

Introduction to Augmented Reality

Fivos DOGANIS



Course contents

- **Theory**
 - Definitions
 - History, 3 Types of AR
 - Technologies, Calibration, Registration, Tracking
 - Rendering, Interactions
- Practice
 - Overview of existing apps and tools
 - AR Web **Programming**
 - **WebXR, AR.js and more!**

whoami

 linkedin.com/in/fivosdorganis

 fivos.dorganis@gmail.com

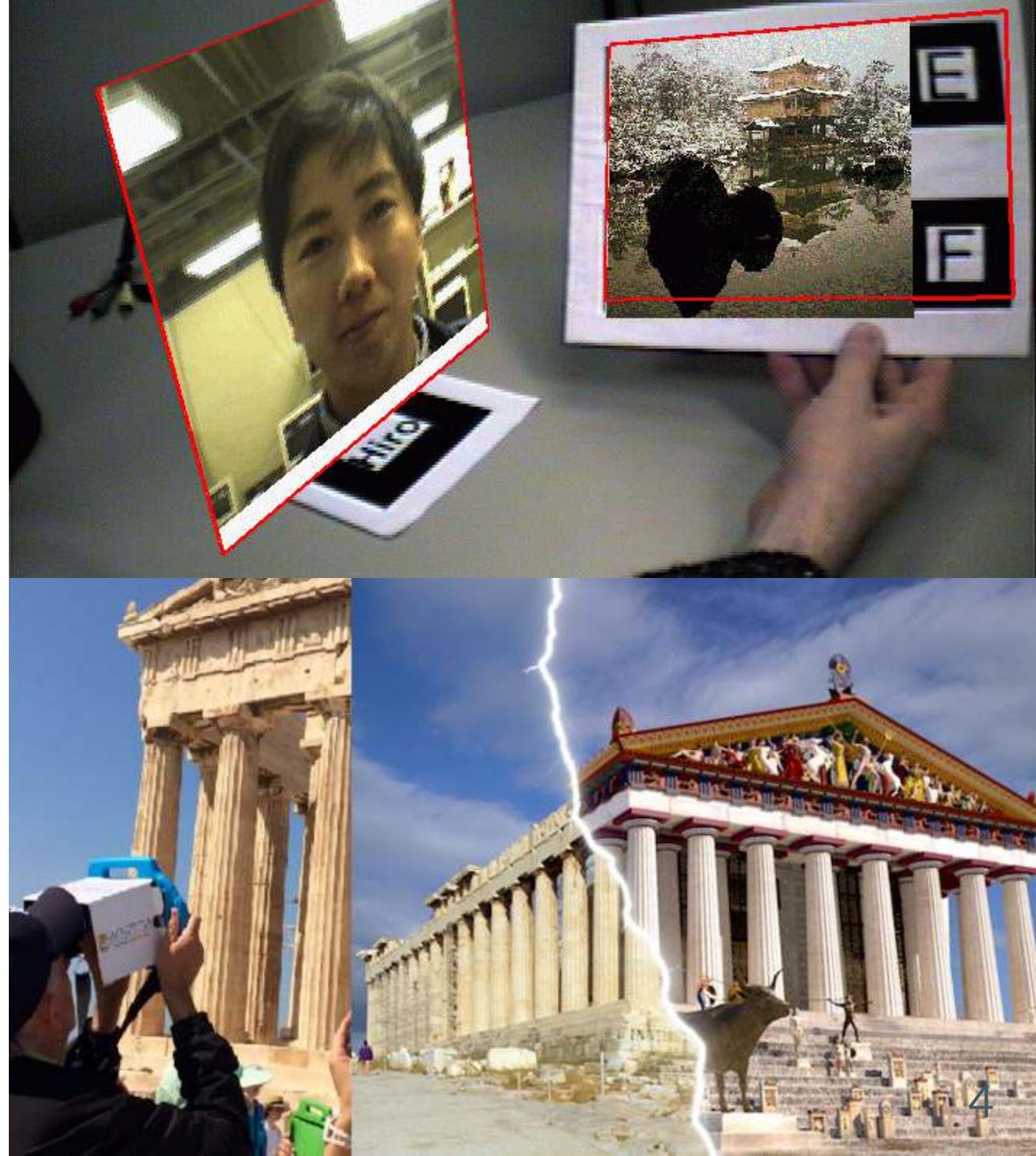
 github.com/fdorganis



UNIVERSITY OF HULL

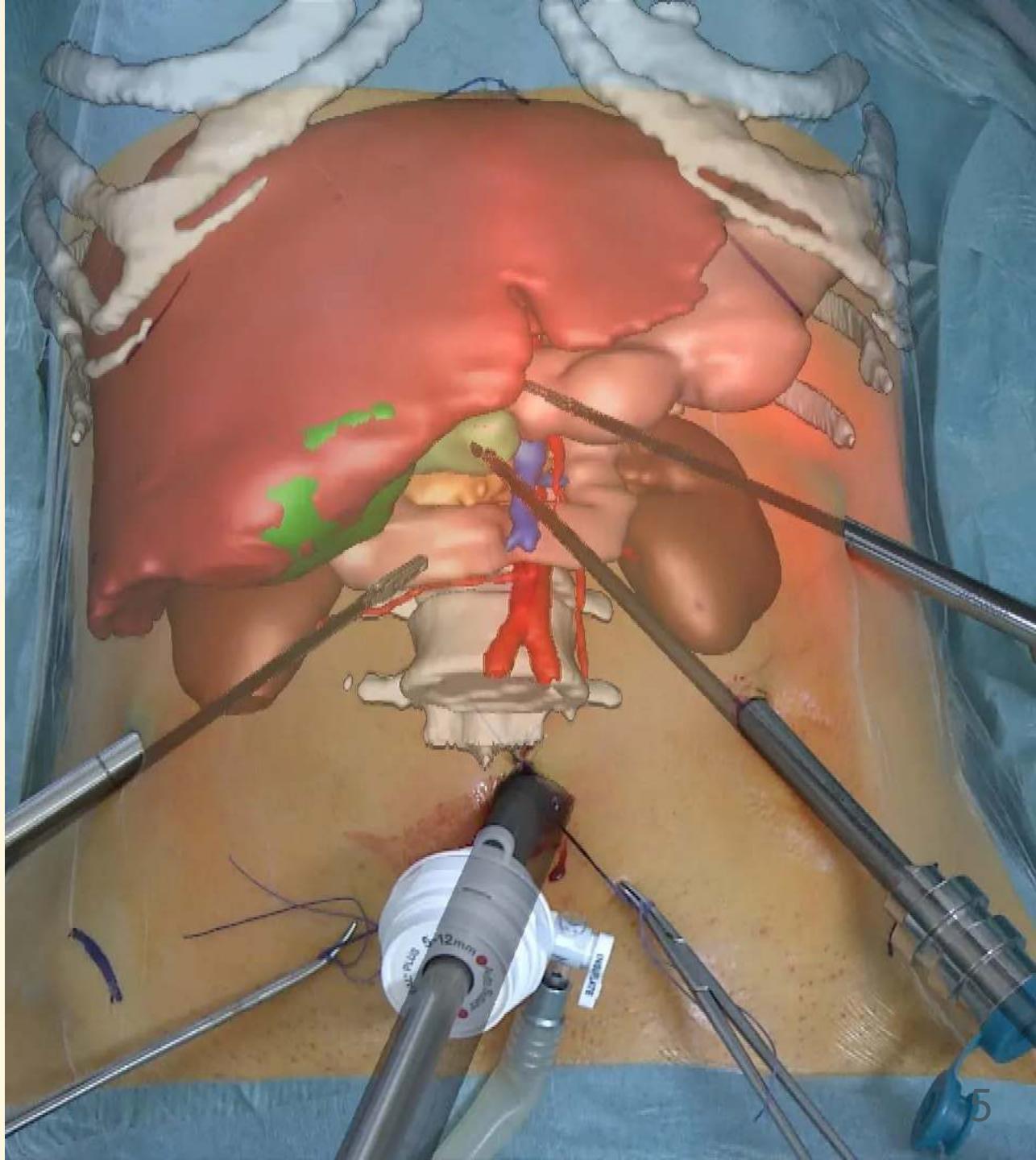
University of Hull

- Master of Science by Research (2001)
Augmented Reality in Archaeology: Registration Issues



IRCAD (2002 - 2003)

- Institut de Recherche contre les Cancers de l'Appareil Digestif
- Startup
 - Virtual-Surg team
- Augmented Reality Research Engineer



Dassault Systèmes (2003+)

- 3D Visualization Engineer
 - Scenegraph, Materials
 - Geometry, Tessellation
- Virtual and Augmented Reality (XR) Engineer
- XR Research Engineer
- XR Research Manager



Dassault Systèmes

From Shape to Life



1981
**3D
Design**

1989
3D DMU
Digital
Mock-up

1999
3D PLM
Product Lifecycle
Management



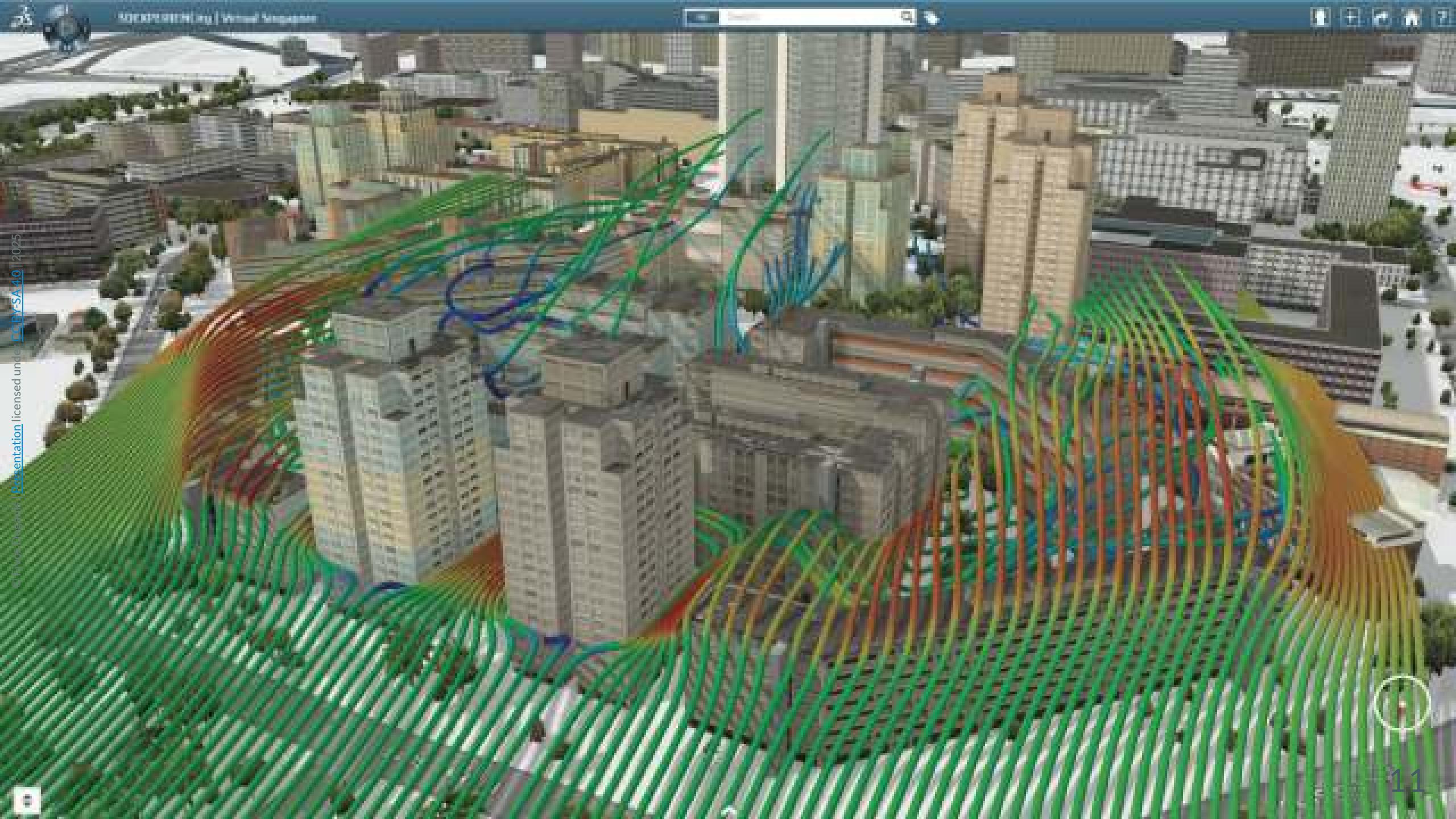
2012
**3DEXPERIENCE®
platform**

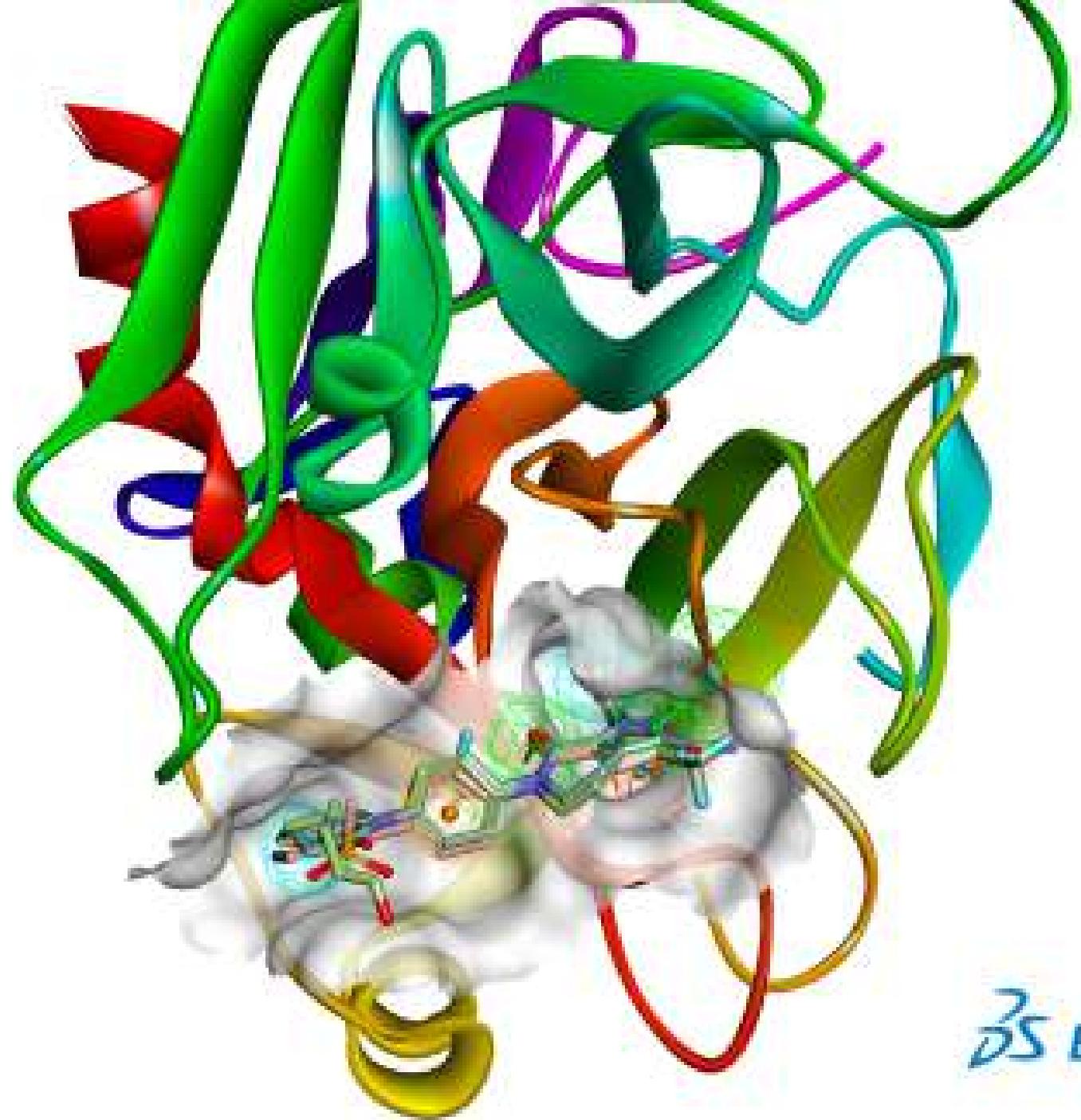
2020
**Virtual Twin
Experience of
Humans**



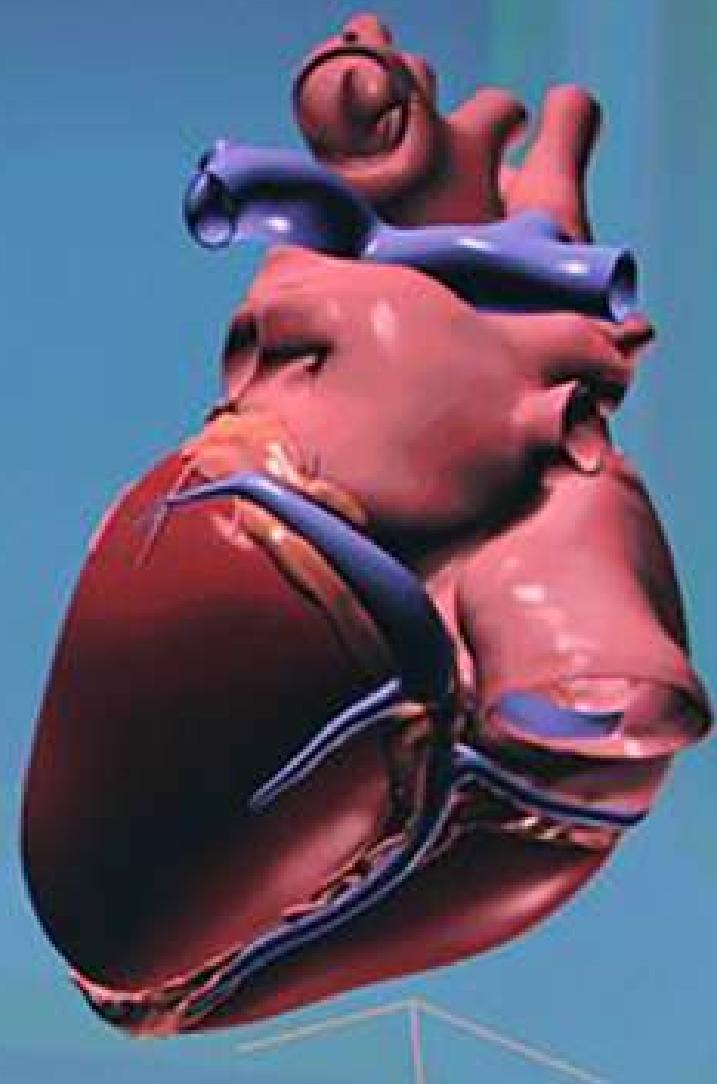








 **BIOVIA**



Course audience

- Anyone looking for a simple introduction to Augmented Reality
 - Computer Science students and engineers looking for a way to create cross-platform XR prototypes or even full-fledged apps
-  Feel free to skim through technical sections and use this course as future reference

Course prerequisites

- **Math** 
 - 3D vectors and [matrices](#)
- **Programming** 
 - **JavaScript** [notions](#), or any similar language (HTML kept minimal)
- 3D Web API 
 - **THREE.js** **notions strongly recommended** (see [Web 3D course](#))
 - alternatives: Babylon.js, WebGL, WebGPU
- Desktop + Smartphone (or, even better, a XR HMD!)  
 - [VSCode](#)

JUL
17

Planning

- **Day 1 (6 hours)**
 -  **Theory**
 -  Lunch
 -  **WebXR Theory + exercises**
 -  Explore examples + choose a personal **project**
- **Day 2 (6 hours)**
 -  Evaluation: **Quizz** 20 questions / ~20 min → 20 points max
 -  **Personal project / game jam** → +5 bonus points!

