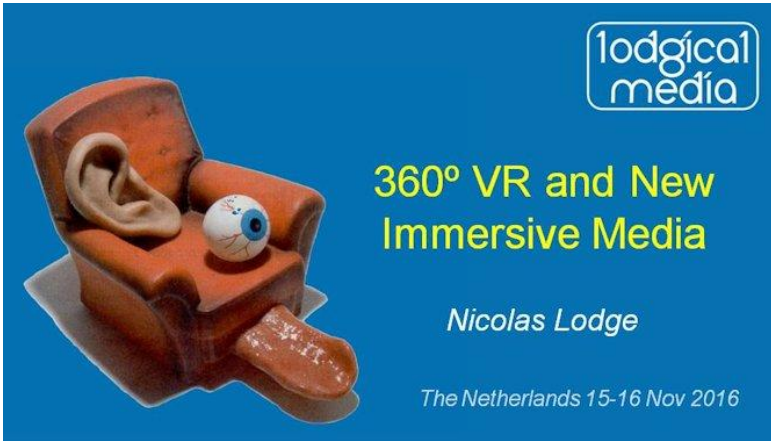


1



- I began with a few definitions: 'VR', 'AR', '360° video', 'cinematic VR', 'telepresence' and 'mediated'.
- Presence is the branch of psychology dealing with the feeling of being present in an immersive environment. It has its own international society, journal and conferences.
- Psycho-social presence is the feeling of being present with another person where one or both are mediated.

2



- Jacob Spoel 1820-1868, Rotterdam. Oil on panel, 41.5 x 48.8 cm . Rijksmuseum, Amsterdam
- Stereoscopy was all the rage from the 1850s to 1920s. This is probably the earliest image in the world depicting use of an immersive medium.

3



- In pure VR an HMD is tracked in 3 spatial dimensions and 3 orientational dimensions - 6D in total.
- VR was developed in the 1960s and has been widely used, mainly for flight simulators, safety-critical training, dangerous environments and, of course, computer gaming..
- VR allows the subject to interact with a computer-generated 3D model of an environment and potentially, to modify that environment.

4



- Here we see a VR tracking zone as demonstrated by Project Syria in 2015.
- Cameras track a head-mounted display in 6D, allowing a user to move around and explore the secure space.
- Views from the HMD viewpoint are computed for each eye from a 3D model representing the synthetic environment.
- Building the model can be time-consuming but it can be captured with photo-realistic scenery. Pure VR cannot be real-time.

5



- Project Syria was an experiment in Pure VR journalism where participants could explore a bombed street and see local Syrian people.
- It was operated by broadcasters led by Al Jazeera.
- Experiencing such a news item in an open area might not be practical in the home.
- Also, would we want a news story having this level of reality – depicting the victims of war?

6



- Current interest is in 360° video (called 'VR' or 'cinematic VR' by some).
- The HMD is tracked in only 3 orientational dimensions; the user 'sees' from the camera position.
- Some 360° cameras allow real-time viewing (ie live event broadcasts). The best ones do not because 360° stitching is computationally heavy.
- 'Telepresence' is a single person remote experience eg flying in a drone, defusing a bomb it may or may not be immersive.

7

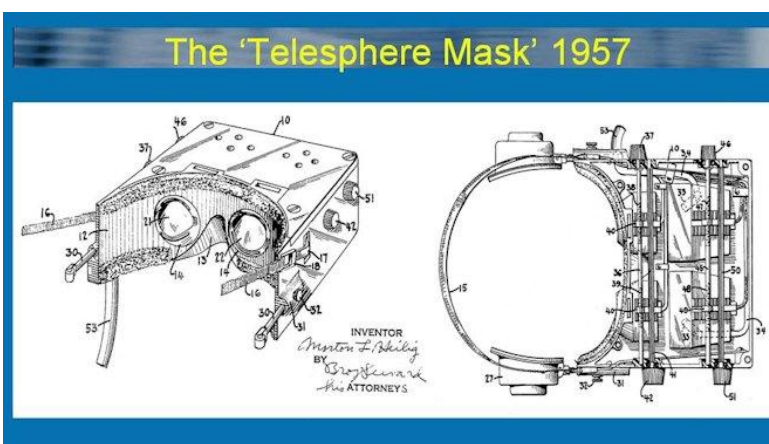
### The Key to Immersive Media is the Human Senses

