



CAPTIONS IN 360 VIDEO

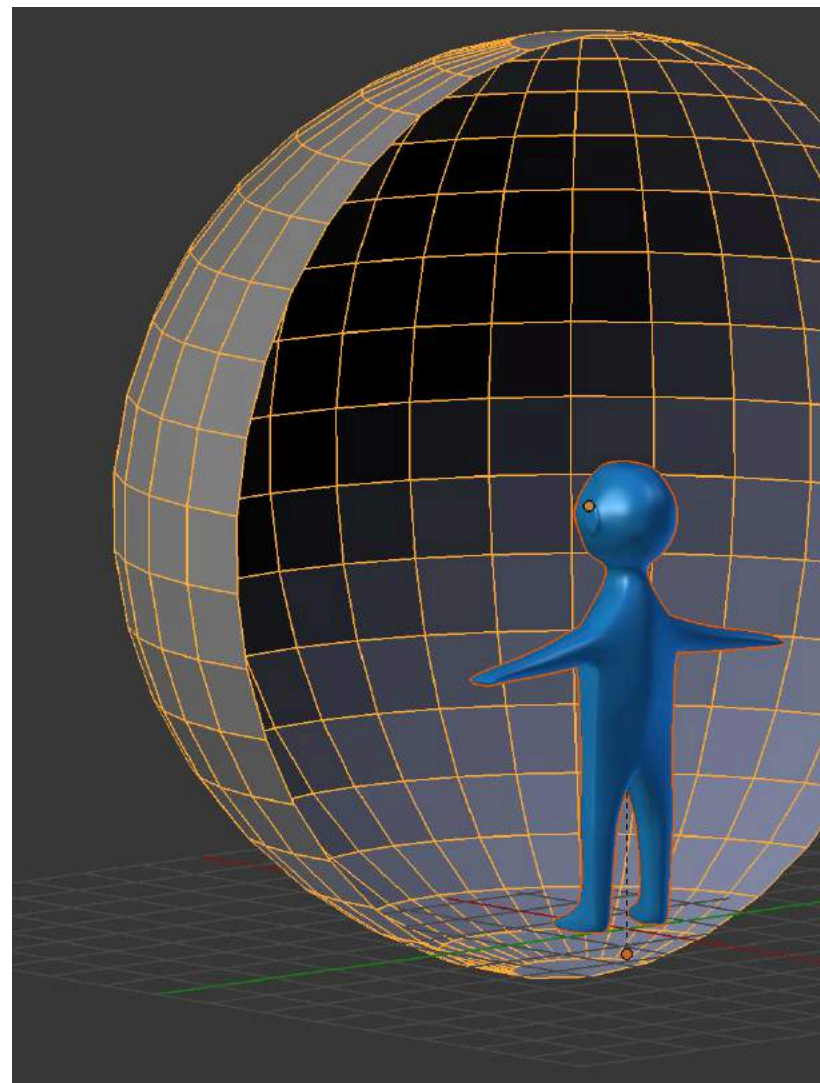
Rapid prototyping for user testing

Marta Brescia-Zapata, Pilar Orero
Universitat Autònoma de Barcelona
[marta.brescia, pilar.orero]@uab.cat

Andrew Duchowski
Clemson University
duchowski@clemson.edu

Chris Hughes
Salford University
c.j.hughes@salford.ac.uk

Krzysztof Krejtz
SWPS University
kkrejtz@swps.edu.pl



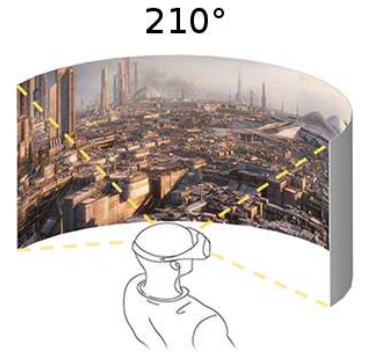
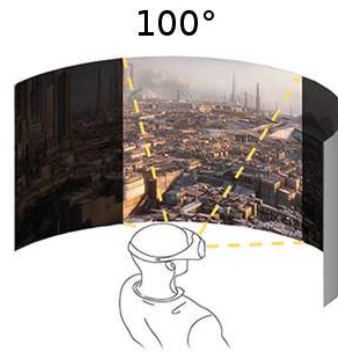
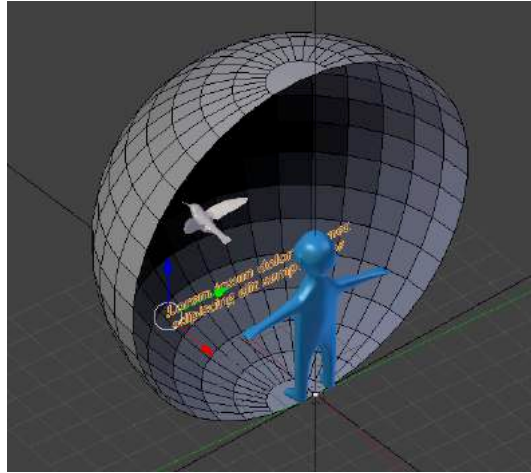
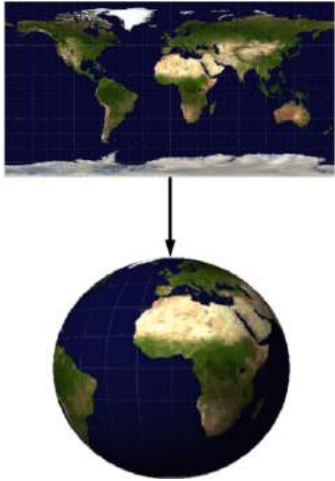
IN THIS PRESENTATION

- Background: 360 video and subtitling
- Framework: prototyping for user testing
- Eye-tracking technology
- Applications
- Conclusions

BACKGROUND

360 video and subtitling

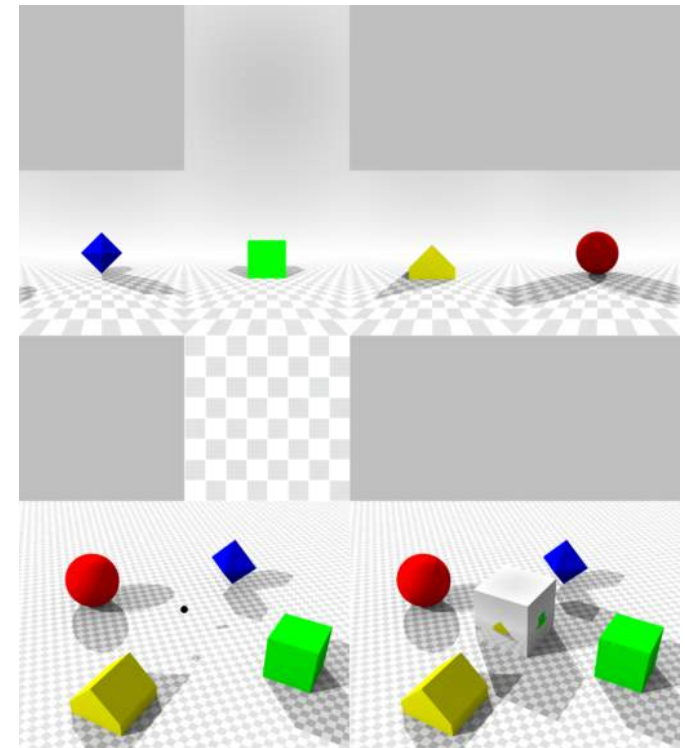
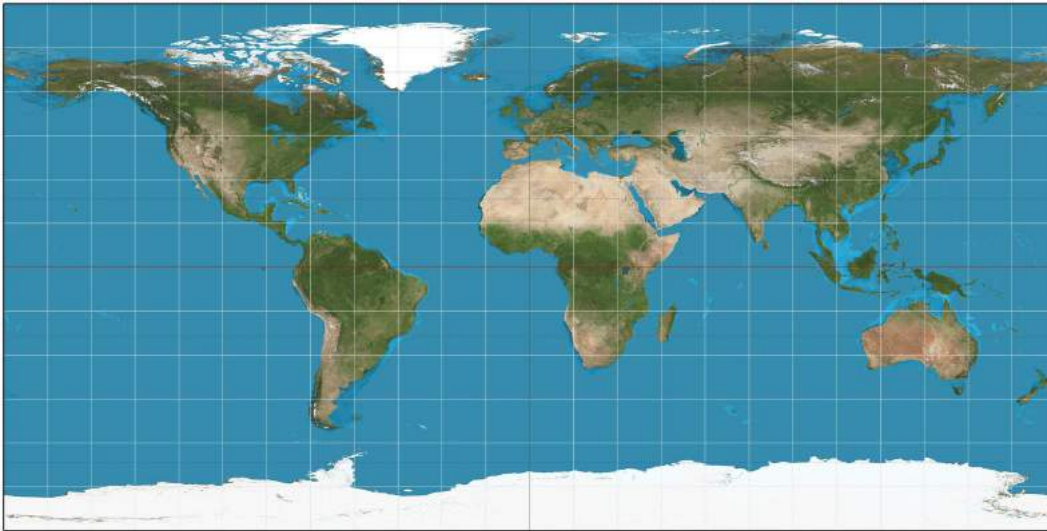
360 VIDEO