Project evaluation criteria

Bonus points for:

- originality 
- interactions 
- physics  / animations  / sounds  / eye-candy 
- GIS 
- code quality , tricks , performance 
- fun 
- **clever use of AR**  

Project grading

- up to **5** bonus points
- choose features from previous slide
- for each implemented **feature**:
 - not done: **0 pt** 🤪 zzz
 - nice try / buggy: **0.5 pt** 😐 🐛
 - basic / good enough: **1 pt** 😊
 - great / polished: **1.5 pts** 😃
 - impressive: **2 pts** 😎 ⭐



AR Applications

Consumer applications

- Entertainment
 - Games: [Pokemon GO AR+](#), Minecraft Earth ([RIP](#) 💀)
 - Social Networks: SnapChat ([City Painter](#)), Facebook
- Interior Design
 - [IKEA Place](#)
 - [HomeByMe AR](#)
- [IGN: Time Machine](#)



Pokemon GO AR+

Mewtwo / CP 2247



9x

Minecraft Earth



CHRISTIAN MARCLAY X SNAPCHAT
PLAYING POMPIDOU
SnapChat

© Fivos Doganis | Presentation licensed under [CC BY-SA 4.0](#) | 2025



IKEA Place



HomeByMe

IGN Time Machine

Professional applications

- Industry
- Healthcare
- Marketing

Renault Trucks



Zeal AR



Alain Affelou

Definitions

Reality

Virtual Reality

Augmented Reality

Mixed Reality

Extended Reality

Metaverse

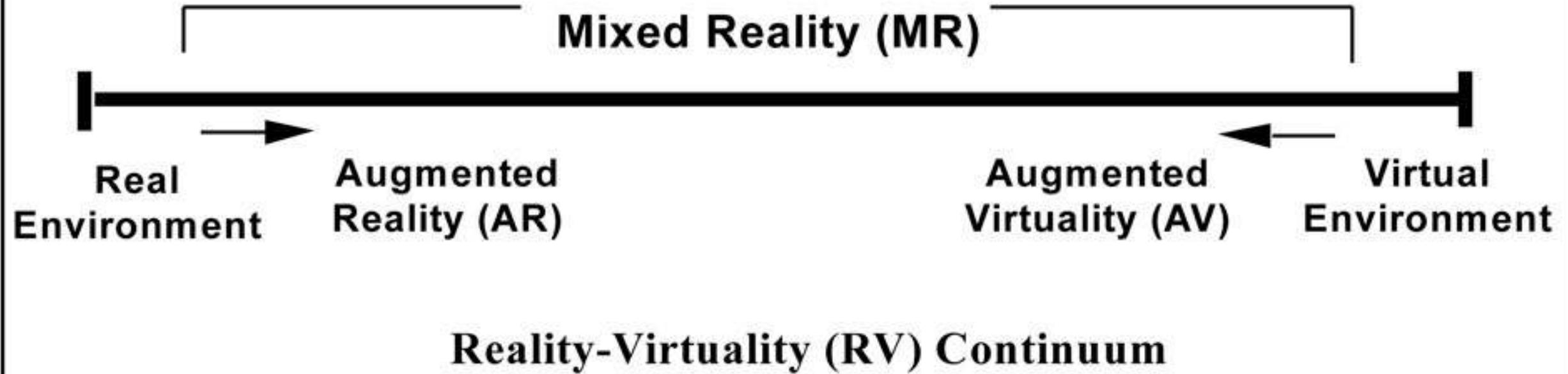
Spatial Computing

...

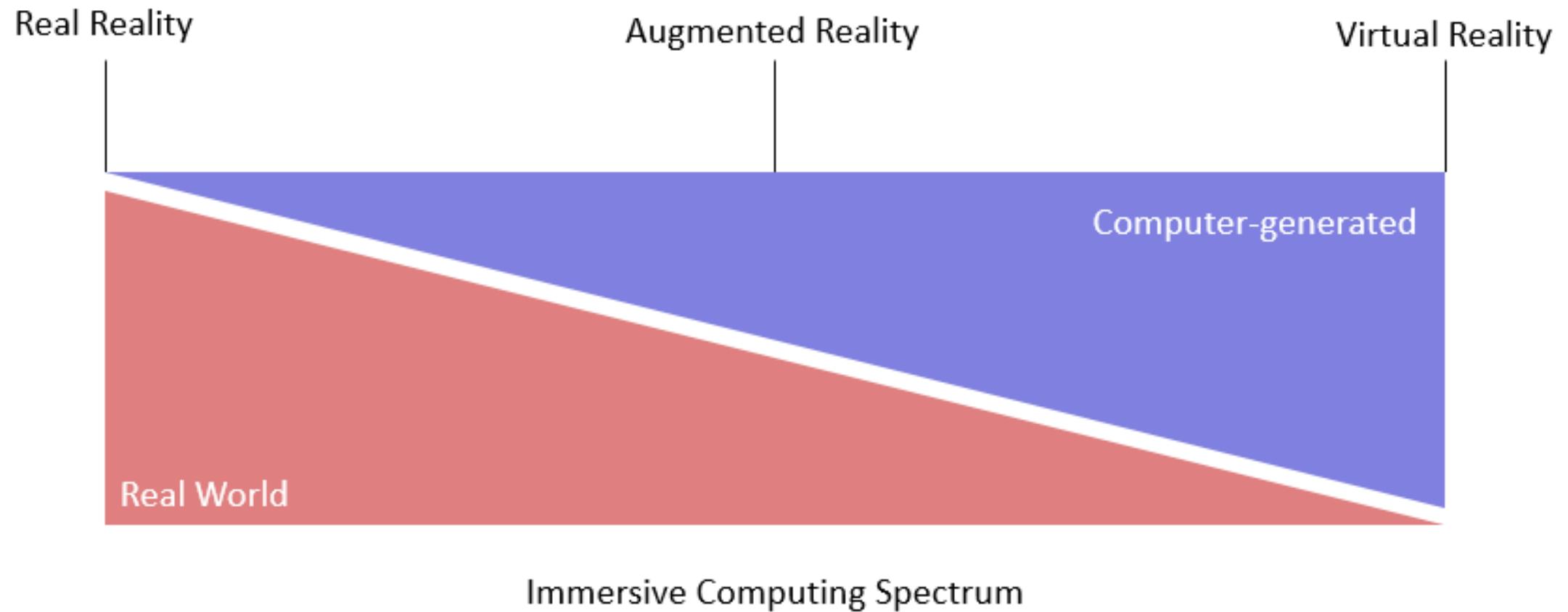
Definitions

- Milgram's **Reality-Virtuality Continuum**

Milgram, Paul; H. Takemura; A. Utsumi; F. Kishino (1994). "*Augmented Reality: A class of displays on the reality-virtuality continuum*".



Google's version



Properties of an AR system ★

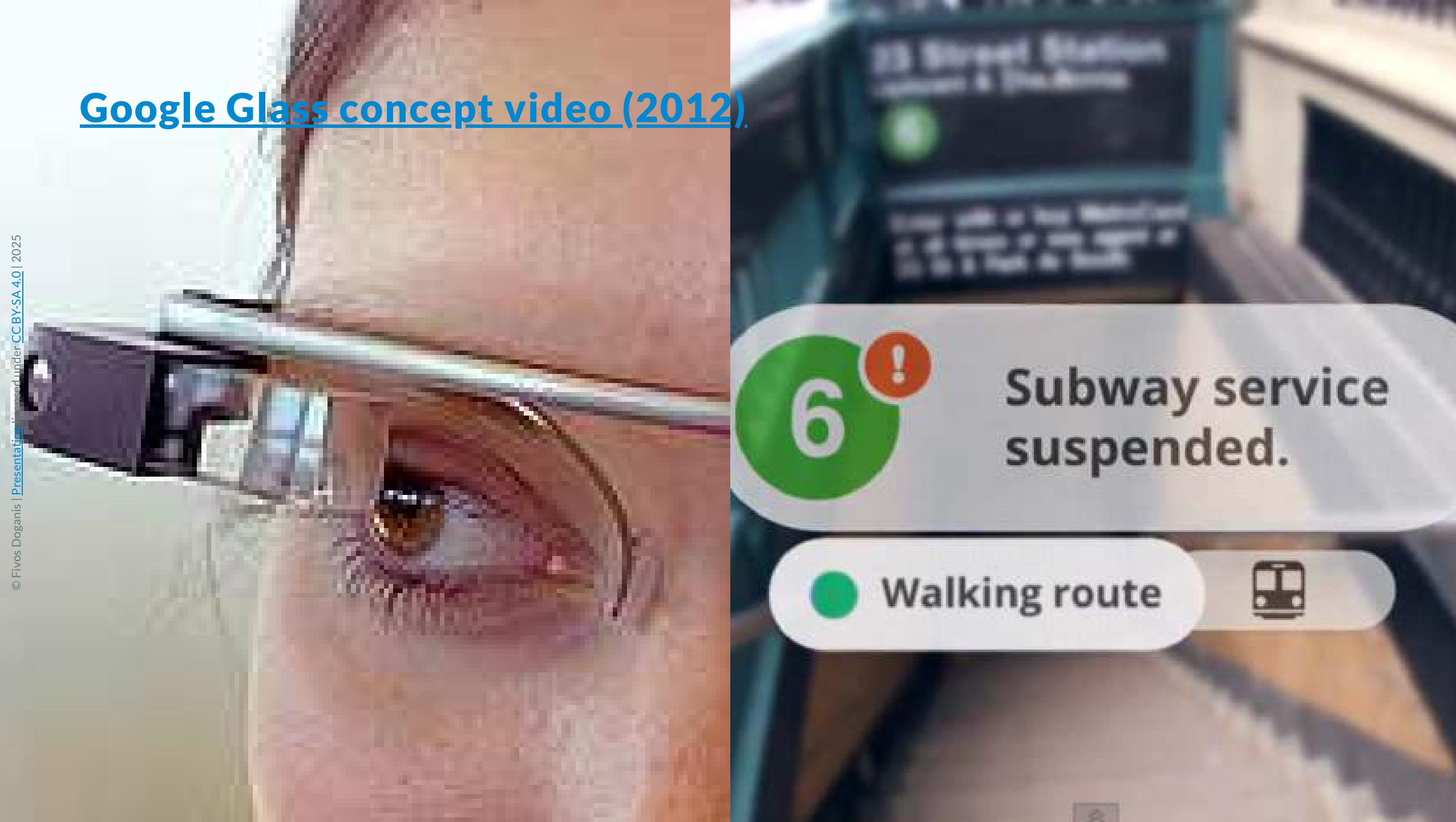
(according to [Azuma, 2001](#)).

- **combines real and virtual objects in a real environment**
- runs **interactively**, and in **real time**
- **registers** (aligns) real and virtual objects with each other.

Not AR:

- special effects in movies
 - technology close to AR
 - **not real time** ✗
 - not in a real environment ✗
- Google Glass
 - combines real and virtual objects in a real environment ✓
 - **no registration** ✗
 - it's a **HUD (Head-Up Display)**
 - can still be useful! (maintenance, sports etc.)

Google Glass concept video (2012)



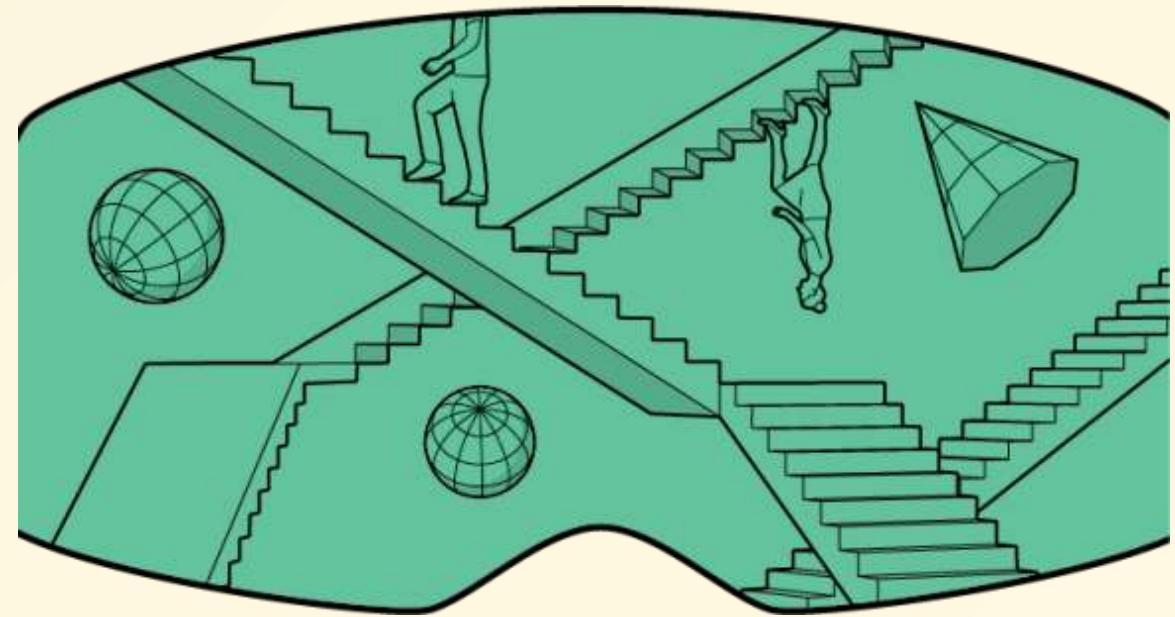
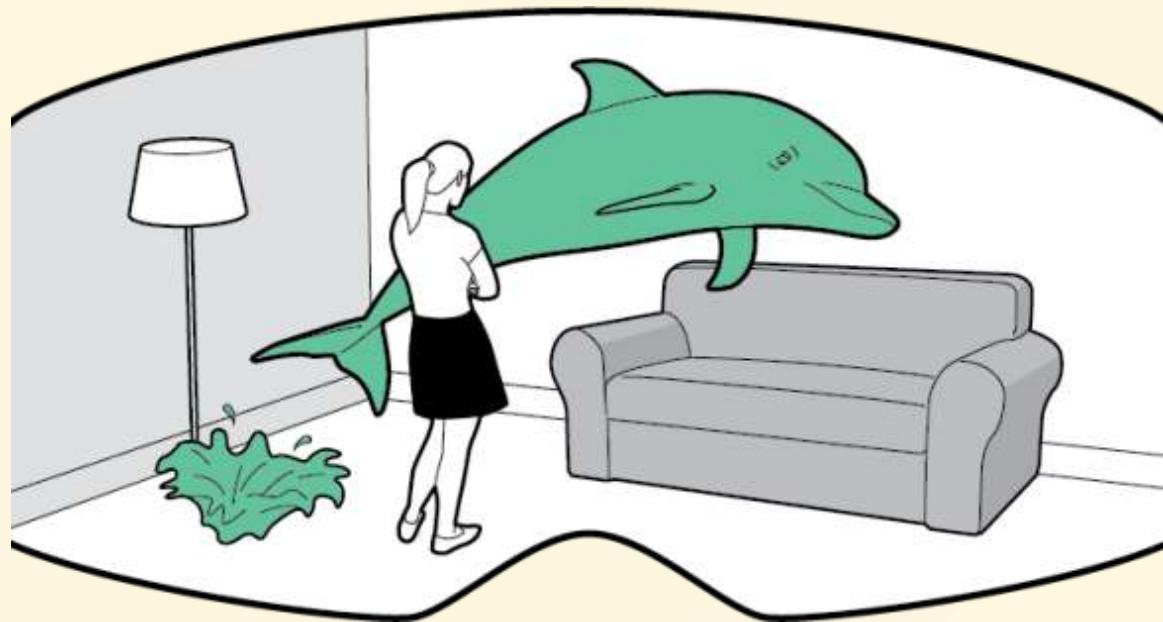


This is not AR!!!

