Marco Cavallo

Merging Worlds: A Location-based Approach to Mixed Reality

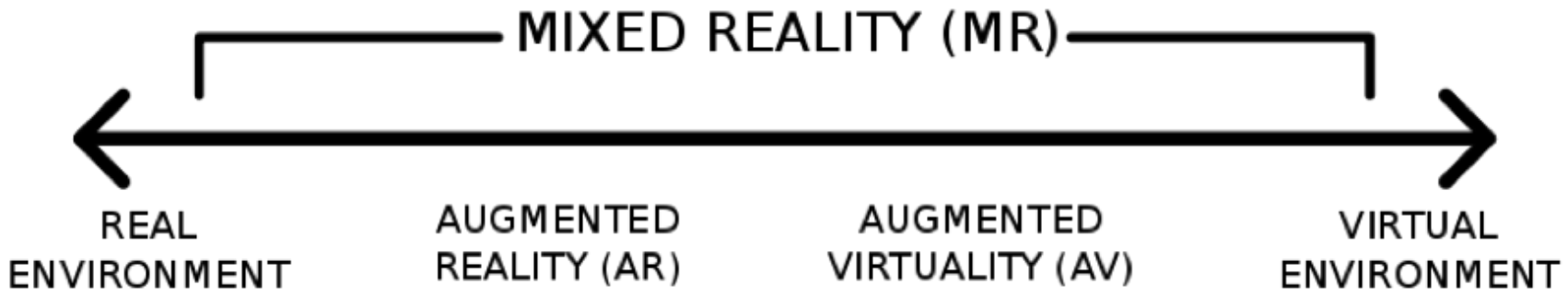
Introduction: **A New Realm of Reality**

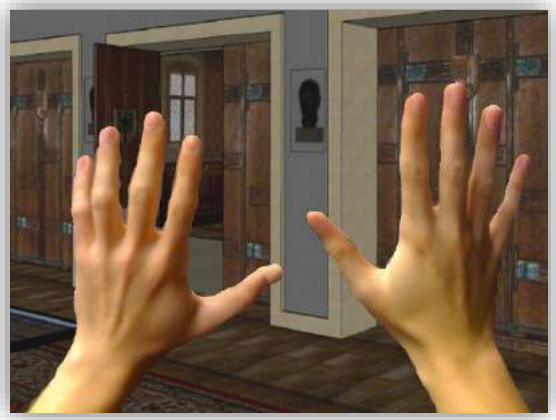
http://www.samsung.com/sg/wearables/gear-vr/

Introduction: **Bridging the Gap Between Real and Digital**

https://www.microsoft.com/microsoft-hololens/en-us

Introduction: **The Virtuality Continuum**



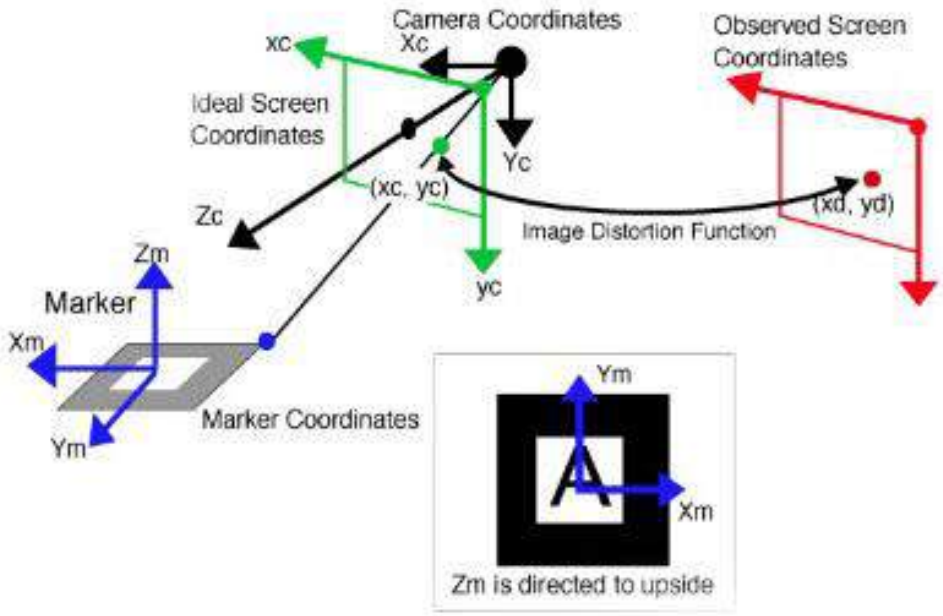


Milgram, P. and Kishino, F.: A taxonomy of mixed reality visual displays. IEICE TRANSACTIONS on Information and Systems, 77(12):1321-1329, 1994

Introduction: **Motivations And Contribution**

General purpose framework for developing MR applications

Let's get rid of the old marker-based approach :)



- Conceive user, content and fiducials **within the same reference system**
- Geolocation of content and **correspondencies between worlds**

- **Abstraction of different techniques** to show content even w/o fiducials
- **Decoupling content definition from tracking**

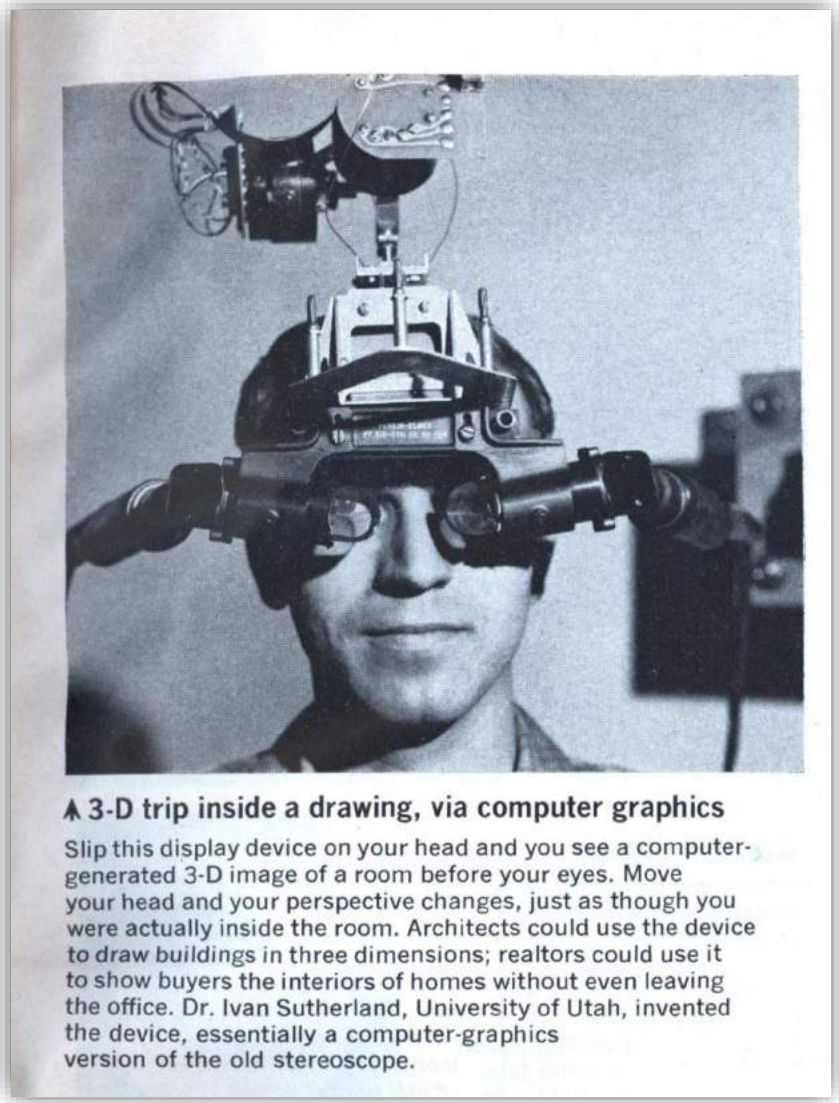
1. **Mobile application**
   >> Estimation of an absolute pose of the device camera
   >> Combining **different tracking techniques** based on the environment

2. **Authoring tool**
   >> Virtual evironment to facilitate the **customization of content**
   >> Operating **remotely** and with **real-time** modifications on clients
   >> Possibility to **preview** the experience and to **interact with users**

Presentation Structure

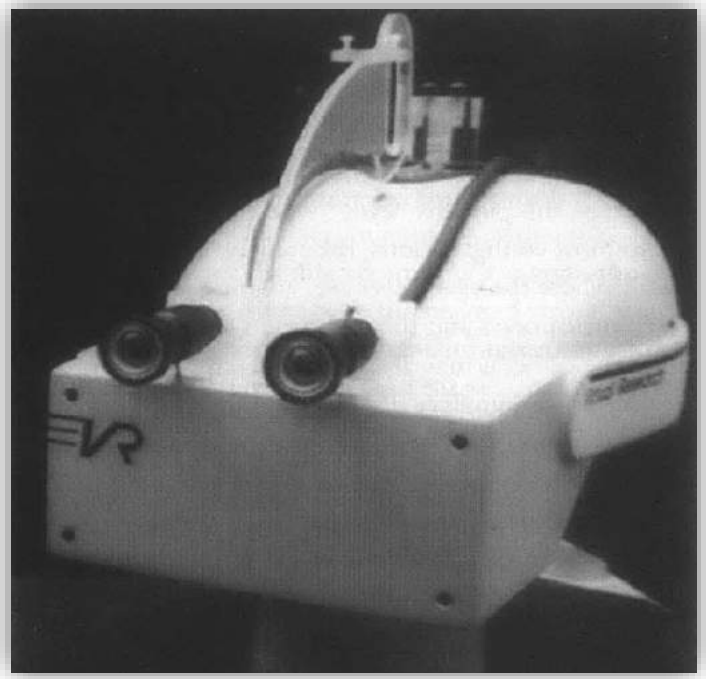
1. Introduction
2. **State Of The Art**
3. Method
4. Case Studies
5. Conclusion

State of The Art: **A Bit of History**



3-D trip inside a drawing, via computer graphics

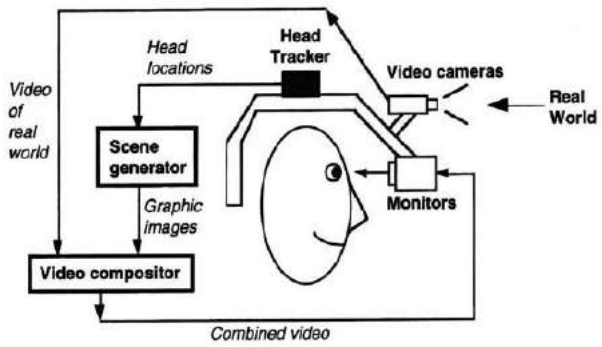
Slip this display device on your head and you see a computer-generated 3-D image of a room before your eyes. Move your head and your perspective changes, just as though you were actually inside the room. Architects could use the device to draw buildings in three dimensions; realtors could use it to show buyers the interiors of homes without even leaving the office. Dr. Ivan Sutherland, University of Utah, invented the device, essentially a computer-graphics version of the old stereoscope.