360 VIDEO

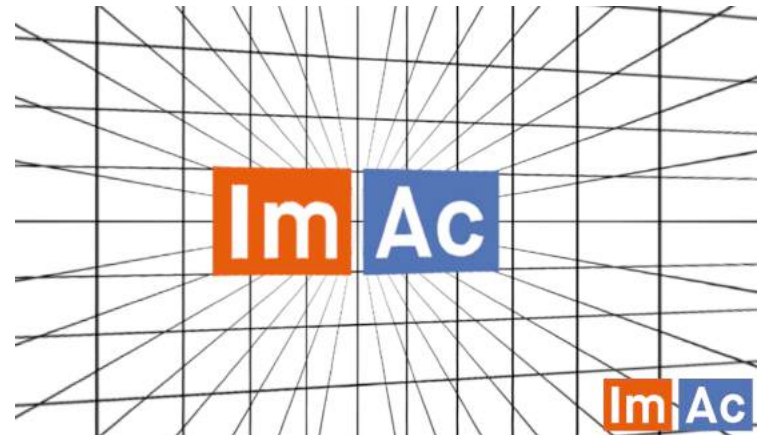
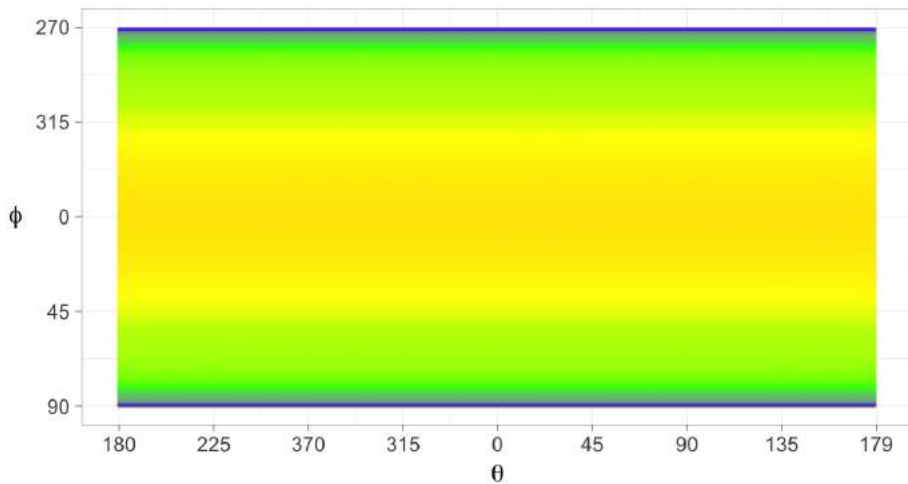
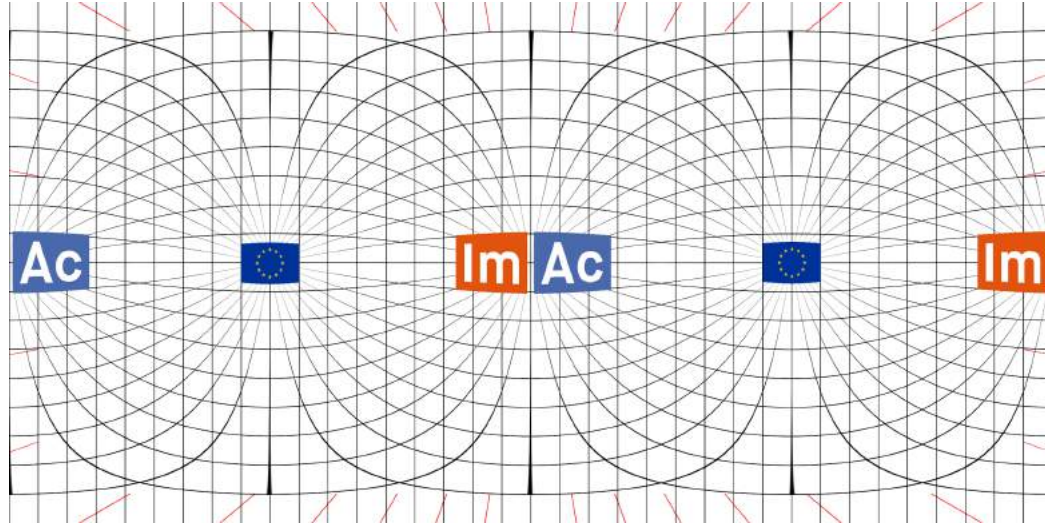
There are two popular formats:

- Equirectangular
- Cubemap



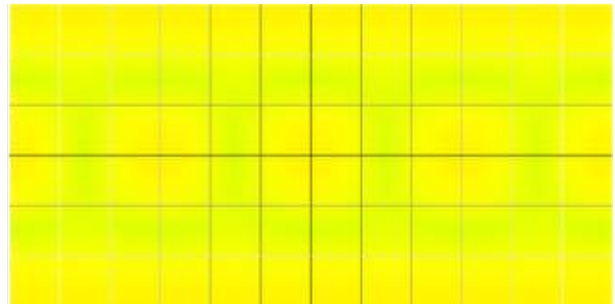
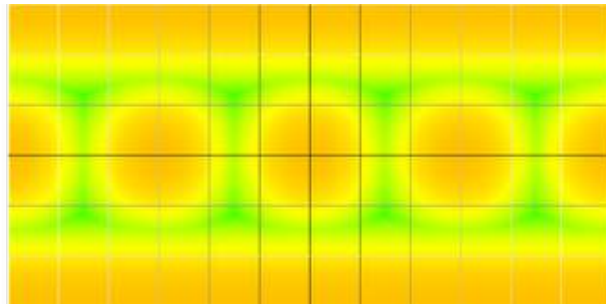
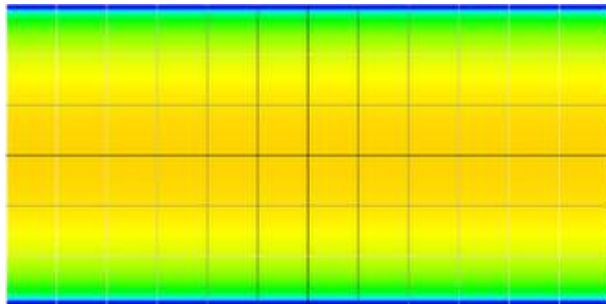
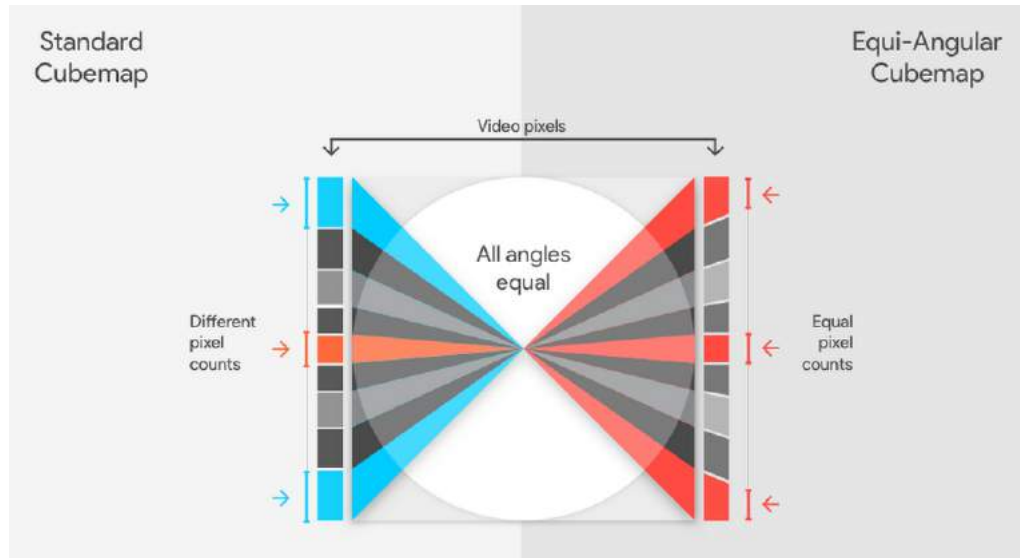
360 VIDEO

- Equirectangular



360 VIDEO

- The Equi-Angular Cubemap (EAC)



SUBTITLING (CAPTIONING)

Comfort & Readability

- Especially for VR glasses
- Where can subtitles be rendered on the screen (safe area)?
- What fonts and text sizes are reasonable?

Speaker identification

- How does the viewer know who is speaking?
- How can the user keep orientation in the scene?