	<ul style="list-style-type: none"> <li>▪ Vision 70%</li> <li>▪ Hearing 20%</li> <li>▪ Smell 5%</li> <li>▪ Feel 4%</li> <li>▪ Taste 1%</li> </ul>
--	--

Heilig 1955

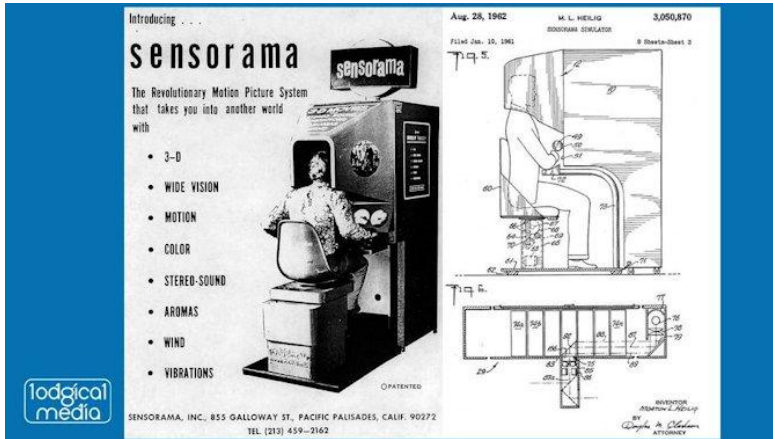
- All immersive media stimulate the human senses.
- If it is compelling, we can 'switch' from the real world to a mediated world. The medium itself then 'disappears'.
- There is no evolutionary precedence for the human to accept the mediated experience as 'real'.
- So which senses do we stimulate and how?
- Morton Heilig's 1955 suggestion.
- He has overrated 'smell' and not mentioned 'balance'.

8



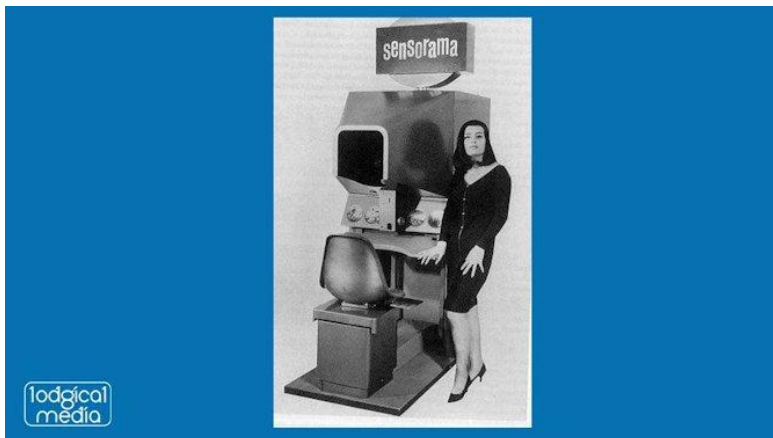
- Senses are complex eg 'feel' can mean temperature, texture, pain, force, location on the body ...
- The first HMD from 1957 - looks familiar! But was heavy because of its CRT displays. It needed a crane to support its weight.
- It had a tube for blowing air with any temperature & smell at the user's face.
- Ivan Sutherland in 1968 first viewed head-tracked VR imagery.

9



- Morton Heilig is best known for his 'Sensorama' system, the forerunner of today's arcade games.
- It was a motor cycle ride through the streets of Brooklyn and was the first such device to use an electronic display.
- Developed in 1957 and patented in 1962.

10



- Publicity shot of the Sensorama product.

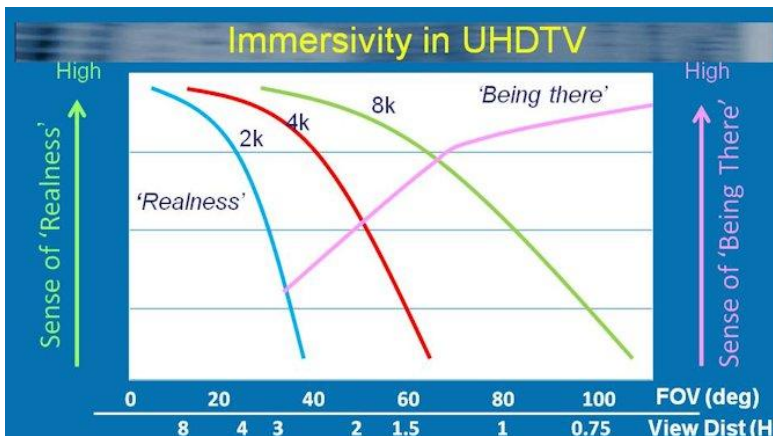
11

### My Top 10 Positive Cues

- Wide field of view (FOV) and head movement
- High resolution (& HDR) imagery
- Depth cues (stereoscopy, parallax, etc)
- Spatial audio
- Ability to move in and modify the environment
- Motion stimuli & forces
- First person viewpoint
- Life-size characters & objects
- Compelling storyline – live action
- Relevant smell