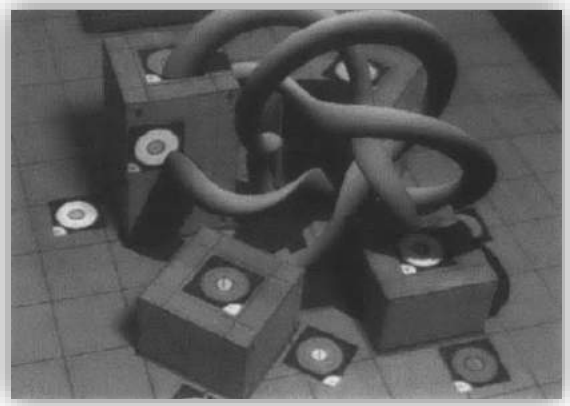


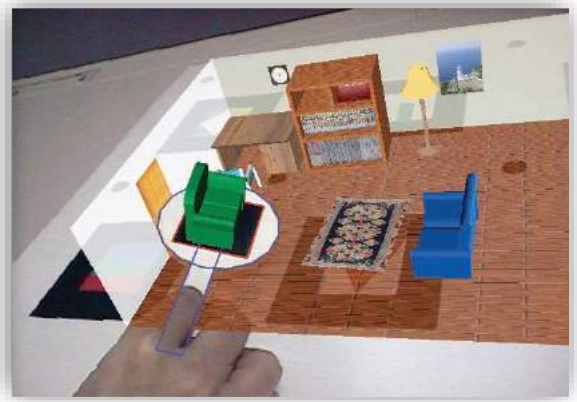
ISMAR

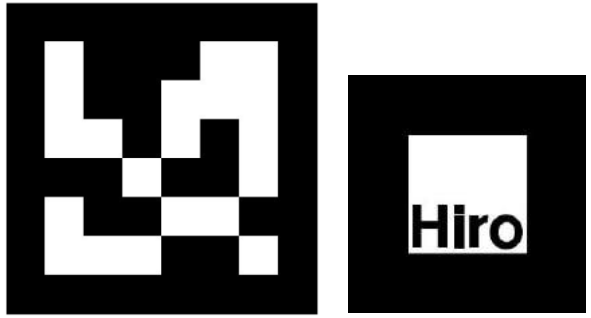
ARTOOLKIT

State of The Art: **Enabling Technologies**









- **Displays**

  >> Head mounted displays, **handheld**, projections

- **Registration and tracking**

  >> Visual and sensor-based approaches

- **Interfaces and visualization**

  >> Interaction, data density, occlusion, mediated reality

- **Human factors**

  >> Registration accuracy, eye strain and fatigue, social acceptance

State of The Art: **What Has Changed**

- Hardware **performance and cost**

- **Display technology**
  >> Resolution, light emission

- **Libraries** availability and **SLAM**
  >> PTAM

- **New interfaces** with public drivers and APIs
  >> Kinect, Leap Motion

- Design of **headsets** and **social acceptance**
  >> Leveraging the trends in VR

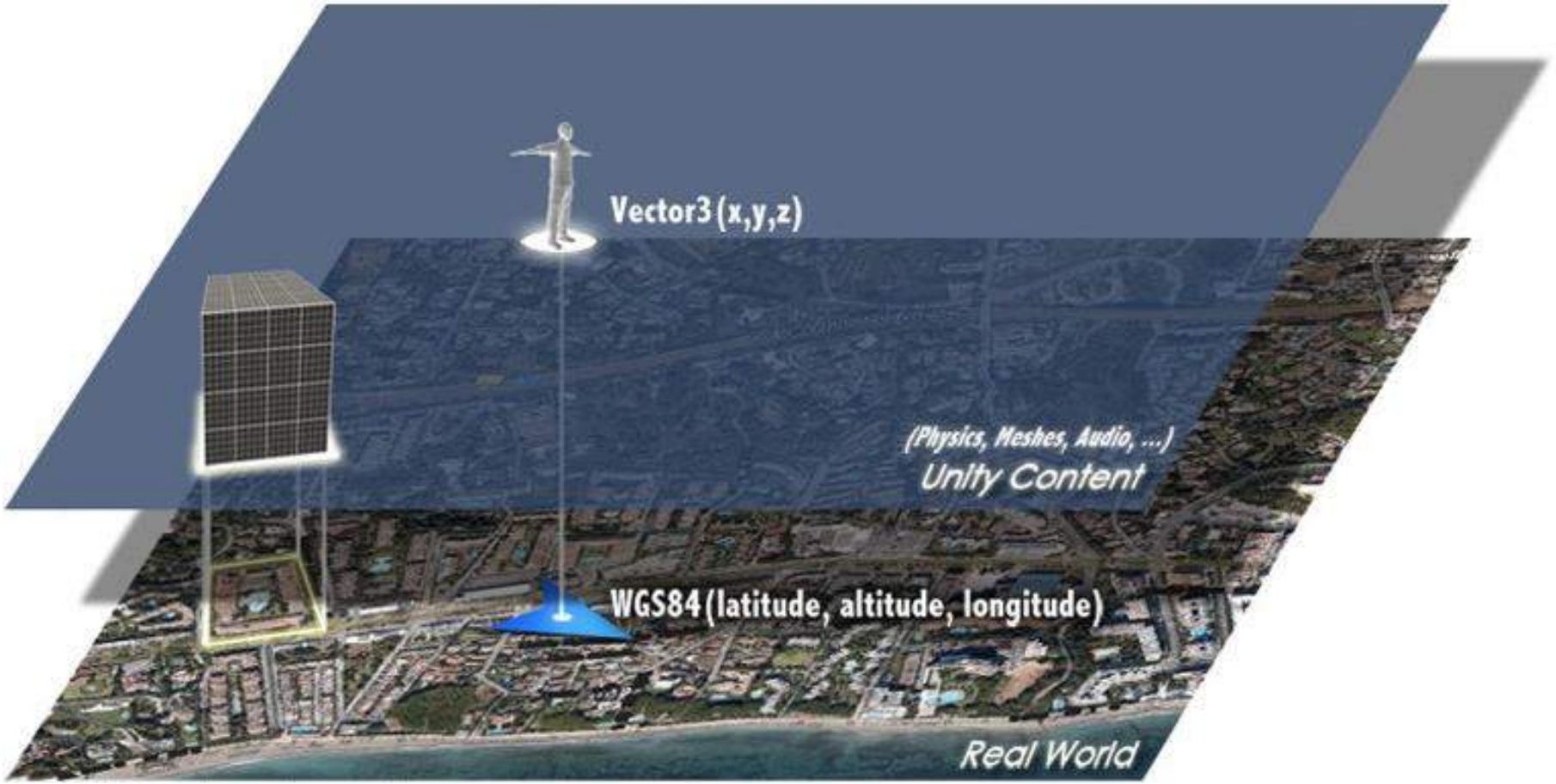
- **New applications**
  >> Mobile, marketing, storytelling, gaming, engineering, architecture

Presentation Structure

1. Introduction
2. State Of The Art
3. **Method**
4. Case Studies
5. Conclusion

Method: **Mapping The Two Worlds**

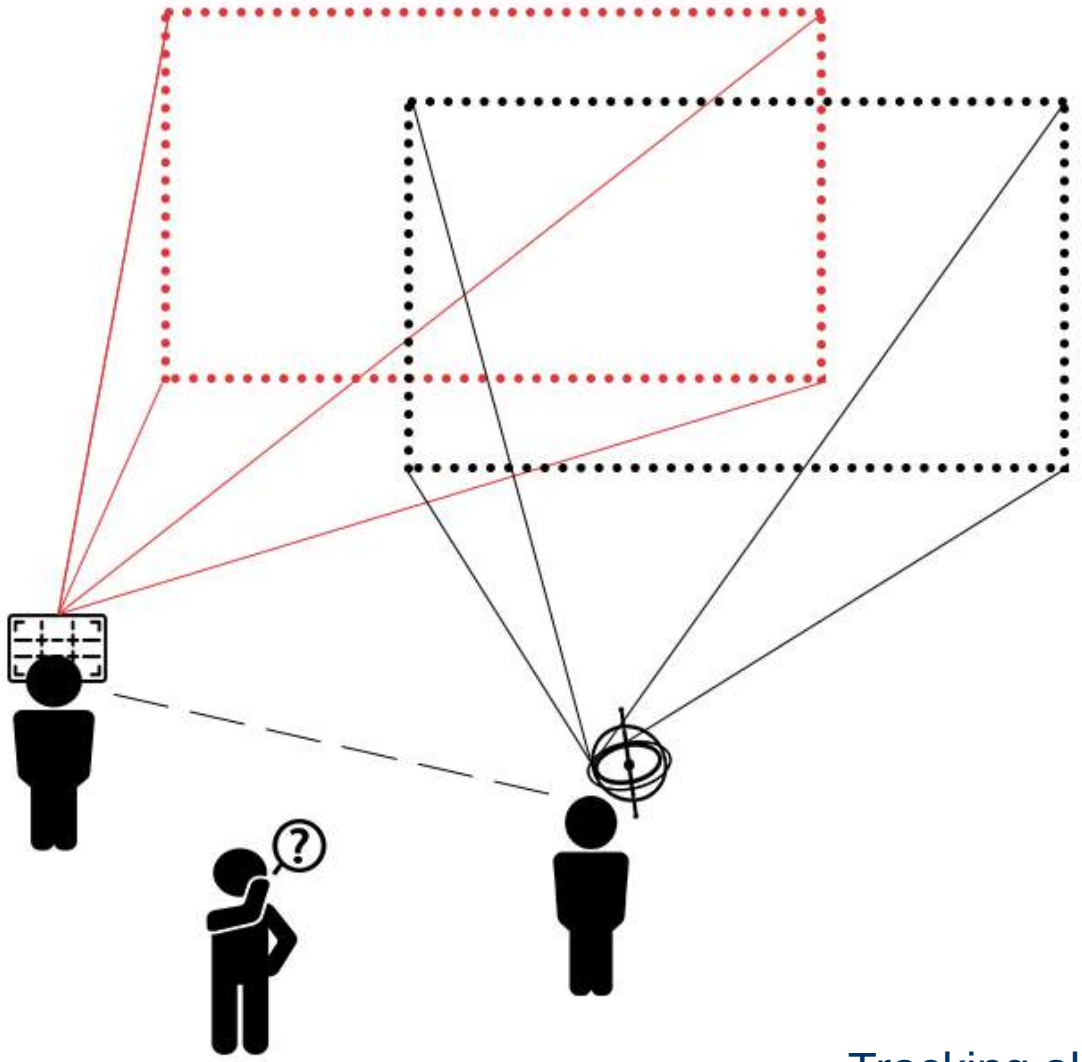
Towards an **absolute camera pose**



Both overlays and fiducials are geolocated!

… we live in a 3D world after all :)

Method: **The Dual Camera Approach**



… an abstraction for
Markerless & Location-based AR

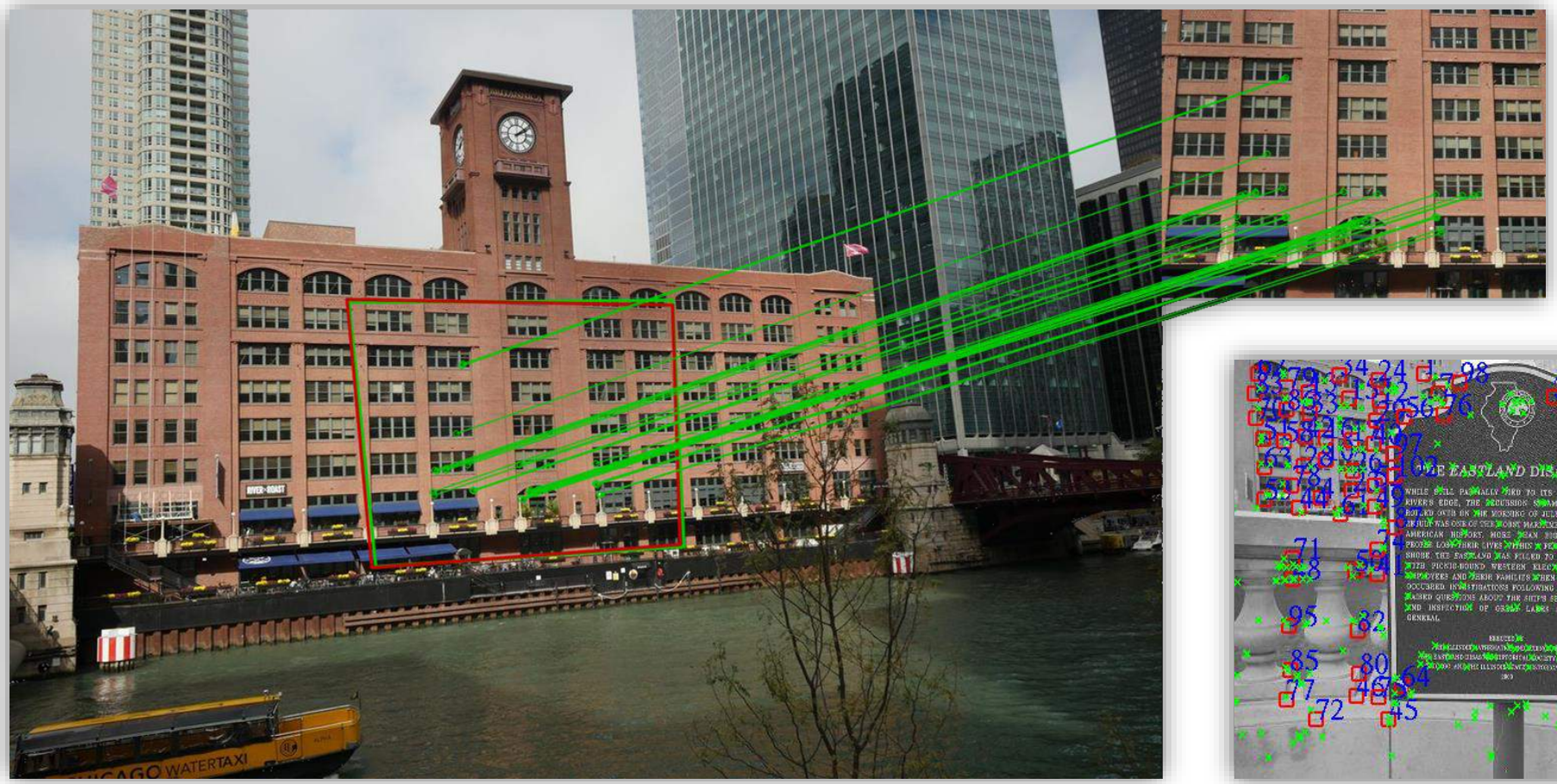
- **ARCamera**
  >> pose estimated through
  pattern-based image tracking

