Audio Description in Immersive Environments

Presented by

Pilar Orero – UAB

John Paton - RNIB

Introductions

Pilar Orero

UAB

John Paton

RNIB

ImAc Project

To cover today

1, Sight conditions

2, Simulation of sight loss

3, Immersive environments

4, Immersive Audio

5, Next Generation Audio

6, AD with NGA

7, The practical

1, Sight conditions

Sight loss is a broad spectrum

- Refractive error

5 main types of obscuring sight loss

- Cataracts

- Glaucoma

- AMD

- Retinitis Pigmentosa
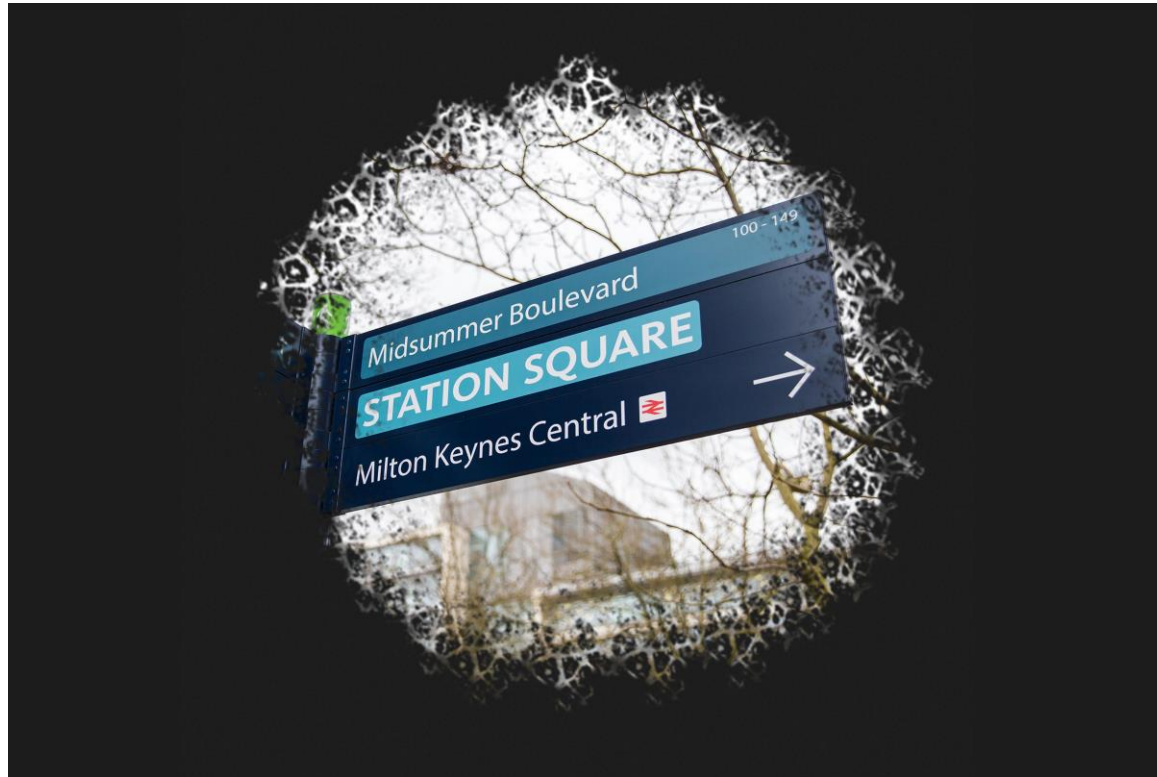
- Diabetic retinopathy

Refractive error

- Image is focused in front of or behind retina

- Often corrected by glasses or contact lenses

- Astigmatism – image is warped often "barrel shaped"

Cataracts

- Lens at the front of the eye becomes slightly opaque and 'milky' due to age.
- Can usually be corrected with surgery – the most common operation in the UK

Glaucoma

- Hereditary condition
- Pressure build-up in the eye leads to distortion and later tunnel vision
- If caught early sight loss can be avoided