COMFORT & READABILITY



Image quality falls off
← towards the edges

Photo through lens of the Oculus Go

(Only exemplary, does not represent real image quality)

SPEAKER IDENTIFICATION

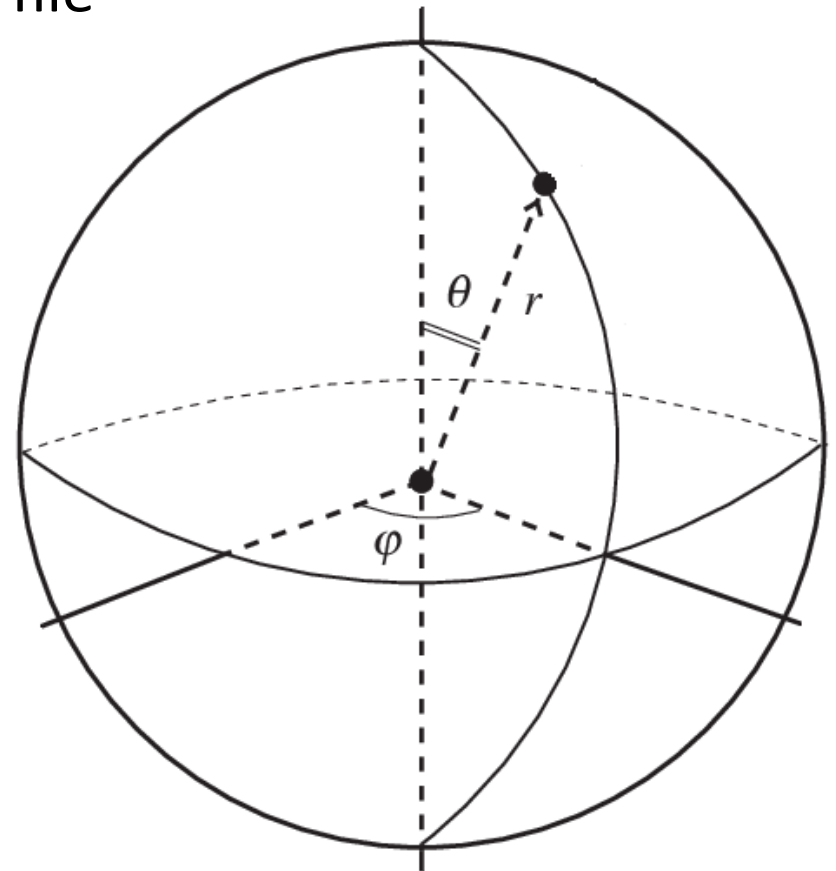


FRAMEWORK

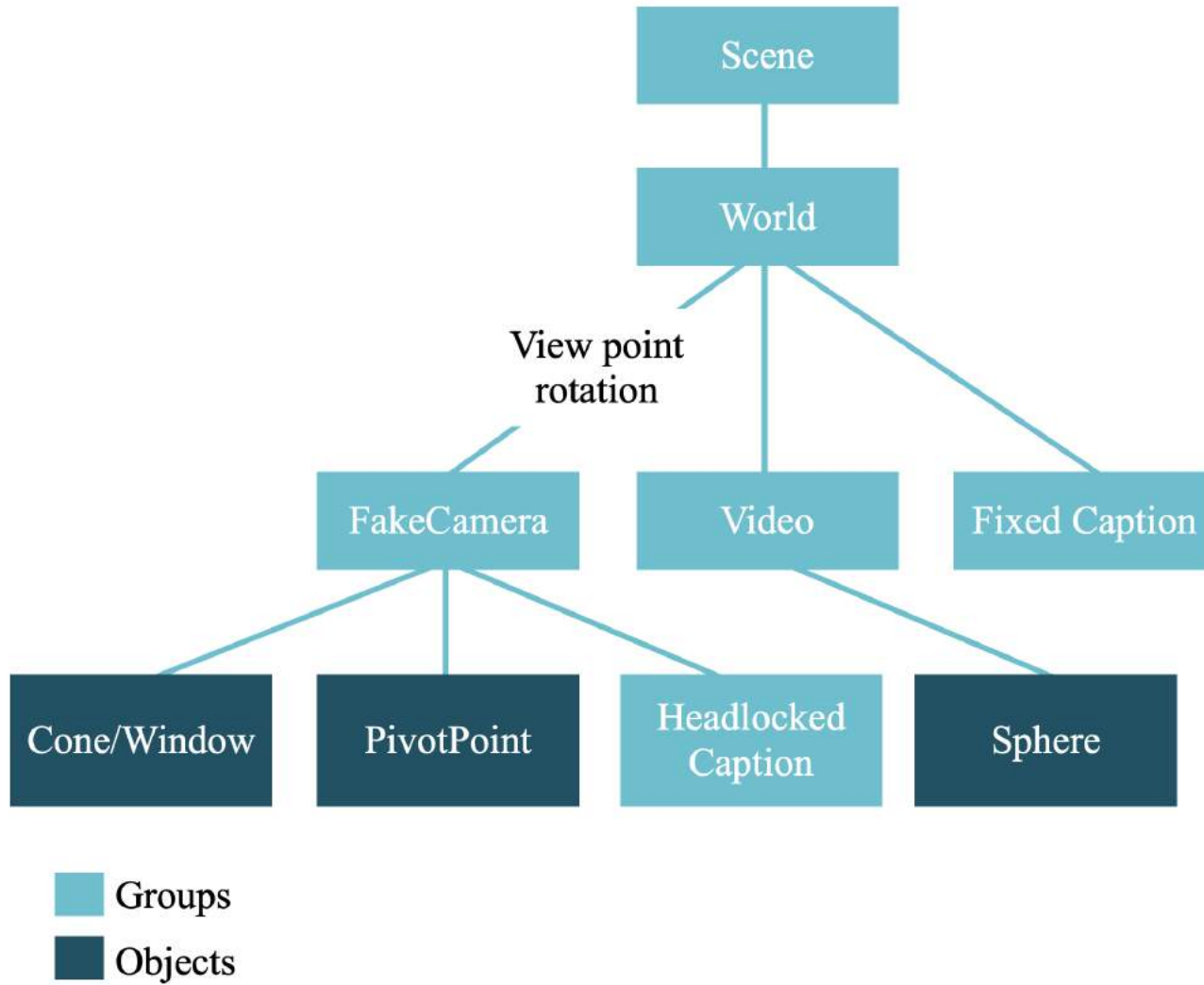
Prototyping for user testing

POLAR COORDINATES

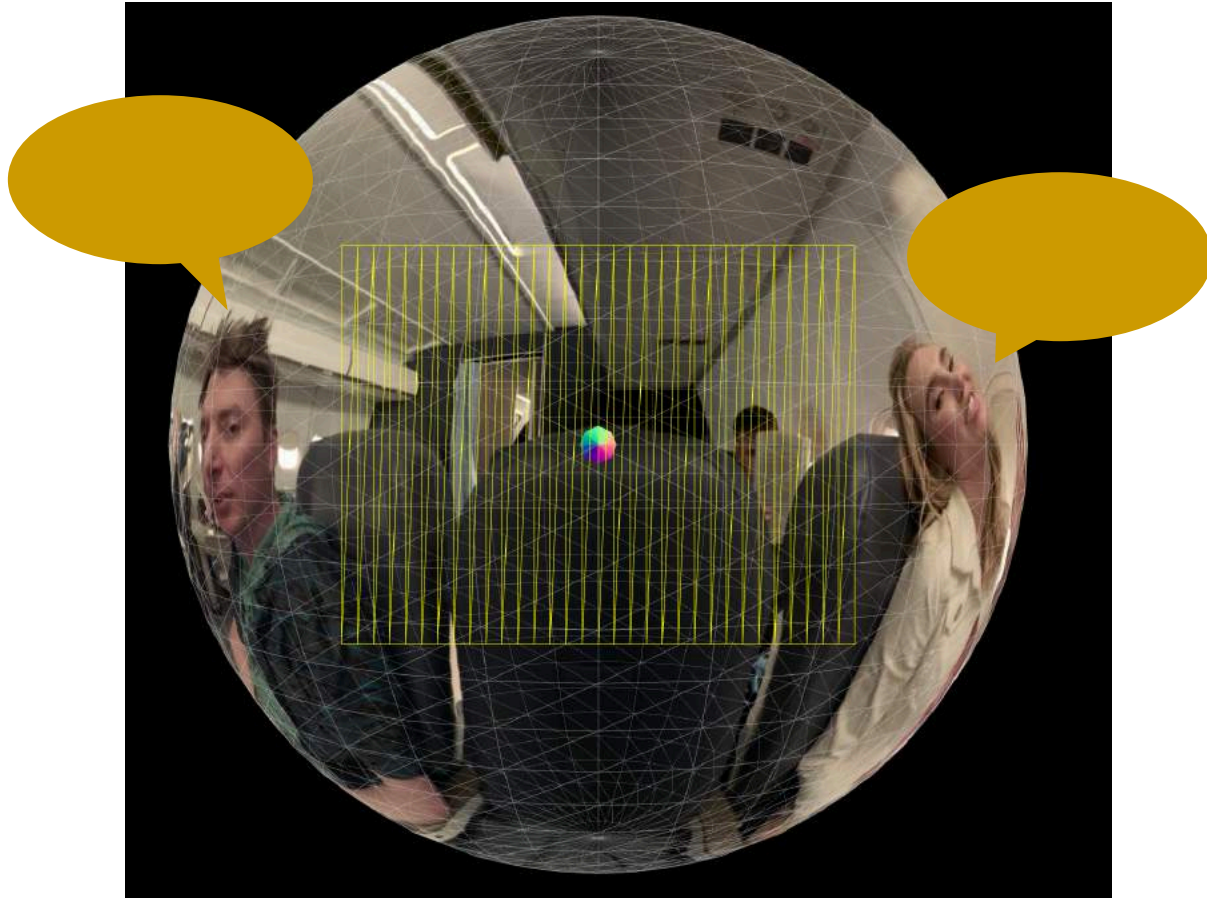
- Position caption anywhere in the scene via polar coordinates
- Radial distance (r)
- Polar angle (θ - theta) and azimuthal angle (ϕ - phi)
- Values are stored in the caption file