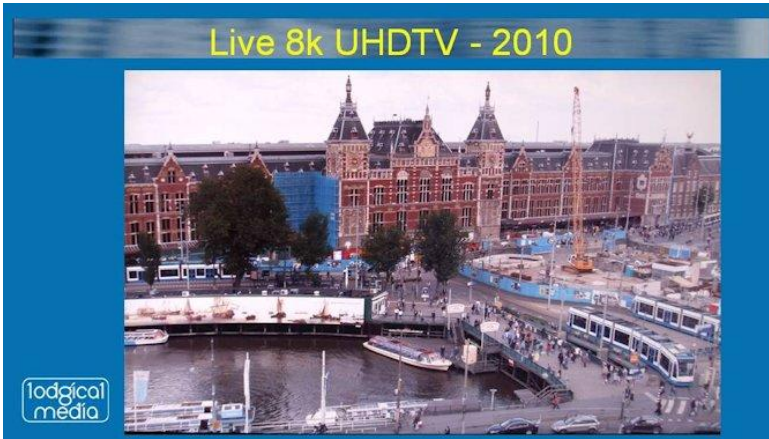
- My top ten sensory cues to encourage presence.
- I discussed each of them.
- The 5th one is not possible with 360° video.
- Life sized people are frighteningly real and evoke psycho-social reactions.
- High dynamic range is a significant factor too, both for enhanced reality and because 360° media are usually shot in available light.

12



- Nishida et al (NHK), 2011, 'Super Hi-Vision System Offering Enhanced Sense of Presence and New Visual Experience', Proc IBC 2011.
- FoV relates to viewing distance (meas in picture heights) by 16:9 aspect ratio.
- Based on subjective tests with real-world comparisons - shows the balance between resolution & FoV for different screen resolutions.

13



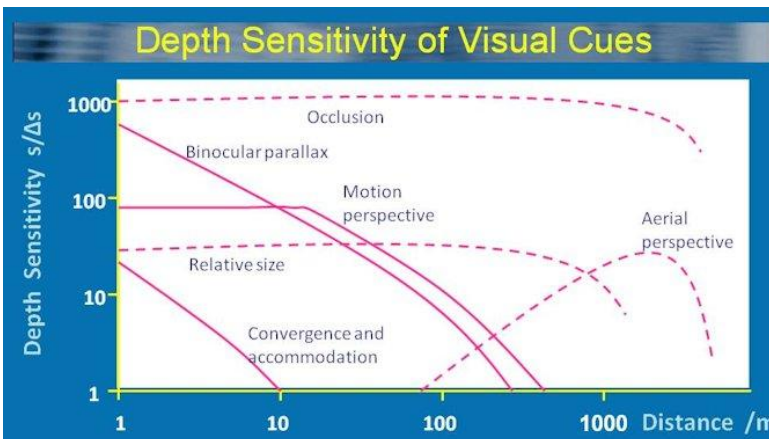
- First transmission of NHK's Super Hi-Vision 8k video, shown on projection display with huge screen.
- The image resolution was almost real-world – previously unknown with electronic imagery.
- Zooming the camera produced a remarkable sensation that the whole room was moving towards Amsterdam's Central Station.