Definitions ★

- **VR** : Virtual Reality [Jaron Lanier, 1987](#)
- **AR** : Augmented Reality [Thomas P. Caudell, 1990](#)
- **MR** : Mixed Reality
 - marketing term used by Microsoft
 - **⚠ no clear definition!** ➡ Term must be defined before use!
 - cf. [What is Mixed Reality?](#)
- **XR** : X = { eXtended / Cross (+) / Any (*) / A+V } Reality
 - **recent** generic term which **encompasses AR and VR**

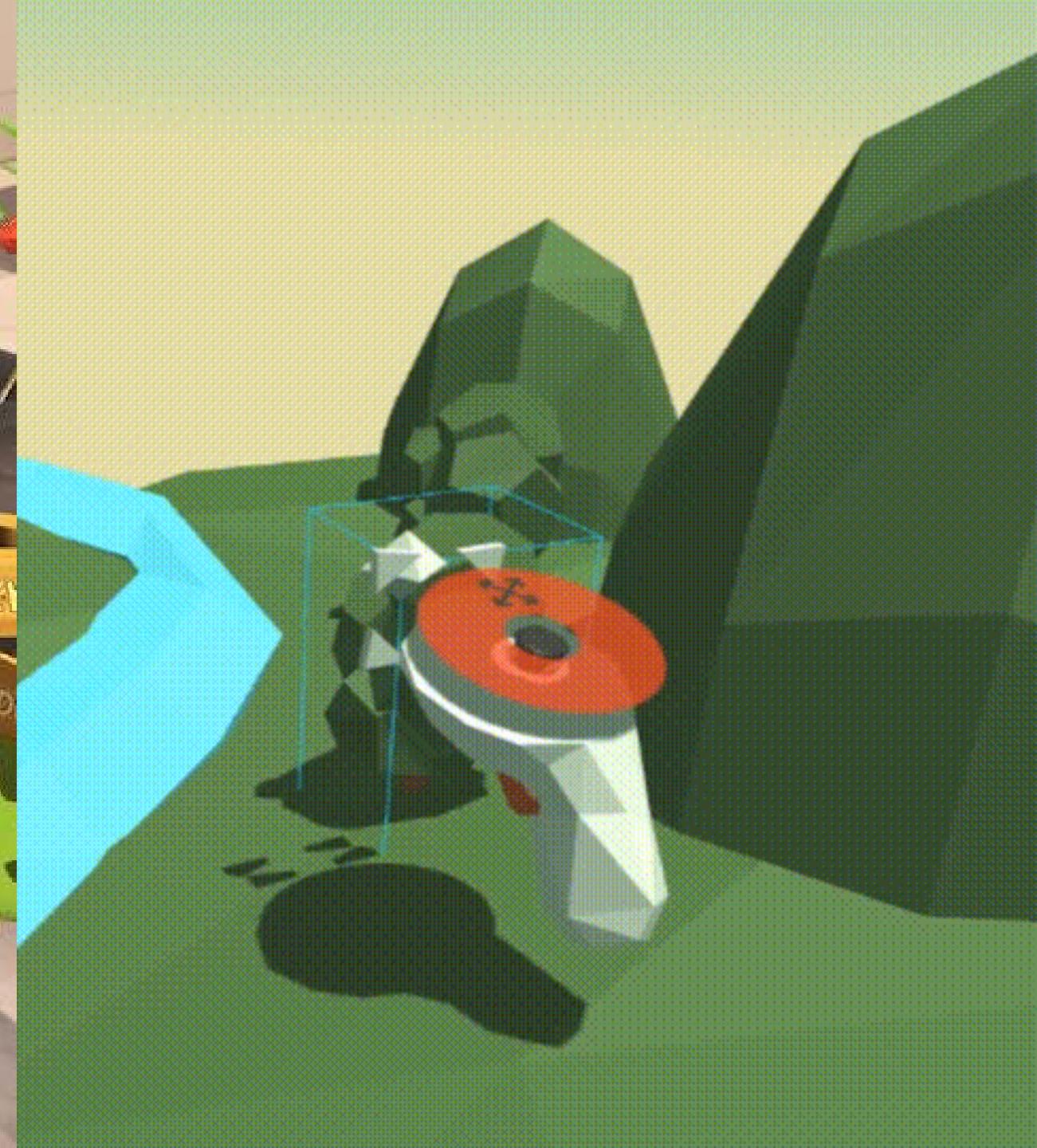
AR | VR



images: <https://www.wired.com/2016/04/magic-leap-vr/>



AR | VR



AR or VR?

- Similar technologies
 - 3D rendering
 - Tracking
 - Immersive interactions
- **Different effects** on the user

Effects of VR

- **Isolates** the user from the real world
- **Teleports** the user to another world, which is entirely virtual

Tiltbrush



Tilt Brush
by Google



The limits of VR

- Reminder: **continuum!**
 - No clear boundaries
- When the whole world is modeled and registered in 3D,
is it still VR?
- Photogrammetry / Lightfields / NeRF / Gaussian Splats
 - VR but immersion in a world entirely rebuilt in 3D

Greg Madison @ Unity

VersaillesVR

La galerie des Glaces

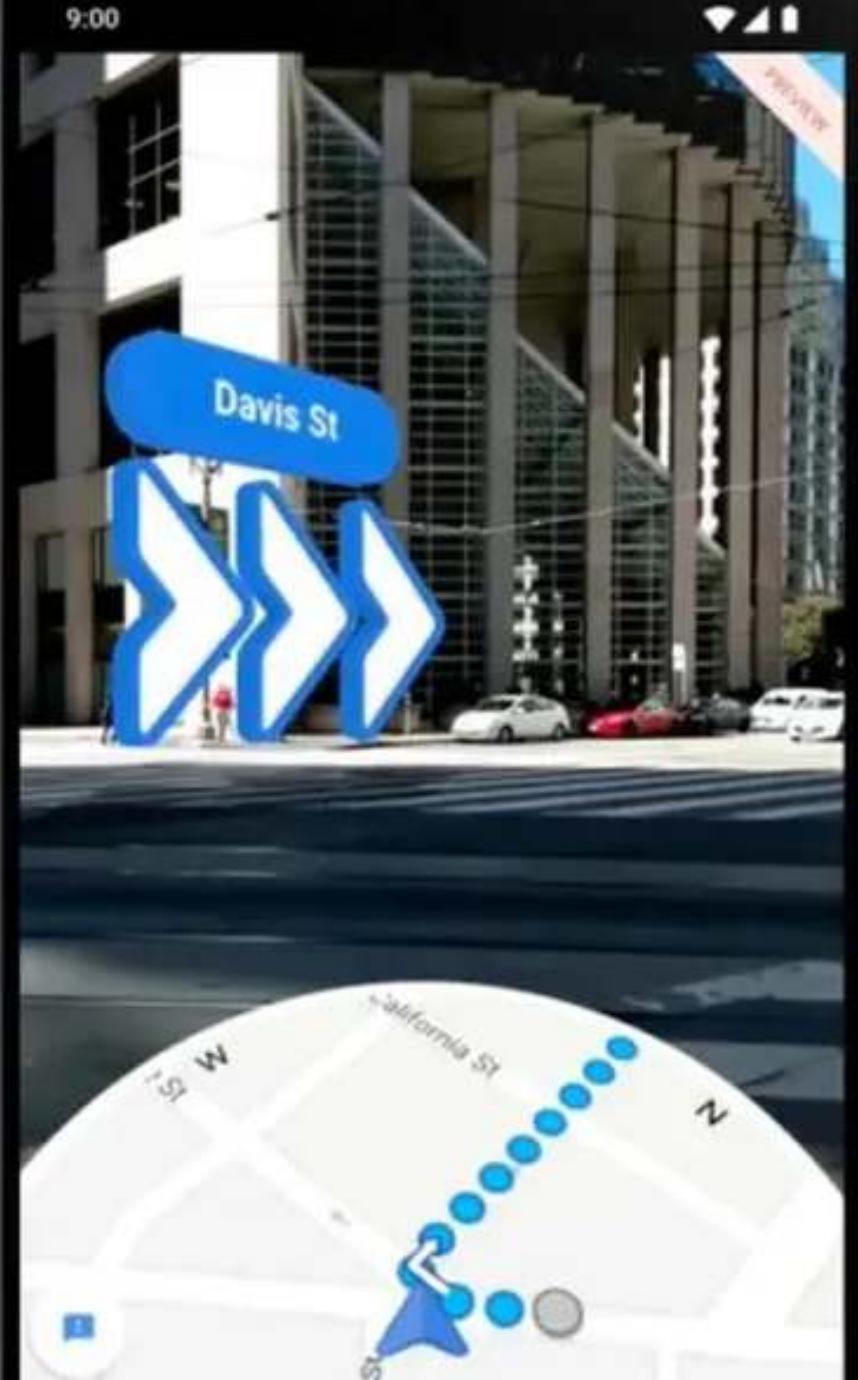


Effects of AR

- The user **stays in the real world**
- AR **enhances** the real world with contextual information
- Augmented user: acquires **new senses!**
- Information becomes visible
 - spatialized **information overlaid on top of the real world**

Google Maps AR

The browser of the future? 🤔



Brake fluid

Audi AR manual (Metaio, 2012)



The limits of AR

- Reminder: **continuum!**
 - No clear boundaries
- When more virtual elements than real ones: **Augmented Virtuality**
 - Window to the real world
 - Real users visible

Augmented Virtuality

Varjo Teleport

video



Dangers of AR

- **Information overload:** Hyper-reality by Keiichi Matsuda
 - Diminished Reality desirable?
- **Excessive assistance, altered behaviors, surveillance**
 - Black Mirror: Nosedive
- **Digital divide**
 - Some people will feel **handicapped**, missing a sense, **daltonians**
- **Privacy:** Cloud Wars
 - MAMAA* vs Open AR Cloud



Hyper-reality (concept)

Black Mirror (fiction)



Jack
3.7₉₇

Scene Responsiveness (Meta 2023)

[paper](#) [video](#)

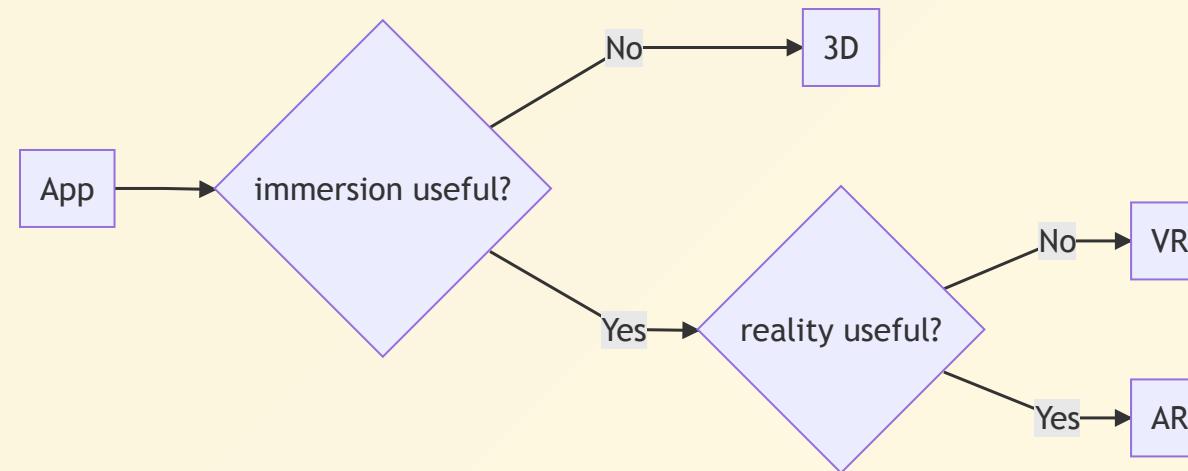
Takeaways

- VR **immerses** the user in a **virtual world**
- AR **brings virtual objects** into **the real world**

Choosing the right paradigm

- Immersion useful ?
 - Yes ➔ VR
 - No ➔ 3D
- Immersion and real environment useful ?
 - Yes ➔ AR
 - No ➔ VR
- Keep in mind continuum to pick the right paradigm to create the best possible experience

Choosing the right paradigm



History

Understand technological evolutions to anticipate the future



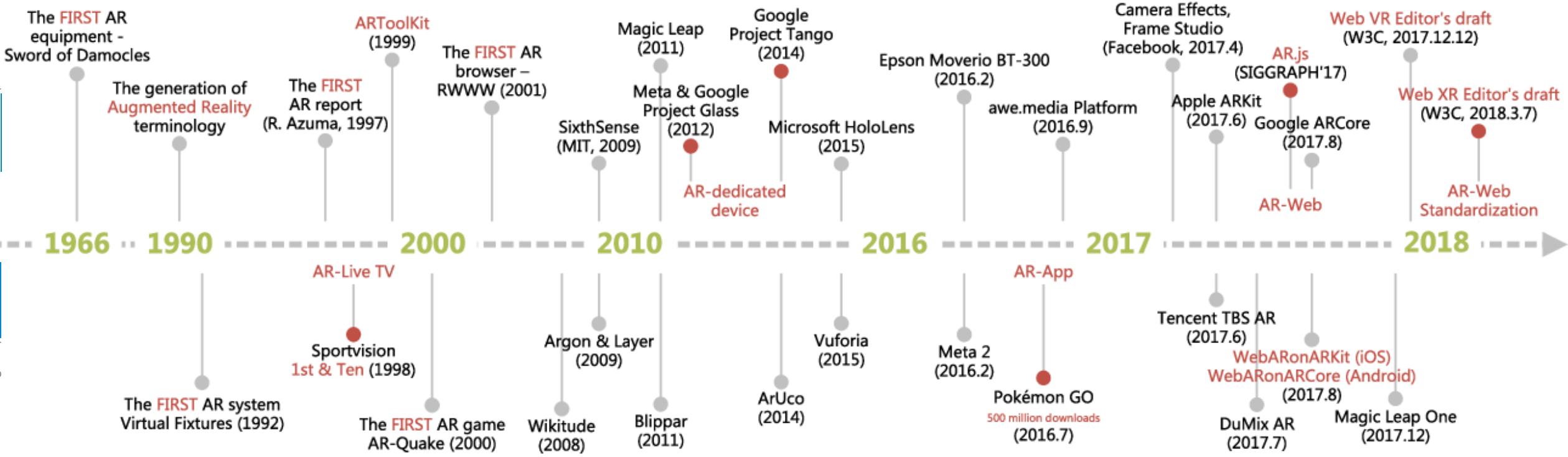


Fig. 1. *Historical evolution of AR.*

History

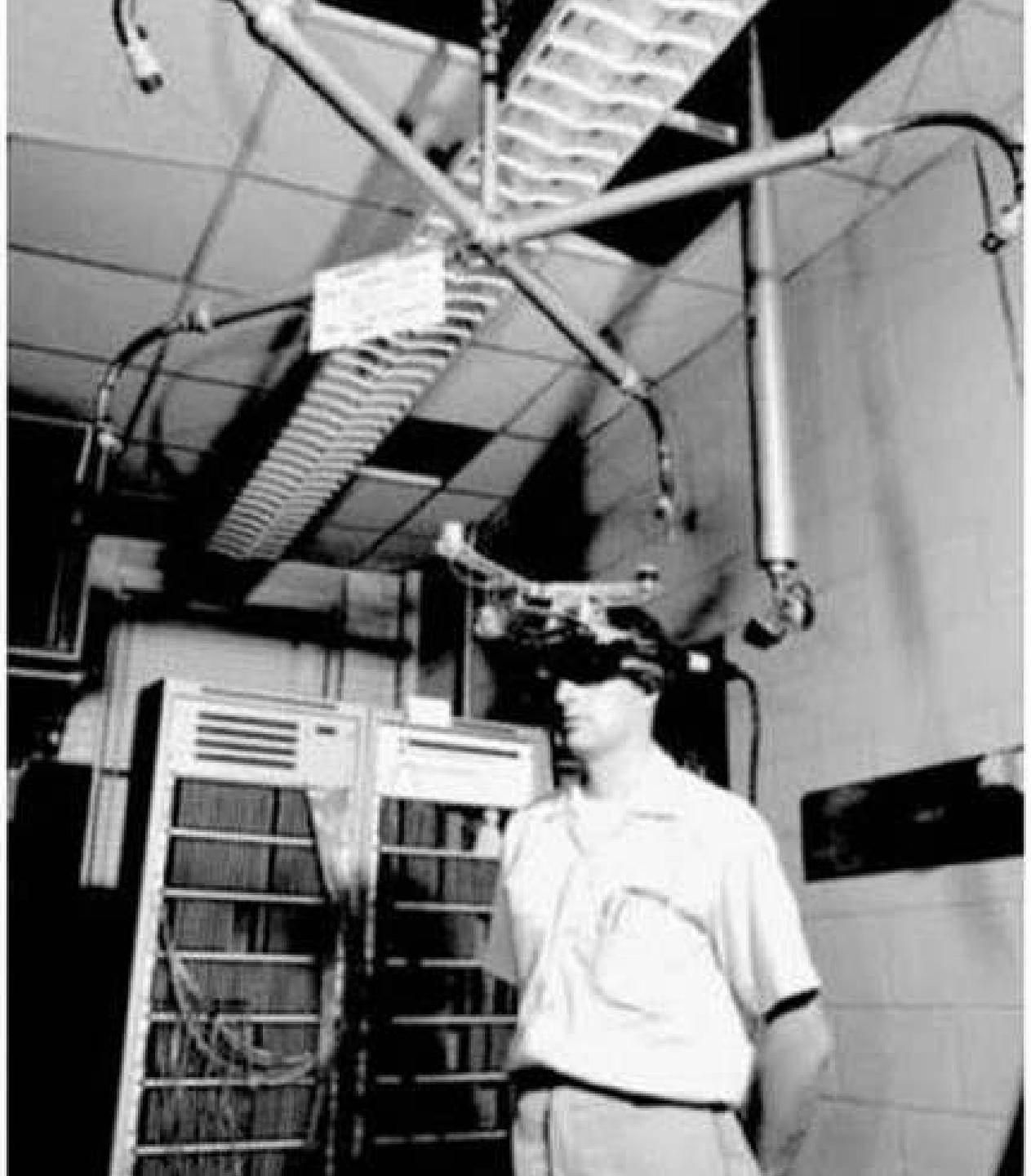
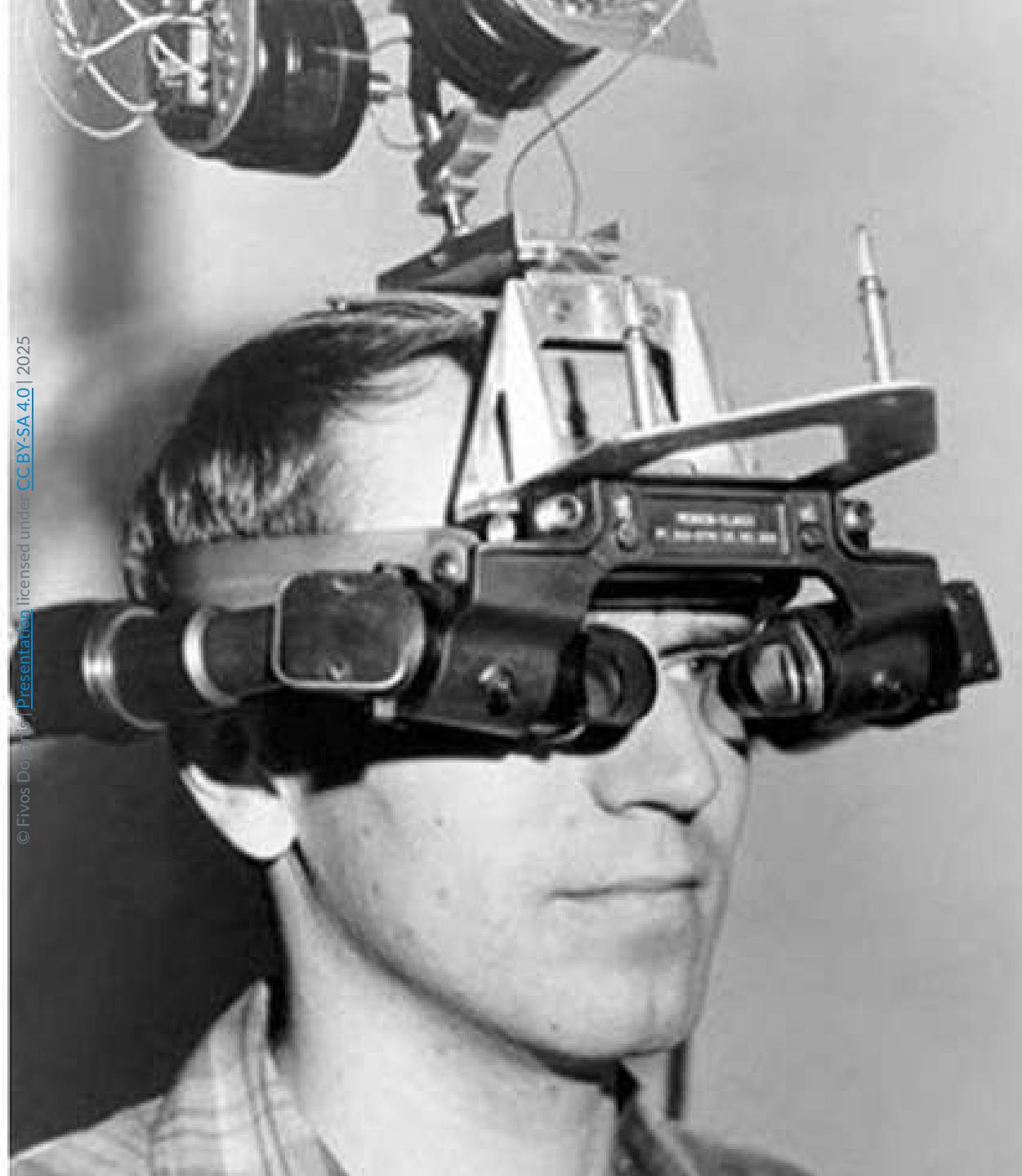
Key milestones

