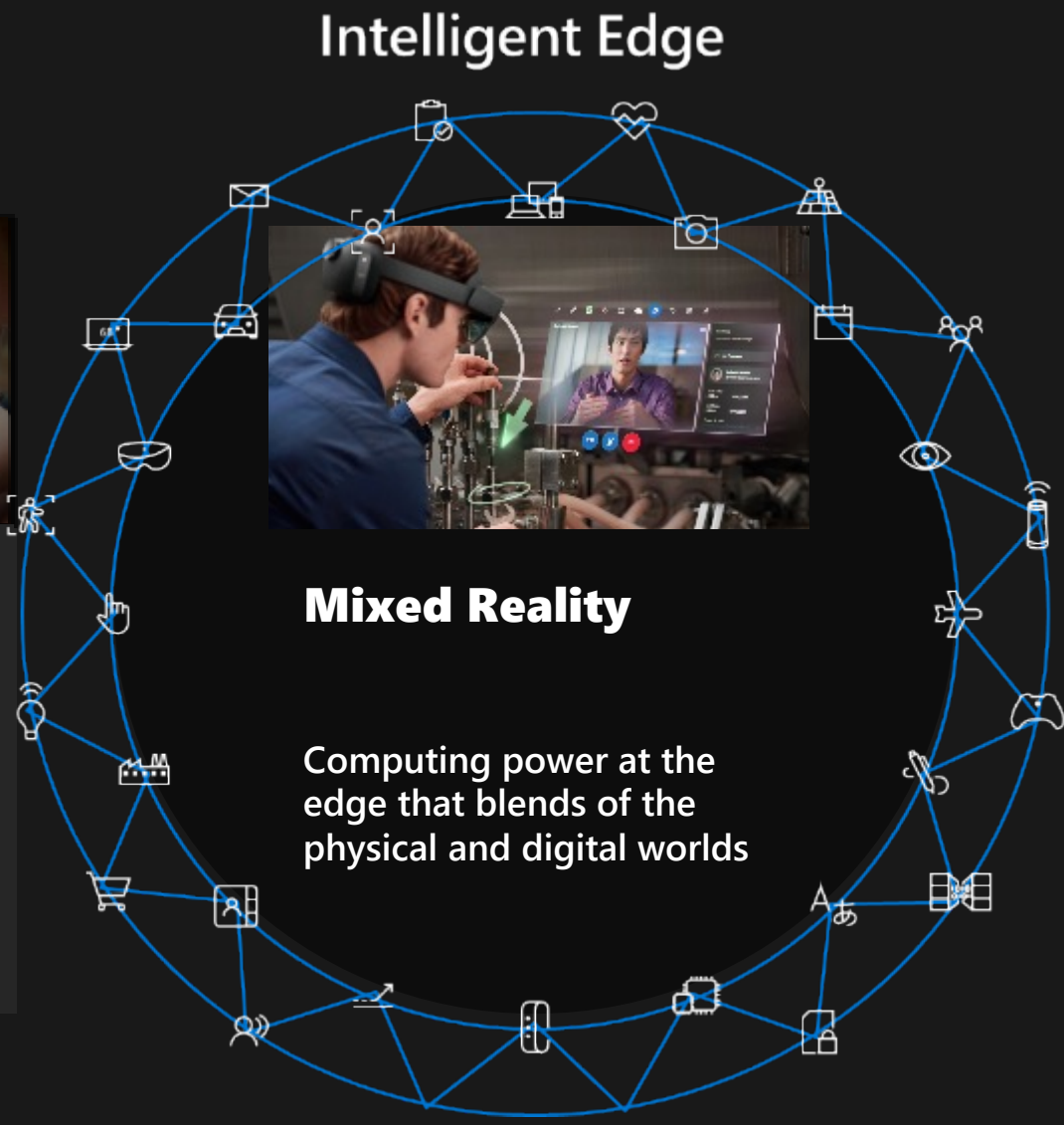
THIRD WAVE OF COMPUTING

Personal Computers

Computing power for many but immobile and no contextual awareness

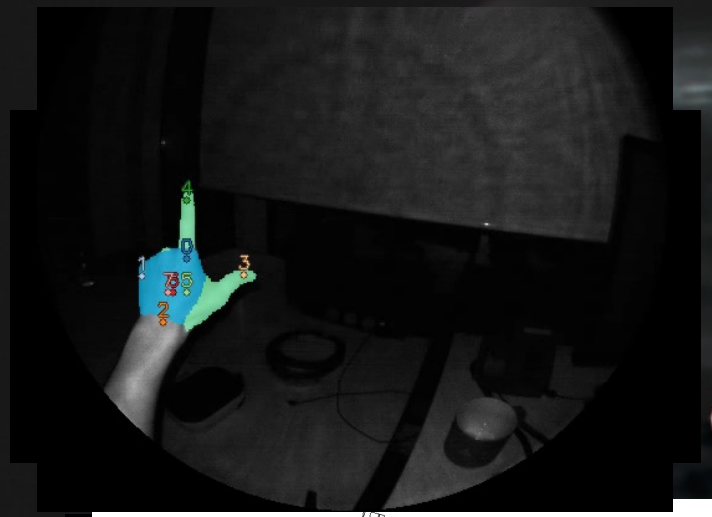
Smartphones

Computing power for all with on-the-go access and some contextual awareness

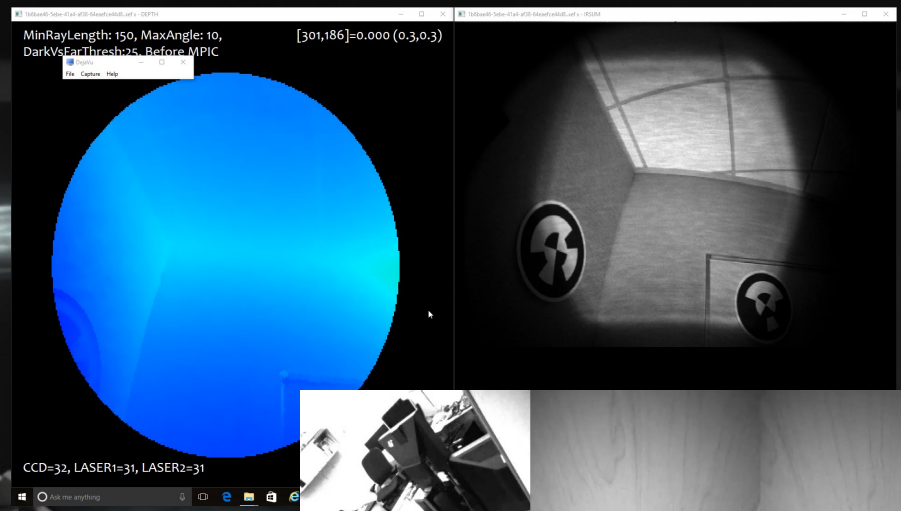


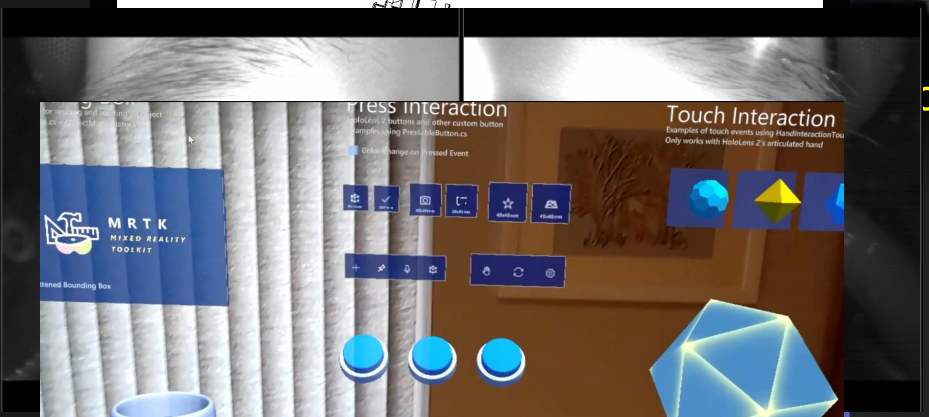
Mixed Reality

Computing power at the edge that blends of the physical and digital worlds

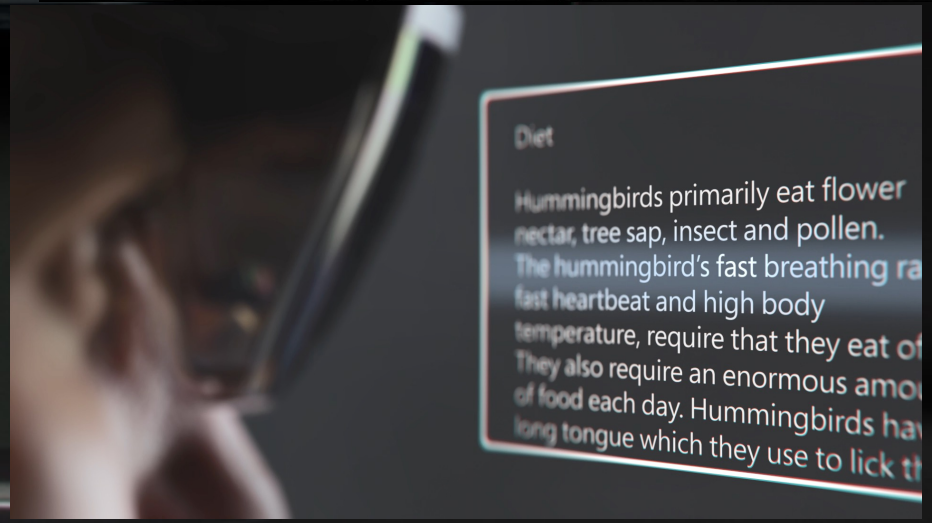


8Mpix RGB camera
(video communication)