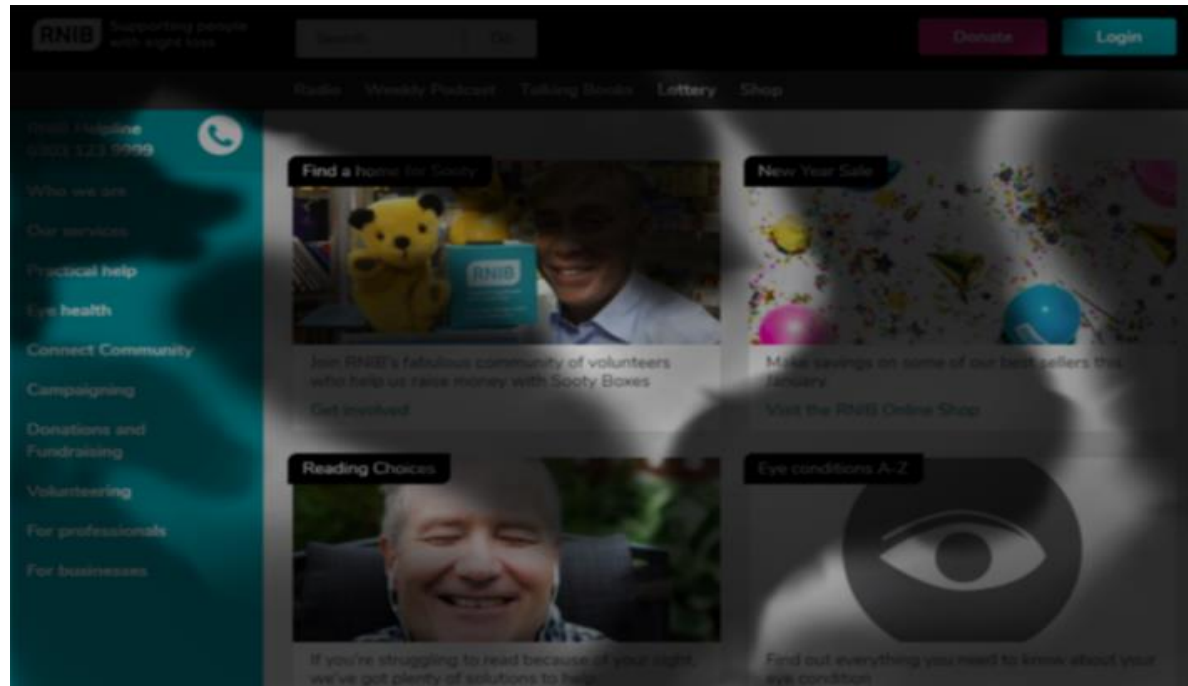# AMD - Age related macular degeneration

Macula

- Central 18% of vision
- Higher concentration of cone cells than rod cells
- Used for 'precision' sight

'Wet' and 'dry' AMD

- Affects the macular
    - Distortion
    - Blurriness
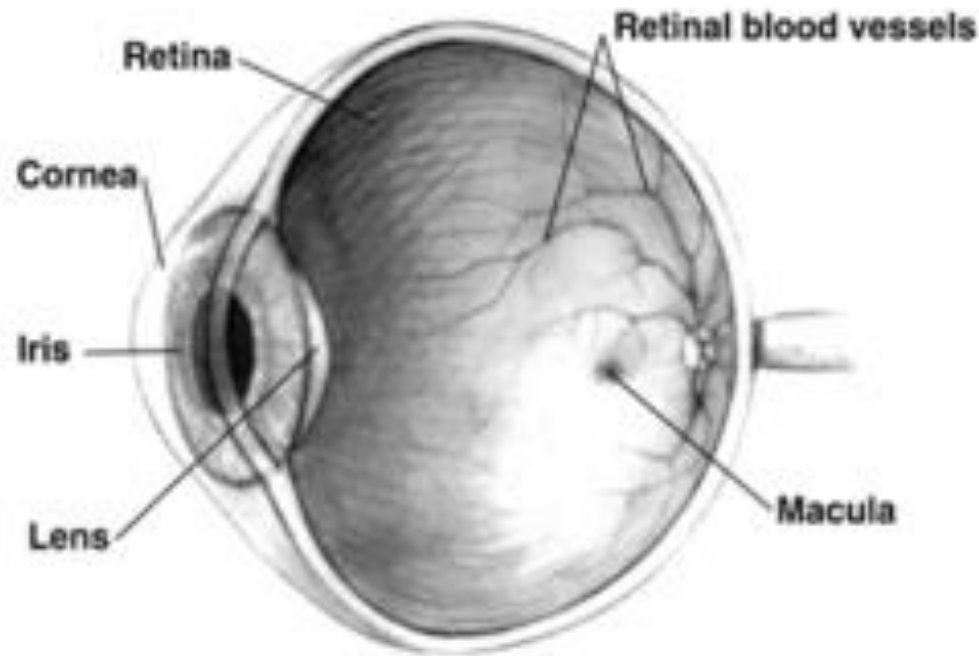   - Loss of central vision