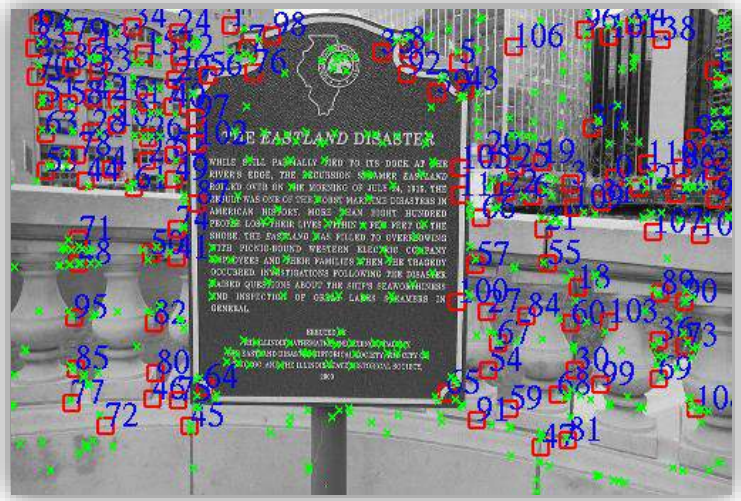
- **SensorCamera**
  >> leveraging geomagnetic field,
  accelerometer and gyroscope

Tracking algorithm independent!!

Method: **The ARCamera (1)**

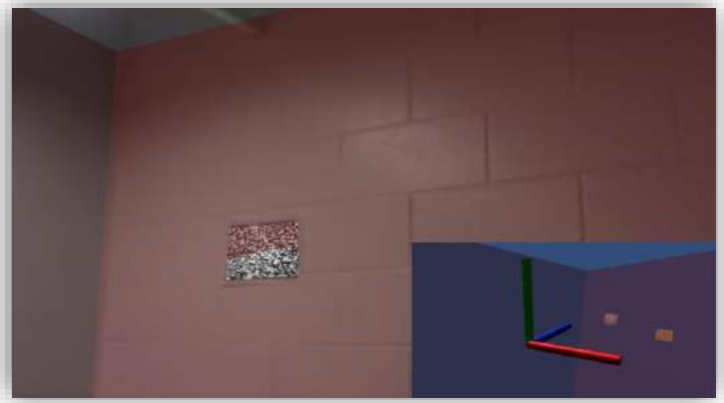
- **Image pattern-based markerless AR**
  >> Estimation of the relative pose of the camera with respect to some known features

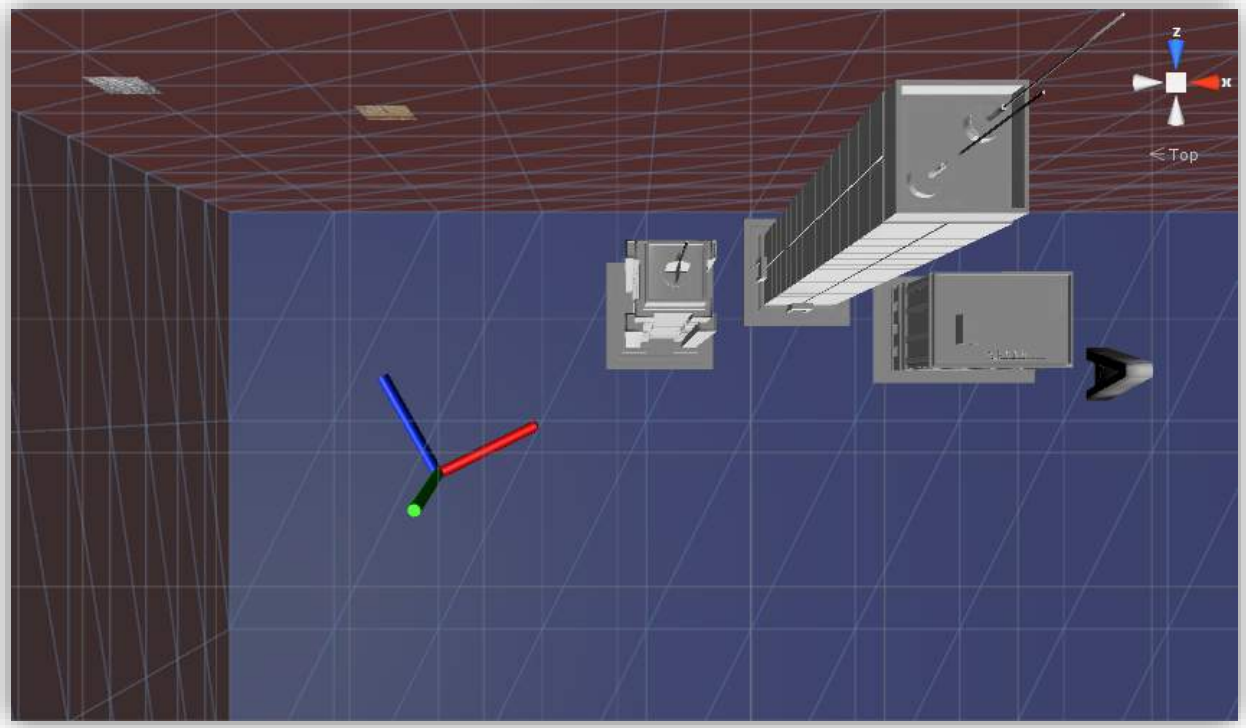




Method: **The ARCamera (2)**

- Geolocation and proper sizing of fiducials
  >> **Absolute camera pose**!

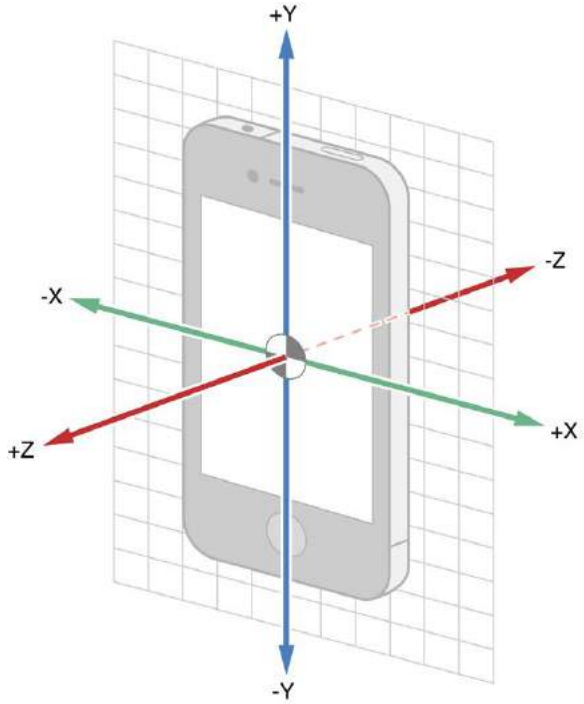


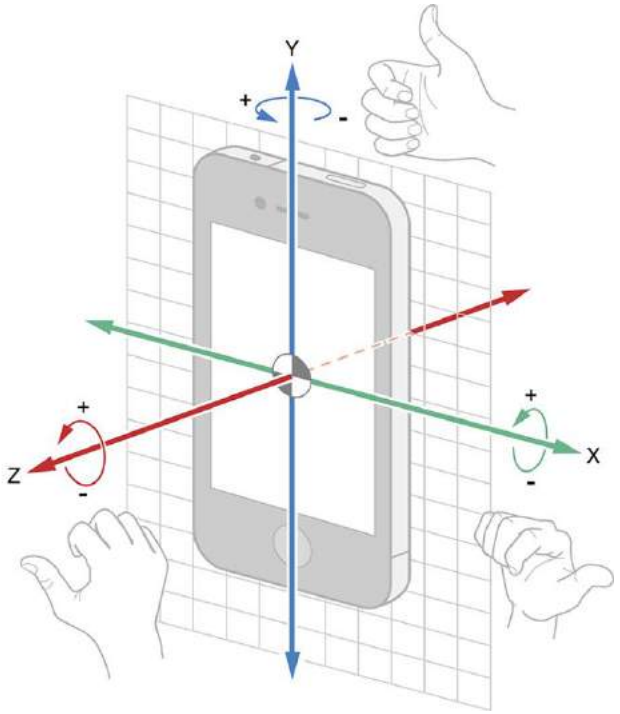


Method: **The SensorCamera**

- **Absolute position: (A-)GPS**
  >> Corrected with step detectors, multi-sensors odometry, visual odometry

- **Absolute orientation: Inertial Measurement Unit (IMU)**
  >> Accelerometer, Gyroscope, Magnetic field





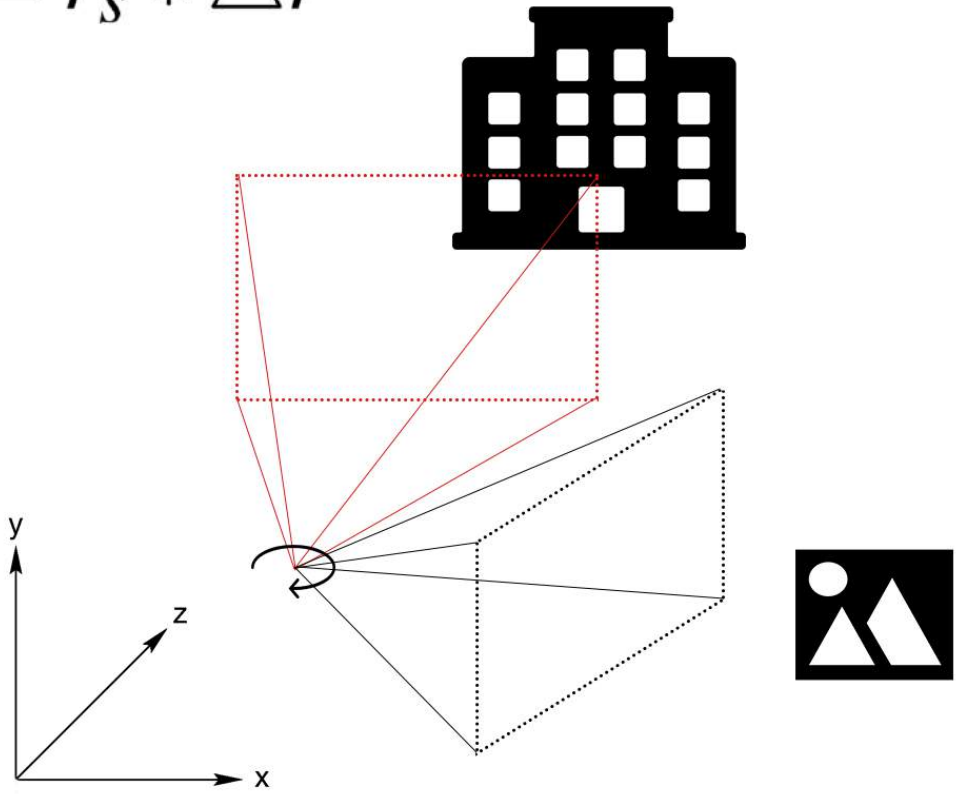
Method: **Combining The Two Cameras (1)**

$$\triangle r = r_{ar}^{-1} * r_s$$

$$r = r_s * \triangle r$$

$$pos_{smoothed} = pos_{old} + (pos_{current} - pos_{old}) * \frac{k_1}{dist} * time_{frame}$$

$$rot_{smoothed} = rot_{old}(rot_{old}^{-1} * rot_{current})^{\frac{k_2}{dist} * time_{frame}}$$