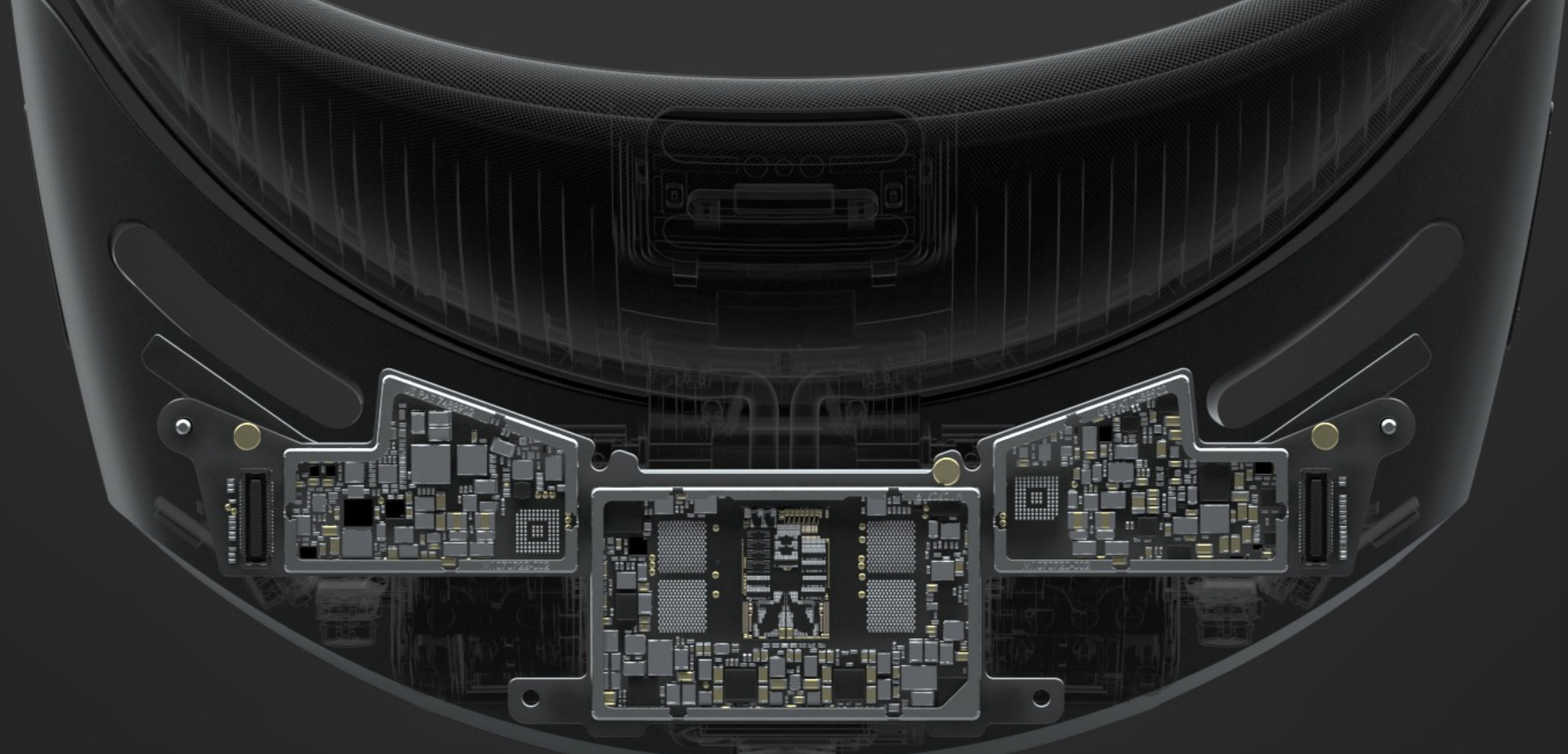


1Mpix depth camera
ong- & short-throw mode)
surfaces and hand-tracking)



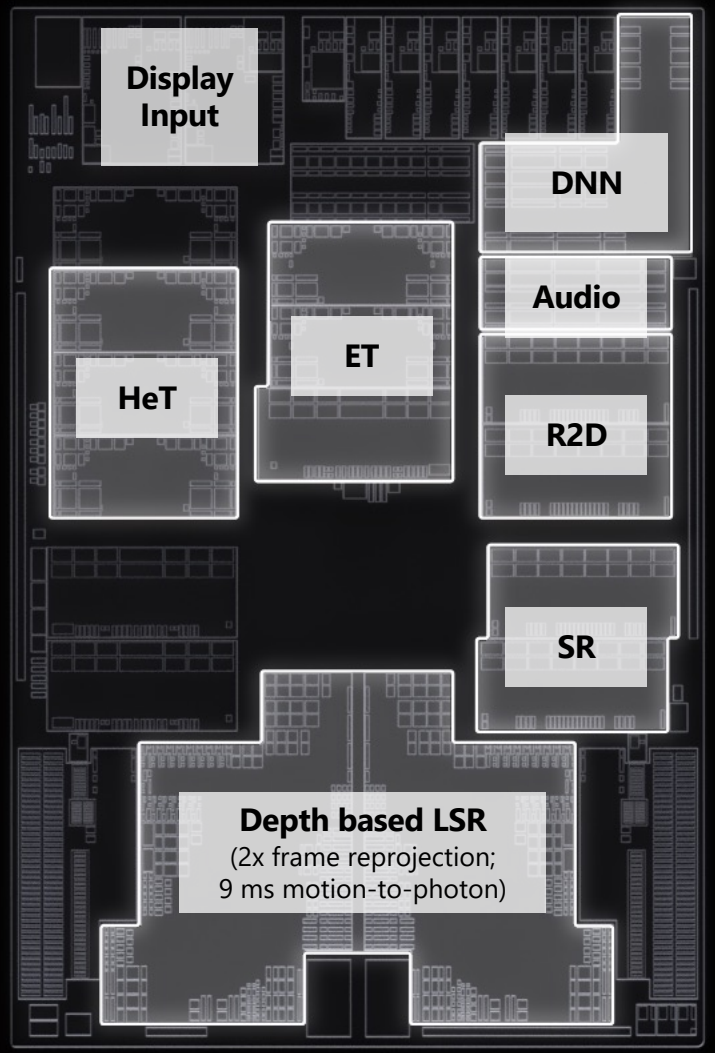
IR eye cameras + IR LEDs
(eye-tracking, iris recognition,
display calibration)