Diabetic retinopathy

- Death of small blood vessels due to untreated (or badly managed) diabetes
- Same process also affects sense of touch leading to reduced sensitivity in fingertips

Retinitis pigmentosa

- Rod cells in the retina gradually stop working
- Night blindness
- Dark spots

Many more

- RNIB website includes a list of 56 eye conditions
- 20 are considered too rare to give detailed information on

Very broad spectrum

- Albinism
- Nystagmus
- Charles Bonnet syndrome

Audio Description

"Audio description (AD) is additional commentary that explains what's happening on screen. AD describes body language, expressions and movements, making the programme clear through sound." - RNIB

- Prioritising important information
- Fit in gaps in dialogue
- Fit in to gaps in the content

2, Simulation of sight loss

SimSpecs

- Simple, effective

EyeWare app

- Produced alongside Transport Systems Catapult

- 5 conditions
  - Cataracts, Glaucoma, Diabetic Retinopathy, Retinitis Pigmentosa, Wet and Dry AMD

- On iOS and Android

Sight loss simulation

3, Immersive environments

Video
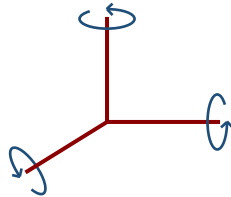
- 360 Video
- Virtual Reality
- Augmented Reality

Audio

- Virtual barbershop
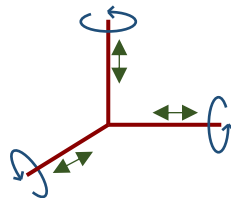- Gaming
- Audio beacons

3 DoF vs 6 DoF

DoF = Degrees of Freedom
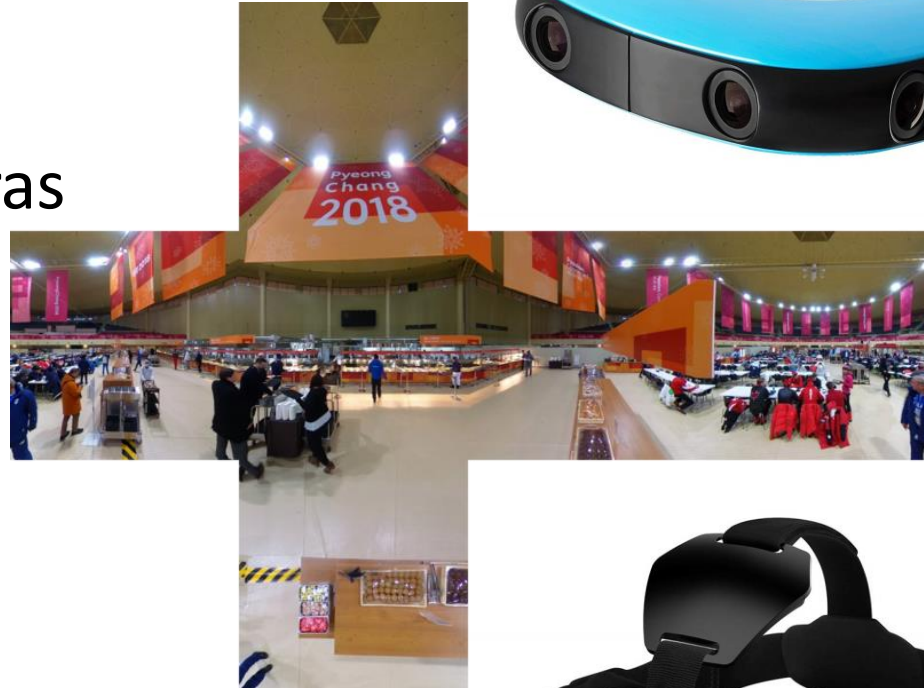
3 DoF = 3 Rotational Axes