14



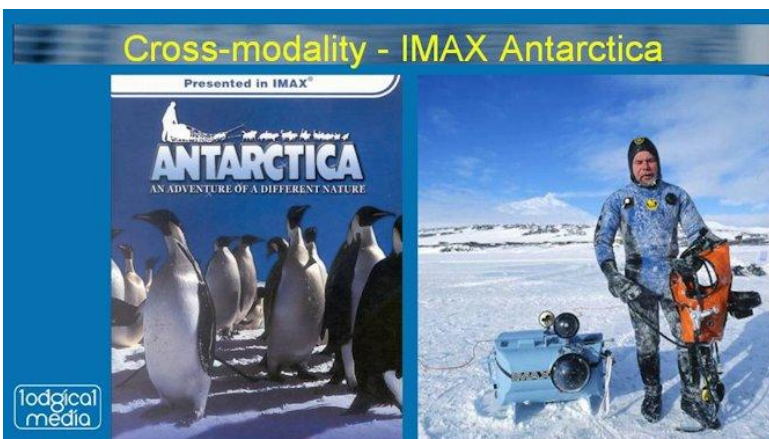
- The launch of the last shuttle in 8k with 22.2-channel sound.
- The sense of presence was staggering.
- The strong sense of presence evoked strong emotional feelings too as the power of the shuttle engines could be felt. Some viewers left in tears.

15



- Vital to a strong sense of presence is depth perception.
- Stereo (shown here as 'convergence & accommodation') is not the strongest or most far-acting cue. Binocular parallax is much better - but harder to achieve technically.
- True VR systems can easily compute L & R eye views. It is not easy to compute stereo in 360° video. Capturing and computing binocular parallax may be possible with light-field camera technology.

16



- Human senses are not independent, their responses are linked by learned experiences. This makes immersive media complex.
- Some 'cross-modal' responses are positive cues, others are negative cues (ie they push you out of the mediated world and back to reality).
- In I-MAX Antarctica, for example, the audience report feeling cold even though the theatre temperature remains constant.

17

## Negative Cues

- Distractions from the real-world (eg discomfort)
- Noise or transmission artefacts
- Processing delays ('lag')
- Sensory dissonance – eg lip sync error
- Lack of expected forces
- Screen flicker
- Edge of screen
- Poor/no parallax rendition



- Here are some negative cues.
- Most negative of all is where visual cues suggest a shift of balance. Unless a motion platform creates corresponding forces on the body, the user can feel sick. Careful production can avoid this.
- Understanding peripheral vision is important.
- Screen flicker has been widely researched as a -ve cue. It gave rise to a new reality film format by Douglas Trumbull: 60Hz Showscan

18

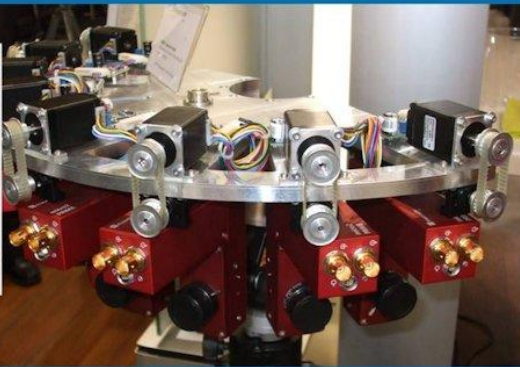
## The Fish-eye Lens



- The earliest form of 360° capture was the fish-eye lens.
- Popular in the 1990s because domestic computers could un-warp the video to produce look-around imagery, albeit with very low resolution.
- With improved sensors the idea has recently been born again in the shape of Ricoh's Theta-S camera.

19

## FascinatE Panoramic Camera



- I will show the development of a 360° camera - the Fraunhofer-HHI Omnicam. Partly funded within the EU FascinatE project.
- An early 180° stereo camera. Individual cameras are shown below, viewing radially.
- Doesn't work well for objects close to the camera because individual images can't be stitched together.

20

## FascinatE 180° Camera 2010



- Their next generation used mirrors. The cameras look upwards. Note that the optical centre of *all* the cameras *appears* to be at a point in the very centre of the mirror array. The cameras all view from the same central point. Very clever. Objects can now get close to the camera without influencing the stitching process.

21

## Real-time 180° Video - 2010



logical  
media

7000 x 1920 pixels

- A real-time 180° panorama shot from a monitor at IBC 2010
- The image resolution qualifies it as 'ultra high definition'.
- Let's look at some of the processes which must be performed in real-time to create this.

22

## Original Camera Views



logical  
media

- The 6 HD cameras are operated in 'portrait' orientation and their fields of view overlap very slightly.

23

## Geometric Correction & Warping



logical  
media

- As each image is projected onto a plane (the camera sensor) the images need to be warped into a curve in order to join them.

24

## Photometric & Colour Correction and Blending



logical  
media

- Colour and sensitivity will vary between the cameras and light levels will vary across the field of view. These must be compensated for.
- Images must then be 'stitched' or matched and blended together. There is scope for some sophistication here to ensure that objects moving across the boundaries do not reveal distortion. The timing of capture by separate cameras is also a critical issue if joins are not to become visible when objects move across them.