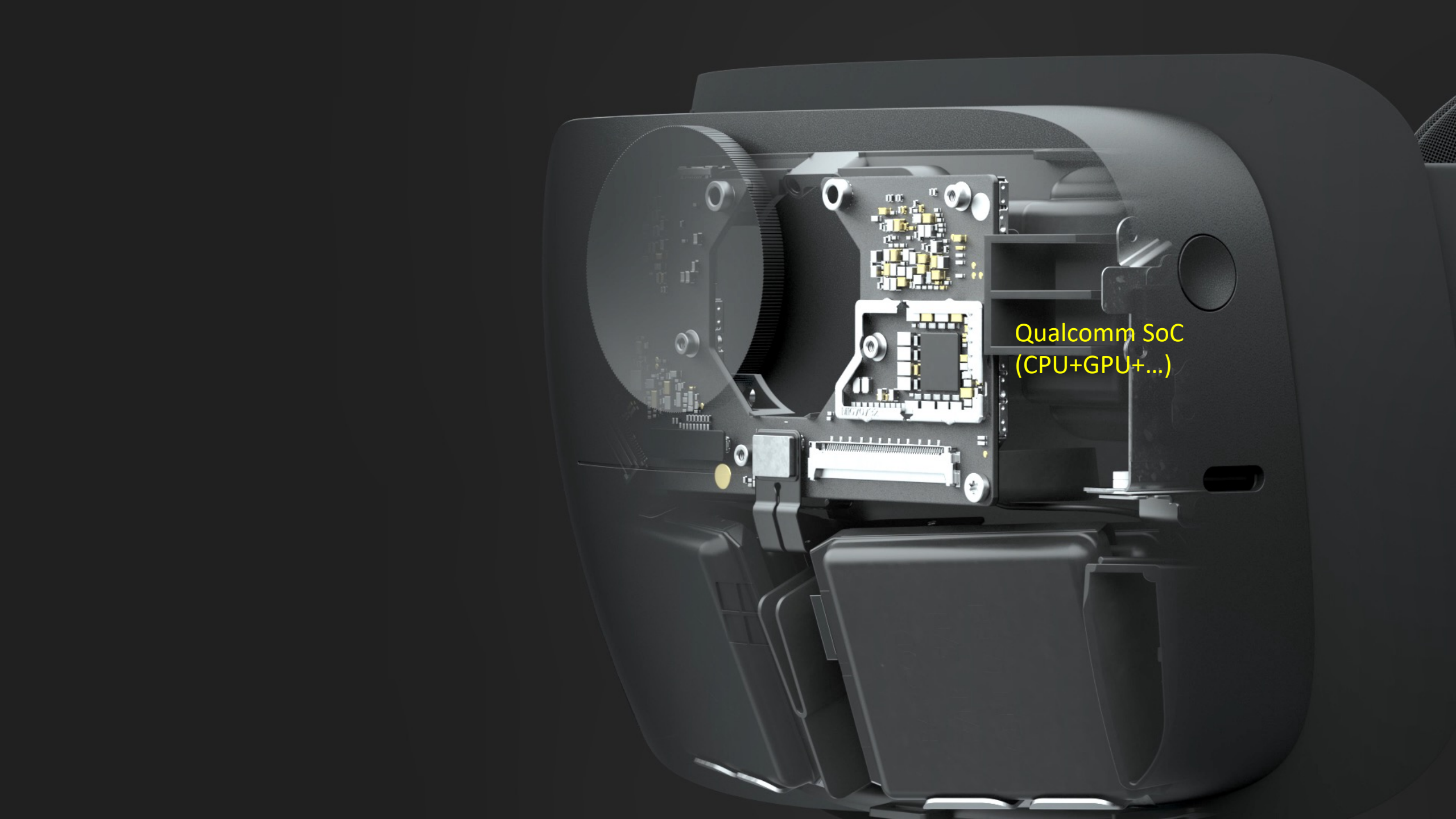
2+3 microphone array
(speech in 90dB noise)

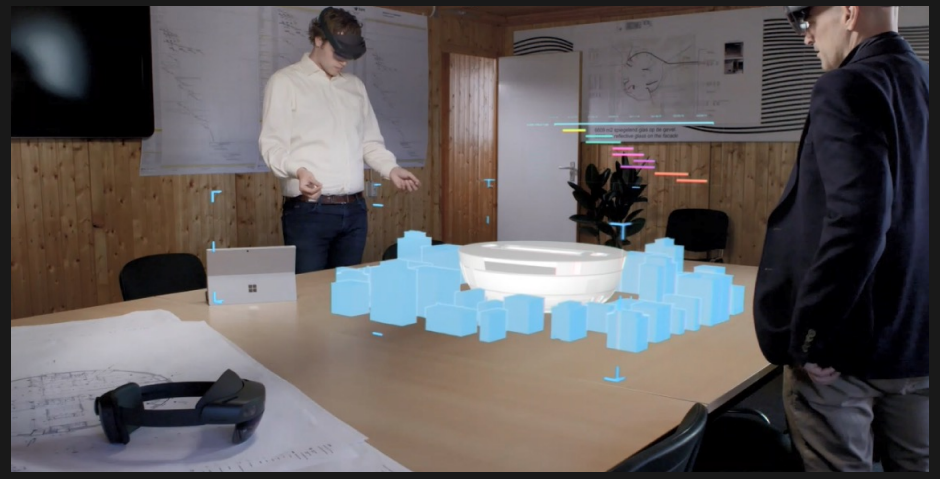


HPU (DSPs, DNN AI core, LSR)

HPU

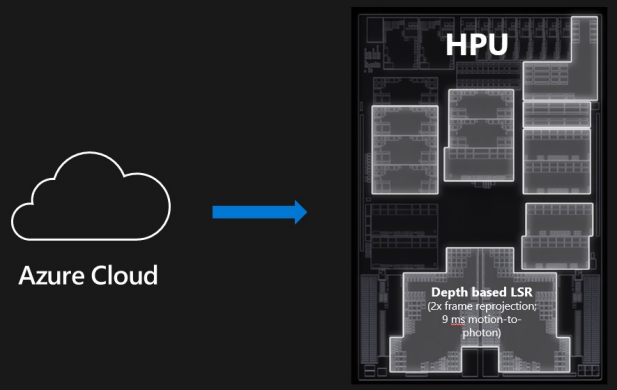

Display Input
DNN
Audio
ET
HeT
R2D
SR
Depth based LSR
(2x frame reprojection;
9 ms motion-to-photon)

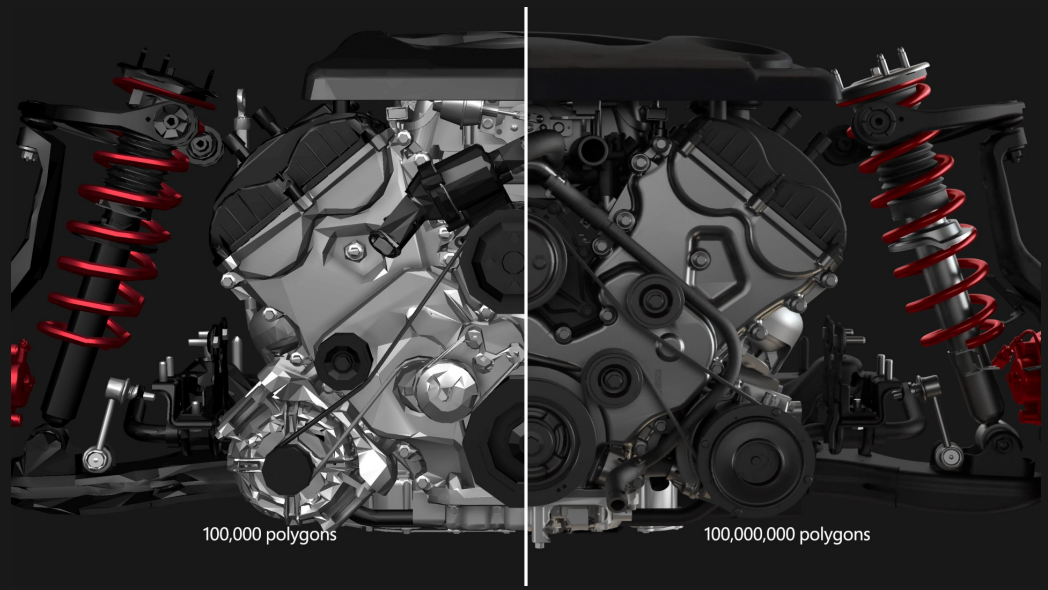

Qualcomm SoC
(CPU+GPU+…)



Planning and design reviews

On-site visualization





Remote rendering


WARNING

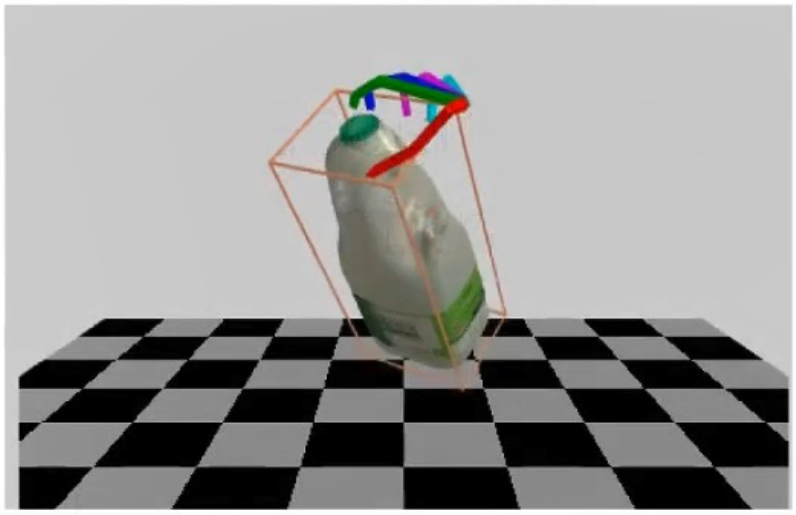

Fuel System
Regulator Line Install

Hands + Objects : Unified Egocentric Recognition of 3D Hands+Object Poses and Interactions