6 DoF = 3 Rotational Axes + 3 Positional Axes

360 Video

- Panoramic video
- Uses special cameras
- Requires HMD

Virtual Reality

Computer generated world

- Gaming
- CAD – Architects, Engineers
- Education - Simulation
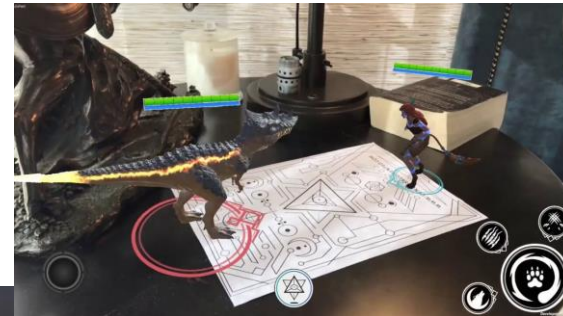- Virtual Galleries

Augmented Reality

Virtual objects placed in the real world

Mixed Reality ?

- Definition seems interchangeable with AR
- Using real-world objects as 'anchors' for virtual objects?

4, Immersive audio

Virtual Barbershop

- Created by Qsound Labs to demonstrate their Cetera algorithm

- A 'Dummy head' recording mimics human ears

Gaming

- Hellblade:Senua's Sacrifice uses Binaural sound to simulate the main characters psychosis

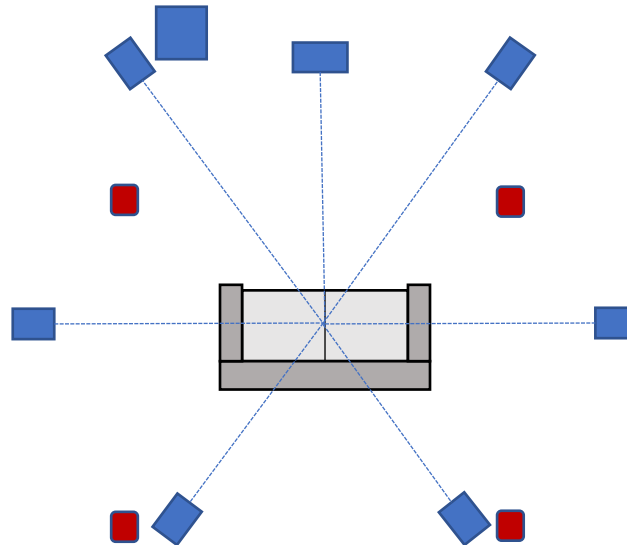# Audio Beacons

Microsoft Soundscape

Next Generation Audio

Places sounds around the user

- Channel Based Audio

- Object Based Audio

- Ambisonics

- Binaural

Channel Based Audio

- Mono, Stereo, 2.1, 5.1, 7.1, 7.1.2, 7.1.4, 9.1 etc
- 1 track per speaker
  - Audio channels need to match speaker setup
  - Speakers placed around user

Object Based Audio

Sound 'objects' contain metadata including

- Position, Label, Category

Personalised audio

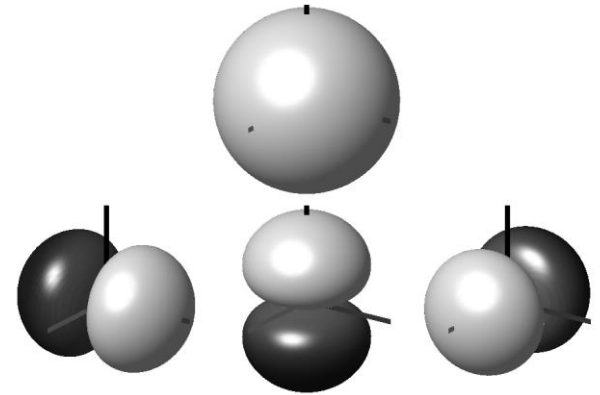- AD/no AD, alternate sports commentary, clear audio, Karaoke track?