SCENE GRAPH



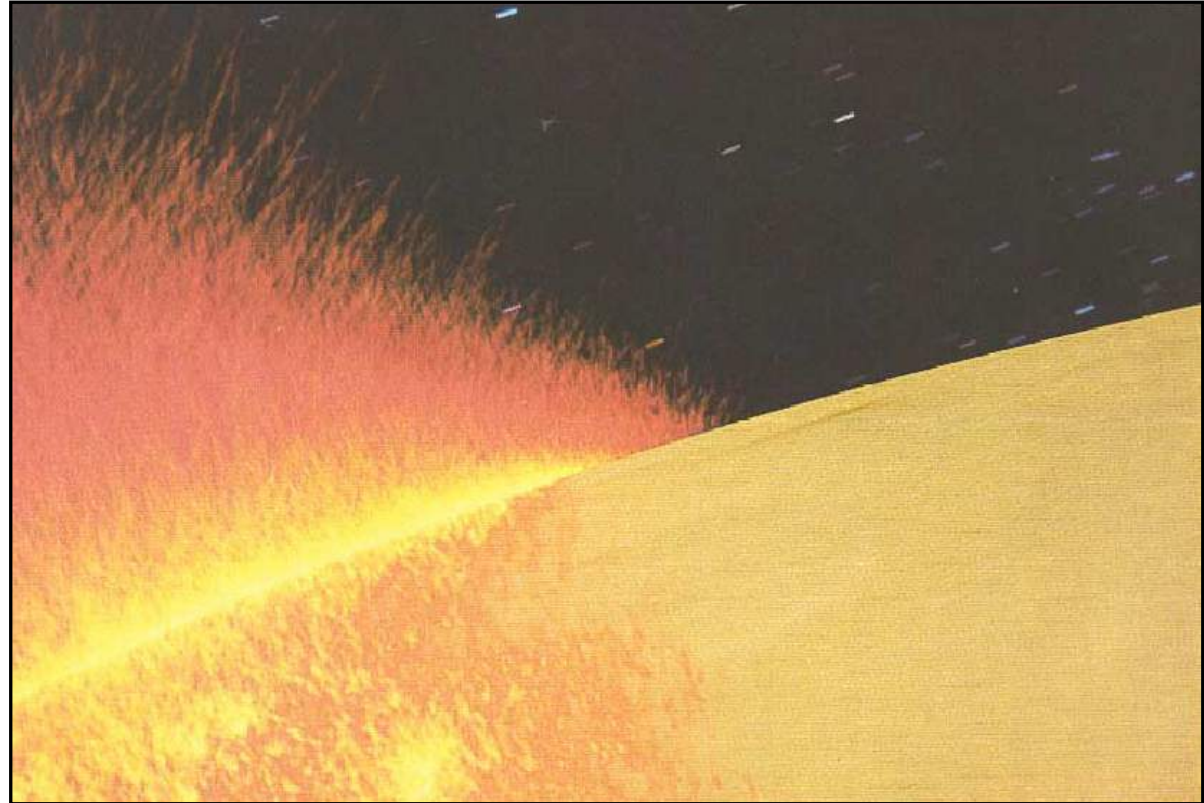
PERFECT SOLUTION?



MULTIPLE CAPTIONS

Particle Systems

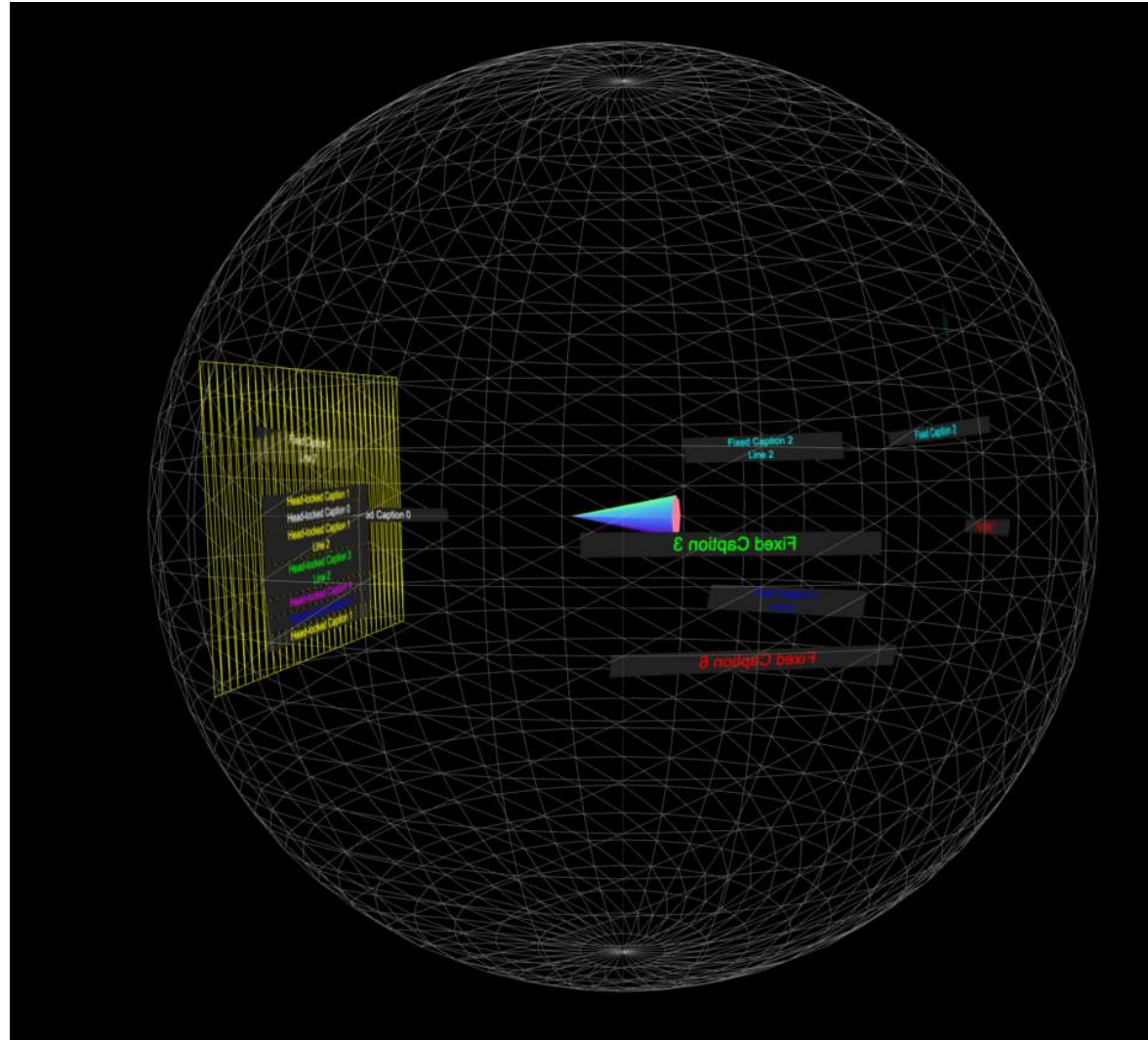
- Position
- Velocity
- Color
- Lifetime
- Age
- Shape
- Size
- Transparency