Tekin, Bogo & Pollefeys, CVPR 2019

2D Hand + Object Pose

3D Hand + Object Pose



Per-frame predictions

Action Recognition for Automated Task Guidance

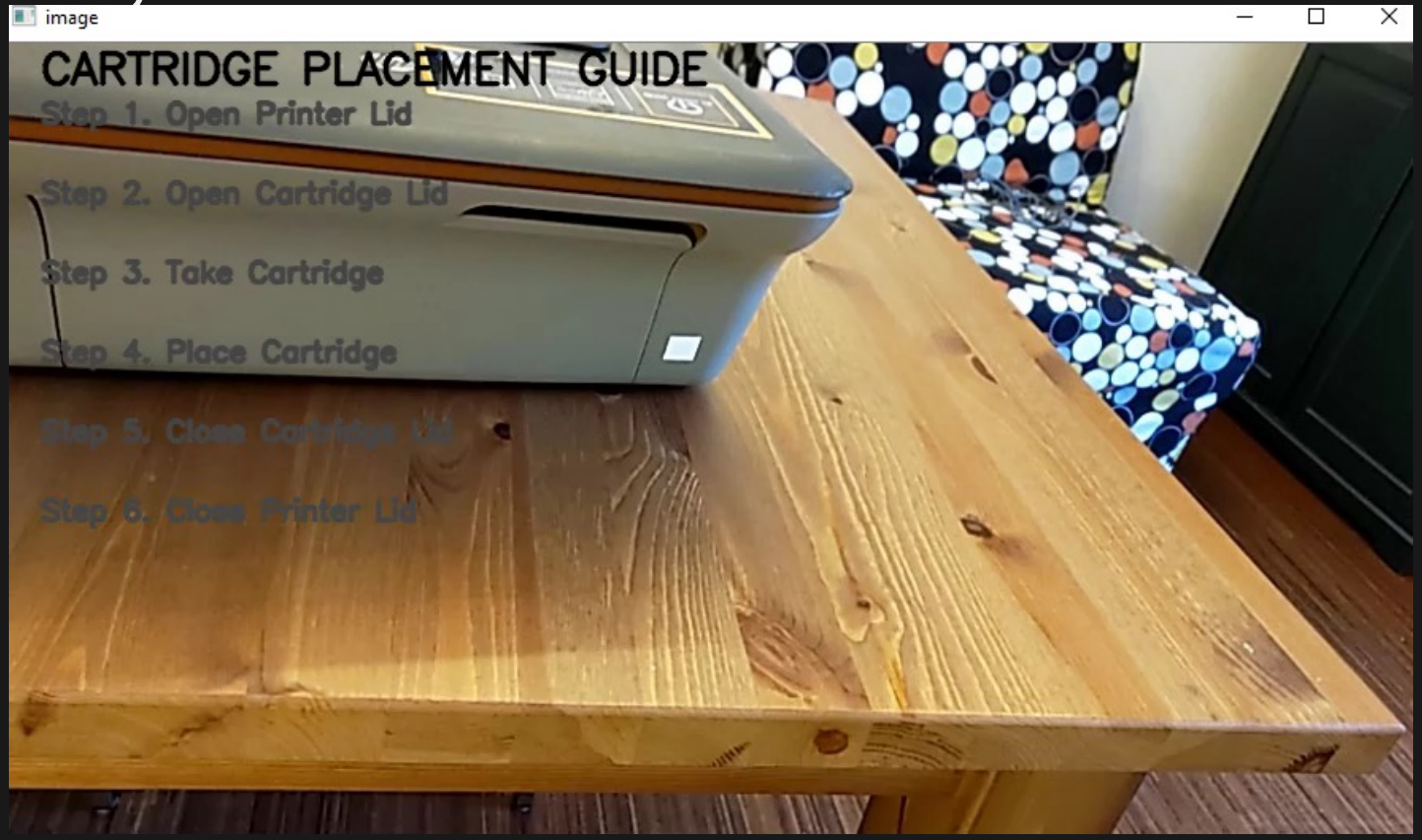
not completed

ongoing

completed

idle

missed



- Detect
  - actions & idle states
  - when we complete an action, when we move on to the next step
  - missed actions, actions not performed in order
  - the duration of the actions

Action Recognition for Automated Task Guidance

not completed

ongoing

completed

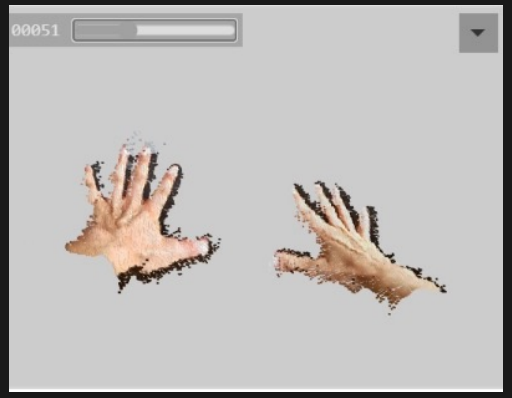
idle

missed

- Detect
  - actions & idle states
  - when we complete an action, when we move on to the next step
  - missed actions, actions not performed in order
  - the duration of the actions

HoloLens 2 as a platform for *egocentric vision*

Research Mode



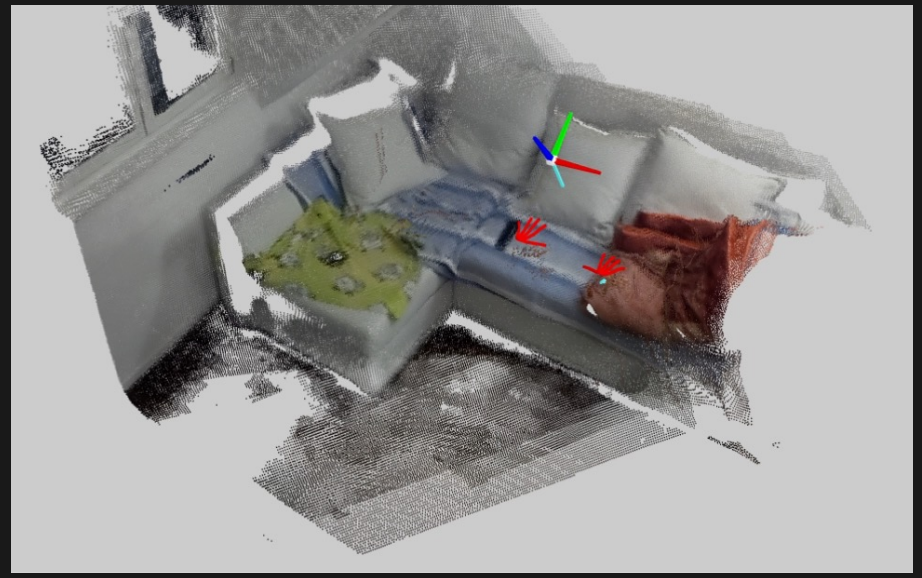
Access to the sensor streams:

- Depth – short and long throw
- Grayscale cameras (head tracking)
- IMU

https://github.com/microsoft/HoloLens2ForCV

Utilities to combine them with:

- Head tracking (6DOF)
- Hand tracking (articulated)
- Eye gaze tracking
- RGB camera



Red: hands
Cyan: eye gaze
Axes: head 6dof

Improved Training and Situational Awareness

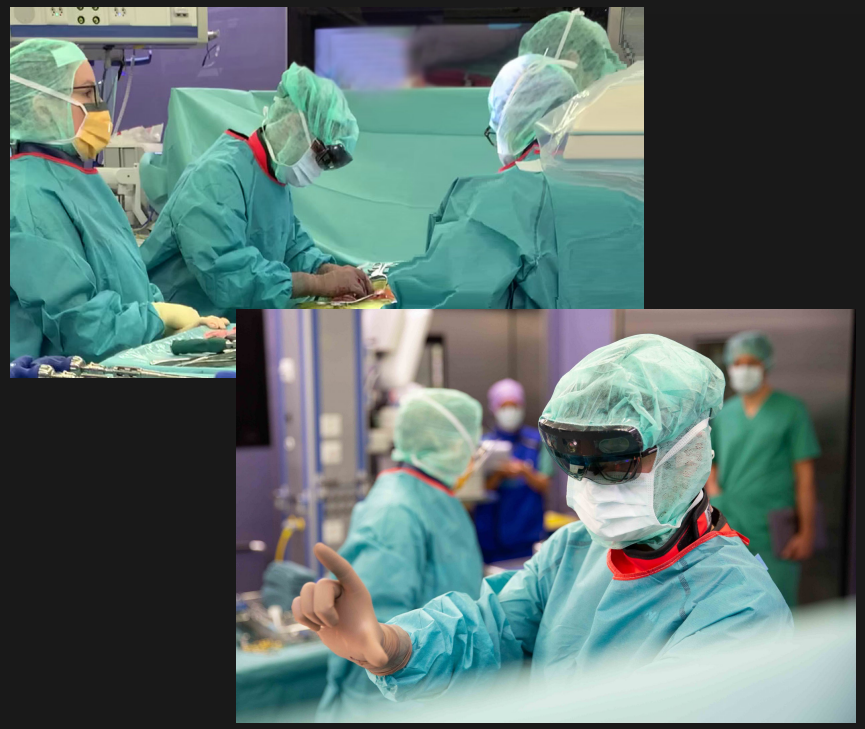
Key benefits:

- Training and mission rehearsal
- Integration of additional sensors
- Improved situational awareness (share information, 3D map, etc)
- Lowers cognitive load to absorb information
- Potential to reduce mistakes in emergency/live-and-death situations

MR technology can support military,
but also police, fire brigades, first responders, etc