Smooth transition between four situations:

1. Fiducial found
2. Fiducial lost
3. Multiple fiducials
4. No fiducials available

Method: **General Advantages**

- Abstraction of location-based and marker-less AR as a single experience

- Know **virtual content and fiducials close to user** or to other virtual content
  >> Load and unload content based on user location
  >> **Prune the dataset** of pattern images based on proximity
  >> **Dynamic resource management**
  >> Display **virtual content even if no tracking is available**

Geolocated content

- Know movements relative to content or to other helper camera
  >> Know how much user is **moving away from a tracked object**
  >> Display **content even if tracking has been lost**
  >> **Know how to orient user** towards other virtual content
  >> Know how to **filter virtual content if overlapping**

Absolute camera pose!

- Signaling **incoherent situations** and false positives / negatives

- **Intelligent camera smoothing** (stabilization)

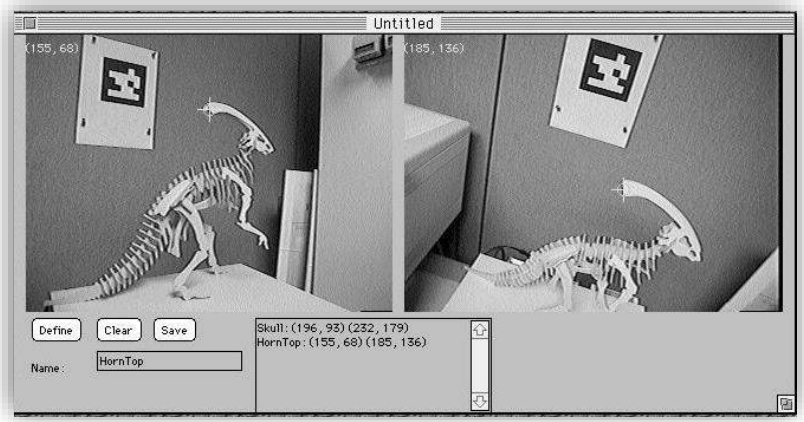
... and definition of an environment suitable for editing

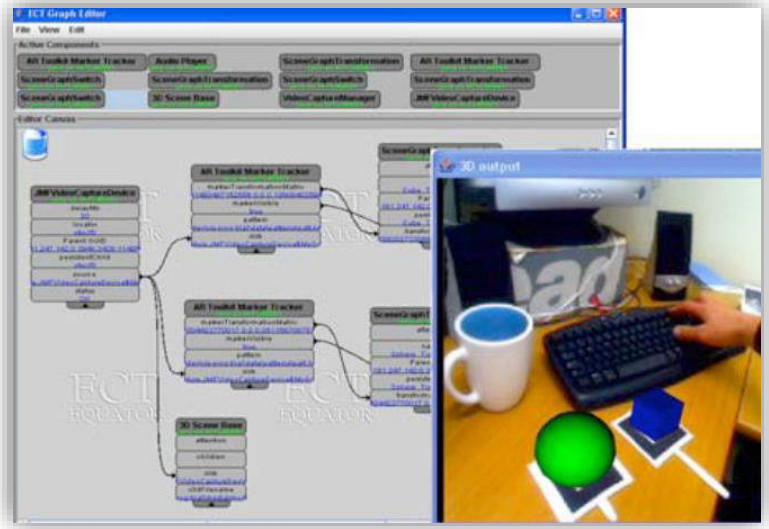
Method: **Towards A Third Camera Approach**

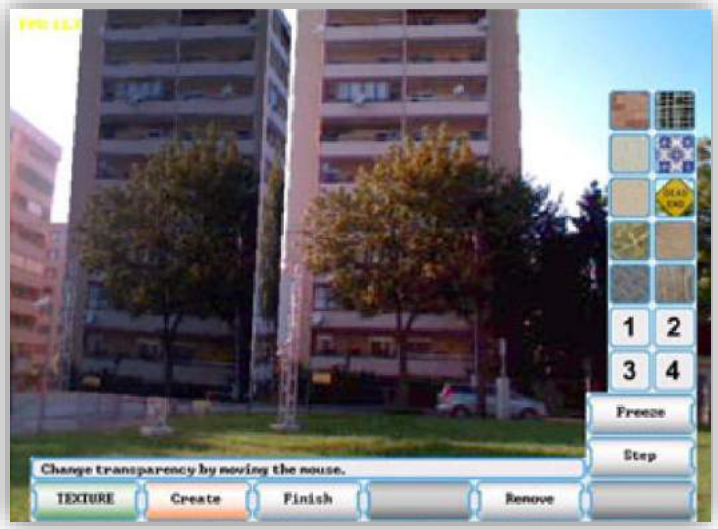
Features extracted at runtime!!

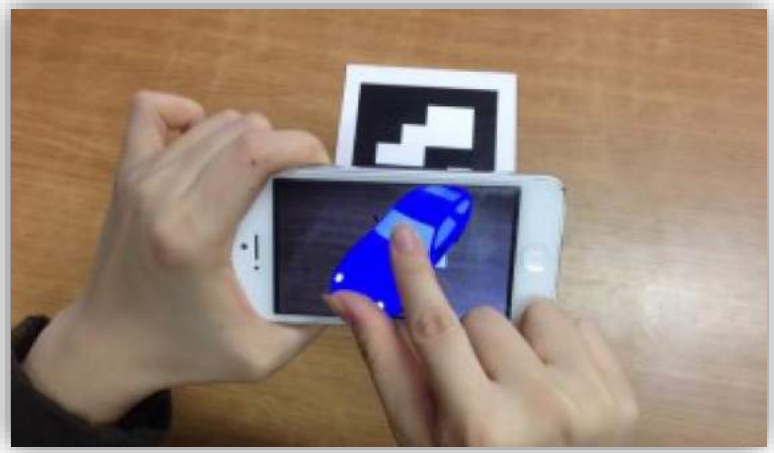
- Small movements
  >> Activate when close to content
- Extended tracking
  >> When fiducial lost
- Light / weather independent

Method: **The Need For An Authoring Tool**









- Technical challenges
- Non intuitive
- Many limitations
- Previewing
- Interfaces?

Editing
MR within MR

Method: **Authoring In a Virtual Environment**

- **Decouple editing from tracking** through a partial environment reconstruction
- **1-to-1 real-world mapping** and geolocation of content

Febretti A. et al.: Cave2: A Hybrid Reality Environment For Immersive Simulation And Information Analysis, IS&T/SPIE Electronic Imaging, 2013.