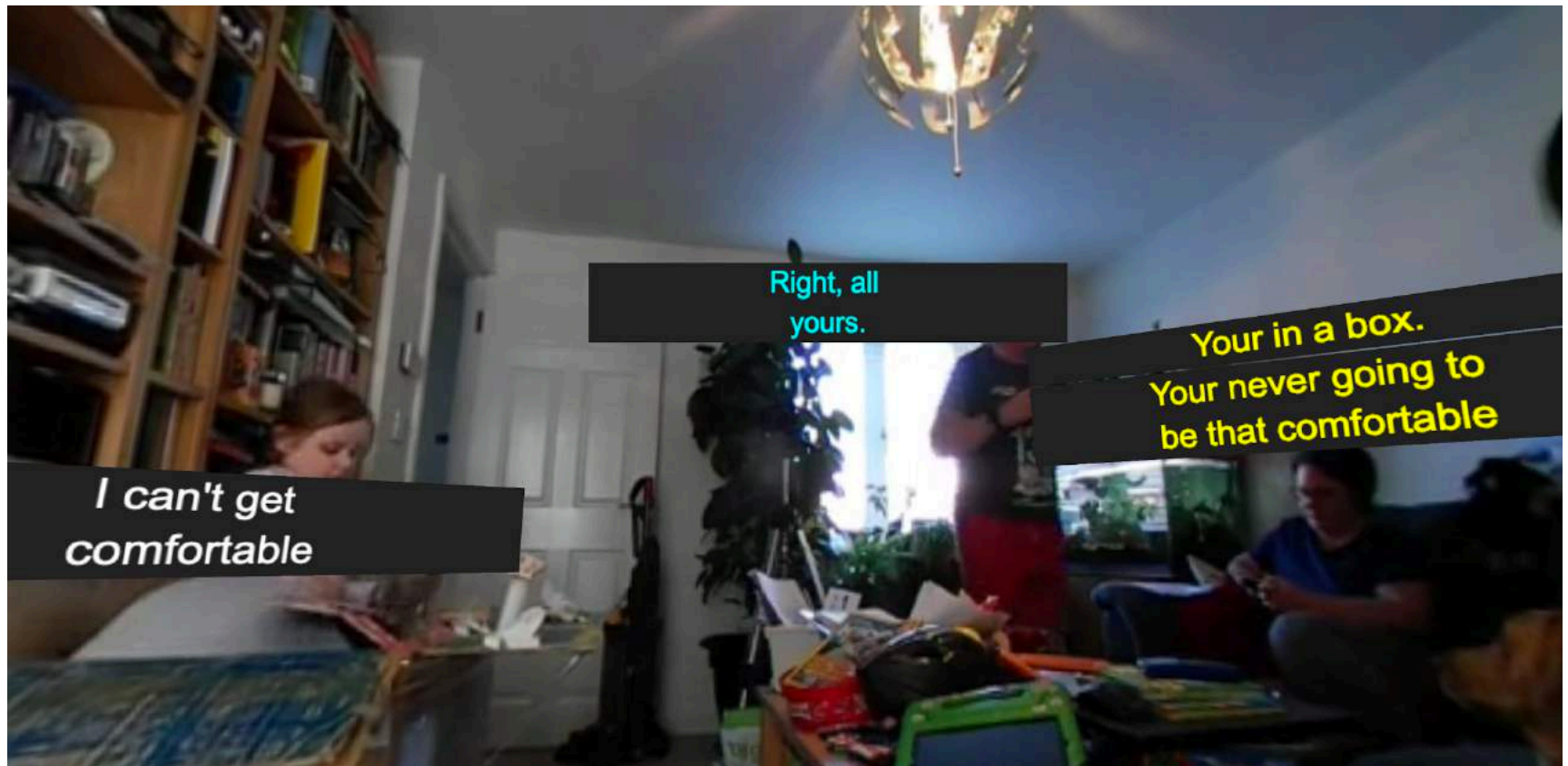
CAPTIONS OBJECT BASED

Particle Objects

- Emitter
- Manager

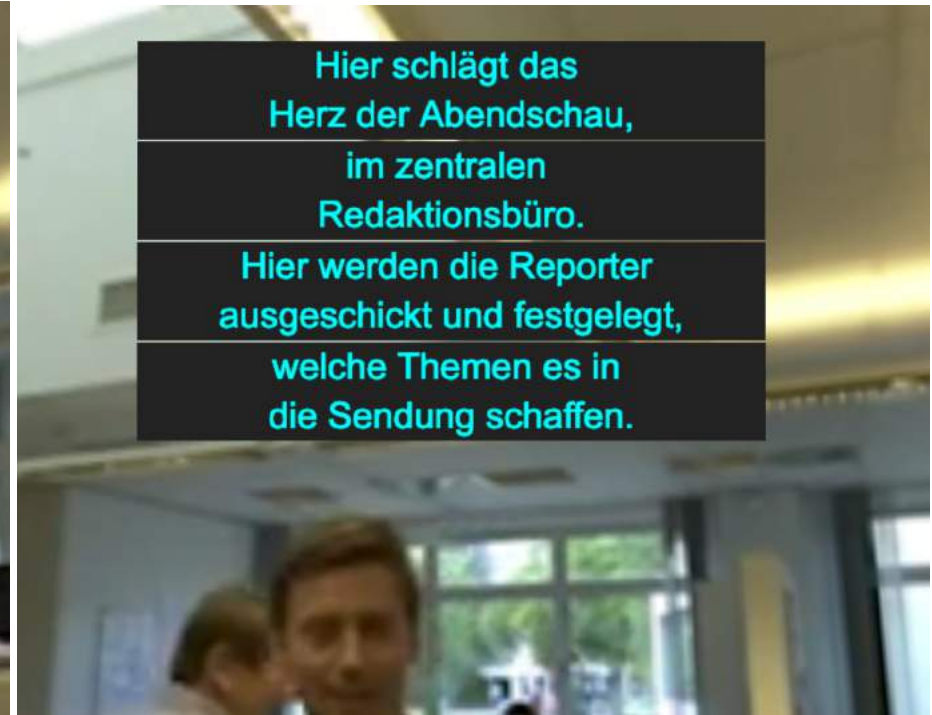
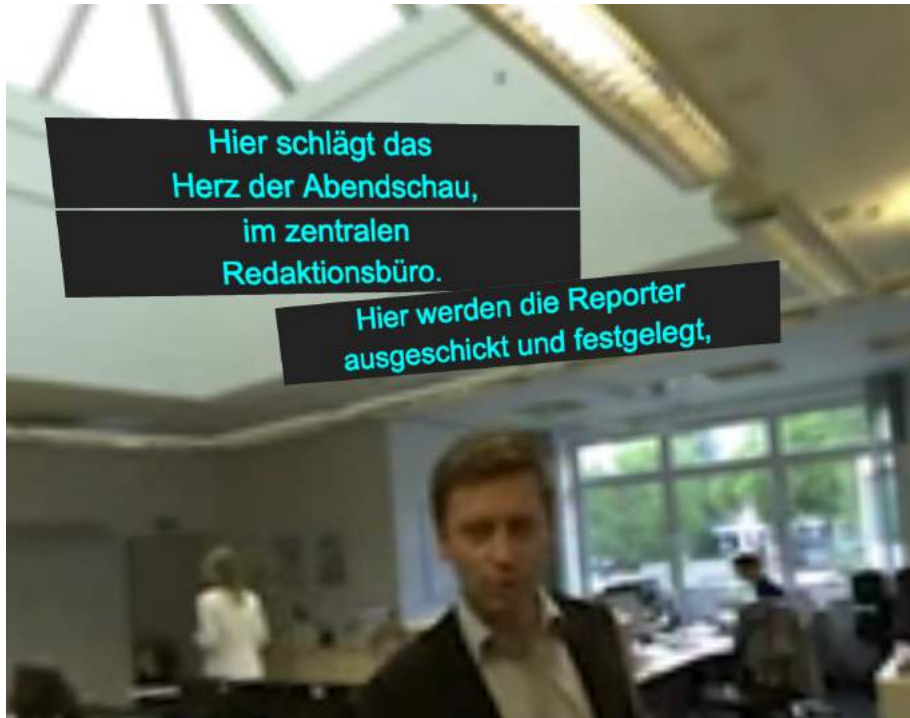


CAPTIONS FIXED TO SPEAKER



<https://www.chxr.org/360/>

STACKING (COLLISIONS)



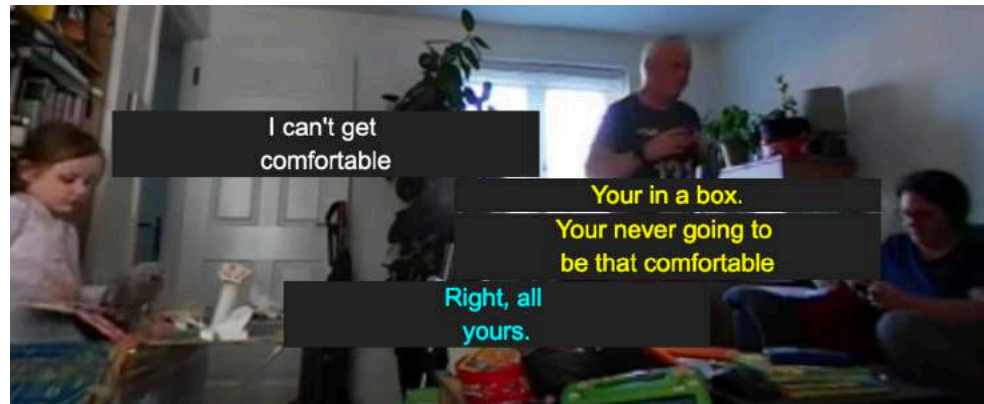
<https://www.chxr.org/360/>

RESPONSIVE (CUSTOMIZATION)