Microsoft

Orthopedic spine surgery supported by MR

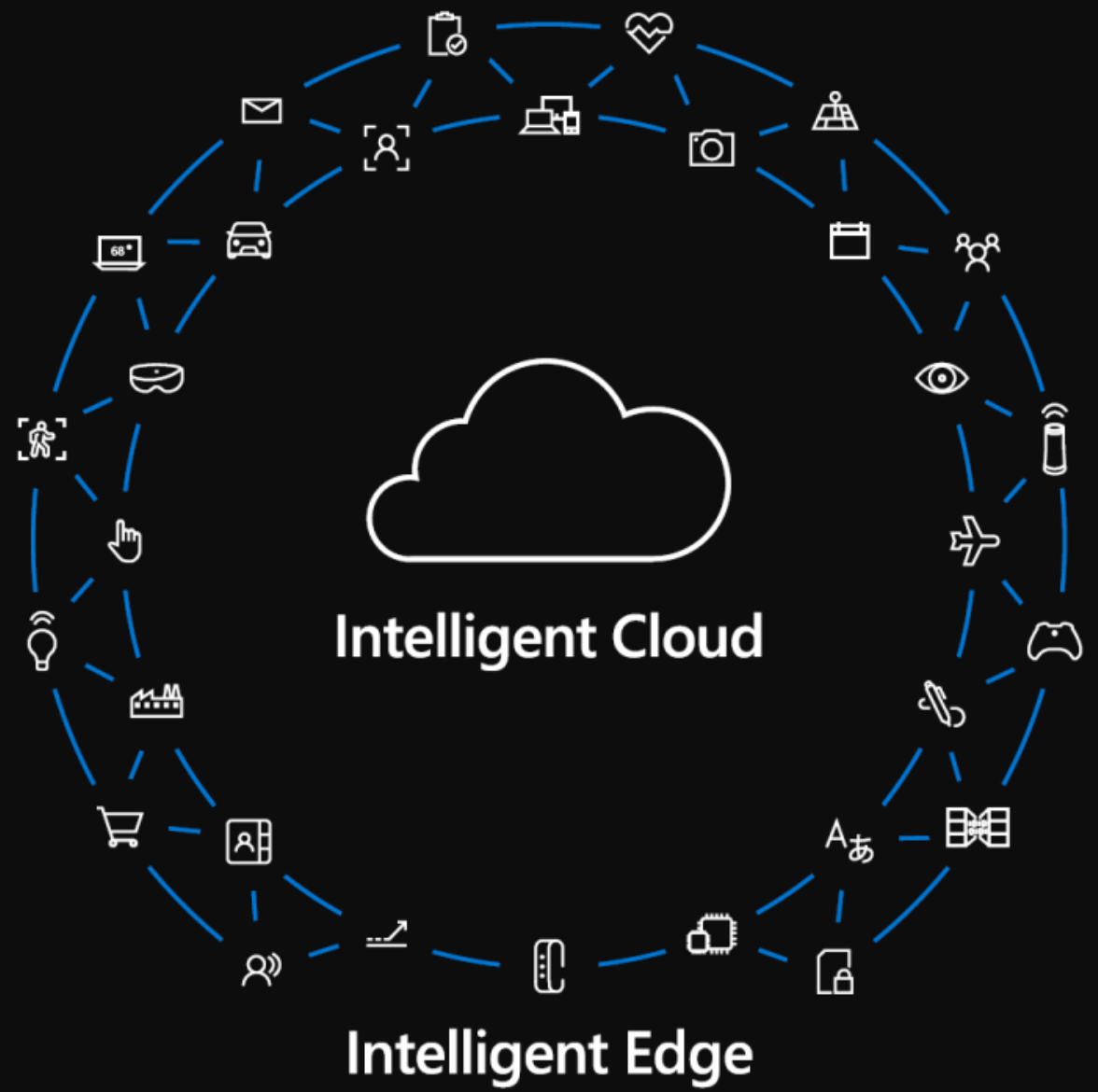


Prof. Mazda Farshad at Balgrist University Hospital Zurich

Ubiquitous computing

Artificial Intelligence

Multi-sense, multi-device experiences



Microsoft Mesh

Microsoft Mesh App (preview)