Method: **Protocol Definition**

- **Virtual content**

  >> Images

  >> Videos

  >> 3D meshes

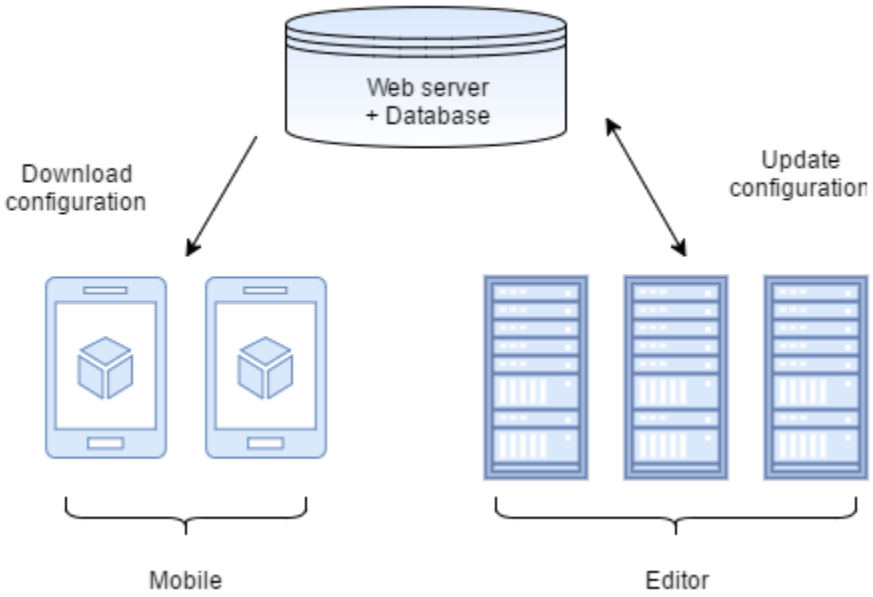
  >> Spatial audio

- **Fiducials**

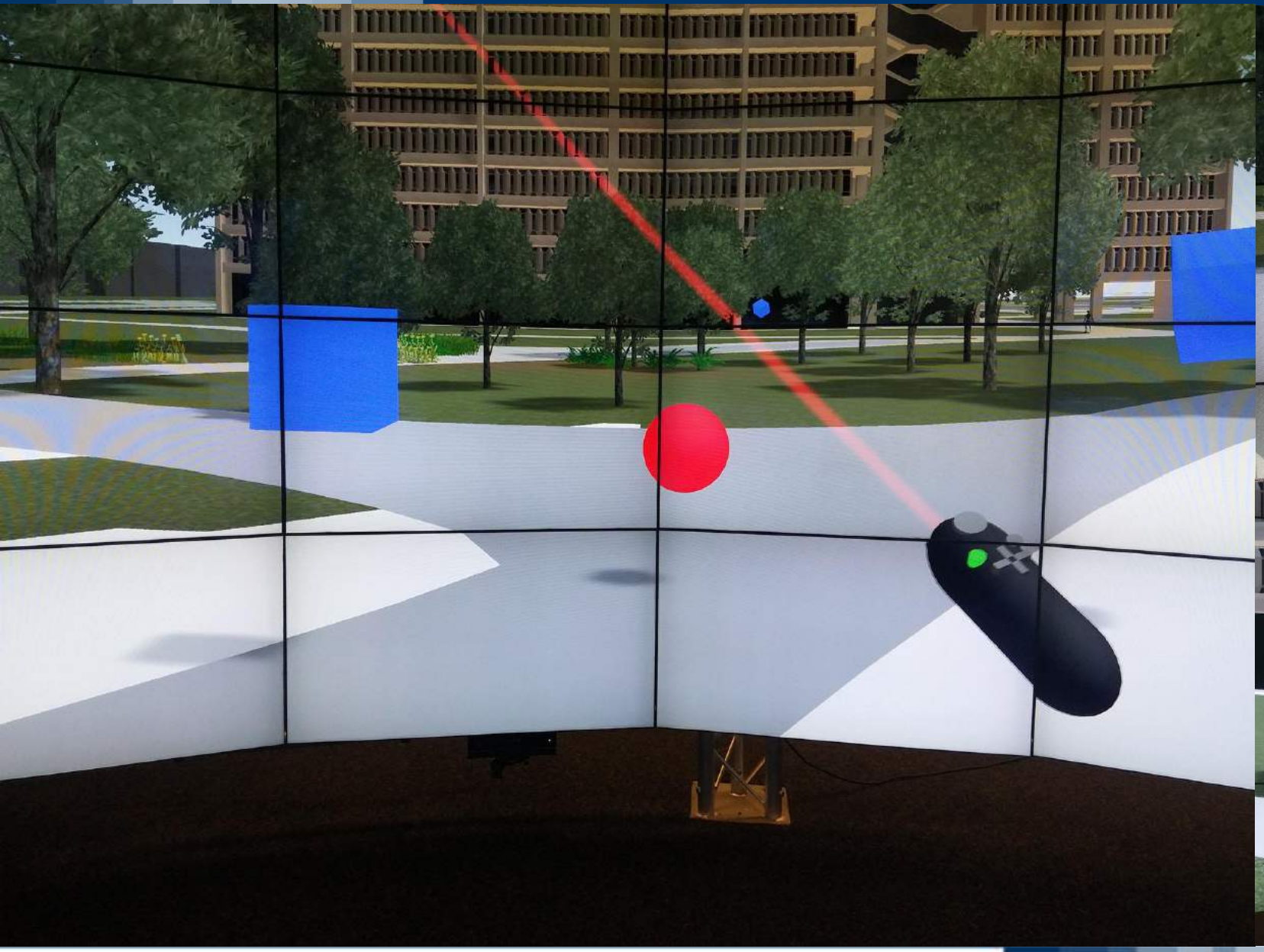
  >> Pattern images

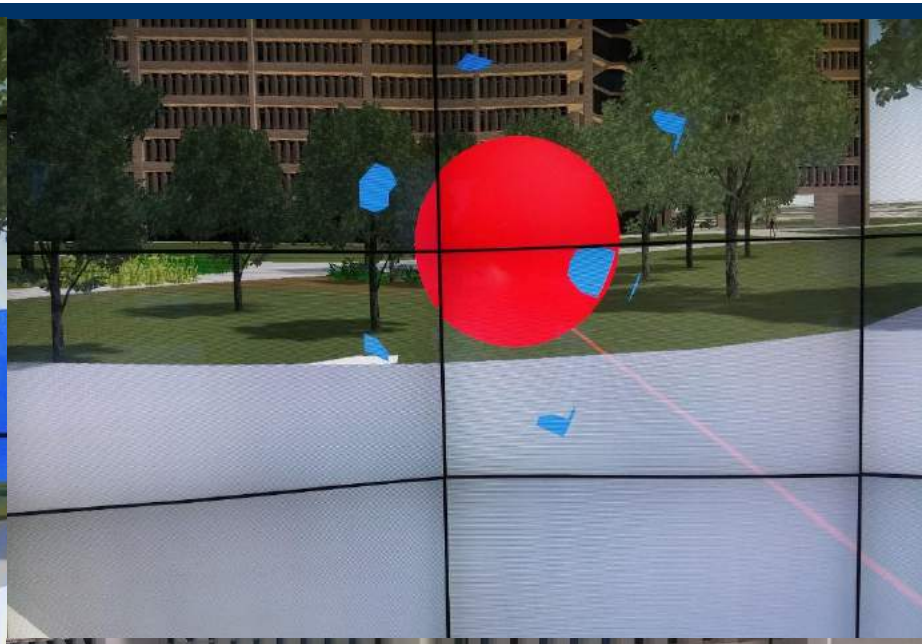
  >> (Bluetooth beacons)

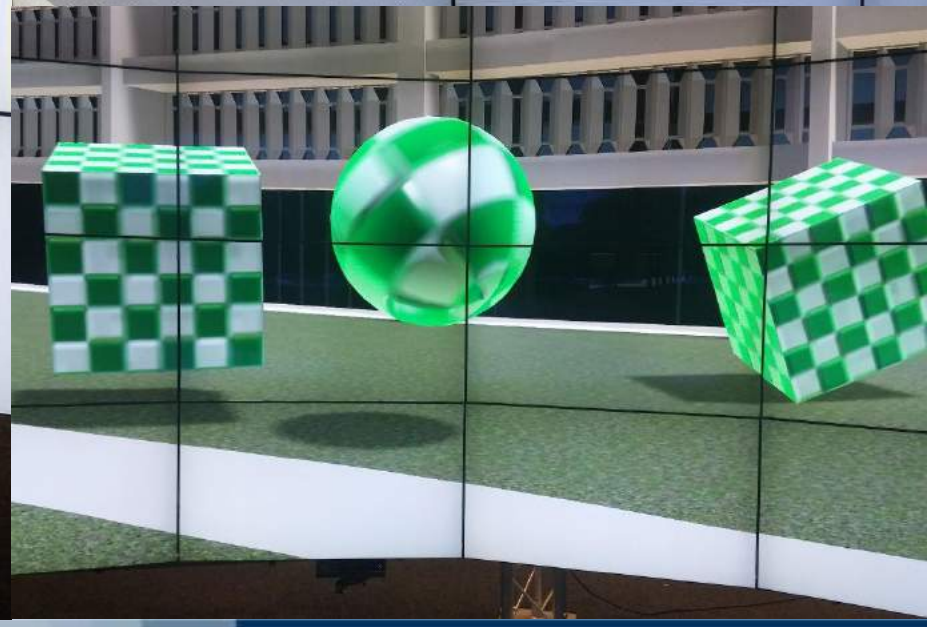
```
"data": {
   "title": "MR Example",
   "last_edited": 1460851200,
   ... //other generic information
   "objects": [{
      "id": 1,
      "asset": "InfoPanel", //content to load
      "type": 1, //bidimensional image
      "position": {41.867272, -87.675434}
      "rotation": {90, 37.4, 0},
      "scale": {2, 2, 2},
      "properties": {
         "angle": 30,
         "area": 10, "visibility": "fading",
         "transparency": 1
      },
      "media": "",
      "cache": "no"
      },
      ... ],
   "fiducials": [{
      "id": 1,
      "asset": "BuildingFacade", //associated fiducial
      "type": 0, //pattern image
      "position": {41.867496, -87.673975}
      "rotation": {90, 22.7, 0},
      "scale": {3.4, 3.4, 3.4},
      "properties": {
         "smoothing": 3.5,
         "reliability": 1,
         "distance": 30 }
      },
      ... ]
   }
```



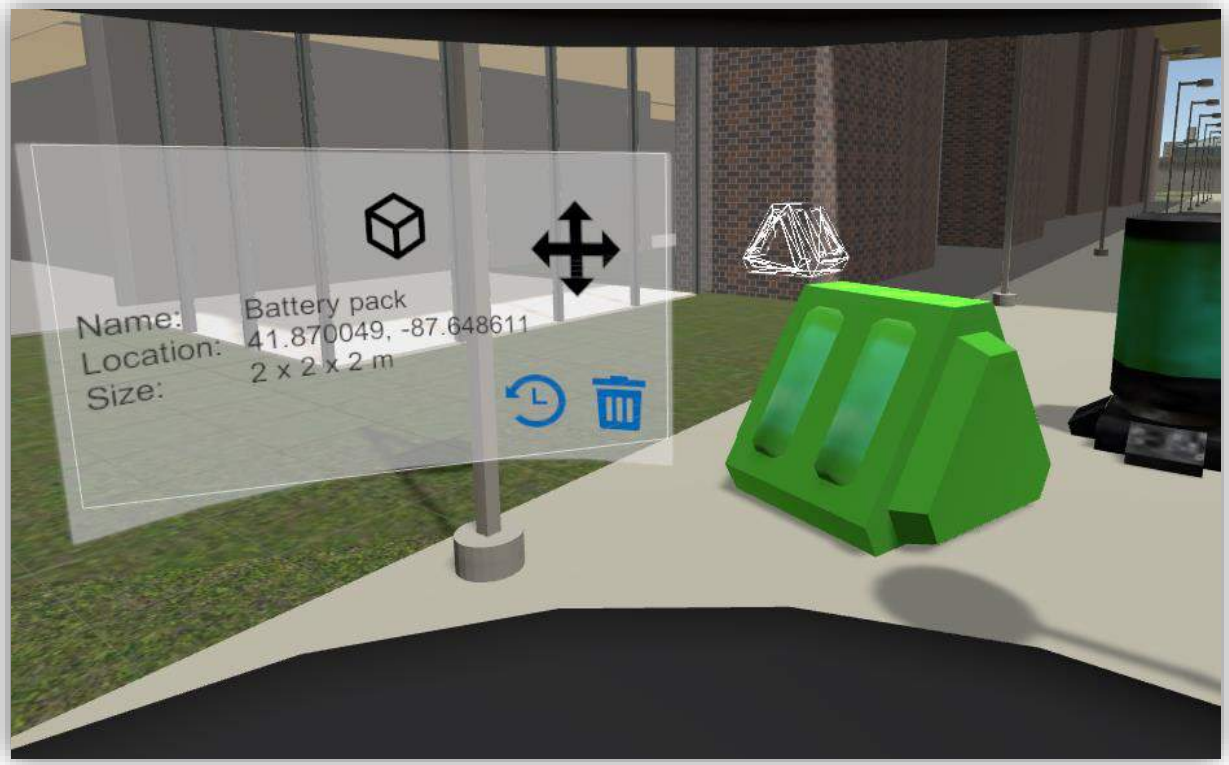
Method: **Interacting With Content (1)**







Method: **Interacting With Content (2)**



- Different **editing modes**
- **Categorization** of virtual elements
- **Visual feedback** to distinguish virtual content
- Detail panels to show **object information**