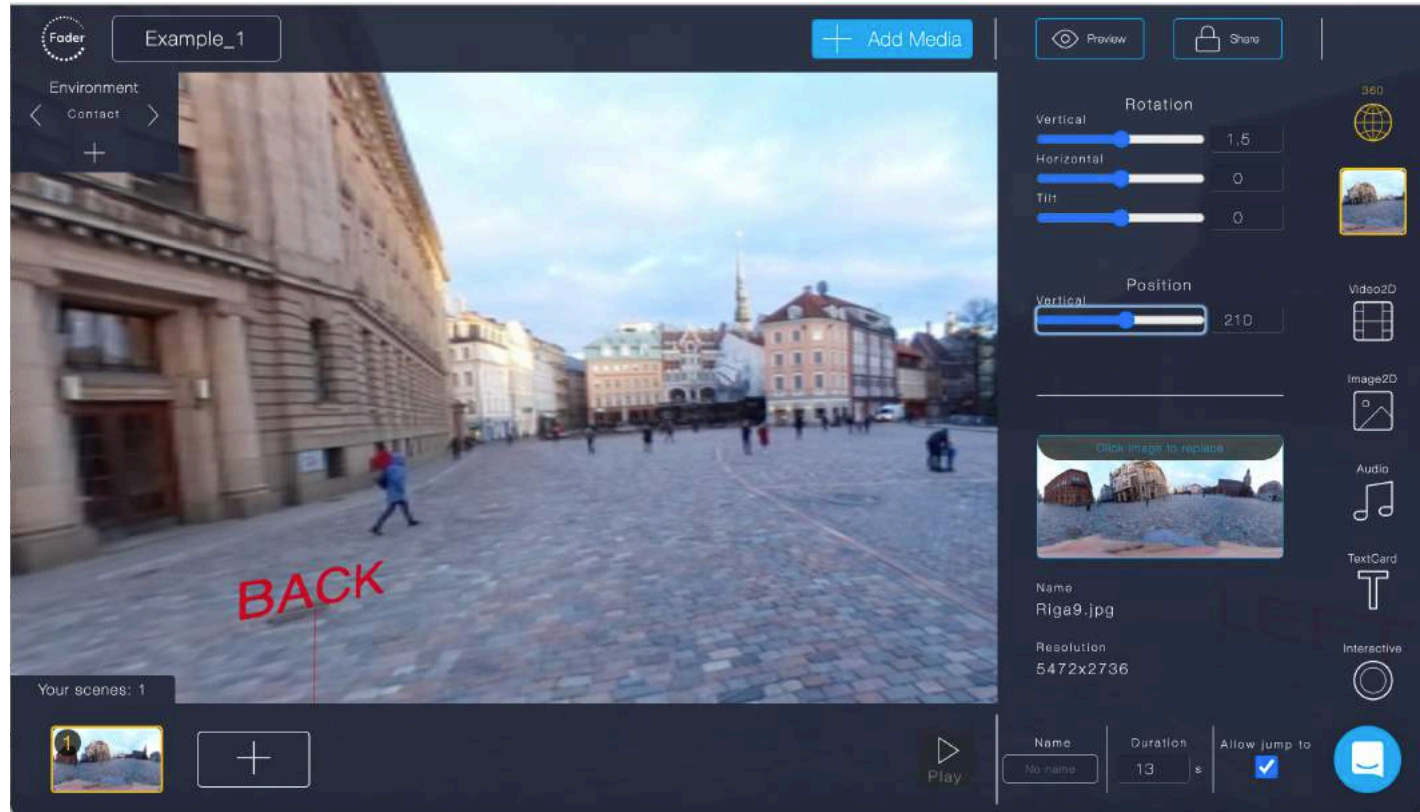


<https://www.chxr.org/360/>

GUIDING MECHANISMS



OTHER EDITORS: FADER



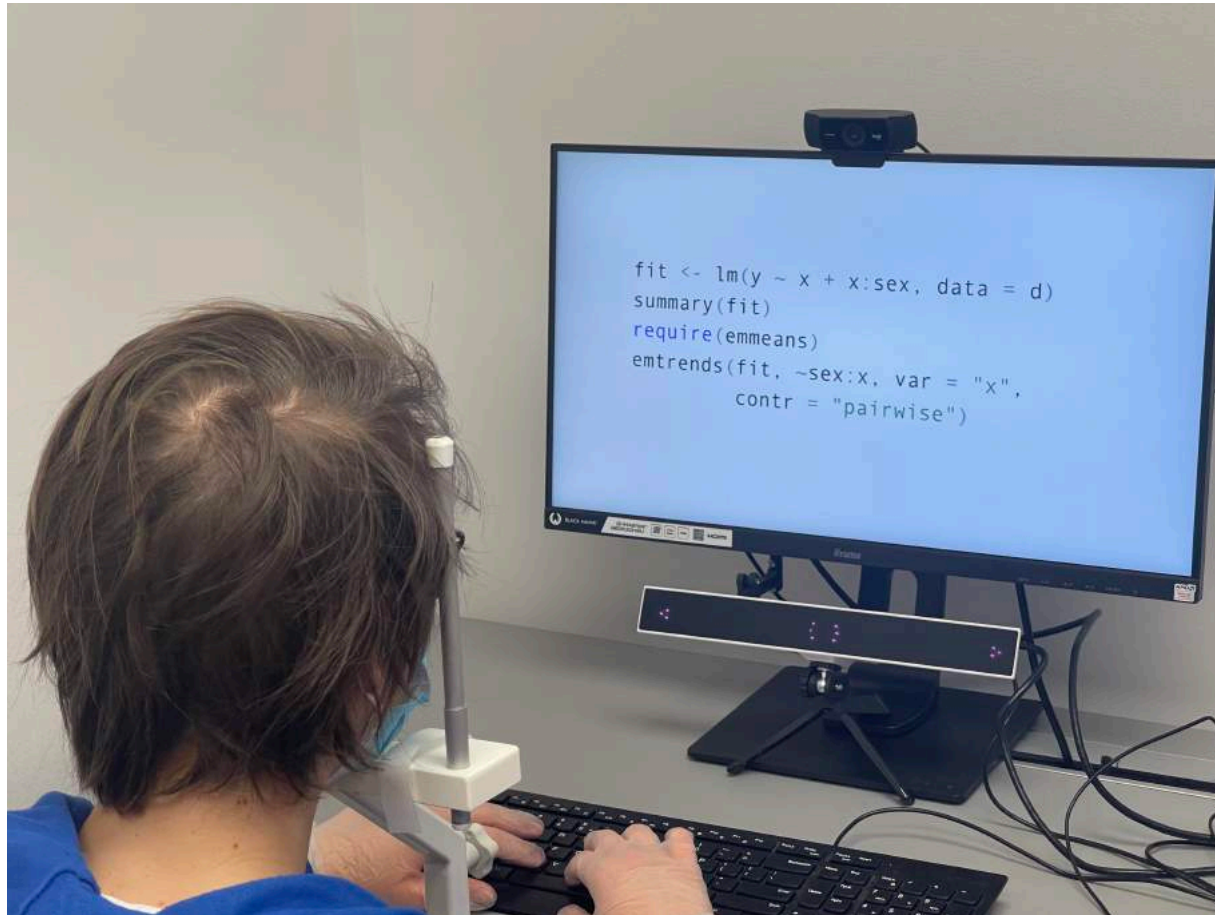
<https://app.getfader.com/>

EYE-TRACKING TECHNOLOGY

Built-in and now available to the consumer

HOW TO DO TRACKING IN 360

- Conventional eye trackers need a clear view of the users eyes
- Traditional eye tracking is done in a fixed position in 2D
- After market eye tracking for HMDs is clunky!