Feel presence

Feel presence

Experience together

Connect from anywhere

Microsoft Mesh

The Mesh Developer Platform

Toolkit

Capabilities

Immersive Presence

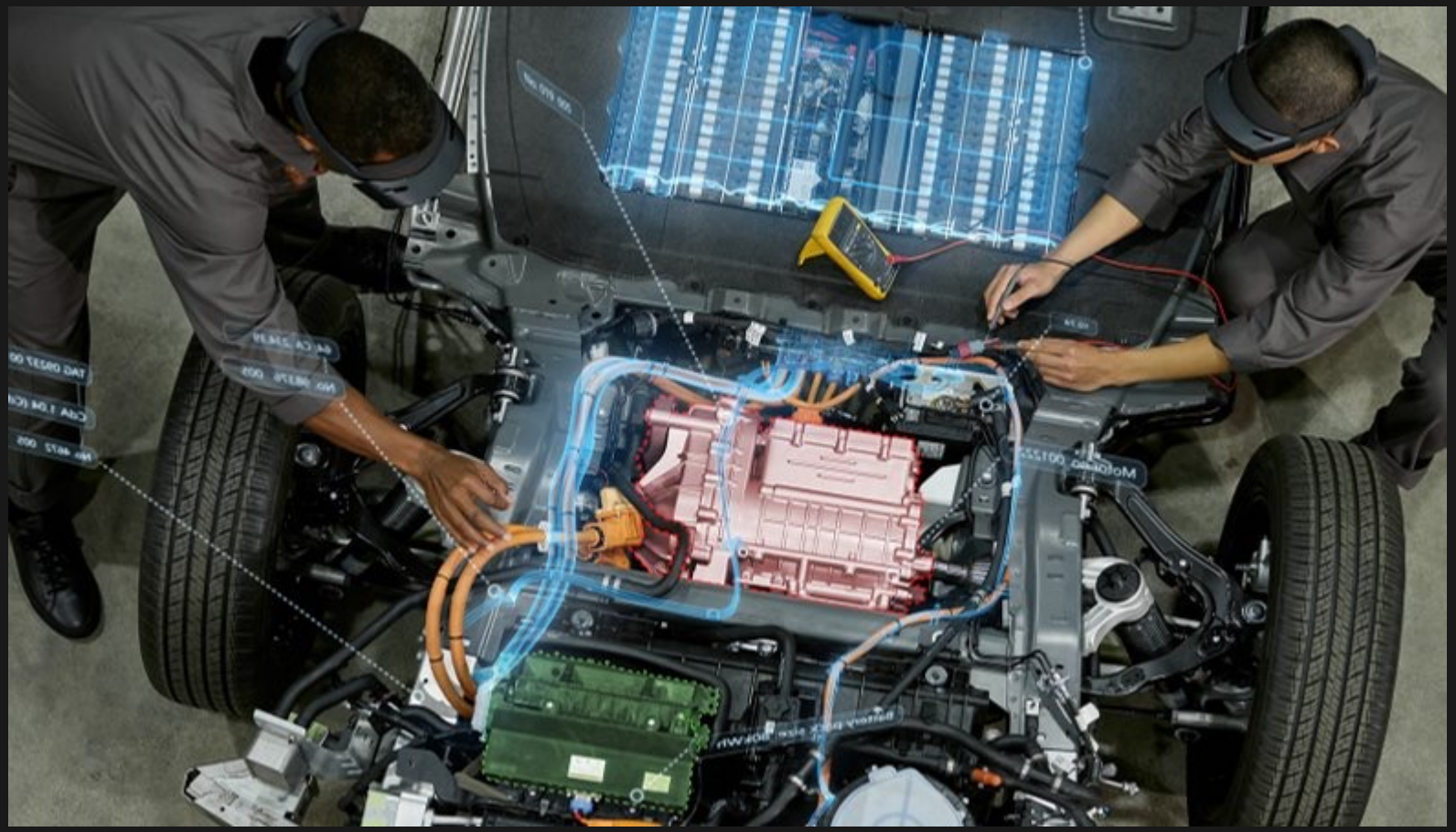
Spatial Maps

World locked holograms
Object locked holograms
Sharing and Persistence

Holographic Rendering

Multiuser sync

Core Platform



AZURE SPATIAL ANCHORS

Enables to *share* and *persist*
3D coordinate systems
across devices

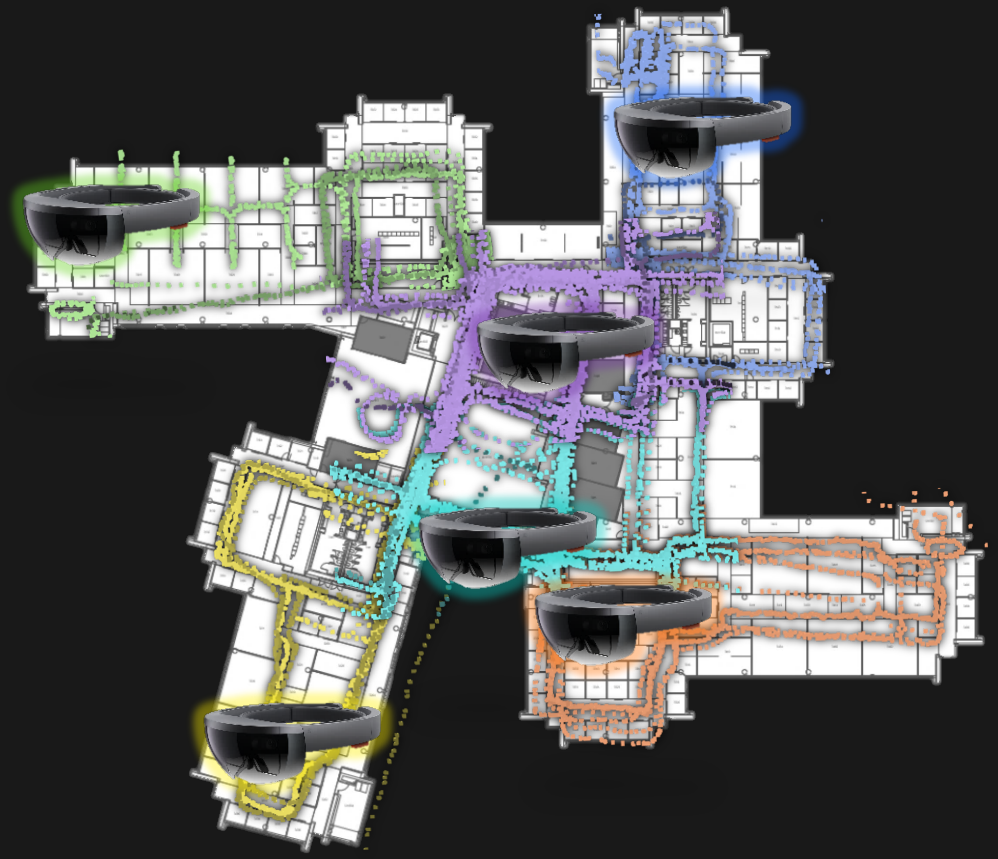

MINECRAFT
EARTH
16
64
52
20


2
Offline
00
3
Online
510
Summary


Trimble
Navigate
Align
Sequence
Model
ToDo
Tools
Spatial Map
Measure
Collaborate
Exit
Trimble
XR10
with HoloLens 2

6DOF relocalization map

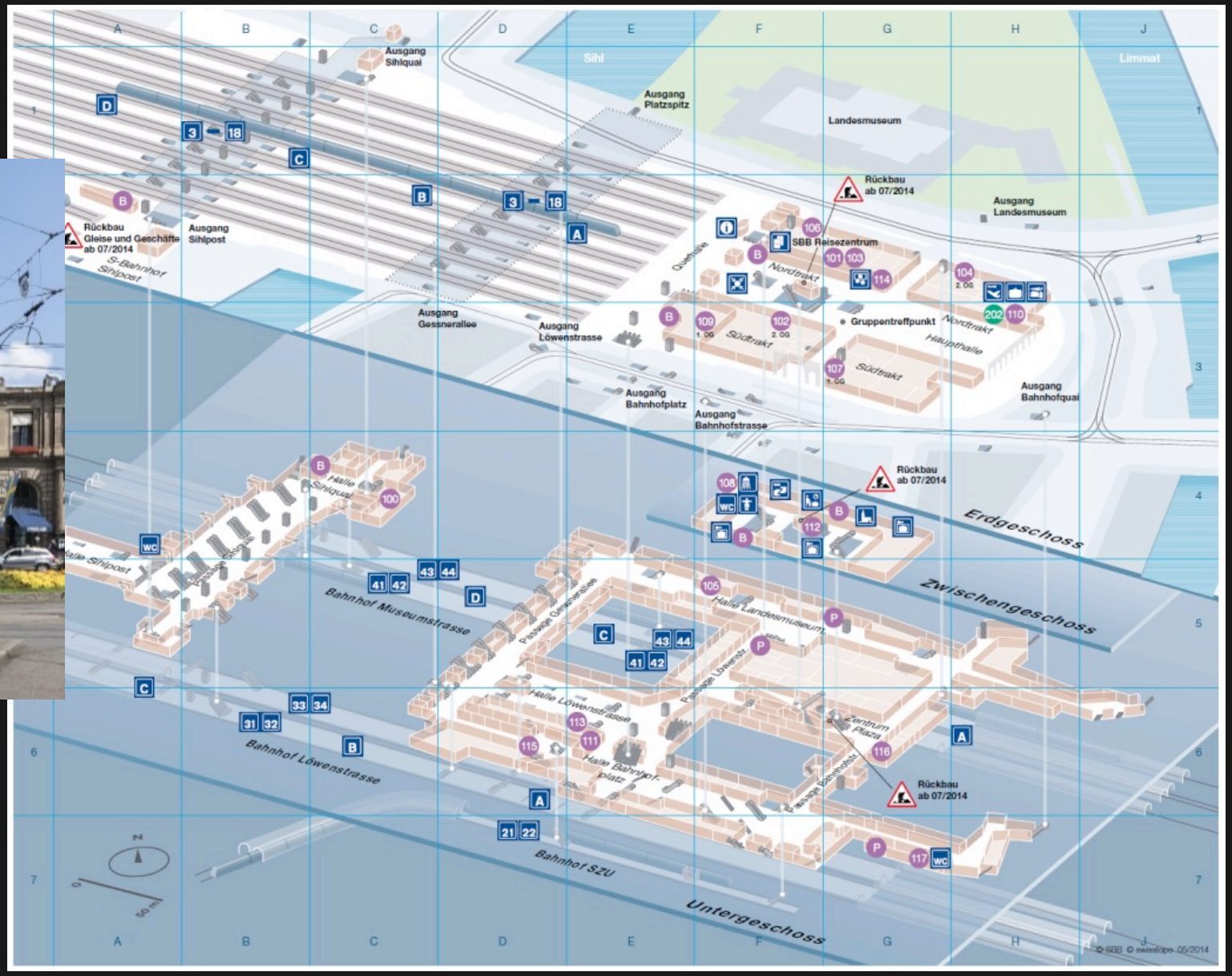
Clients

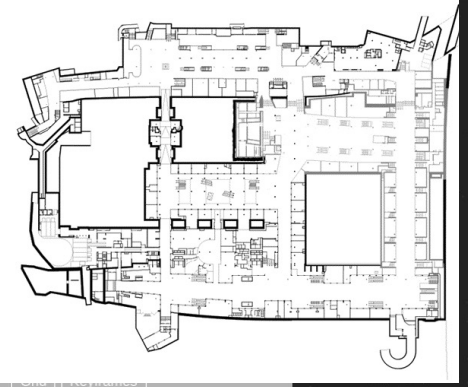


Cloud

combine multiple maps in the cloud

SBB Zurich HB





KNs

00142b2e-6c3d-4c7e-836a-d14
0029f983-fc08-4e2c-9e14-bc171
0067a516-c353-4be2-a14c-bfe3
006df6fd-3d78-4e14-b785-e5f85