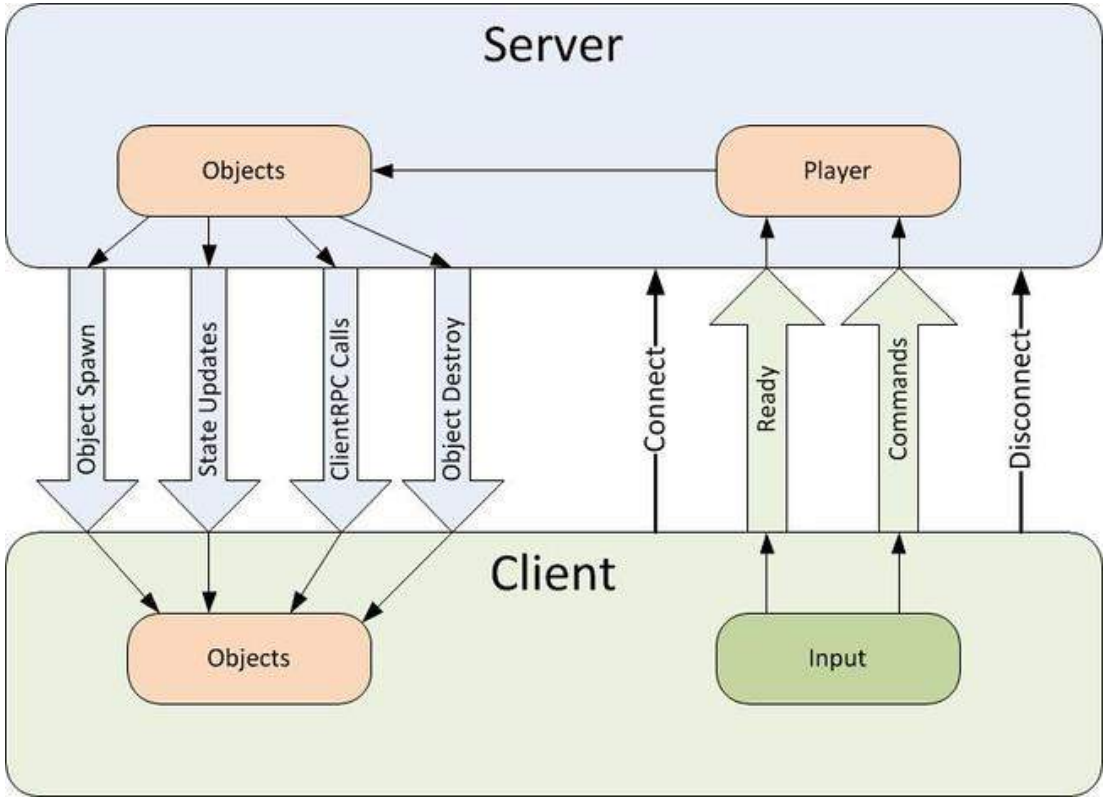
Method: **Real Time Editing**

What if we could…

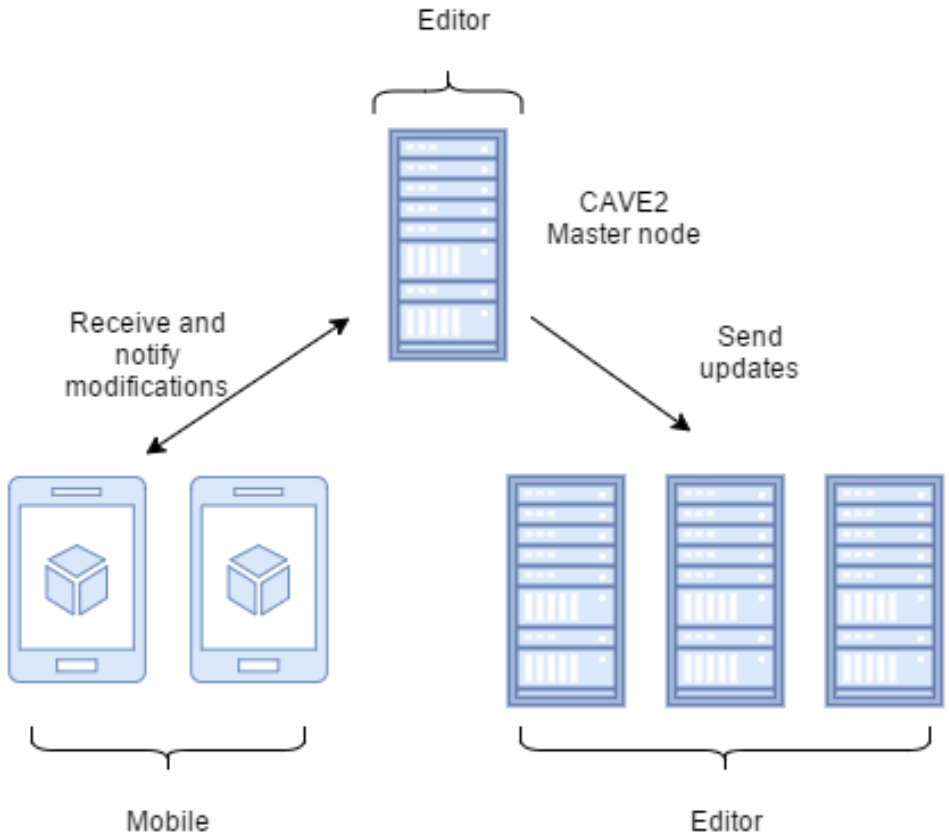
- Modify at **runtime** the MR experience
- Monitor **how users behave** while using our application
- **Interact selectively** with them

… real time functionalities need to be added!

Method: **System Architecture**

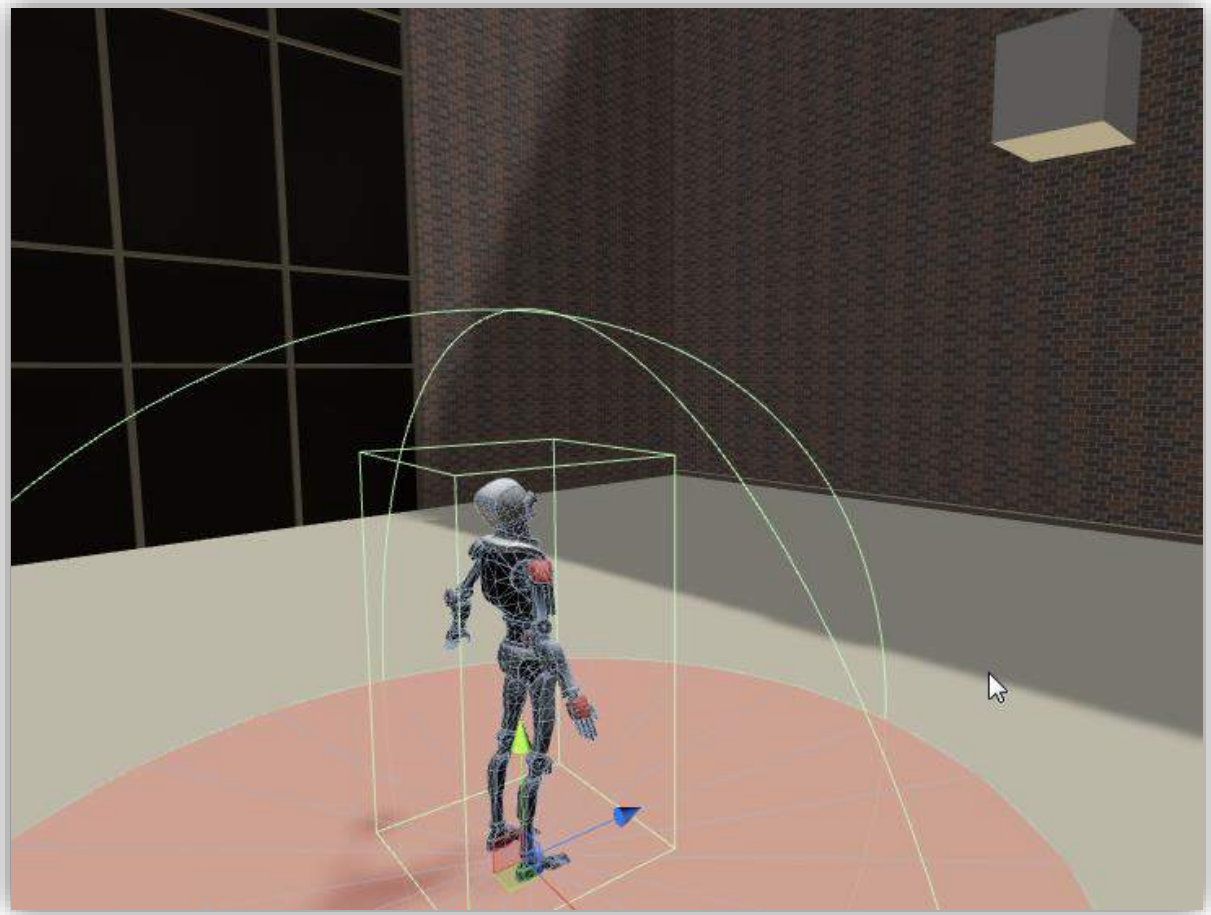


http://docs.unity3d.com/Manual/UNetActions.html



- Only the master node is connected to mobile devices
- Need to maintain consistency!
- 3 – 150ms delay during our tests, 10-30 FPS updates

Method: **Representing Users**

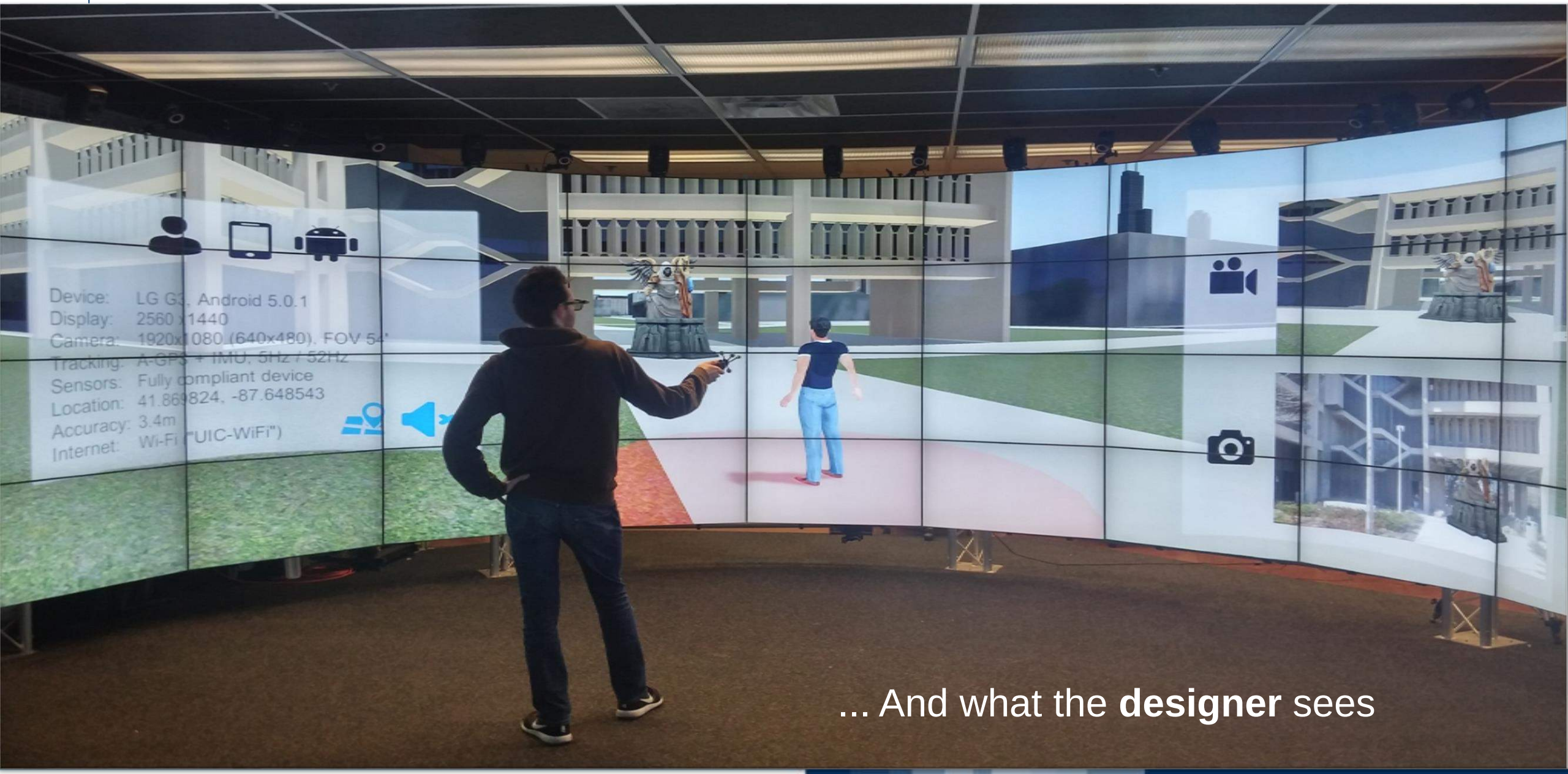


5+2 DOF representation: geolocation and head orientation + accuracy visualization

+ lots of other information gathered from their mobile device :)