HTC VIVE PRO EYE



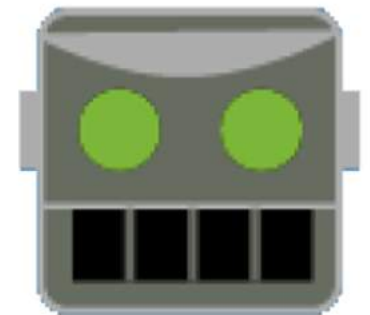
HTC VIVE PRO EYE

120Hz Eye tracking



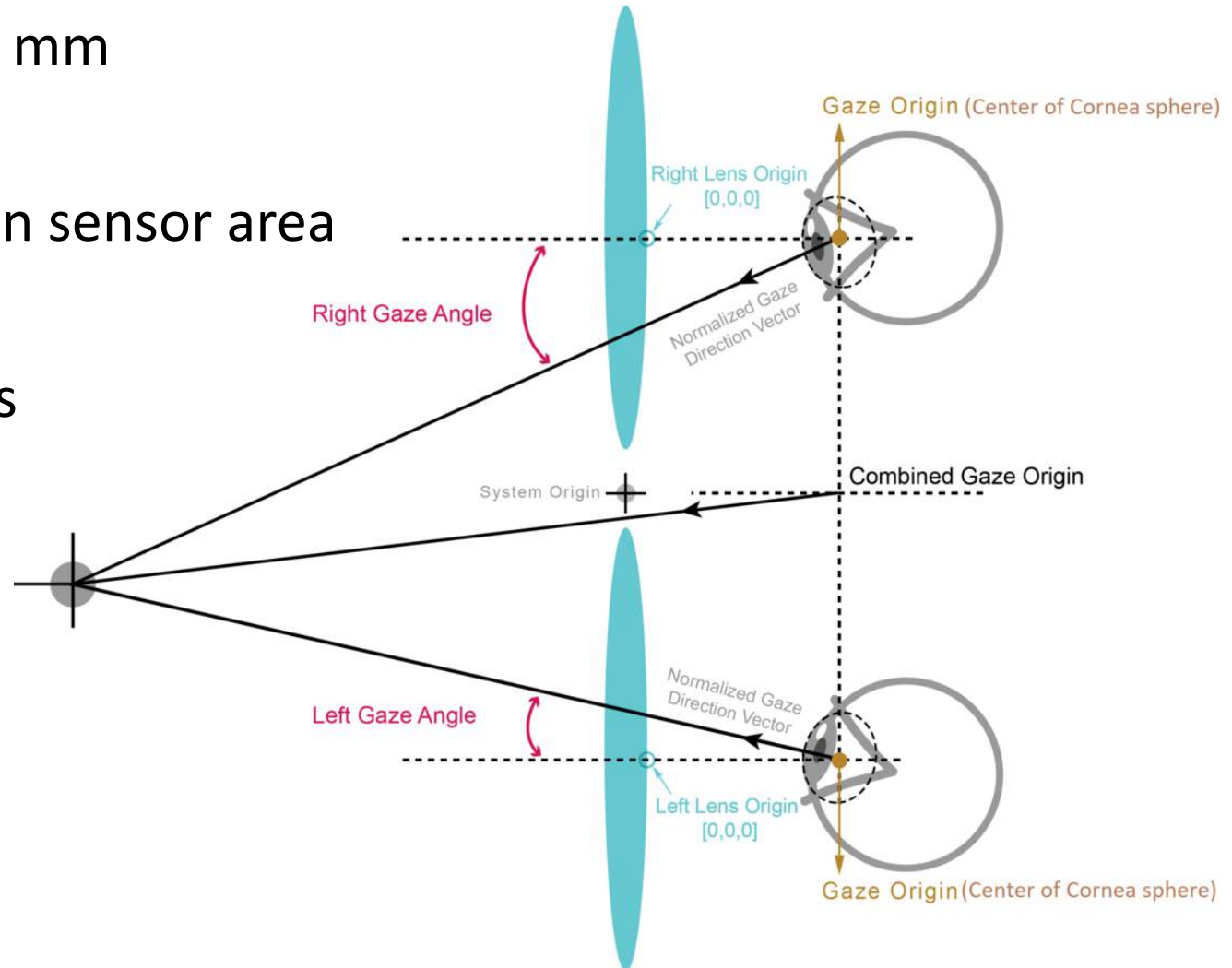
SRANIPAL SDK & API

- SDK from Vive
- Track eye and lip movements, mainly for avatar animation
- SDK icon in notification shows current eye tracking status
- Plugins for Unreal and Unity are included



KEY DATA

- Gaze origin mm
- Gaze direction normalized
- Pupil diameter mm
- Eye openness
- Pupil position in sensor area
- ... for both eyes





The screenshot displays the Unity development environment. On the left, the Hierarchy panel shows a scene with a yellow rectangular base and a white rectangular top with several red spheres. The Project panel shows a folder structure with assets like 'New Game Object'. The Console panel at the bottom shows a message: "Stanking by 1.4 on your desktop to make it".

The main window shows a C# script named 'Stanking' with the following code:

```
using System.Collections;
using System.Collections.Generic;
using UnityEngine;

public class Stanking : MonoBehaviour
{
    //Define an object to keep each of the data elements
    //-----
    //-----
    public class headInfo
    {
        public long rx;
        public Vector3 angle;
        public Vector3 position;

        public headInfo(Vector3 angle, Vector3 position)
        {
            this.angle = angle;
            this.position = position;
        }
    }

    public headInfo[] head;

    var e = GetComponent<T>();
    while (e != null)
    {
        this.head = new Vector3[] { e.transform.position };
        this.position = new Vector3(e.transform.position.x, e.transform.position.y, e.transform.position.z);
    }

    public string GetInfo()
    {
        return "
        head: " +
        "angle: " +
        "position: " +
        "-----";
    }

    public void AddInfo(Vector3 angle, Vector3 position)
    {
        head.Add(new headInfo(angle, position));
    }

    public void GetInfo()
    {
        foreach (var h in head)
        {
            Debug.Log(h);
        }
    }

    public void GetInfo()
    {
        Debug.Log(head);
    }

    //-----
    //-----
    public class infoInfo
    {
        public long rx;
        public double time;
        public long frame;

        public infoInfo(double time, long frame)
        {
            this.time = time;
            this.frame = frame;
        }
    }

    public infoInfo[] info;

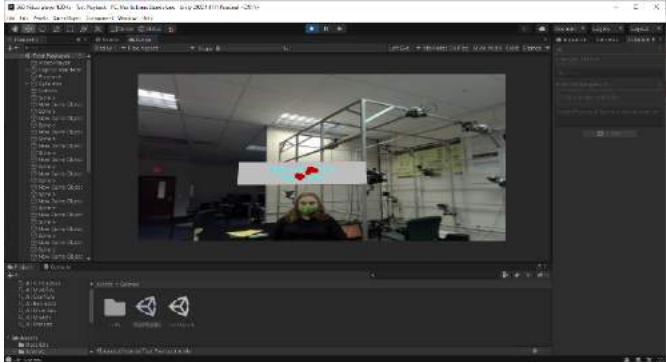
    var e = GetComponent<T>();
    while (e != null)
    {
        this.info = new double[] { e.transform.position };
        this.frame = new long[] { e.transform.position };
    }

    public string GetInfo()
    {
        return "
        info: " +
        "frame: " +
        "-----";
    }
}
```

ARCHITECTURE



120Hz Thread
Capture Eye
Tracking Data



90Hz Unity Test
Recorder

Head Orientation
Video Position
Caption State

Data Manager

Head Orientation
Video Position
Eye Data
Caption State

Unity Playback /
Analyser



Caption
Manager

File
Management

Eye Data

APPLICATIONS

Testing with users

EXPERIMENTAL DESIGN

- Define and obtain the material (360° videos)
- Design the questionnaires
- Elaborate subtitles (2 different visualisation modes)
- Chris Hughes web-based editor and player
- Test the experiment set-up in Unity

Back to Project Area History holidays_2 - 360 Caption Editor v0.11 (martabrescia)

Start Time	End Time	Text	WPM	CPS	Progress
00:00:01,160	00:00:03,020	Two rainy days...	1.61wps	8.60cps	<input type="checkbox"/>
100:00:03,180	00:00:05,400	It's so sad that the Summer is over...	3.60wps	16.22cps	<input type="checkbox"/>
200:00:05,560	00:00:08,970	Yes, although this year we haven't done much.	2.35wps	12.32cps	<input checked="" type="checkbox"/>
300:00:09,470	00:00:10,630	Well, you're right.	2.59wps	15.52cps	<input type="checkbox"/>
400:00:10,790	00:00:13,220	I wanted to go to Paris to visit a friend.	4.12wps	16.46cps	<input type="checkbox"/>
500:00:13,380	00:00:15,380	I've never been to Paris, you know?	3.52wps	16.50cps	<input type="checkbox"/>