Remove all

Curve size

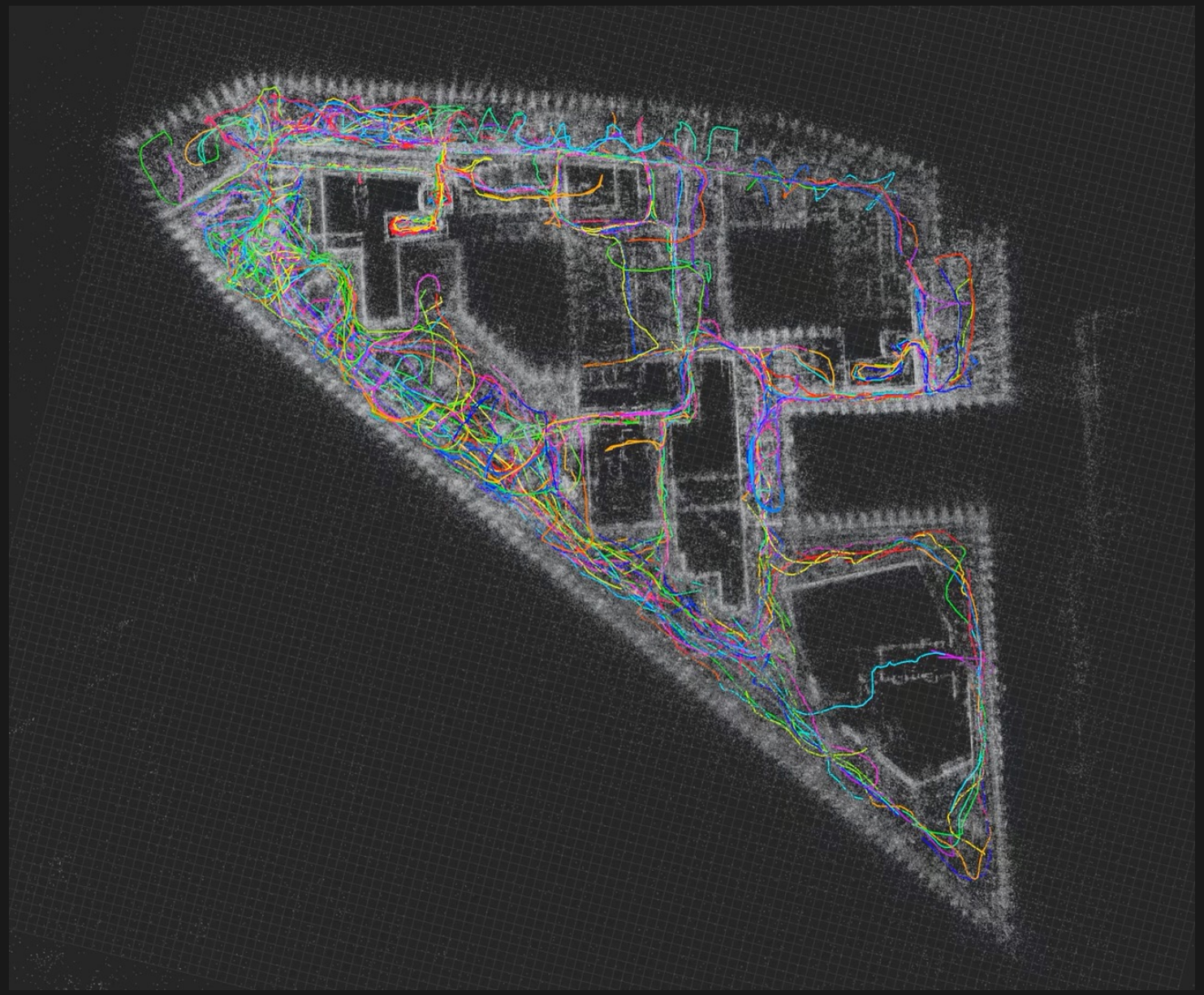
Curve opacity

Point size

Point opacity

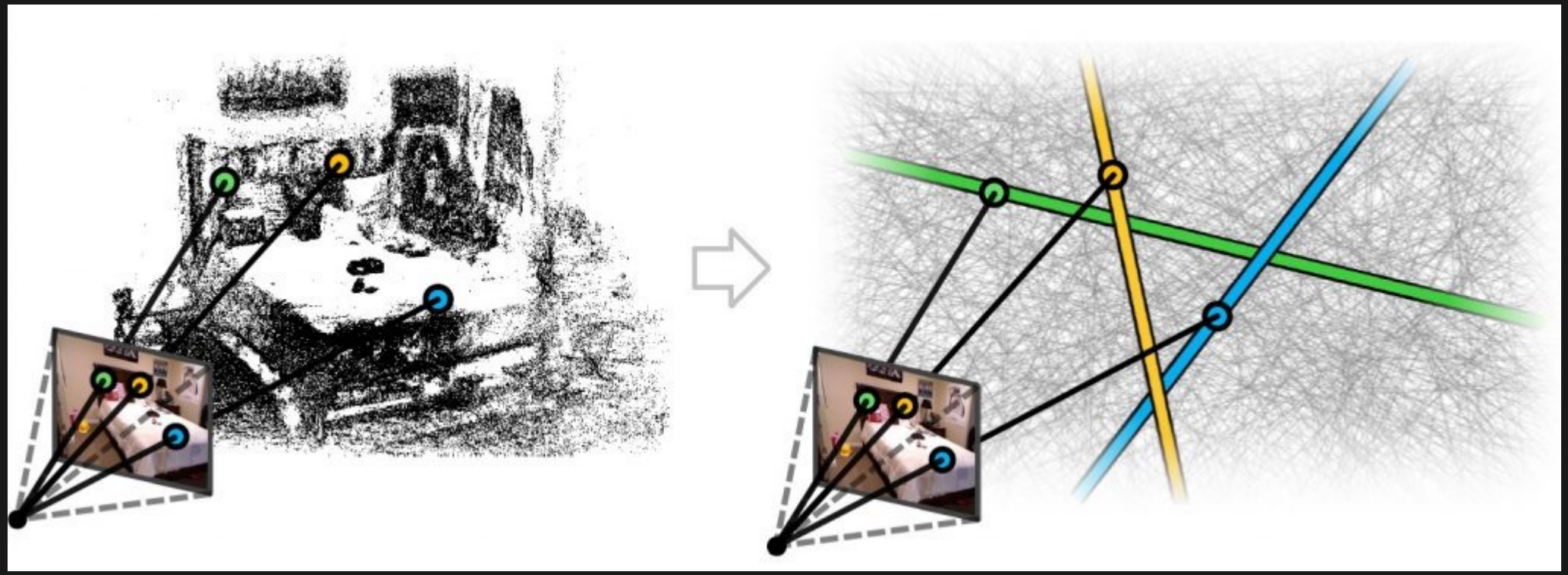
Orthographic | Perspective

The Circle (Zurich Airport)



Privacy-Preserving Image-based Localization

[Speciale et al. CVPR 2019]



3D Point Cloud
(Traditional)

3D Line Cloud Map
(Proposed)

Standard SfM [1]

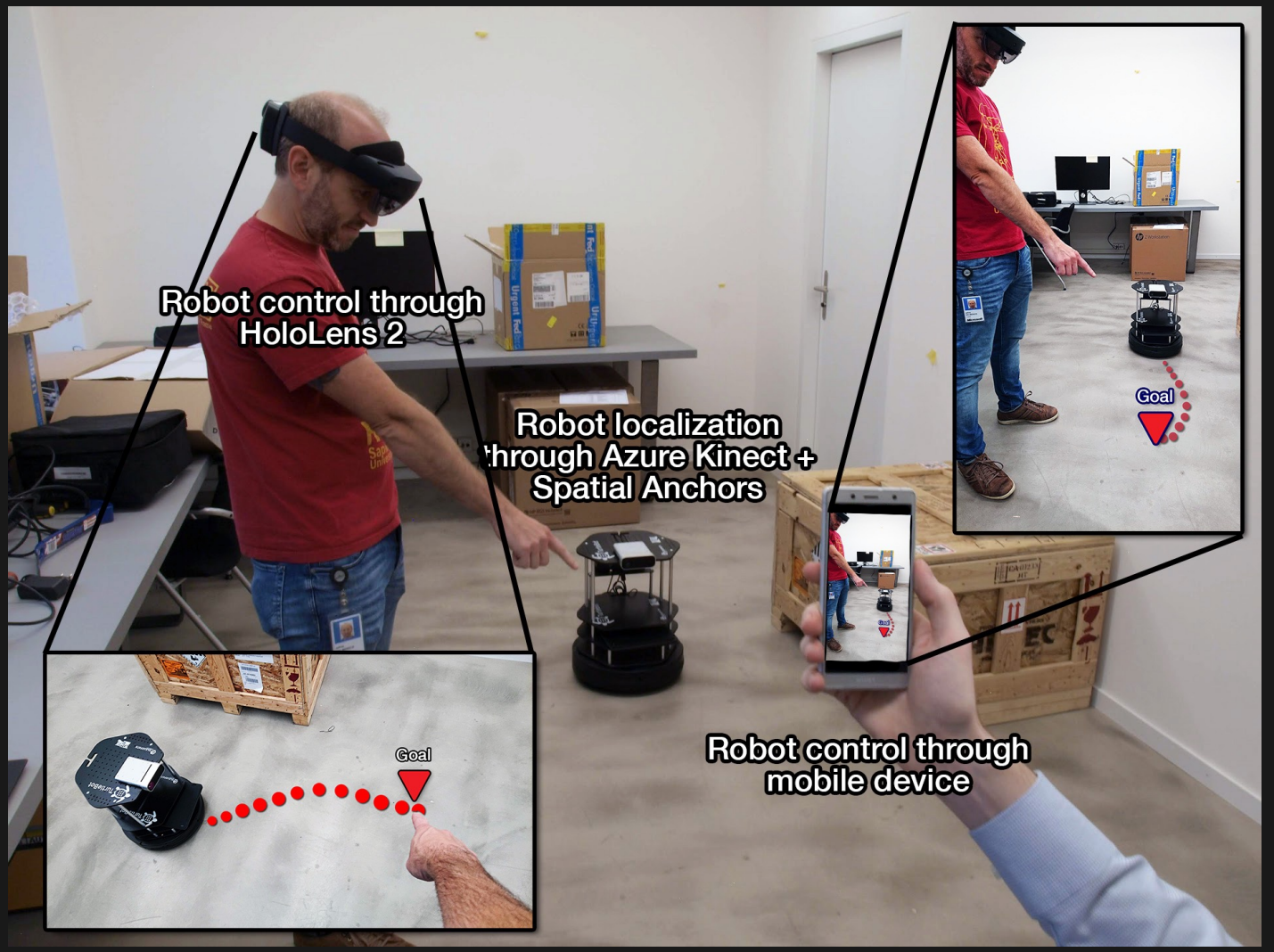
Image reconstructions [2] from keypoints