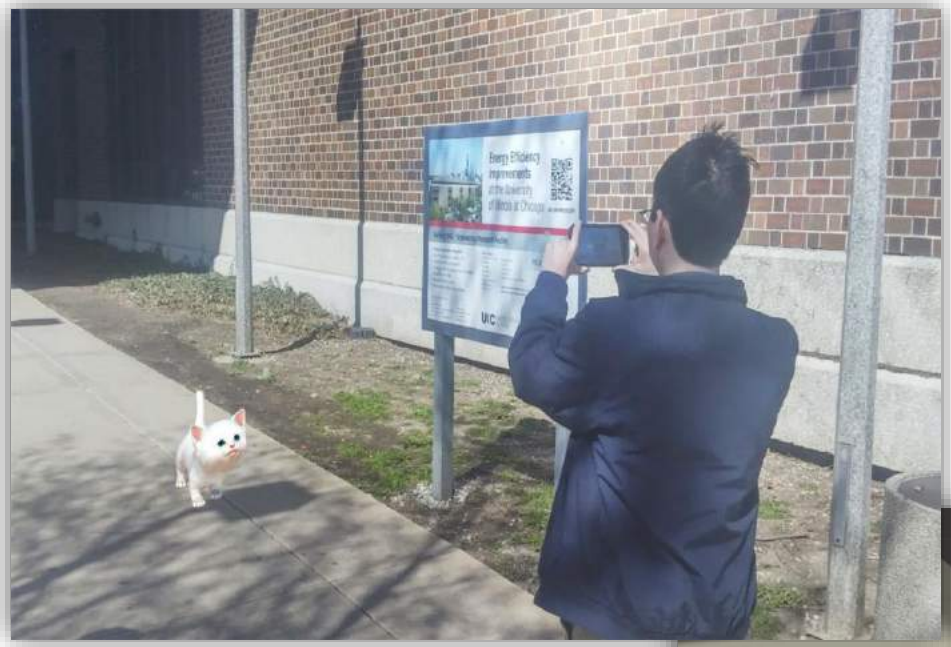
Method: **Interacting With Users (1)**

Method: **Interacting With Users (2)**



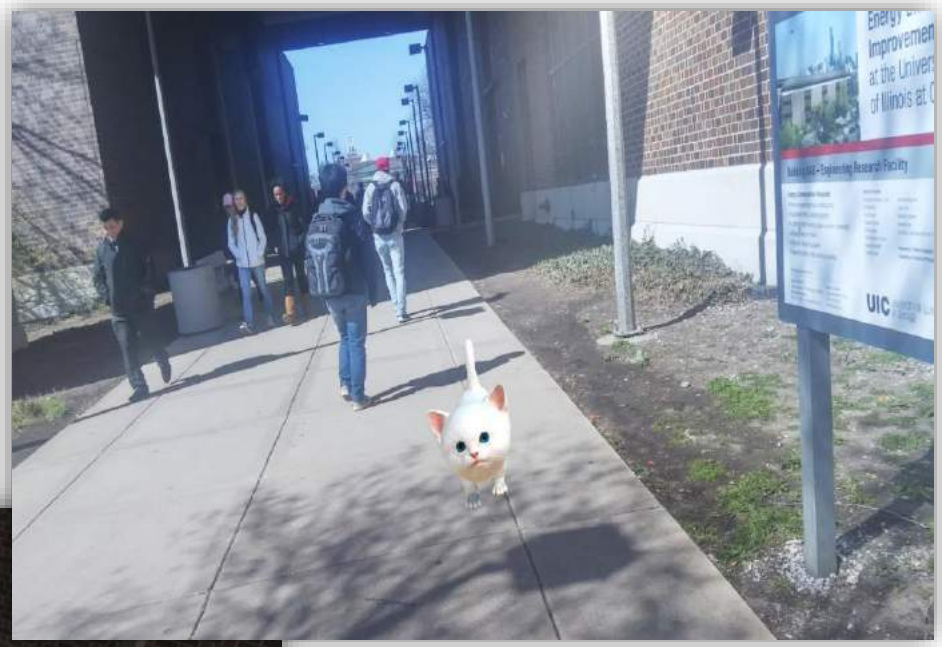
... And what the **designer** sees

Method: **Interacting With Users (3)**



External view

Designer view



User view

+ ‘’user perspective’’ mode...

Method: **Real-time Applications**

- **Detecting design flaws** and **correcting registration errors**
  >> Adjusting the location of virtual content

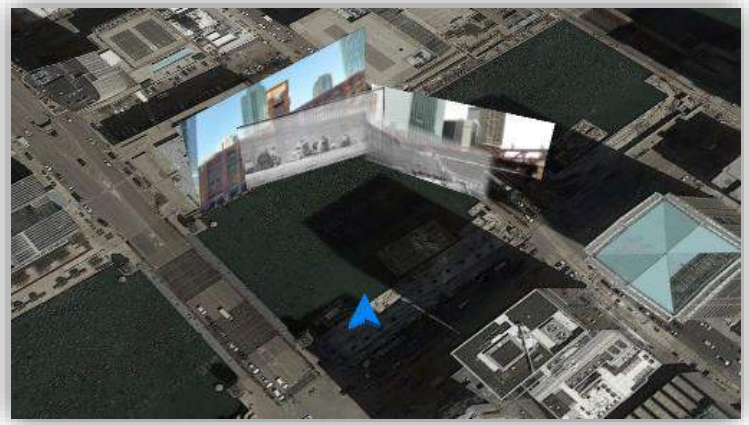
- **Monitoring users' behavior** and previewing the environment surrounding them

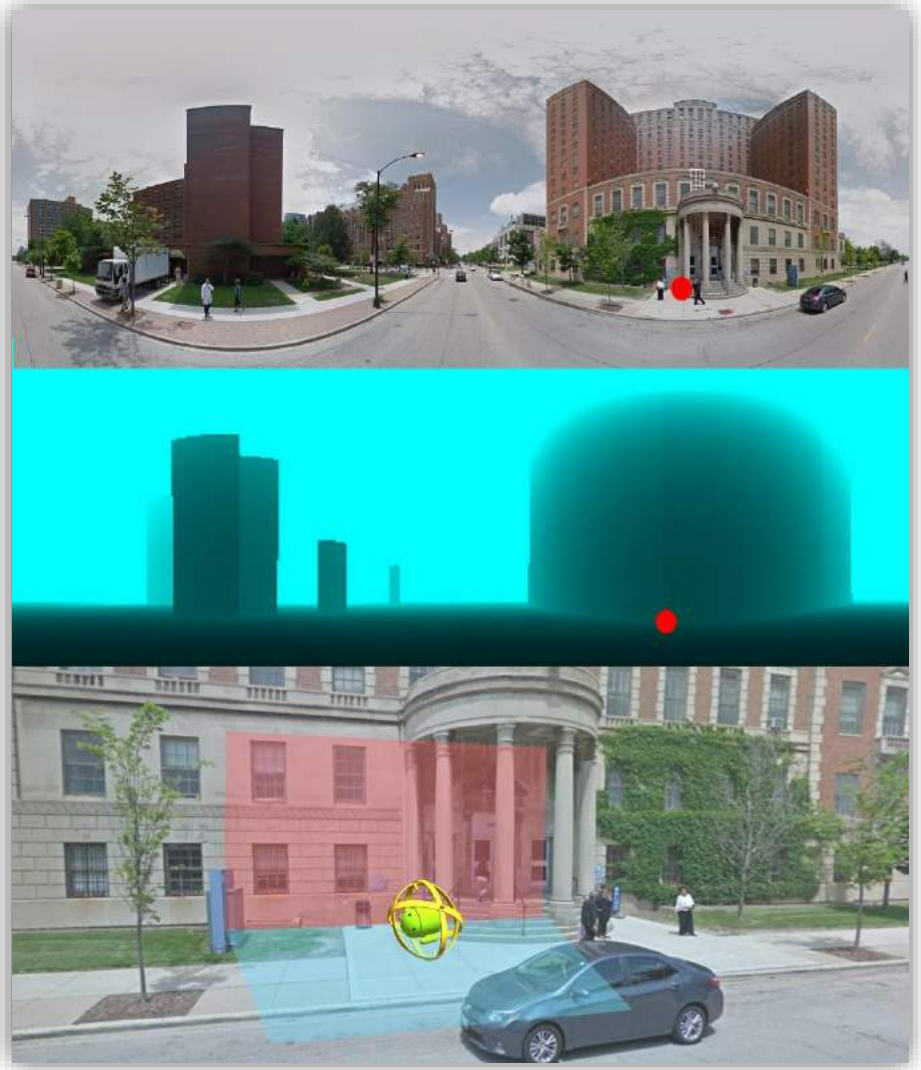
- **Adaptation** to external events or environmental conditions

And also...

- Allowing users to contact the designer and notify anomalies
- Possibility to **provide assistance to users remotely**
- Assignment and **centralized control** of individual and collaborative tasks
- **Allowing users to customize themselves** the environment
- Enabling dynamic content and narratives
- **Inserting the designer himself as an avatar** inside the MR experience!

Method: **Alternative Implementations**





Presentation Structure

1. Introduction
2. State Of The Art
3. Method
4. **Case Studies**
5. Conclusion

Case Studies: **Chicago 0,0**

Case Studies: **Chicago 0,0**

Case Studies: **Overlaying Historical Photos (1)**

- Urban Environment
- **Low sensor accuracy**

- **2D content** (transparencies)
- Precise overlay on current views of the city

Case Studies: **Overlaying Historical Photos (2)**

In collaboration with the **Chicago History Museum**

>> Huge archive of historical photos never published before

Case Studies: **Estimating The Camera Pose (1)**

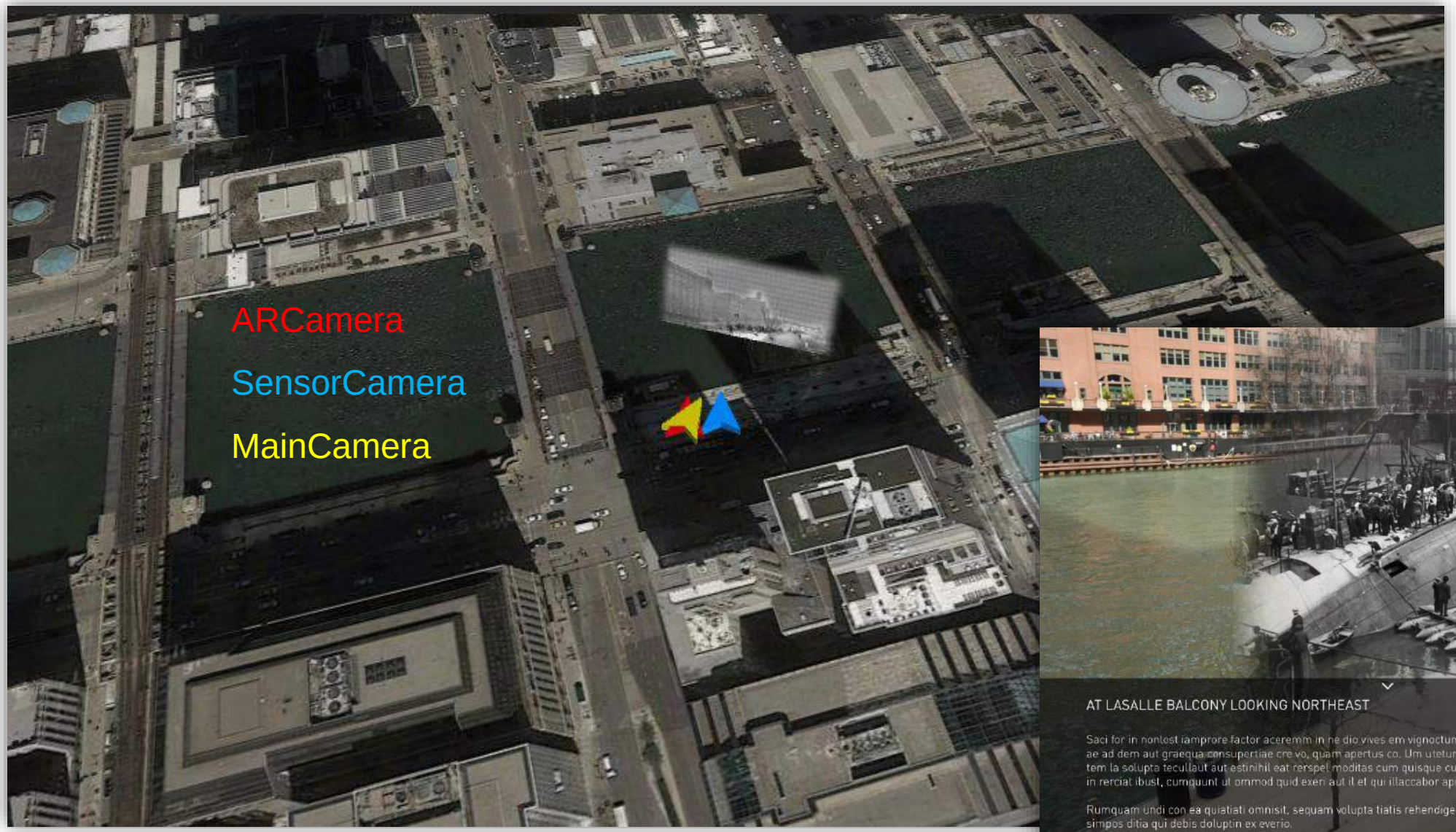
- Repetitive patterns
- Uniform color
- Lighting & weather

Oversampling :-/

Unreliable sensors: use fiducials to start tracking

>> Big enough and available from all viewpoints...