360 Equi Help Console

Select: All, None, First, Last, Next, Previous, Next, Previous, 1-8

Set Character: 1-8

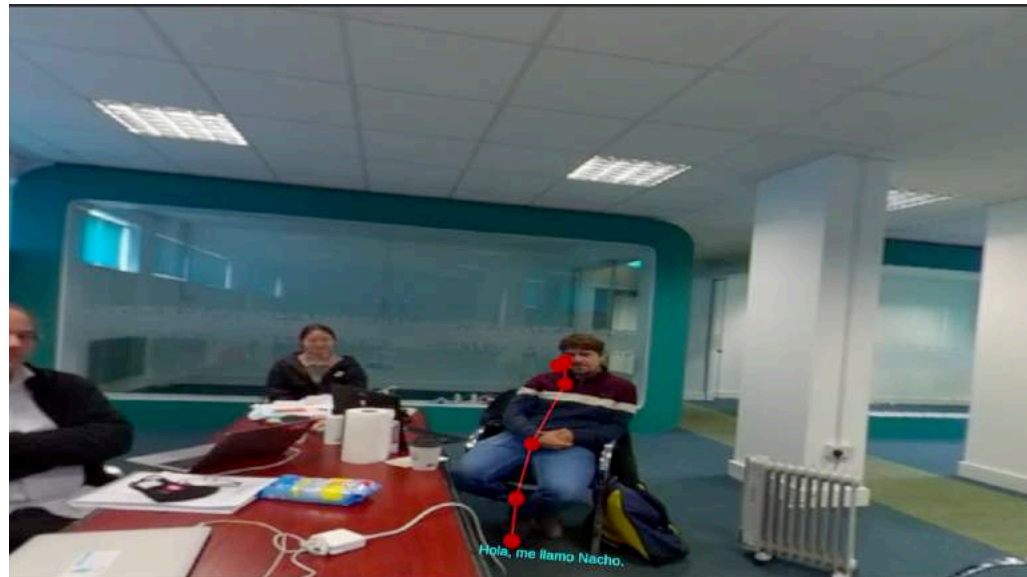
Captions: Clear All, Delete, Trim, Divide, Combine, Split

Play Scene Pause << < > >>

Two rainy days... It's so sad that the Summer is over... Yes, although this year we haven't done much.

2 X 2 X 2 DESIGN

- 2 different videos
 - Arab video
 - Multi-language video
- 2 subtitle position
 - Head-locked
 - Fixed
- 2 colour
 - B&W for all speakers
 - Unique color for each speaker



DATA COLLECTION METHOD

1. Inform user and sign consent form
2. Answer demographic questionnaire
3. Watch the first video
4. Answer brief after-movie questionnaire
(task-load and comprehension)
5. Watch the second video
6. Answer brief after-movie questionnaire
(task-load and comprehension)
7. Answer brief preference questionnaire
comparing both videos



THE MATERIAL

Arab video - head-locked - Colour



THE MATERIAL

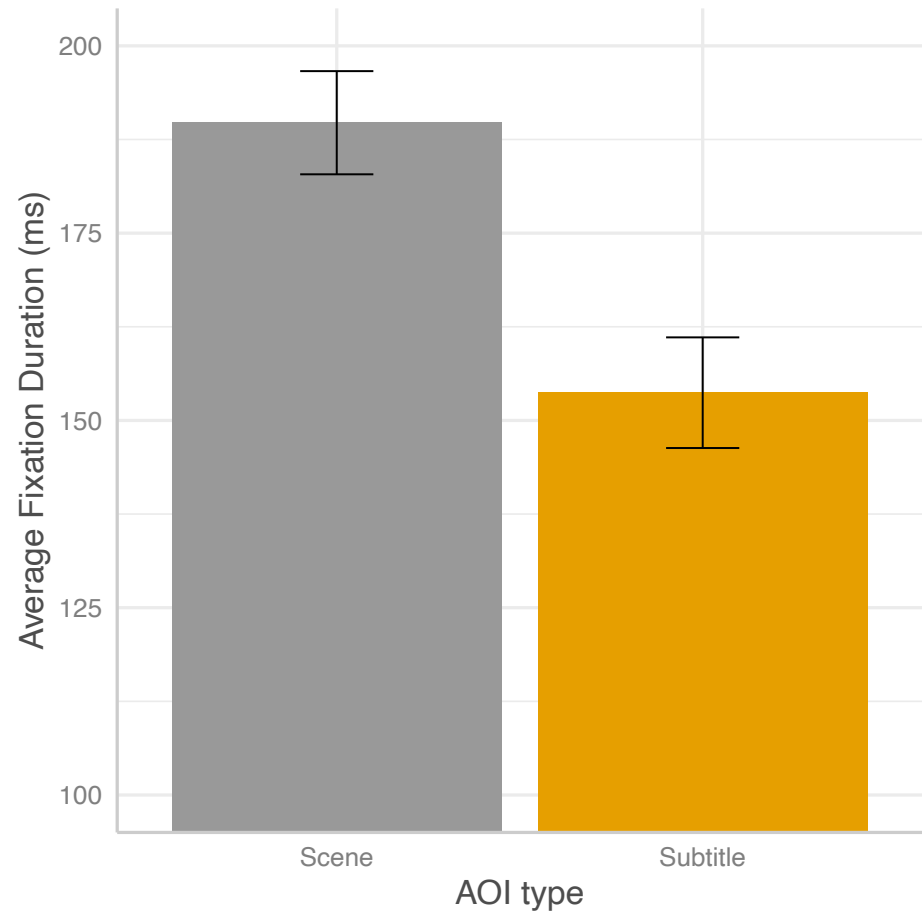
Arab video - fixed - b&w



EXAMPLE (PRELIMINARY) DATA ANALYSIS

Scene vs. Subtitles

- Participants spent more time exploring the scene than reading the subtitles

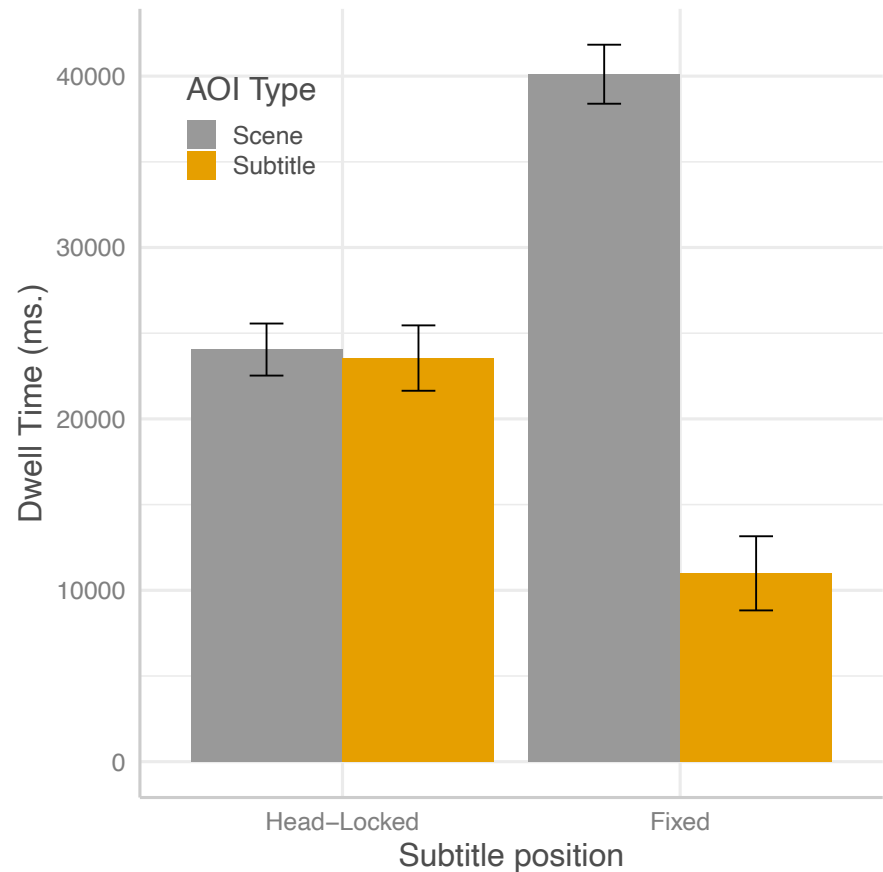


$F(1,14) = 13.24, p < 0.01$

EXAMPLE (PRELIMINARY) DATA ANALYSIS

Head-Locked vs. Fixed

- Head-Locked subtitles draw about as much visual attention as scene
- Fixed subtitles afford greater scene exploration



$F(1,14) = 35.54, p < 0.01$

CONCLUSIONS

What's next?

WHAT HAVE WE LEARNED?

User tests yield **limited results**

Unless you can put a working product in front of them

- Paper prototypes may cause confusion
- Often lead to users saying they prefer what they already have

In this area many technologies pose a **learning challenge**

- Difficult to ask a user to evaluate prototype when learning

Need to combine questionnaires and/or focus groups with new techniques like **eye tracking** to measure user behavior

This work was partially funded by the European Union projects

TRACTION (<https://www.traction-project.eu>) and

MEDIAVERSE (<https://mediaverse-project.eu>)



and partially funded by the US National Science Foundation
(grant IIS-1748380)



**QUESTIONS?
COMMENTS?
THOUGHTS?**