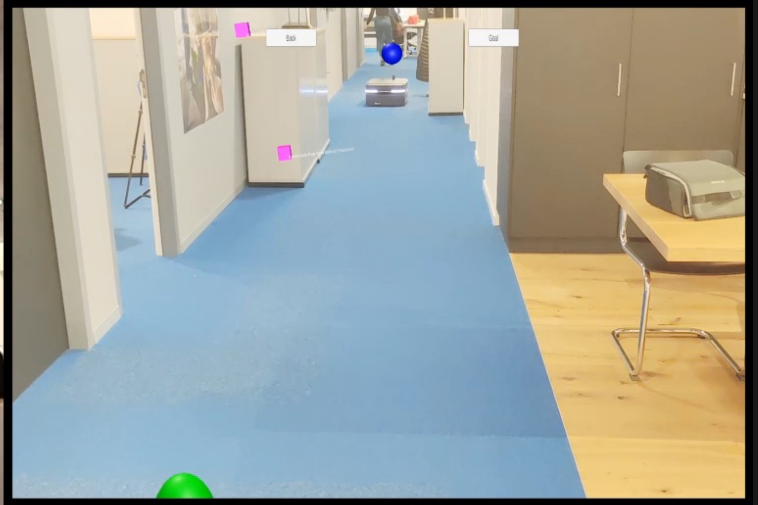
Privacy Preserving SfM

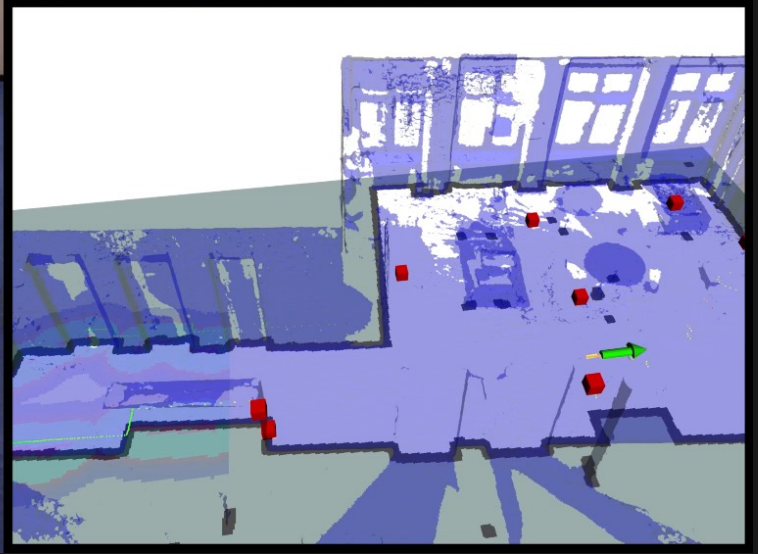
Image reconstructions [2] from pointcloud

Tower of London [3]

[1] Schönberger, Frahm CVPR '16 [2] Pittaluga et al. CVPR '19 [3] Wilson, Snavely ECCV '14


Robot control through
HoloLens 2
Robot localization
through Azure Kinect +
Spatial Anchors
Goal
Goal
Robot control through
mobile device



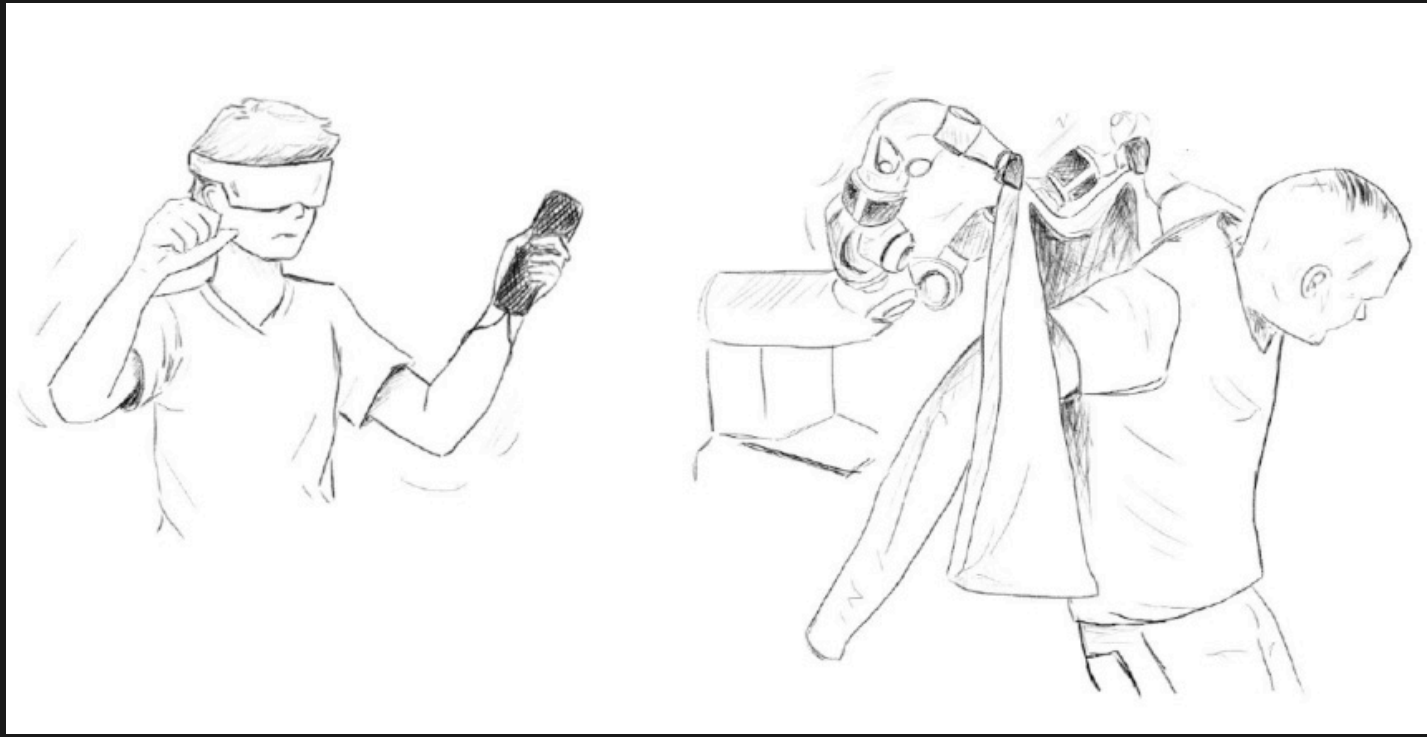


Microsoft Mixed Reality & AI lab Zurich

aka.ms/ASALinuxSDK

Mixed Reality Robot tele-manipulation

- Collaboration with Stelian Coros, Roi Poranne et al. from ETH Zurich

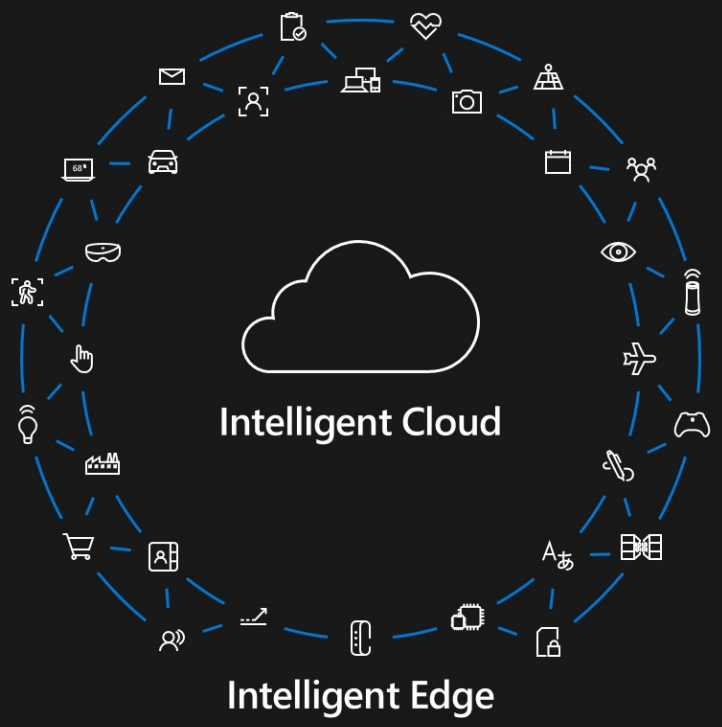



ABB

Mixed Reality & AI

- Mixed Reality headsets have potential to have much more user/task context to assist user
  - Observe user actions
  - Understand environment
  - Access relevant digital information
  - Natural user interface
  - Display information in spatial context

Opportunities & open issues



- Combine edge and cloud computing
- Tele-presence
  - Project remote presence of user
  - Allow user to access remote places
- Synergies between HMDs, robotics, ambient sensing and IoT
- Privacy

Microsoft