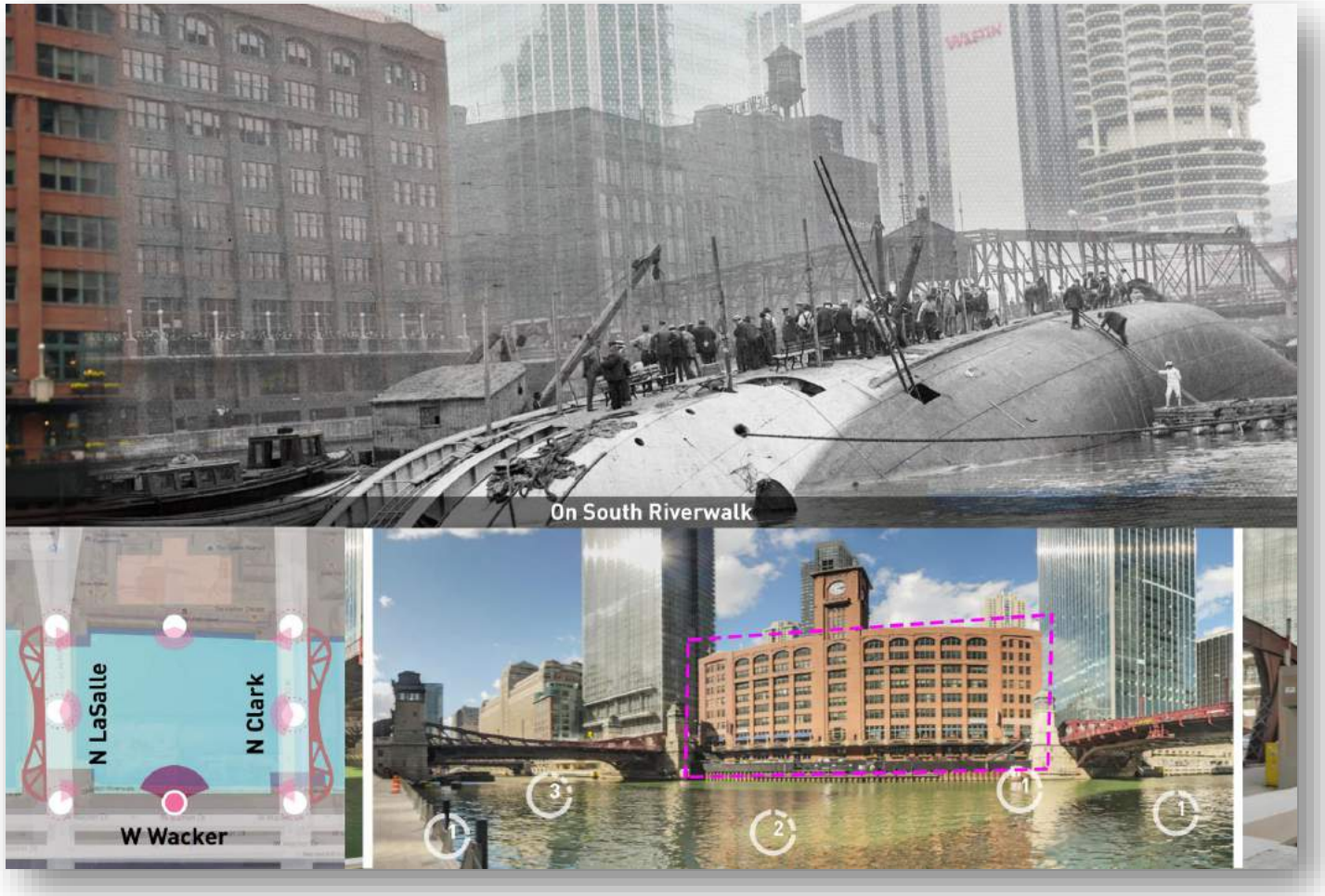
Method: **Estimating The Camera Pose (2)**



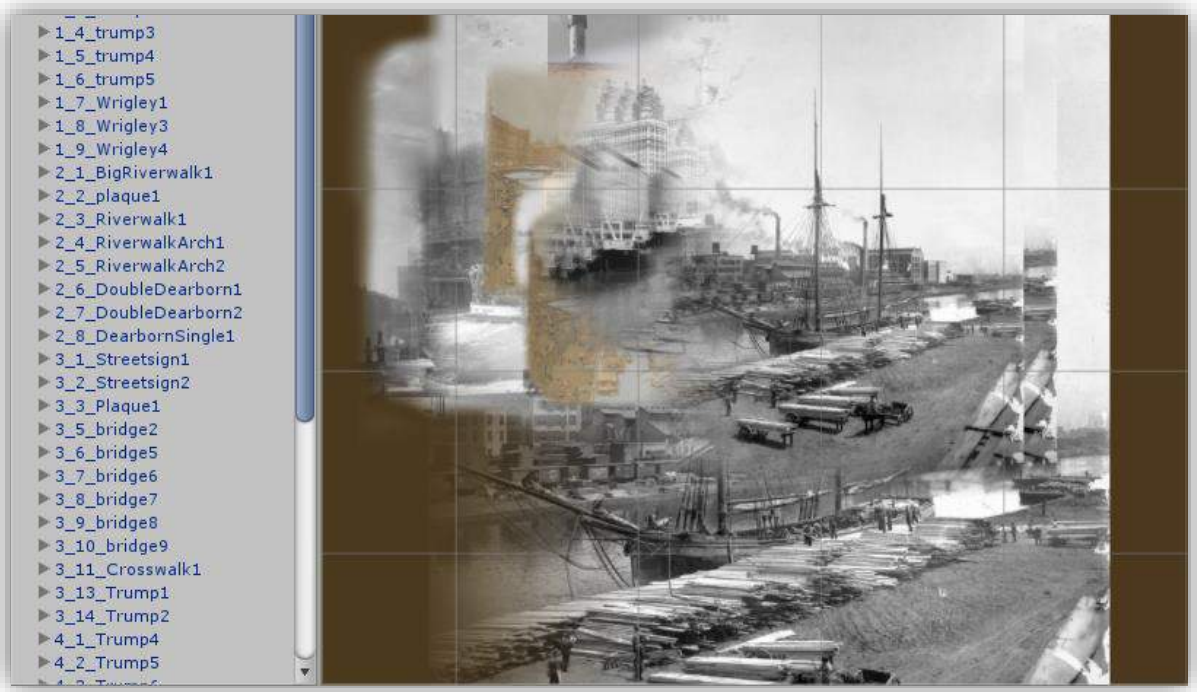
1-4 m accuracy

Case Studies: **Navigation And Information Browsing**

- **Color coding** and showing available content
- **Showing target** where the user has to aim
- Indicating how to reach **next object**
- Indicate when device needs **calibration**
- Handling of adjacent / overlapping content
- Suggest **particular view points**
- **Map** + optional navigation
- Allow user to **correct by himself**?



Case Studies: **Authoring Comparison**



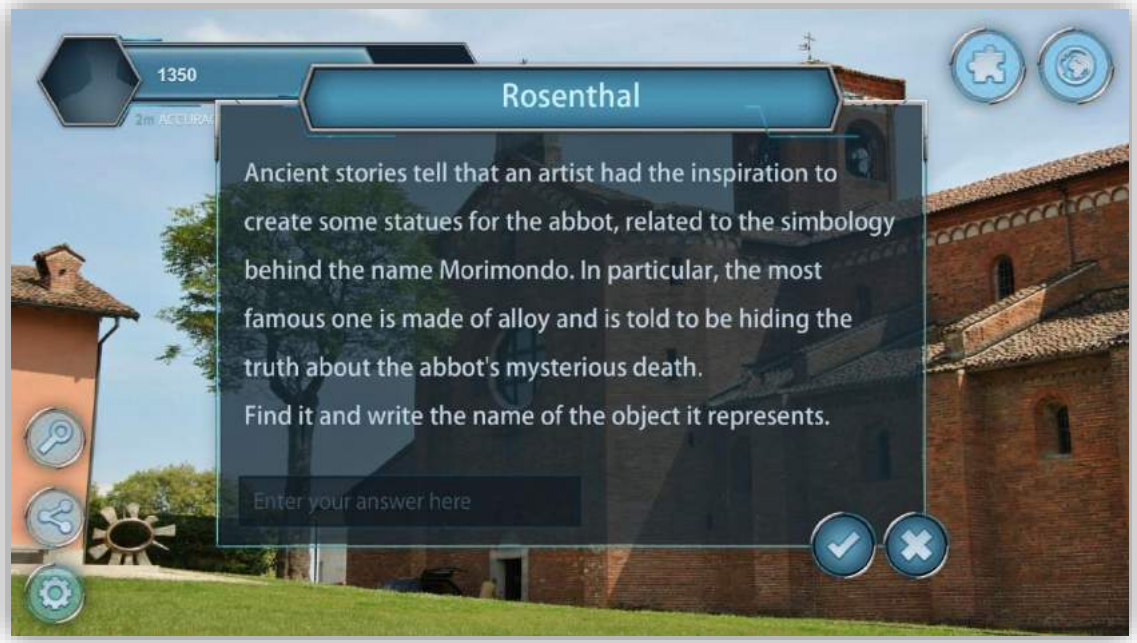
Without our authoring tool...

- No concept of **tridimensional space**
- No **knowledge about the environment**
- Virtual content in **isolated reference systems**

Case Studies: **Digital Quest**

- Mostly outdoors
- **Higher sensor accuracy**
- 3D virtual content
- Custom behaviors
- **Lower requirements**

Case Studies: **Digital Quest**



2-9 m accuracy

- An innovative version of classical real-world **scavenger hunts**
- Players compete with other participants in **finding virtual objects and solving puzzles**
- The way virtual content is displayed varies based on the challenge to be solved
- Definition of **relationships and precendences** among virtual objects

Case Studies: **Extending The Protocol With Behaviors**

```
"data": {
        "title": "DigitalQuest Example",
        "last_edited": 1460851200,
        "start_time": 1460973600,
        "end_time": 1460977200,
        ... //other event information
        "objects": [
                { ... },
                {
                "id": 4
                "asset": "DarkSkull", //content to load
                "type": 0, //3D model
                "position": {41.867272, -87.675434}
                "rotation": {-12, 76.4, 0},
                "scale": {2, 2, 2},
                "properties": {},
```

```
                "attributes": {
                        "name": "Rosenthal",
                        "precedences": "((1&2)|3)!4",
                        "pre": "Ancient stories tell that...",
                        "answer": ["sun", "planet"],
                        "post": "The sun is strictly related...",
                        "score": 250,
                        "type": 0, //no additional content
                        "content": "",
                        "distance": 4,
                        "map": "visible",
                        "icon": "self", //representation on the map
                        "animation": "SplashDefault",
                        ...
                }

        },
        ...],
        "fiducials": [...]
}
```

Case Studies: **Real-time Editing And Guidance**

- **Finding flaws** in the design of the MR experience
- **Correcting how users see content** at runtime
- Studying the **behavior of users**
- **Dinamically change** the ‘’rules’’ of the game

Presentation Structure

1. Introduction
2. State Of The Art
3. Method
4. Case Studies
5. **Conclusion**

Conclusion: **A Brief Recap**

A **spatial approach** combining **different types of tracking**…

- *Chicago 0,0*
- *Digital Quest*

…in order to obtain an **absolute camera pose**

- Defining user, content and fiducials within the **same reference system**
- Creating a **1-to-1 mapping** between the two worlds
- Creating an **abstraction for different tracking** techniques
- Allow **remote and real time** editing
- Possibility to **preview** a MR experience and **monitor the behavior** of users

Thanks For Watching :)