Prehistory

(1966) ★

- Ivan Sutherland invents the first AR headset
- Can display a cube
- Follows the movements of the head (**6 DOF**)
- Nickname : The Sword of Damocles
- paper: *The Ultimate Display*







Markers (1999)

- Monochrome markers
- [**ARToolkit**](#) created by Hirokazu Kato

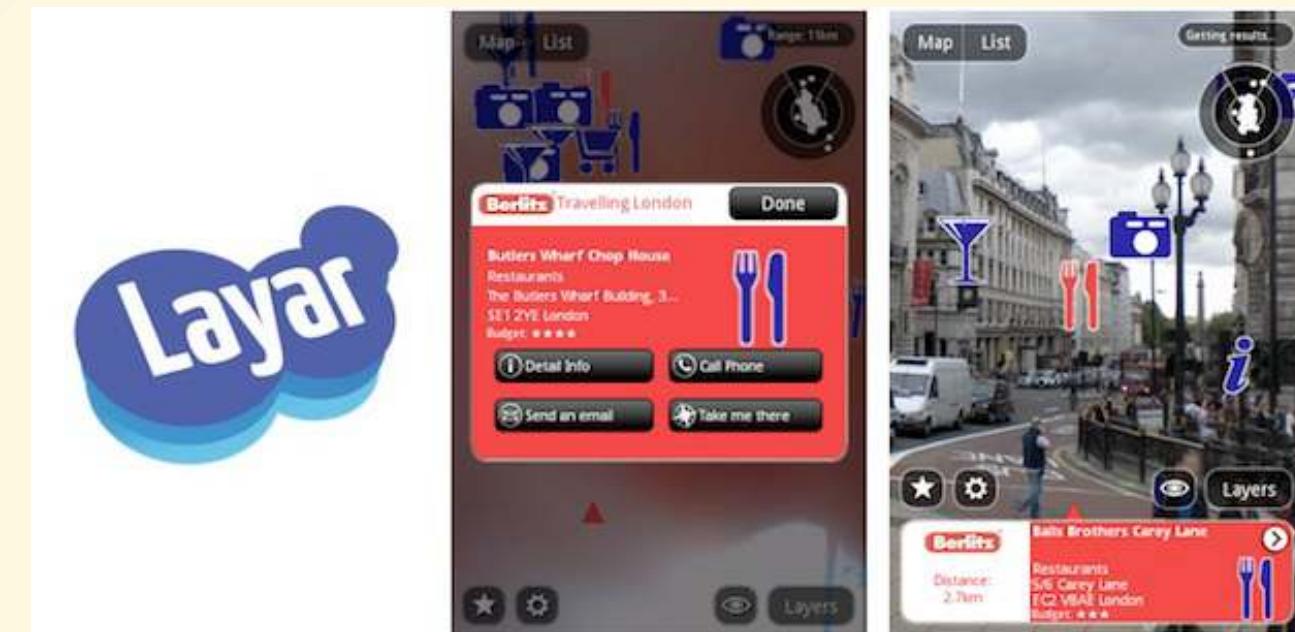


- Alternatives: ARTag, ArUco
- PC + Webcam



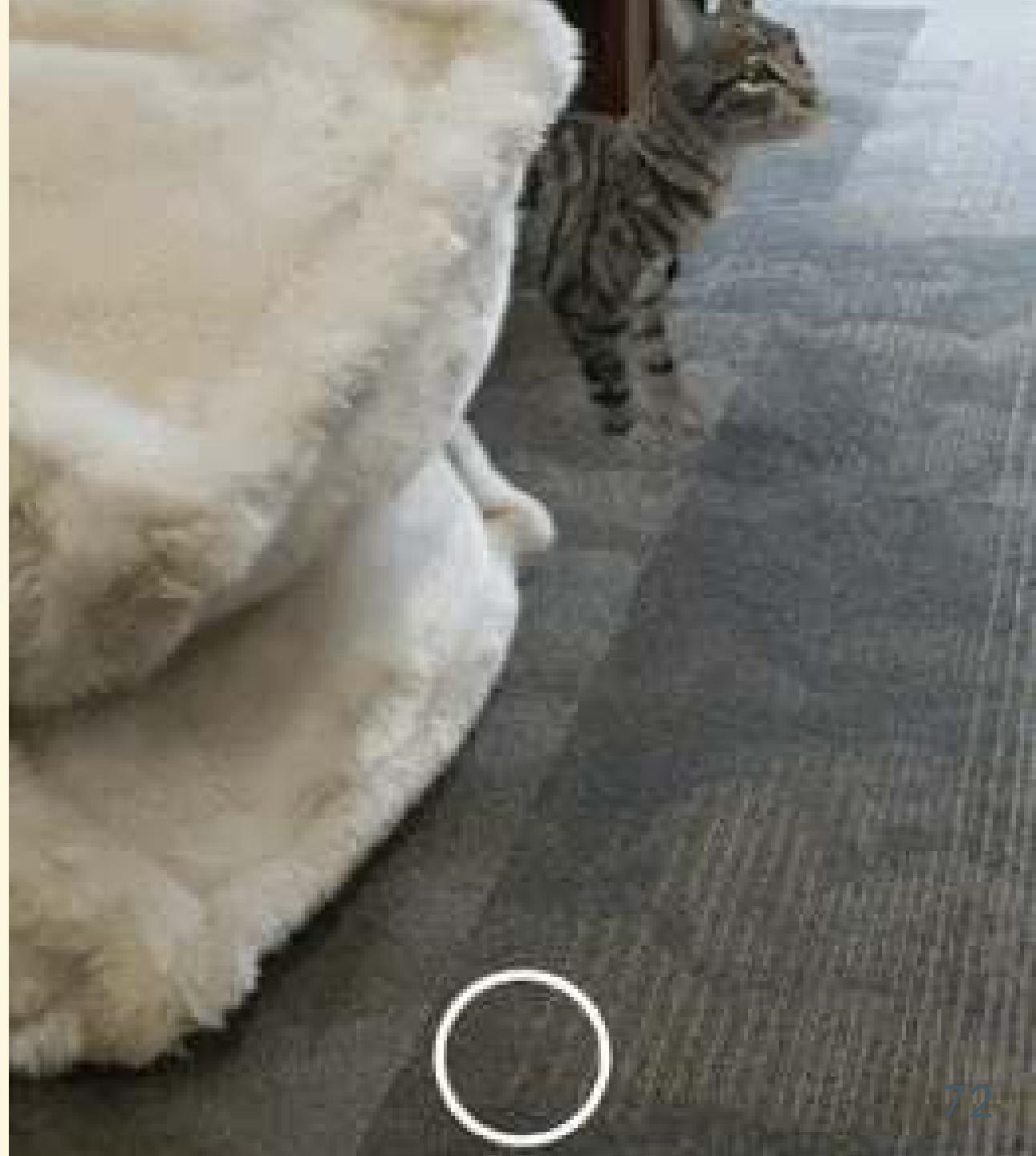
NFT, GPS (2005)

- NFT: Natural Feature Tracking
 - Color photo tracking
- Wikitude, Layar (GPS)
 - no image processing needed with GPS!
- Vuforia
- Marketing use-cases
- PC, mobile phones, tablets



SLAM, 3D (2015)

- **3D environment tracking**
- **SLAM:** ★ **S**imultaneous
Localization **A**nd **M**apping
- **3D object tracking**
- Deep Learning
- Occlusion 3D
- ARKit, ARCore
- Smartphones, HoloLens,
Azure Kinect



Azure Kinect + HoloLens 2



HoloLens 2



Apple LiDAR

iPad Pro 2020, iPhone 12 Pro

Apple LiDAR vs FaceID



Near future

- Form-factor: glasses 😎
- AI
 - contextual assistance
 - understands both environment and user
- 5G
 - application and information streaming
(Edge Computing)
- Spatialized Web: **AR Cloud**





Original Source (Cloud)



High bandwidth low latency content delivery

Local Source

Base station/
Network edge

- AR Object cache
- AR Data cache

Application point

Object IO

Application
(CSP)

Application
(OTT)

Core network

- Central AR cache

Internet

Google Glass

Telco Edge: 1 to 2 ms

Telco Cloud: 20 to 50 ms

Public Cloud: 50 to 100 ms

Far future

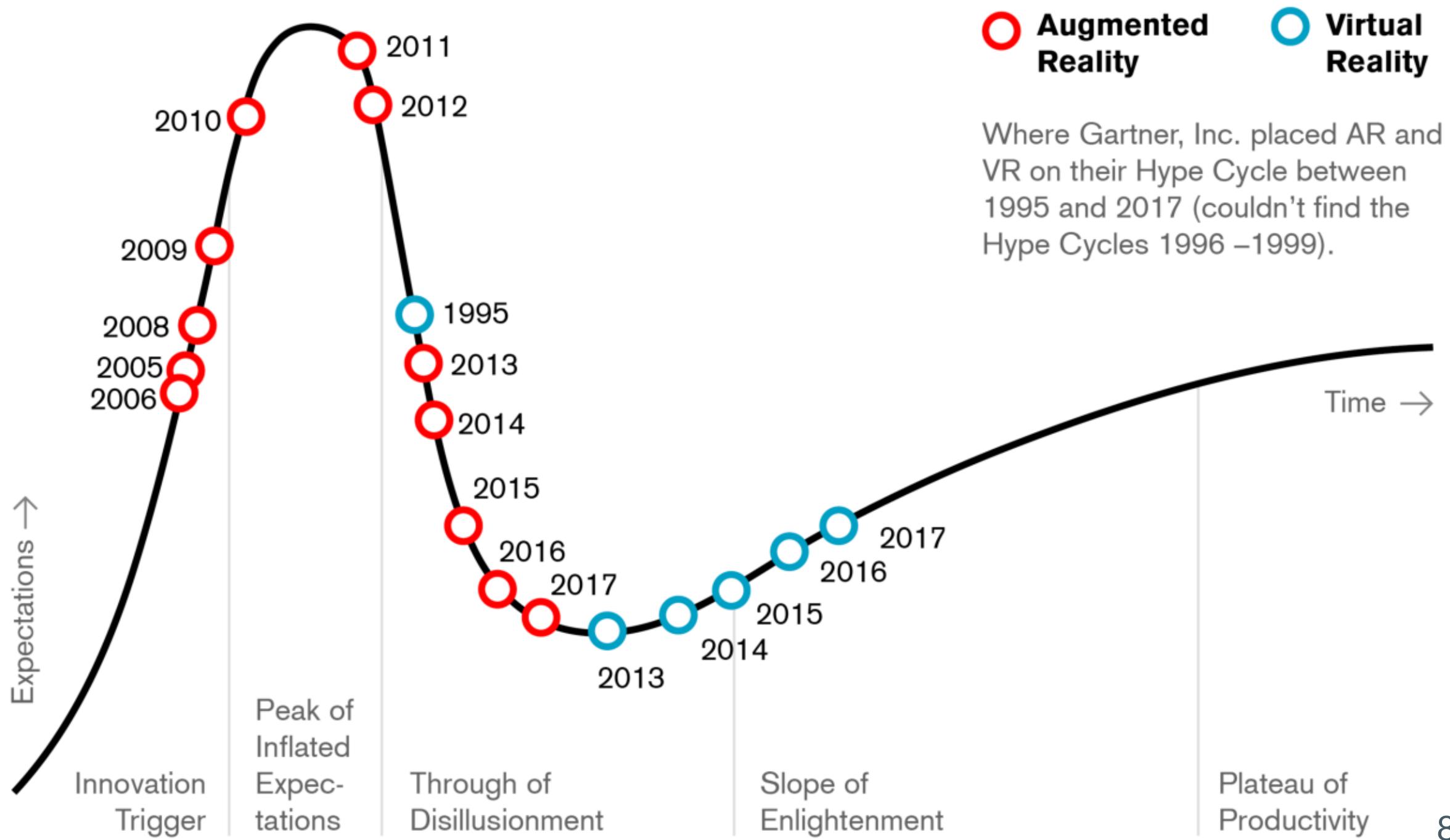
- AR will **replace or complement smartphones**
 - users will raise their heads again
 - but will they see better?
- Contact Lens (Mojo Vision)
 - RIP (2023) 
- **Ambient Computing**
- Ubiquitous Computing
- **Smart Cities**





Gartner Hype Cycle

- Technology trends evolution
- Gartner
 - 2005 - 2017 period



Where do we stand now?

- We are getting close to productivity for AR and VR
- Or AR is now considered as "productive" since 2020?
- Future trends:
 - AR Cloud
 - "Metaverse" and "Digital Humans"

2023

Magicverse Layers

The Layers of Spatial Computing

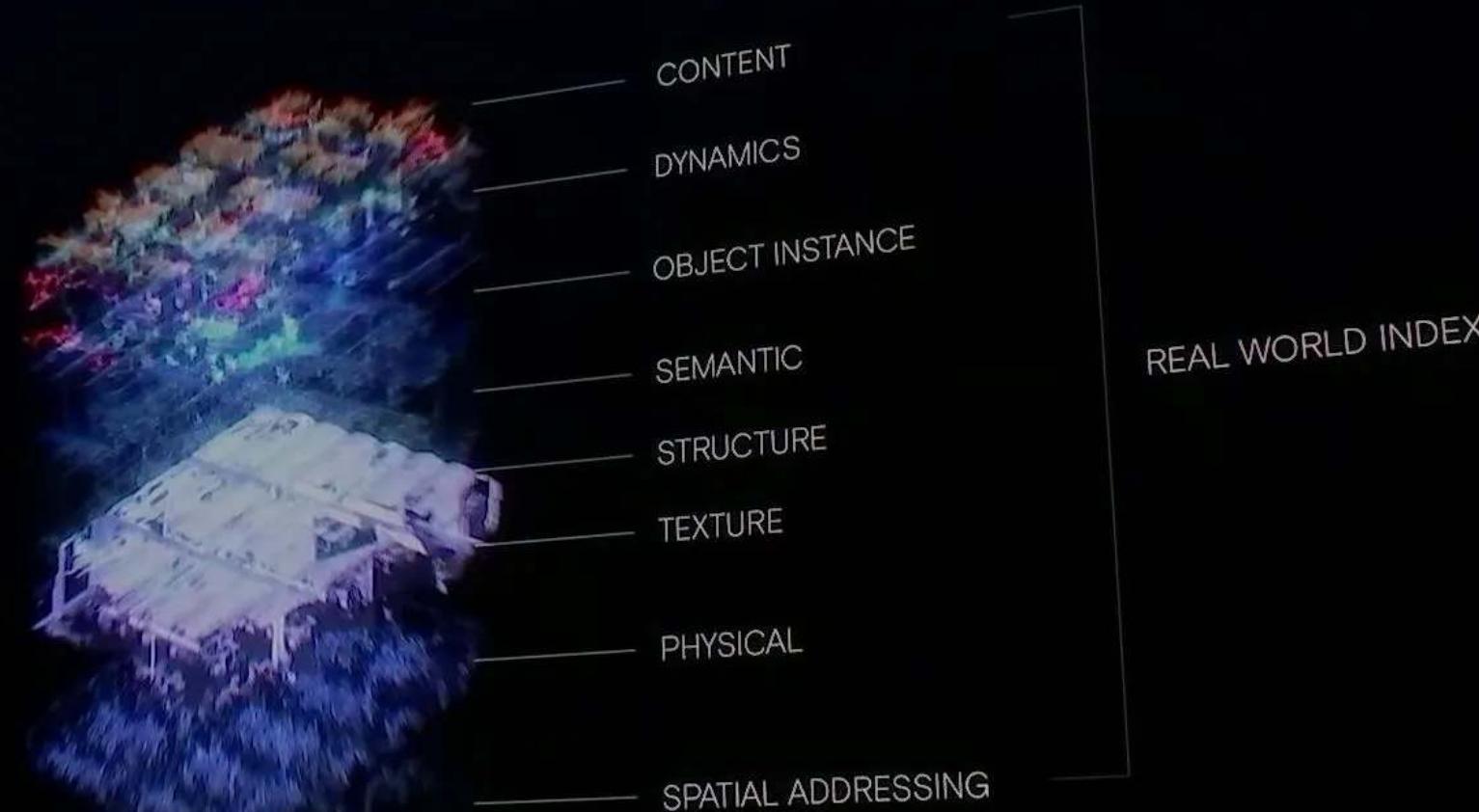
A nearly infinite number of application layers make up the city-scale Magicverse.

User choices and contextual AI power this system of systems, navigated through a spatial computing interface.





facebook
Reality Labs



MAMAA Strategies

**Meta, Alphabet, Microsoft, Apple, Amazon
and others!**



“ AR [*is for*] adding shared meaning in the interaction between people.

Johnny Lee, [Google I/O 2017](#)

”

- Dropped mobile VR (Cardboard 💀, Daydream 💀)
- Dropped Tango 💀, to reach more devices: rely on **RGB camera + AI**
- API [ARCore](#), competes with Apple's [ARKit](#)
- Google wants to provide **cross-platform AR services**
- Android XR + Samsung XR Headset coming in 2025

“ The Web connects the world's information, and AR connects information with the physical world. So together they can be applied to solve real life problems.