25

### Final Image Cropped



logical media

- After warping, the images may have slightly curved edges at top and bottom (see previous slides). These must be cropped off.
- This results in the final panorama.

26

### Viewing a 180° Panorama



- Viewing a panoramic live event evokes a strong feeling of presence.
- The camera is fixed. The viewer moves his or her head to follow the action, as if viewing the real event. Moving one's head to follow the action is a positive cue - whether it takes place in front of a screen or while wearing a HMD.

27

2012

2013

6x ALEXA-M  
180°  
1920 x 6000



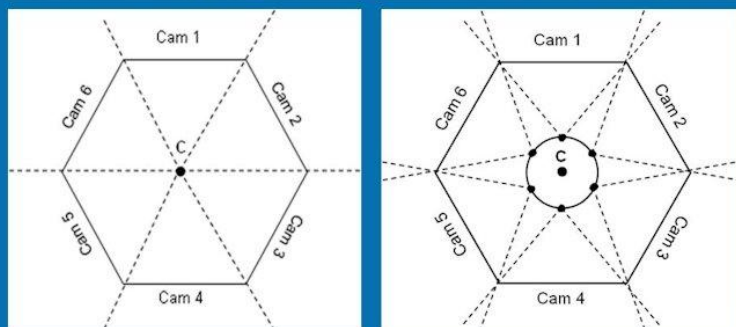
10x INDIECAM  
360°  
1920 x 10000



- Further generations of the camera extended it to 360° operation.
- It now has a further camera physically sited at the centre of the mirrors and looking upwards. With this, the camera array can capture a complete dome view. But not currently in real-time.

28

### Overlapping Views to Assist Stitching



- The cameras looking into the mirror would have a non-overlapping field of view if their optical centre were exactly at the centre of the mirror array. (left-hand fig). This would make stitching difficult.
- The designers cleverly offset the cameras slightly (right-hand fig) to generate a 10-pixel overlap between adjacent images. This allows excellent stitching.

29

### Berlin Philharmonic – December 2014



- One of many trial recordings. The Berlin Philharmonic Orchestra and choir shot from above the head of conductor, Sir Simon Rattle.

30

### 360° Cylindrical Video



A rectangular format – convenient for editing, compression, transmission and Streaming

- Live 360° cylindrical video from IBC 2014.
- You can see me taking this picture 2/3 of the way across. There is nowhere to hide in 360° media!
- Notice that a rectangular format is good for digital processing and compression. It is also convenient for 'tiled' streaming, where only the viewer's current region of interest is transmitted.

31

Nokia 'Ozo'



Google 'Jump'

Sphericam 2



Jaunt 'One'

- Now there are many 360° cameras with different features and performance specifications.
- If you think that all this is clever, spare a thought for the prolific inventor and mechanical genius, Fred Waller ...

32

### Vitarama 1938

Fred Waller 1886-1954

11 x 16mm cameras

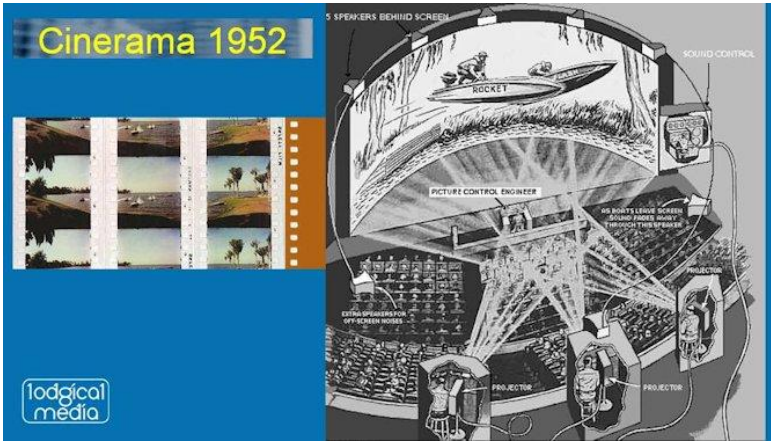
180° x 90°



- In 1938 he linked together 11 16mm film cameras to record and reproduce a 180° H x 90° V image on a dome for the New York World Fair in 1939.
- Fred was the first to study the peripheral vision and to understand the psychology of immersive entertainment. He was later to create the Cinerama immersive experience which enthralled people throughout the 1950s and was the forerunner of formats like IMAX.