Position your own AD?
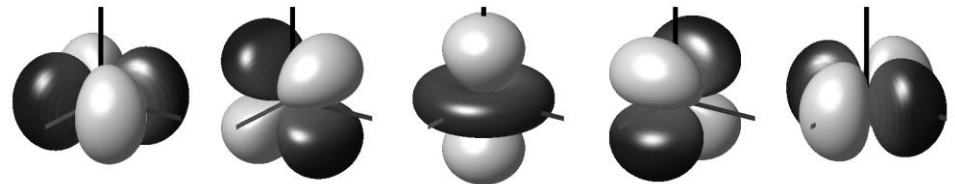
- Fraunhofer MPEG-H demo - pan and fade AD
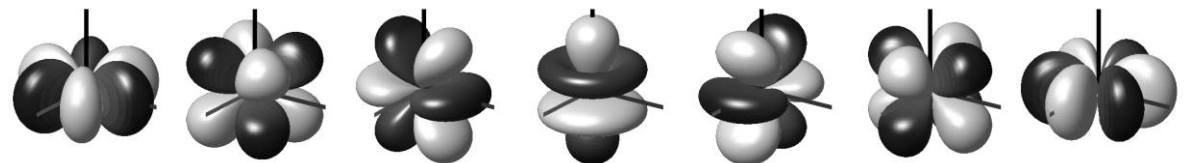
Ambisonics

1st Order (4 channels)

- Mono channel + Left/Right + Front/Back + Top/Bottom

2nd Order (9 channels)

3rd Order (16 Channels)

Binaural Recording

- Using 'dummy head'
- Mimics the way our ears work

Binaural Audio

- Means 'two ears'
- Attenuations and echos 'encode' position of sounds
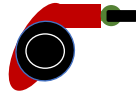- Sounds 3D through headphones
- Can be simulated with algorithm

5, AD with NGA

Different placement of audio description

- Classic (VOG)

- Static (Friend on Sofa)

- Dynamic (AD on Action)

Classic (Voice of God)
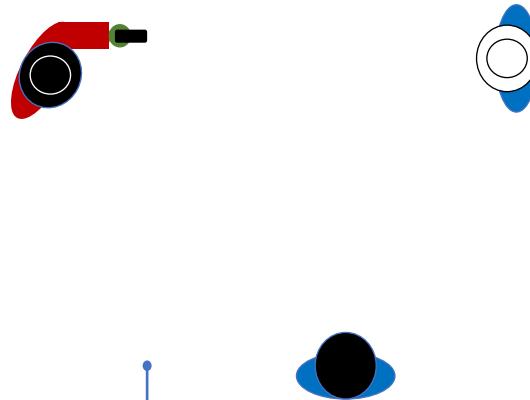
- AD has no position associated
- AD style is more traditional

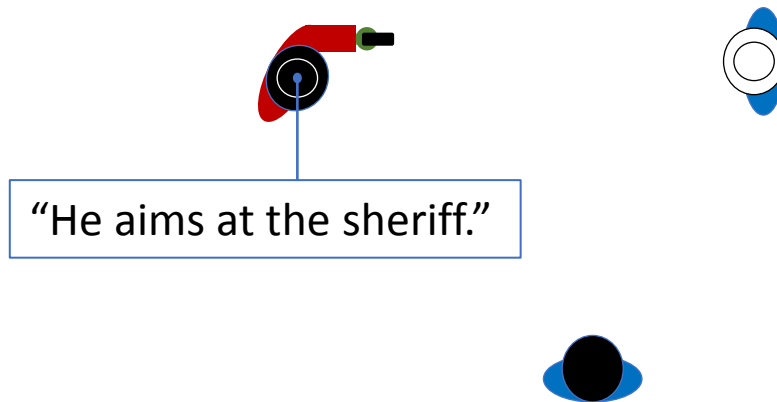"Mal pulls a gun on Sheriff Nemo."

Static (Friend on Sofa)

- 1st person

- AD comes from a static point in the scene

- Reminiscence ("that's my friend...", "We saw this...", "The shop I was in was full of beads...")

"That's when he pulled a gun on me."

Dynamic (AD on Action)

- Short, "to the point" sentences
- 3D audio can mean less description

"He aims at the sheriff."

6, The Practical

Audio Describe a video

Classic

- when, where, who, what

Static

- 1st person, past tense, make it personal

Dynamic

- Just the facts

Thank you!

WWW

http://www.imac-project.eu/

@ImAcProject