*Andrey Doronichev,
Google I/O 2017*

”



Google Glass Enterprise Edition 2

Google Glass killed, for the second time, in 2023 💀

ist

Shipping Tasks

wi

Gorillaz mobile AR app

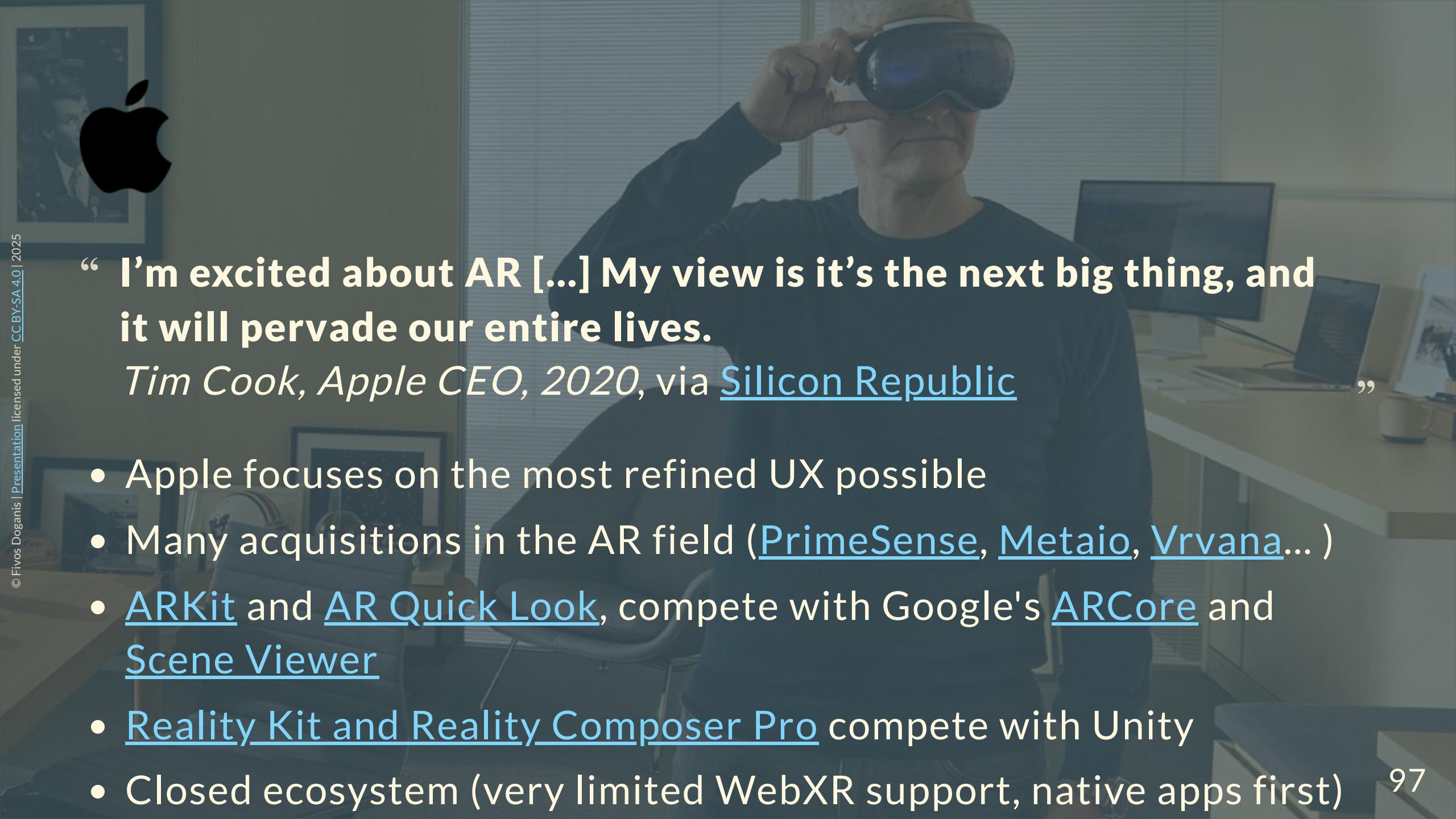


Android XR



Samsung Project Moohan





“ I’m excited about AR [...] My view is it’s the next big thing, and it will pervade our entire lives.

Tim Cook, Apple CEO, 2020, via [Silicon Republic](#)

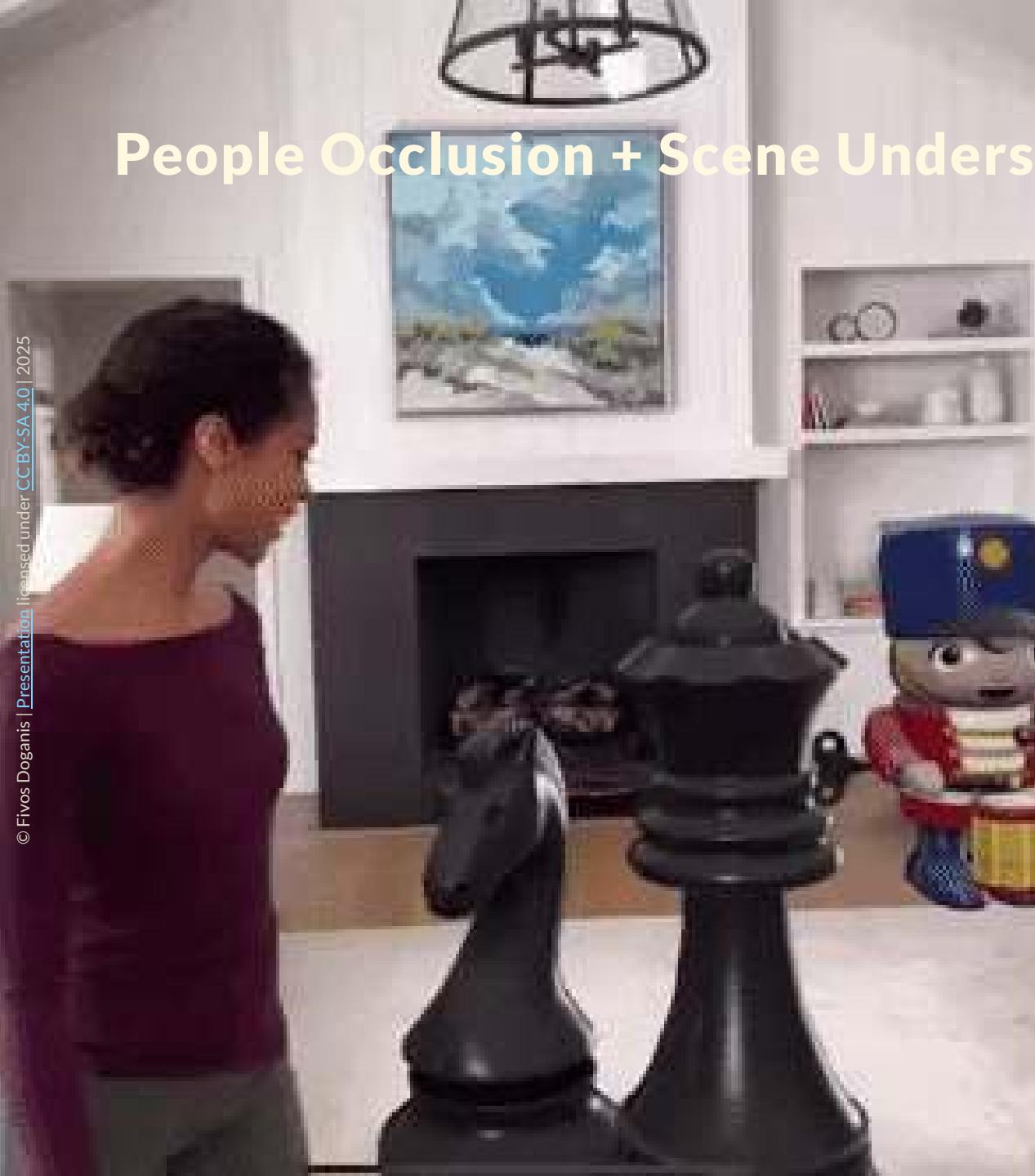
- Apple focuses on the most refined UX possible
- Many acquisitions in the AR field ([PrimeSense](#), [Metaio](#), [Vrvana...](#))
- [ARKit](#) and [AR Quick Look](#), compete with Google's [ARCore](#) and [Scene Viewer](#)
- [RealityKit](#) and [RealityComposer Pro](#) compete with Unity
- Closed ecosystem (very limited WebXR support, native apps first)

Hardware

- Adds **LiDAR** for a robust **SLAM** (e.g. white walls scenario)
- Extends its 'wearables' category
 - AirPods
 - Apple Watch
 - **Apple Vision Pro**
 - unveiled in June 2023
 - released in February 2024



People Occlusion + Scene Understanding (iOS)



Eye tracking (visionOS)

"Spatial Computing", "EyeSight", Real Virtual Continuum



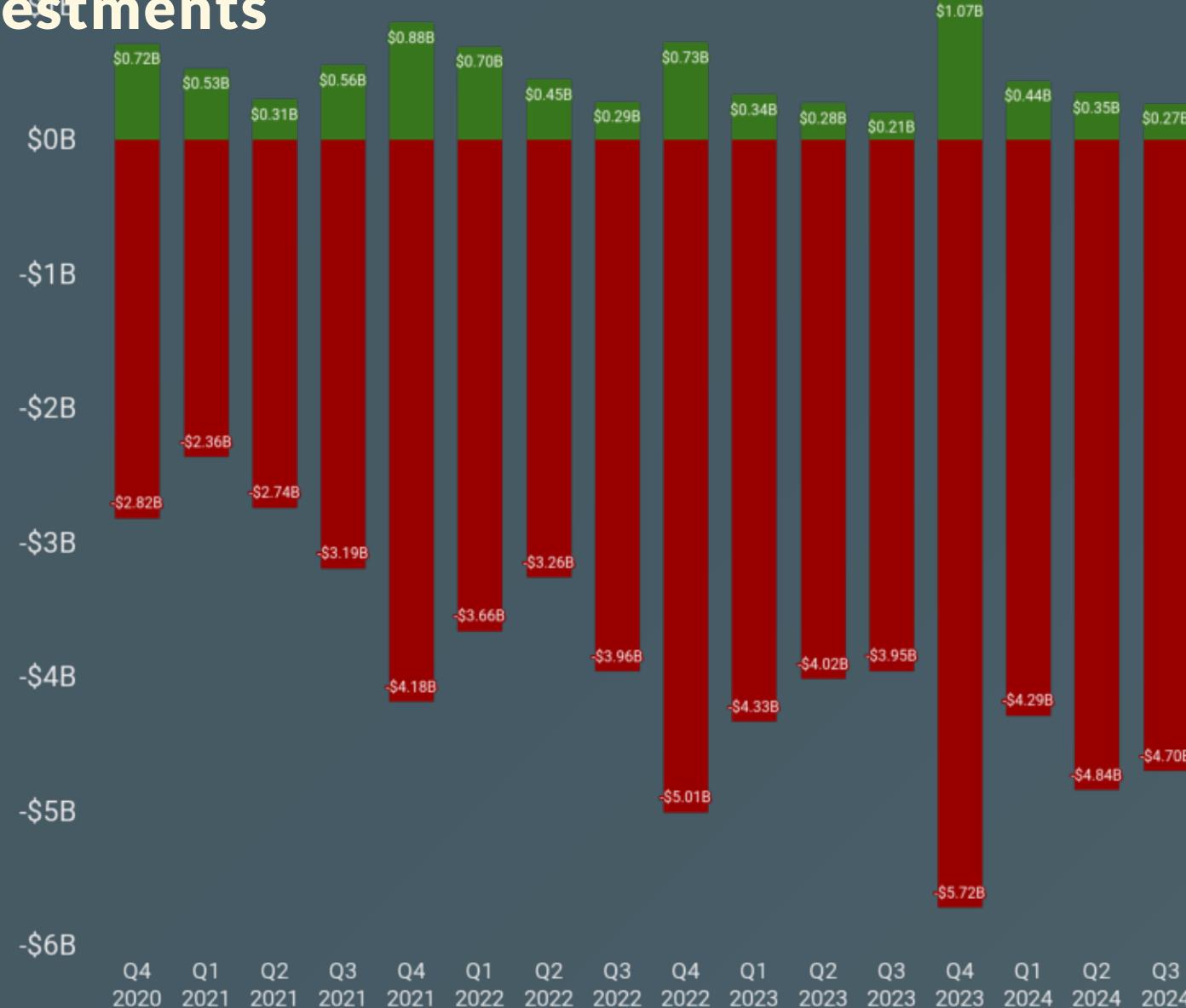
Avatars

Collaboration

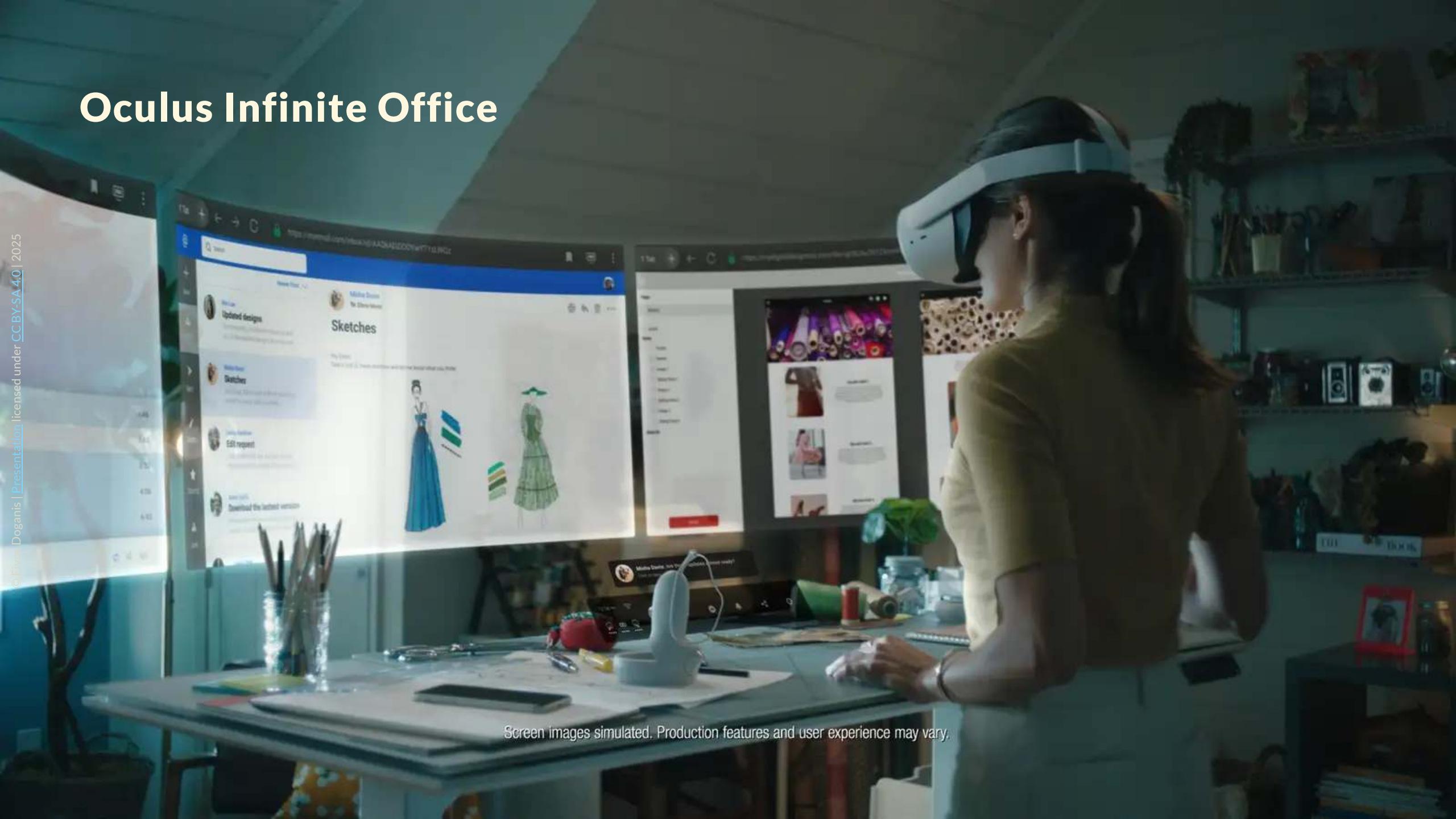
- Created **Horizon Worlds** social network for virtual encounters, and **Horizon Workrooms** for remote collaboration
- Believes that **AR will replace smartphones** and every screen **in 10 years**, does not want to miss this revolution: **Project Aria / Orion**
- Interested in **personal data**, centers of interest of their users (eye-tracking, scanned environment), **although they claim the opposite**
- VR with **Oculus**, but they now **focus on AR**, cf. **Infinite Office**
- ~~AR filters Facebook and Instagram via **Spark AR RIP**~~ 
- Great **WebXR supporter**

Revenue Costs

Meta XR Investments



Oculus Infinite Office



Screen images simulated. Production features and user experience may vary.

horizon worlds

Meta



Meta Quest 3: focus on AR

Reverse Passthrough prototype (CAD render)

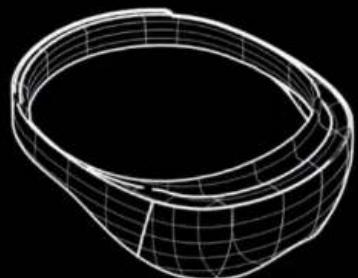
video

Michael Abrash in 2019



VR

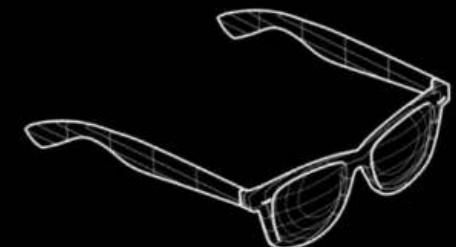
AR



FULL CONTROL OF PIXELS

FEWER CONSTRAINTS

BEST MIXED REALITY



SOCIALLY ACCEPTABLE

WEIGH 70 GRAMS

DISSIPATE 500 MILLIWATTS

5

Project Aria



Next?

“ I might get myself in trouble for saying this; I think it might be the **most advanced piece of technology on the planet in its domain**. In the domain of consumer electronics, it might be **the most advanced thing that we've ever produced as a species**.