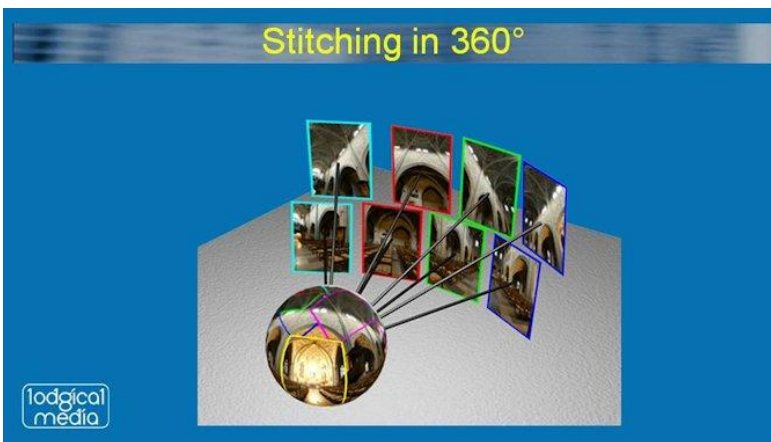


33



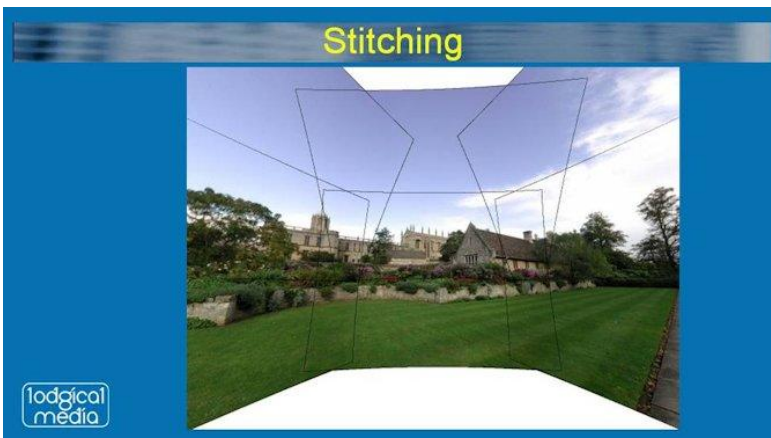
- The Cinerama immersive cinema. It had a 146° curved screen 28m x 12m.
- The film format consisted of 4 x 35mm films. 3 held the image (extra wide because the usual magnetic audio track had been removed) and the fourth held 7 audio tracks.
- Stitching the 3 images was done with a slight overlap and mechanical interpolators – a picture control engineer manually adjusted for drift!
- Launched on 30 Sep 1952. There are now 4 left: 3 in the US 1 in the UK.

34



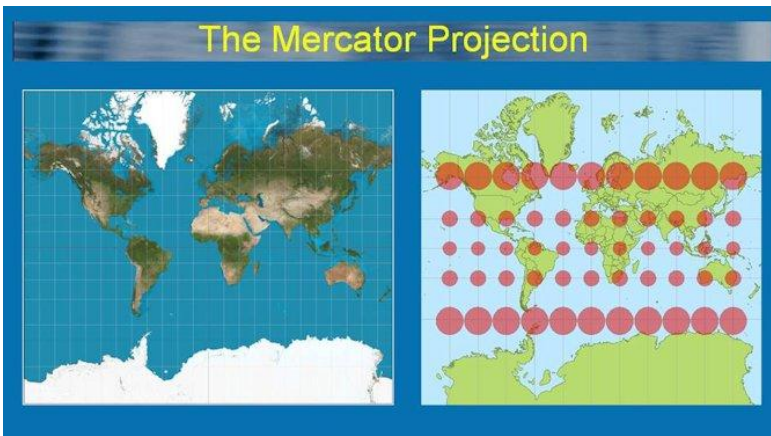
- The extension of cylindrical to spherical image capture follows the same principle.
- Images are captured from radial cameras and correctively warped so that they appear to be part of a sphere

35



- Stitching is more difficult.
- Handling a spherical array of pixels in a digital memory or processor would be a huge inconvenience, however. We need a 'flat' representation and preferably one that is rectangular.
- This is far from a new problem. Cartographers have had to deal with the representation of a sphere on paper for hundreds of years.

36



- Mercator's projection essentially wraps a cylinder of paper around a sphere and projects points onto it. This happens to preserve the angles needed for navigation, so it was popular.
- However it distorts, as the red circles show - these are the same area on the surface of the sphere but grow towards the poles. And look how Antarctica is distorted!

7

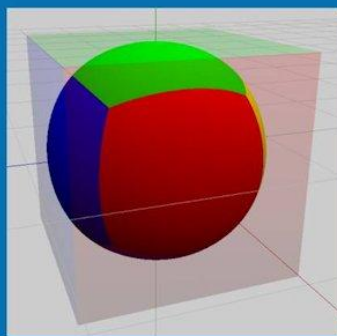
### 360° Equirectangular Representation



- This is what it does to an image.
- It's a poor representation because constant image resolution on the surface of the sphere is not equally disposed across the rectangle.
- The representation is still used within the processing of most spherical cameras because of its convenience. It is also used in editing (usually treating it as if it were a regular HD frame). It is not good enough, however, as the starting point for digital compression and transmission.

38

### Facebook's Cube Solution



- Bodies such as MPEG and DVB are still deliberating on their technical approach to 360° video. Facebook has proposed an efficient approach to a flat image representation.
- Imagine a cube surrounding our image sphere. The cube's surface is divided into 6 regions, each of which is projected onto a face of the cube.

39

### Facebook's Cube Solution



- Let's see this work. We start here with an equirectangular image.

40

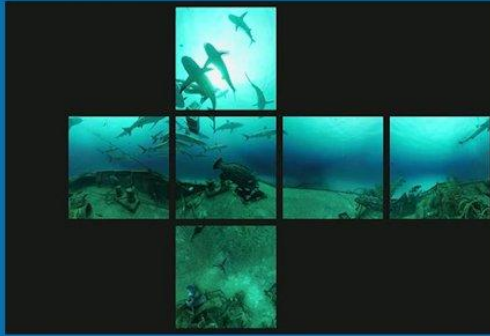
### Segment into cube faces:



- In this equirectangular world, the equivalent of projecting onto Facebook's cube looks like this.

41

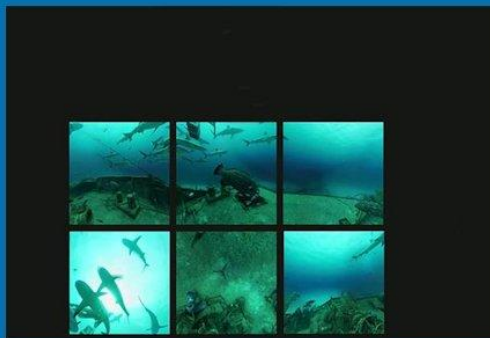
### Warp into squares



- Now we can squeeze or warp the representation into the faces of a cube. We have a flat arrangement of squares which is a much more faithful representation of resolution as it was in the spherical image. It is also a much more efficient representation and therefore good for compression.

42

### Rearrange



- The squares can be shuffled to form a 2x3 rectangle - more convenient as a starting point for image compression prior to transmission.

43

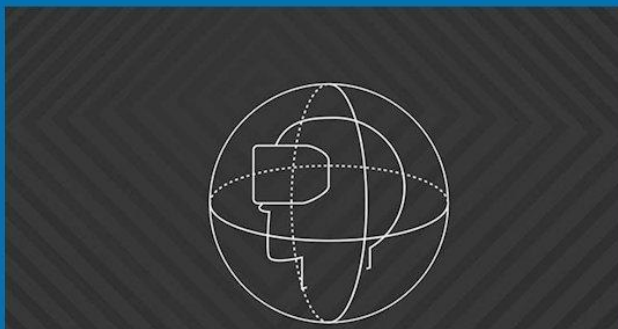
### Facebook's approach to efficient streaming



- When streaming 360° data to a HMD it is not necessary to convey parts of the image that the user is not currently viewing. Standards already exist for streaming a 'tiled' image representation but Facebook has published a different approach.