Andrew 'Boz' Bosworth, Meta CTO, January 2024

”



Project Orion (2024)

\$10 000 prototype





- Amazon focuses on e-commerce and its Web Services
- [AR View](#) to see a product at home before buying it
- Offers [Sumerian](#) as a paid tool via AWS (Amazon Web Services) to create XR experiences
- Pushes machine learning, smart assistants (Alexa)
- Bets on AR on demand via 5G with its [Wavelength Project](#)
 - **5G + Edge computing**
 - AWS

View in Your Room

Design and Decorate Your Home with Augmented Reality





Microsoft

- Consumer: Minecraft
- Focuses on AR and industry, to maximize added-value
 - assistance and training
- ~~HoloLens 1 et 2, Kinect~~ RIP 2024
- VR: OEM Partners
- ~~Windows Mixed Reality~~: RIP 2023
 - now provides popular software to competitors!

Other players

- Hardware
 - [Magic Leap](#), [Lynx](#), [XReal](#),
[Qualcomm](#), [Snap](#)
- SDK
 - [Wikitude](#), [Kudan](#)
 - PTC [Vuforia](#): IoT
- Web
 - [Firefox Reality](#) [Wolvic](#)
 - [Chrome](#)



Takeaways

- Big tech companies invest massively in AR, which they see as a **promising technology** evolving fast
 - hardware
 - algorithms
 - services, **data**
- Many players try to bring their users into their **closed ecosystem** (hardware, app store, cloud)
- Others focus on the openness of the **Web** to **create and share open AR experiences**
 -  **ultimate goal of this course!** 

Further reading

- History and future of Web AR
 - [Web AR: A Promising Future for Mobile Augmented Reality – State of the Art, Challenges, and Insights.](#)

Qiao, Xiuquan & Pei, Ren & Dustdar, Schahram & Liu, Ling & Ma,
Huadong & Junliang, Chen. (2019).

Proceedings of the IEEE. 107. 1-16.
[10.1109/JPROC.2019.2895105](https://doi.org/10.1109/JPROC.2019.2895105).

3 Types of AR ★

- **Video**

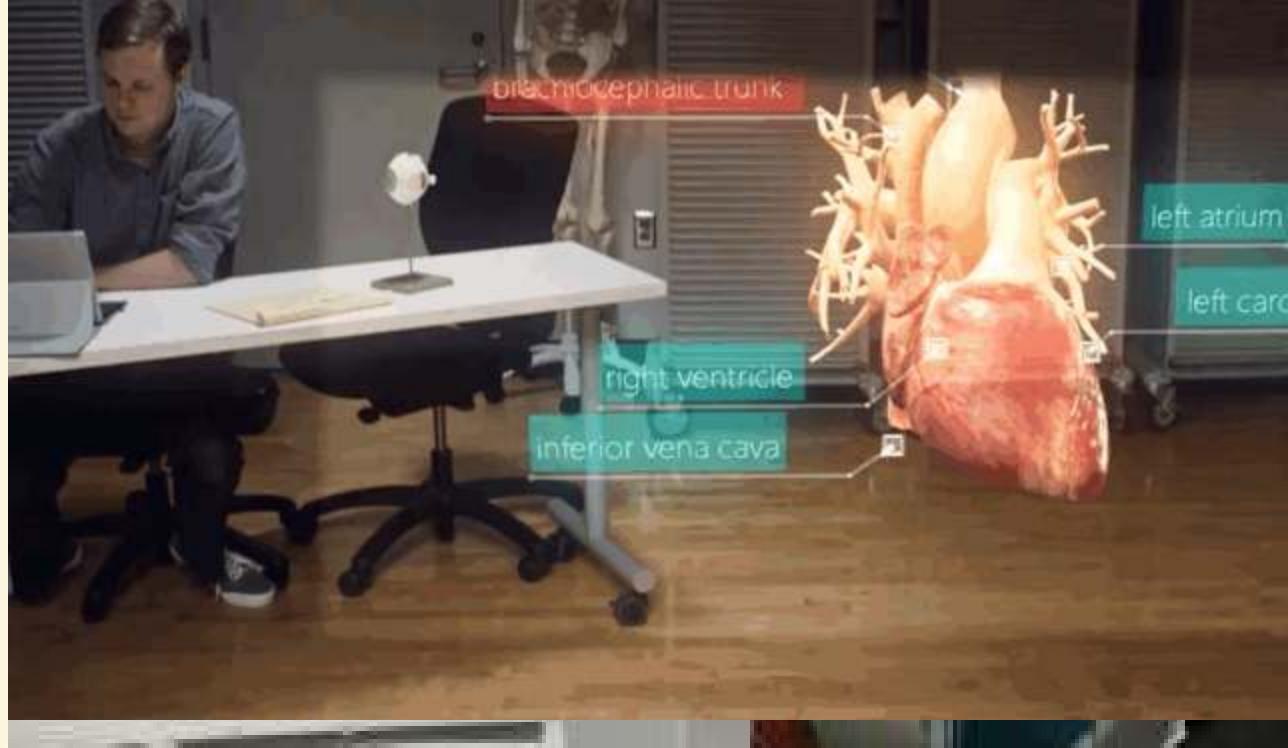
- e.g.: smartphone,
Meta Quest 3, Apple
Vision Pro, [Lynx-R1*](#)

- **Optical**

- e.g.: [HoloLens](#)

- **Projective**

- e.g.: [DIOTA](#) ➔



Lynx-R1 ([video](#))



Technologies Required for AR

Calibration
Tracking
Interactions
Rendering

Calibration

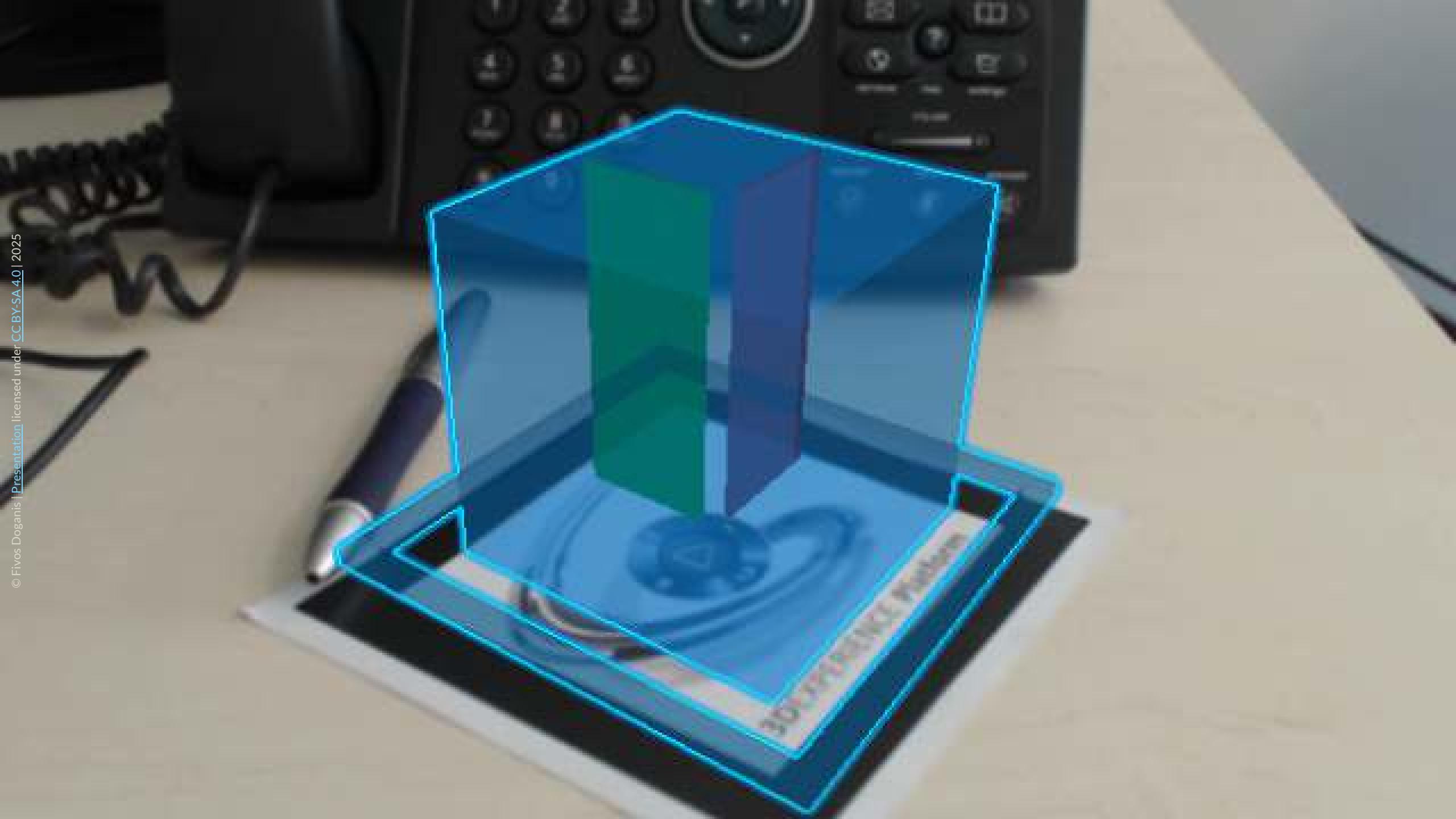
Goal: overlay accurately **the virtual rendering and the real image**

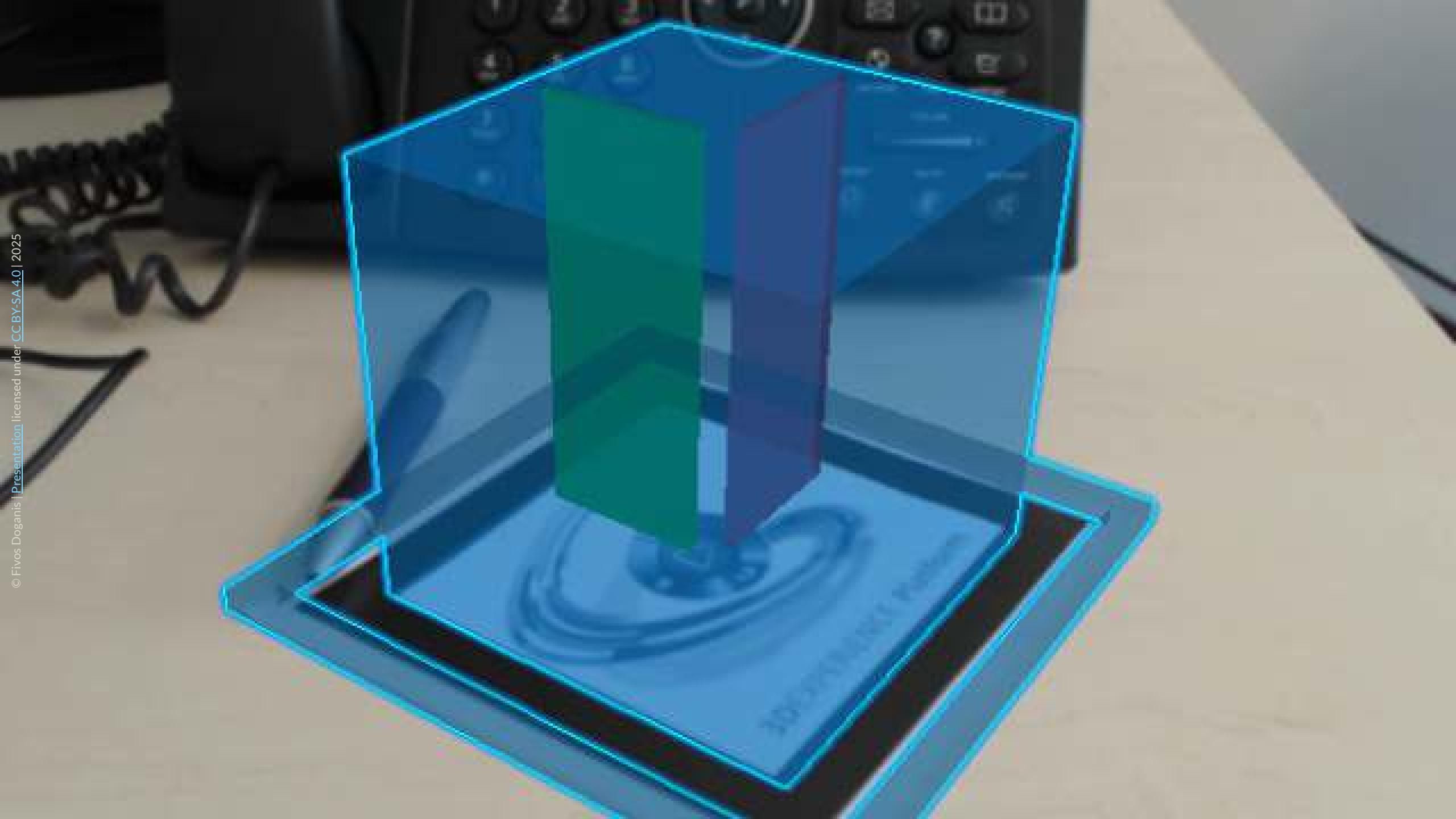
Optical AR calibration

- very complex
- hardware dependent
 - projection and image formation systems
- depends on the body metrics of the user
- **made and provided by the AR hardware manufacturer**
 - possible adjustments for each user, cf. eye calibration in HoloLens

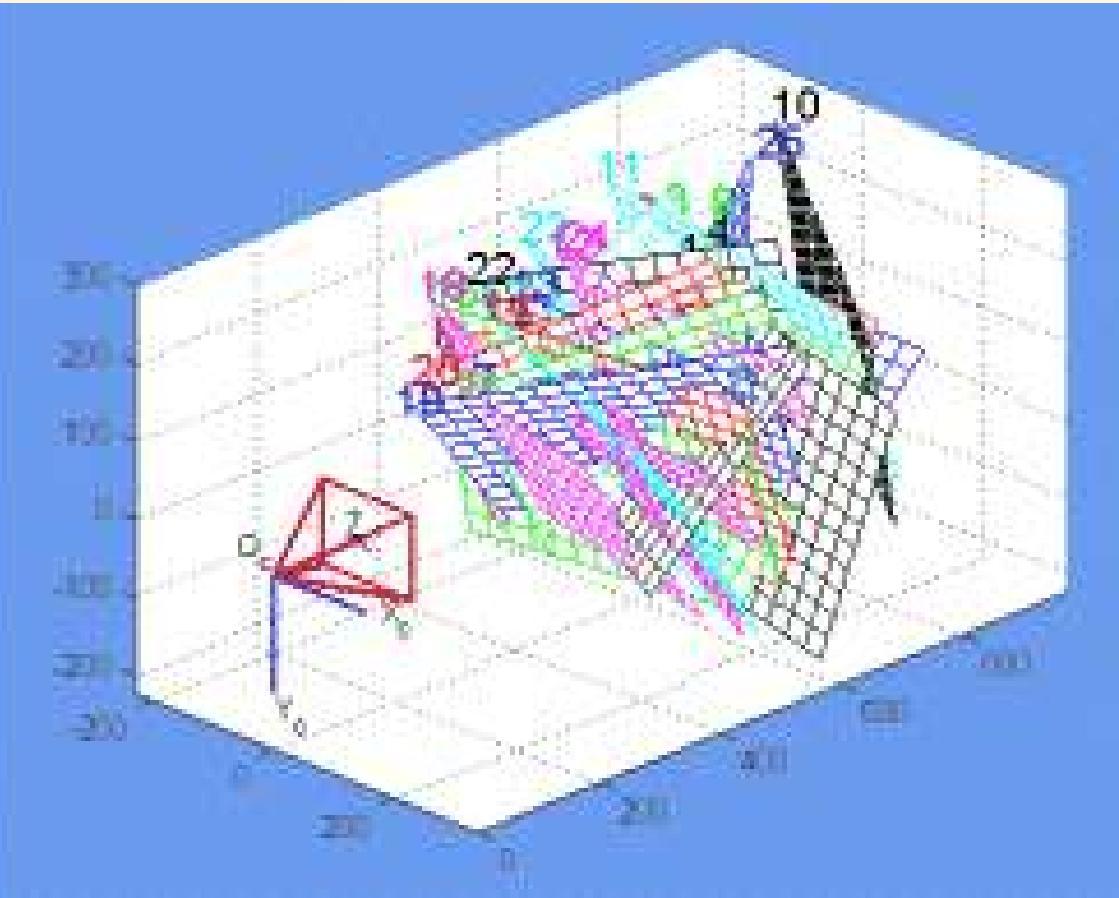
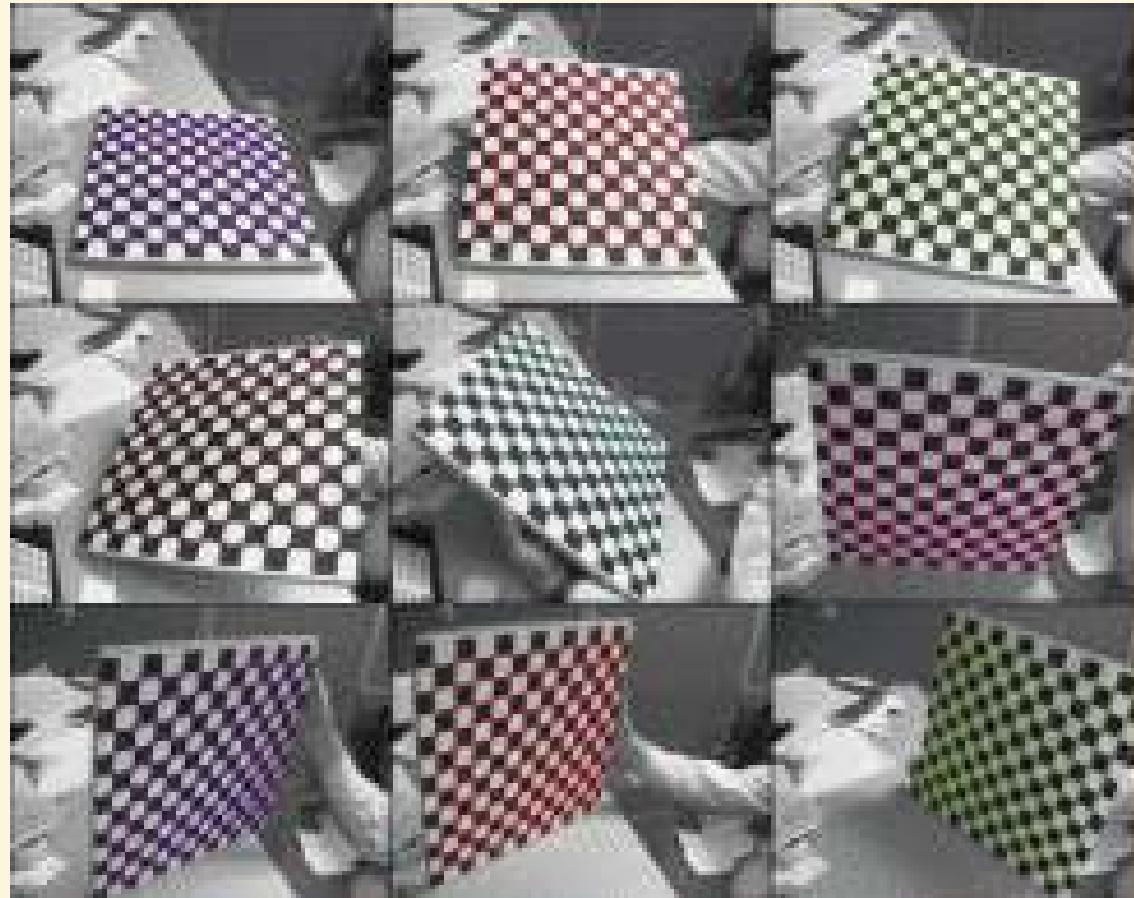
Video camera calibration

- Goal: compute the **optical parameters of the real camera**
 - **focal length**
 - radial distortion, lens imperfections
- Method:
 - capture images of known patterns (grids, **calibration patterns**) with a real camera
- **⚠ the focal length may be variable** (autofocus)
 - update calibration data for each frame
 - calibration data is provided by the API (ARKit, ARCore, WebXR)





Video camera calibration method



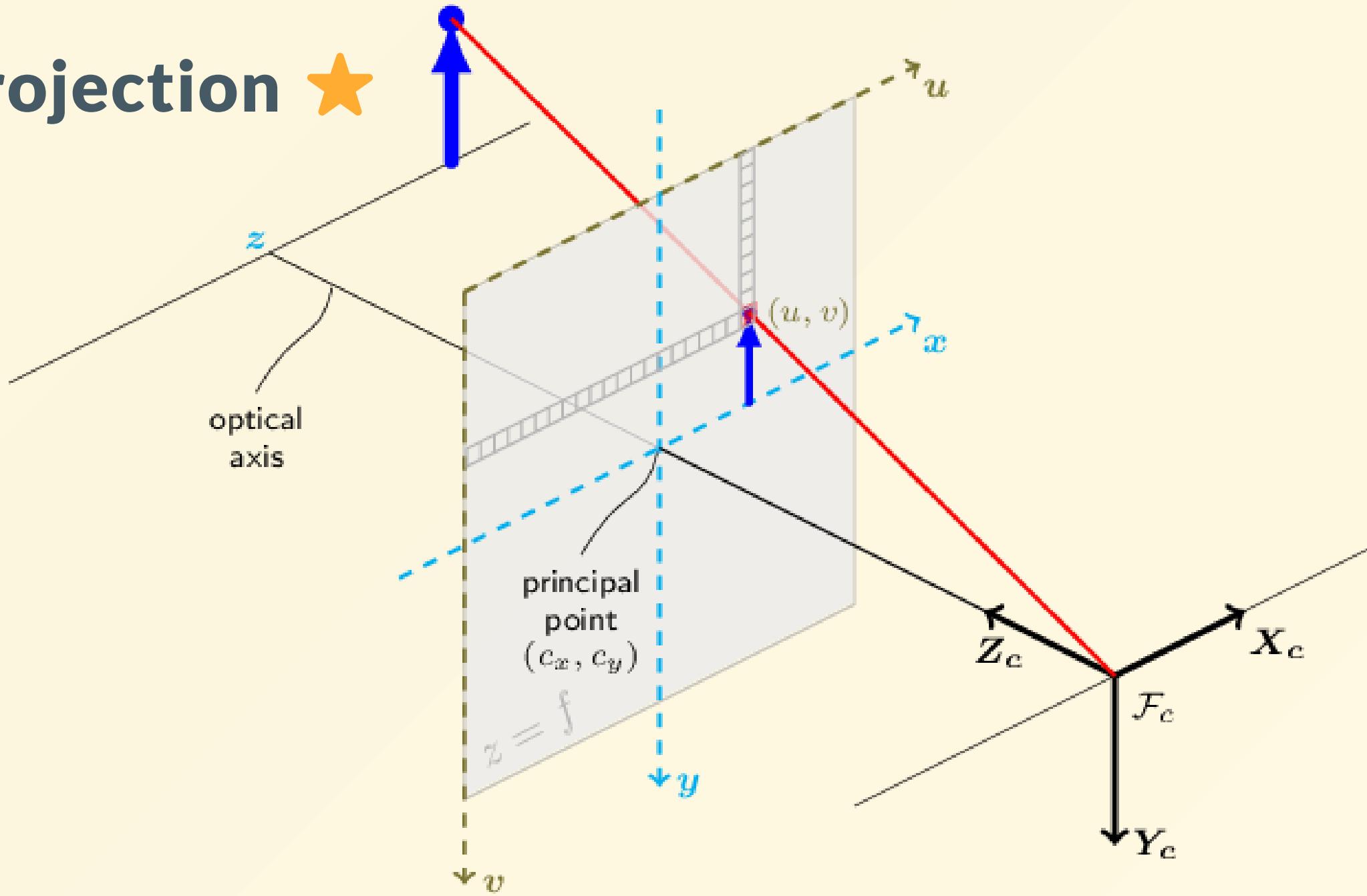
Pinhole camera model

Extrinsic and intrinsic parameters

3D coordinates → Camera 3D coordinates → Image coordinates

$$P = (X_w, Y_w, Z_w)$$

Projection ★



$$s \ p = A[R|t]P$$

$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

(X_w, Y_w, Z_w) 3D world coordinates O_w

(u, v) projected coordinates (pixels)

$[R|t]$ **extrinsic** matrix, A **intrinsic** matrix

(cx, cy) principal point (pixels), **center of the image** in the ideal case

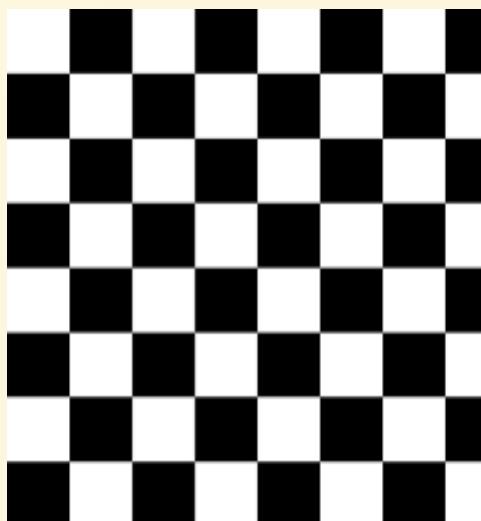
f_x et f_y focals along x and y (pixels), **equal** in the ideal case

Non linear radial distortion

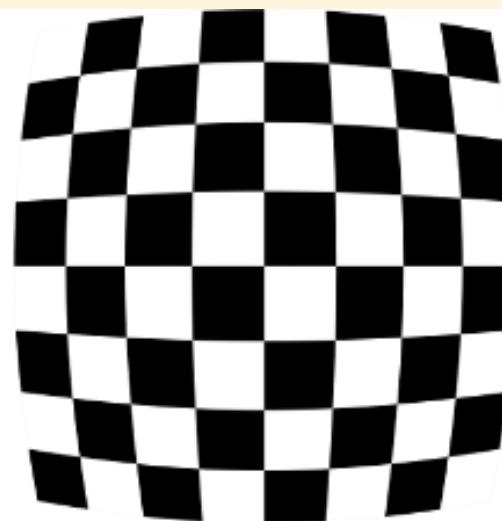
- due to the lens, approximated by a polynomial expression

$$x_{distorted} = x(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$$

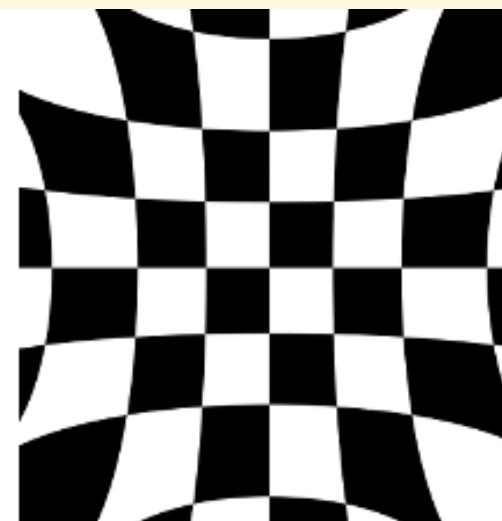
$$y_{distorted} = y(1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$$



No distortion



Negative radial distortion
(Barrel distortion)



Positive radial distortion
(Pincushion distortion)

Registration

Goal: find the **rigid transformation** $[R|t]$ between a 3D point in the world and the center of the camera

Pose estimation

- Computed from 2D/3D pairs of points
- Optimization: projection error minimization between transformed 3D points V_i et image 2D points v_i

$$\arg \min_{R,t} \sum_i ||P(RV_i + t) - v_i||$$

P : projection function

R : rotation matrix

t : translation vector

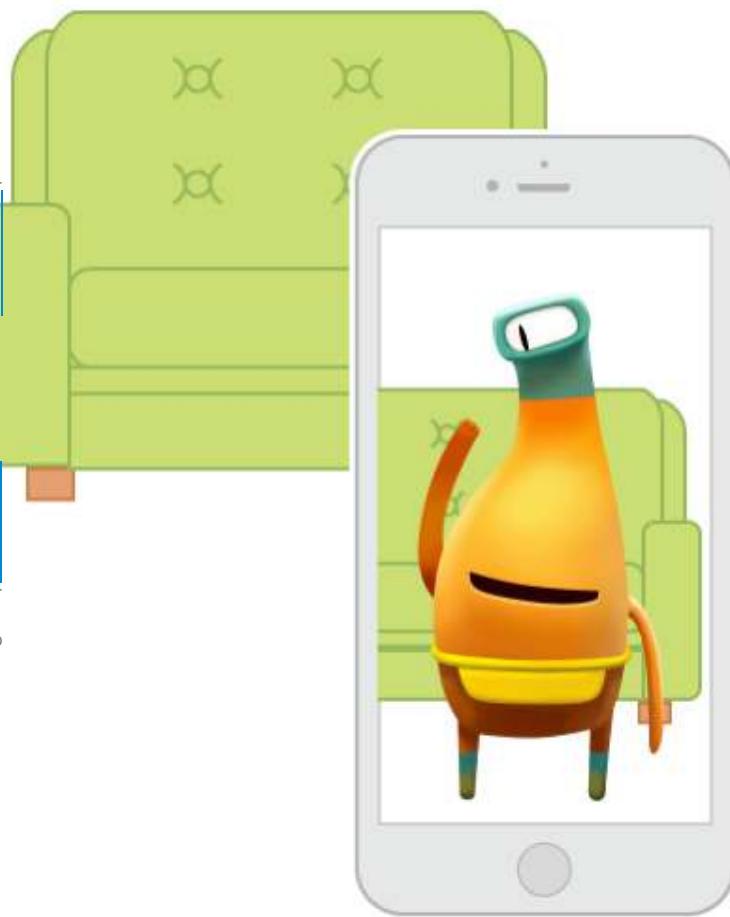
Tracking

after initial registration

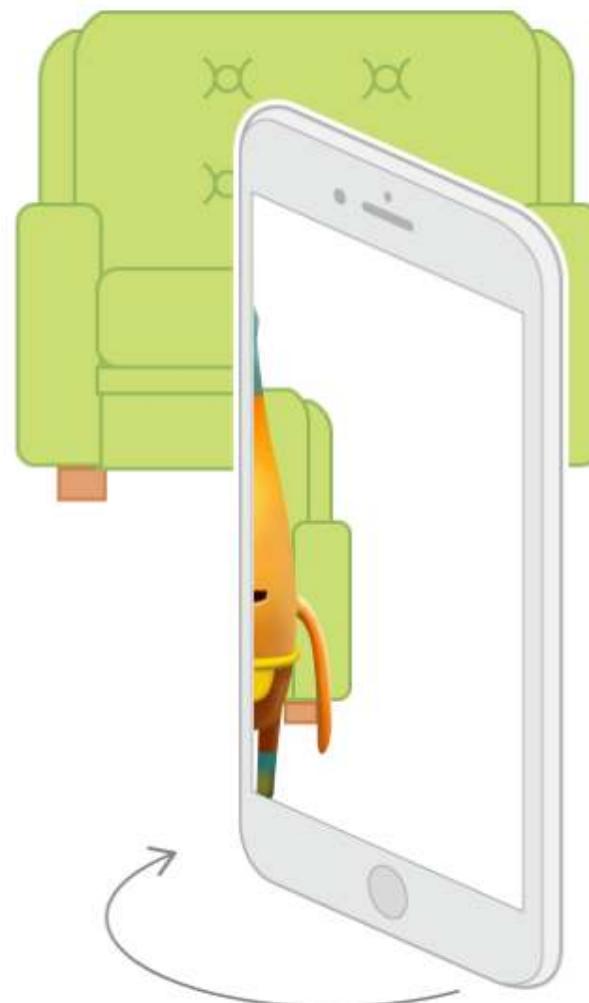
Tracking

- Degrees Of Freedom (**DOF**):
 - **0 DOF**
 - no tracking!
 - simple information overlay, cf. [HUD](#)
 - **3 DOF**
 - **rotation** only (gyroscope, accelerometer, compass)
 - limited experience (can be good enough, cf. planetarium)
 - **6 DOF**
 - **rotation + position**

3 DOF



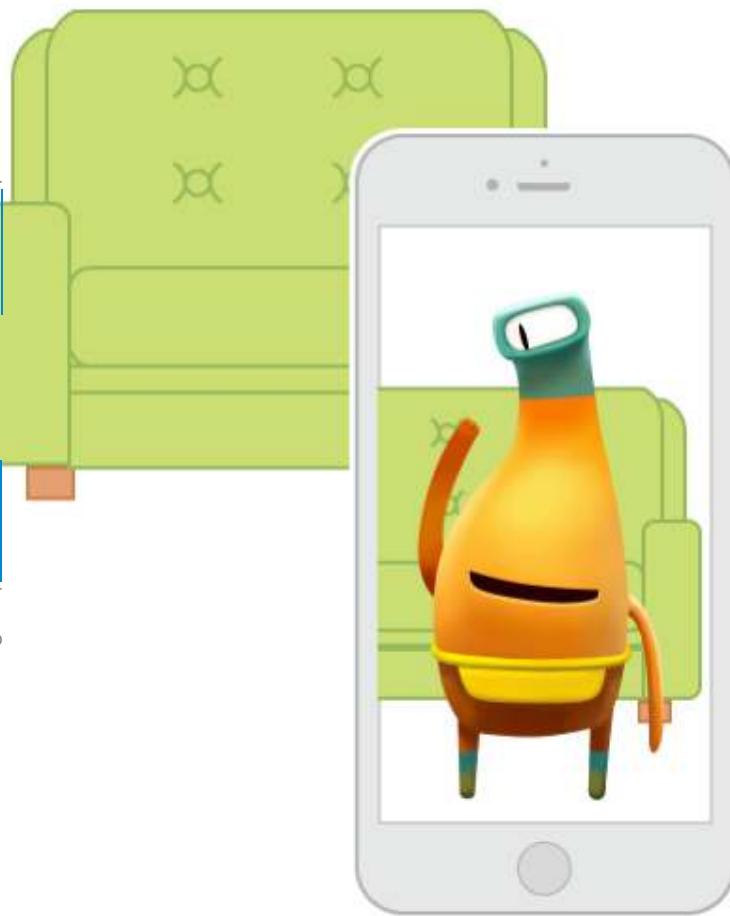
✓ Device rotation



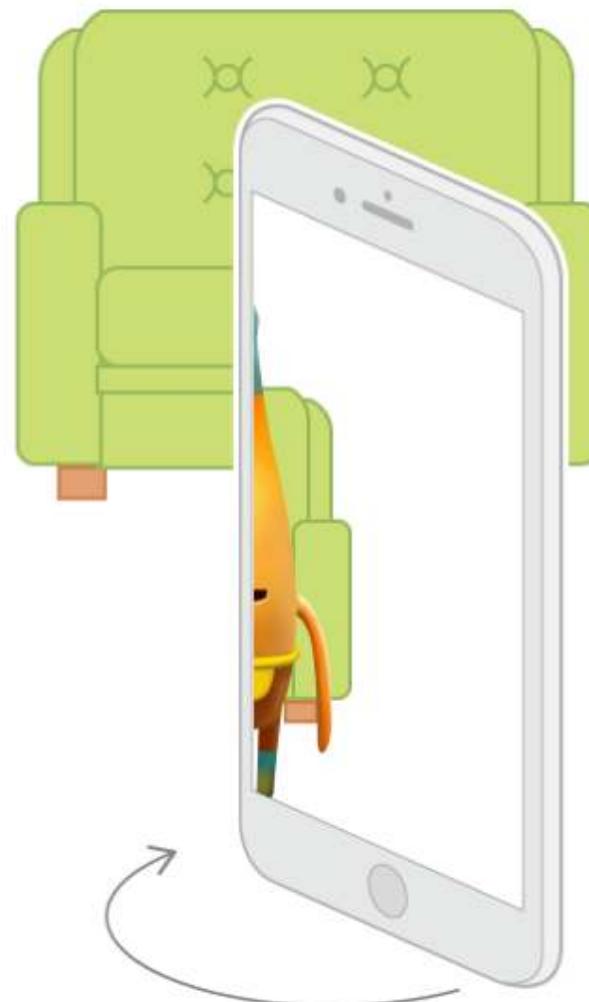
✗ Device position



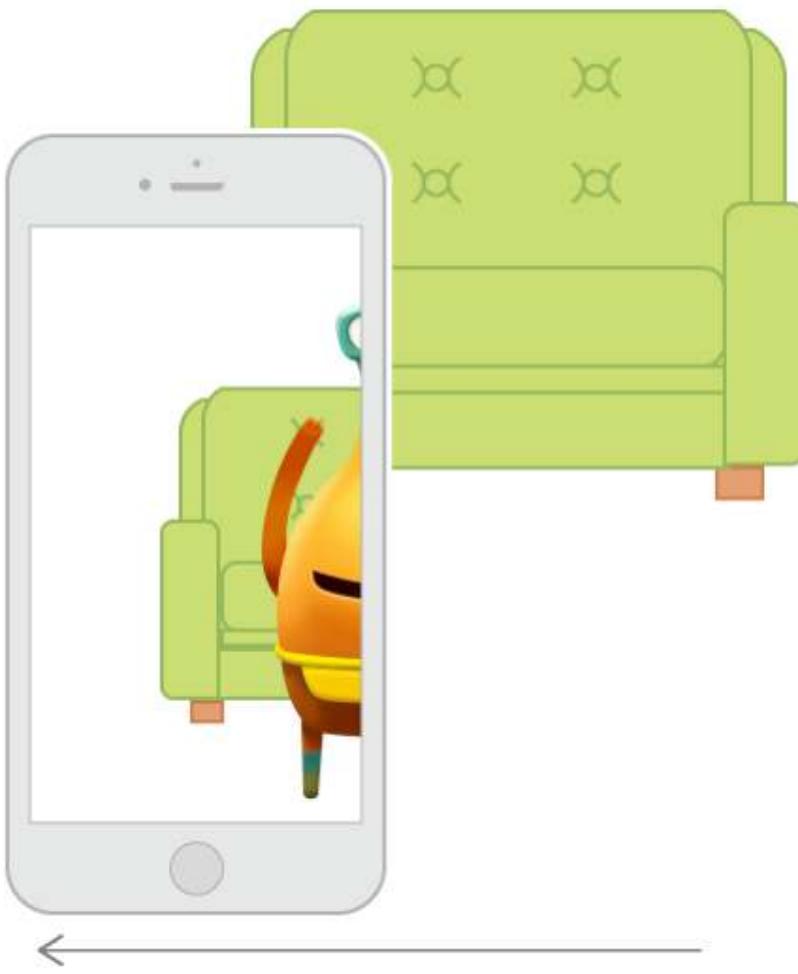
6 DOF



✓ Device rotation



✓ Device position



Tracking techniques ★

GPS

Marker

NFT

SLAM

3D

GPS

- global, satellite based, **no network connectivity required** ✓
- **no image processing** ✓
- **outdoors** only ✗
- **slow** ✗
- **not very accurate** ✗

Marker

- accurate, **fast** ✓
- **tangible**, printable ✓
- need to display a marker to enable AR ✗
- **non-aesthetic** ✗
- **can be hard to detect** (low lighting, motion blur, occlusions) ✗

Valve VR HMD early prototype



Natural Feature Tracking

Same as marker but

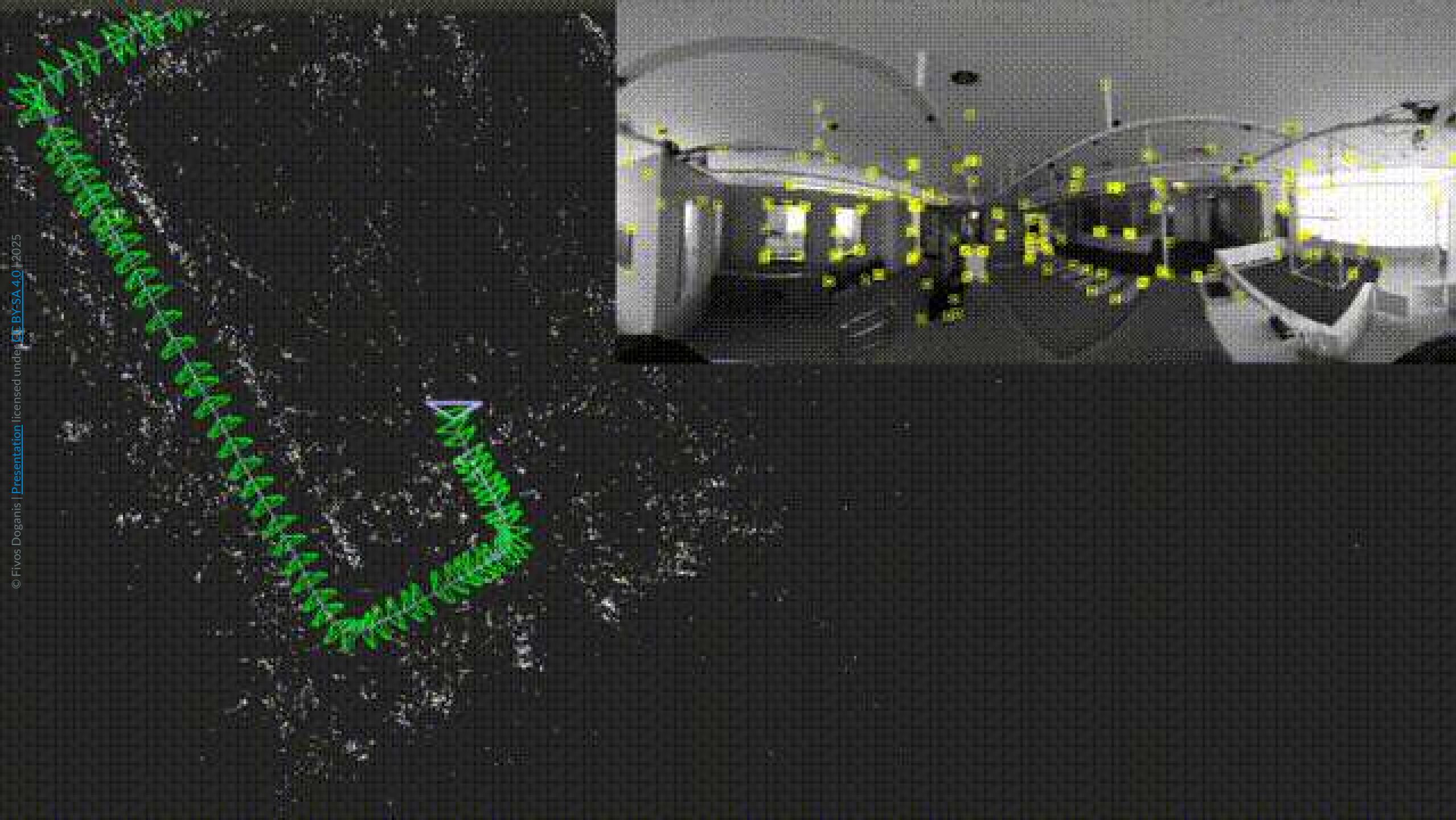
- more **aesthetic**, easier to embed in the real world (ads) 
- **more robust to occlusions** 



SLAM

NFT evolution + reconstruction

- more **natural** markerless experience 
- partial **scene reconstruction** 
 - allows advanced functionalities (occlusions, collisions etc.)
- not very accurate  (not ideal for medical uses)
 - **drift**, loop closure
 - scene reconstructed and refined in real-time
 - difficult to define the origin of the scene
 - stable **anchor** points required



3D object detection in a real scene

- using **computer vision** (lighting, edges, silhouette)
 - generic algorithm 
 - but **slow**, especially during initial registration 
- using **Deep Learning**
 - faster initial detection 
 - more robust regarding occlusions and lighting changes 
 - **not generic**: requires per model training 



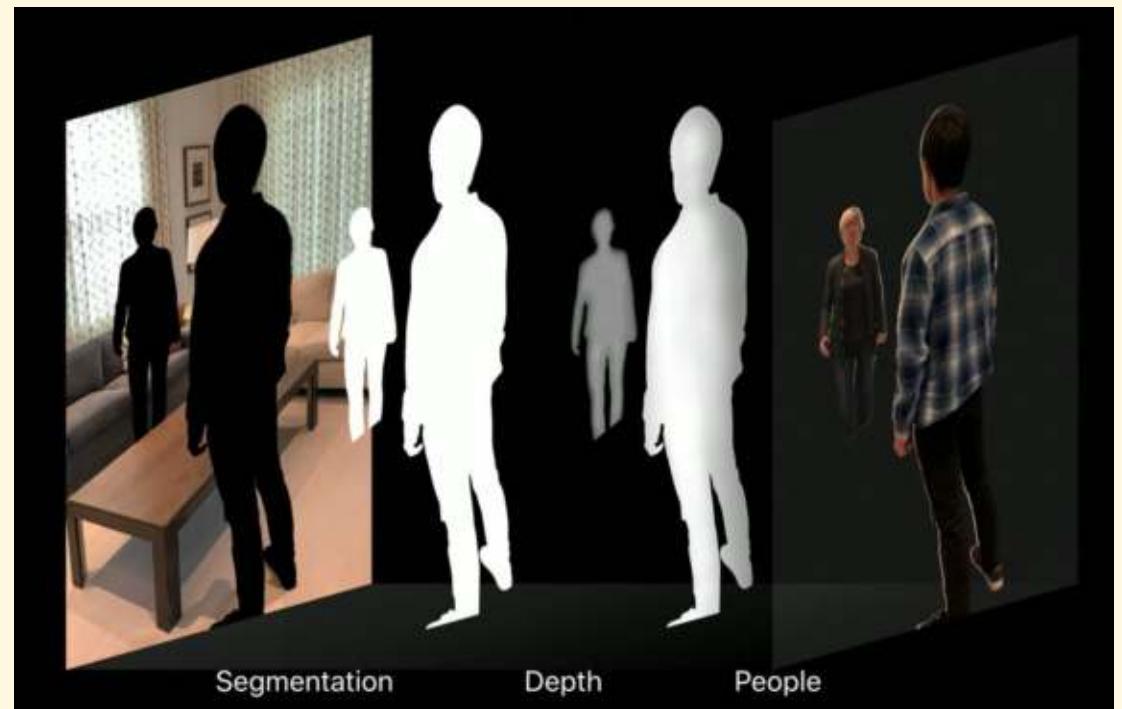
Tracking techniques : Conclusion

- No tracking technique is ideal
- Keep them all in mind and choose the right one according to:
 - the **scenario** of the AR experience
 - industrial context, consumer, generic or specific
 - **constraints**
 - indoor, outdoor, mobile

Rendering

Rendering

- Realistic or not
- Lighting
 - detect the direction and intensity of real lights
 - fast environment reconstruction to simulate reflections (SLAM + AI)
- Occlusions
 - people, objects



Interactions

- The missing part of the equation
- Often neglected (cf. NReal)
- [Myth of the dying mouse \(p. 17\).](#)
 - each form factor has an optimal interaction technique
 - most headsets handle hand tracking, but also offer controller, keyboard and mouse support!
- **The XR equivalent of the mouse has not been invented yet!**

Touch

Mouse+Keyboard

Motion / Remotes

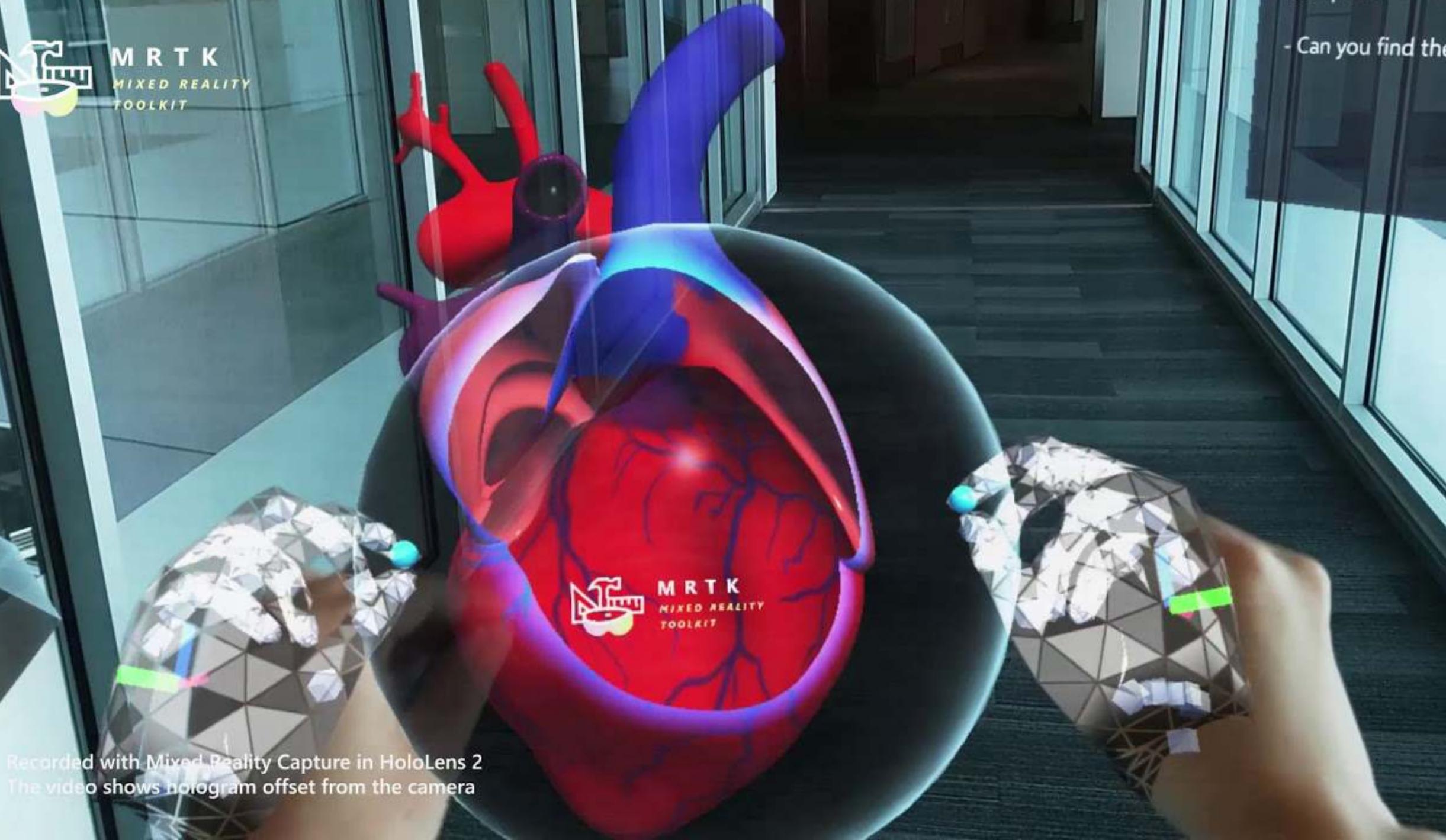
Interaction techniques

- **Screen**, when using a smartphone 
 - not very immersive but accurate, and provides tactile feedback
- **Controllers with buttons** 
 - great haptic feedback but not immersive
- HoloLens **GGV** : Gaze, Gesture, Voice   
 - natural interactions, with no external hardware
 - great but tiring, lacks privacy ("hey Cortana!"), and accuracy
- **Tangible interactions**  
 - markers or accessories to add some tactile feedback 



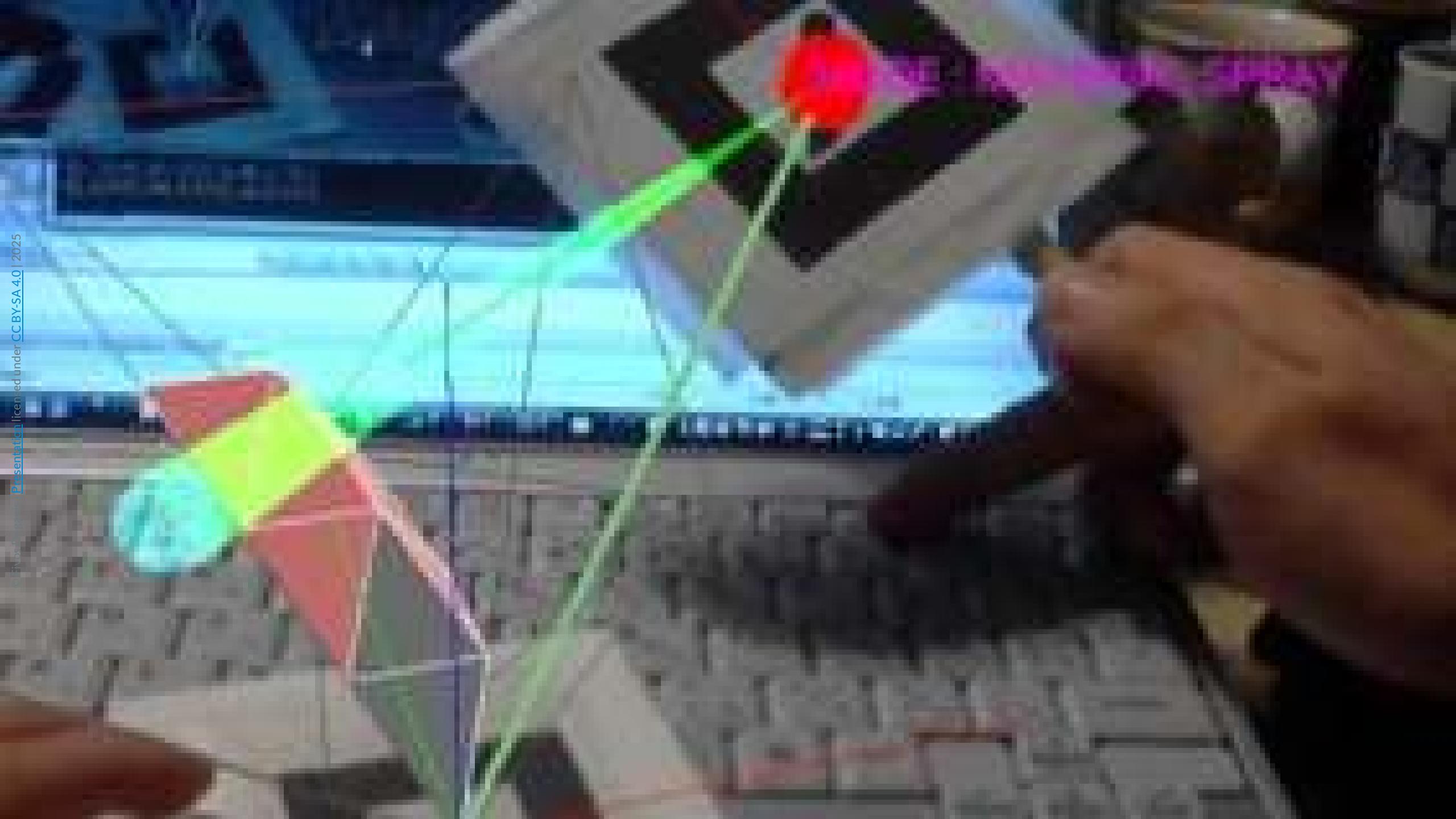


M R T K
MIXED REALITY
TOOLKIT



Recorded with Mixed Reality Capture in HoloLens 2
The video shows hologram offset from the camera

- Can you find the
manipulate them



Interactions

Conclusion

- Immersive AR interactions **have yet to be invented!**
- No interaction paradigm has become a standard yet
- We must **guide the users** and try to **understand their intent**

End

of part 1!

Questions?

Extra :)





Reconstruction

3D

- Colmap
- AliceVision
- Capturing Reality
 - RealityScan for iOS (link)
- MicMac
- Scaniverse ★



Links

<http://www.ign.fr/institut/innovation/minecraft-a-carte>

[http://lsc.univ-evry.fr/~didier/home/lib/exe/fetch.php?
media=cours:ra:ra.pdf](http://lsc.univ-evry.fr/~didier/home/lib/exe/fetch.php?media=cours:ra:ra.pdf)

Photo Credits

<https://unsplash.com/photos/RgPVZvA4wBM>

<https://unsplash.com/photos/r2CAjGQ0gSI>