44

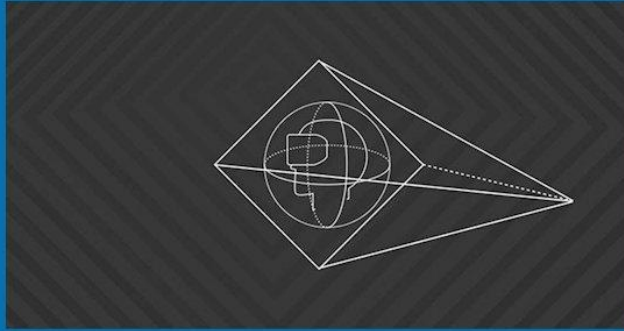
### Observer with 360° view



- Imagine a viewer watching a 360° image (here shown as a sphere at a distance from his head).

45

### View is split into ROI and peripheral

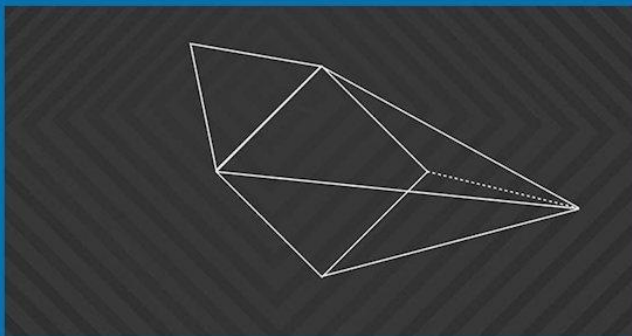


1odgical media

- In front of the viewer is the area he is watching (the region of interest or ROI).
- The viewer's sense of presence is enhanced if we fill more of his field of view - but we may not need to represent these peripherals with full resolution video.
- So in Facebook's diagram they show the remaining part of the 360° image mapped onto a stretched pyramid at increasing distance from the viewer (ie the more it is stretched, the lower the resolution).

46

### Unwrap ...

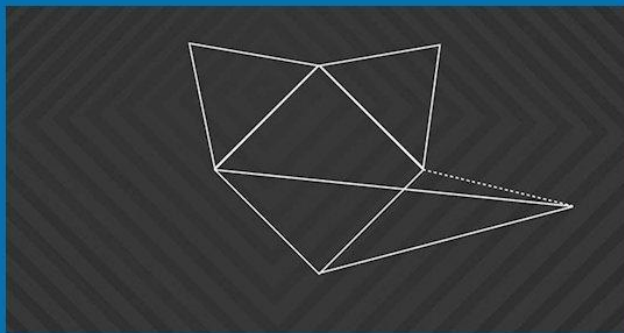


1odgical media

- So, for this particular gaze direction, the image data we want to stream is mapped onto a square-based pyramid.
- If we unwrap this to get a flat representation...

47

### Unwrap ...

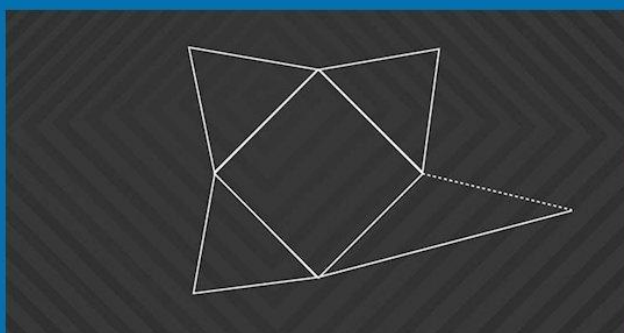


1odgical media

•

48

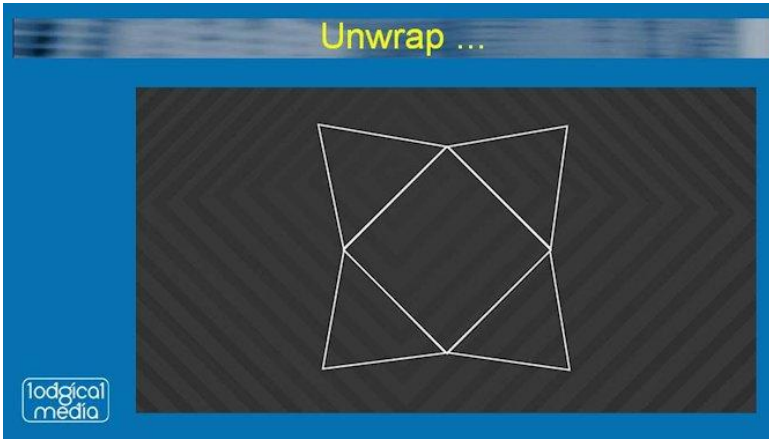
### Unwrap ...



1odgical media

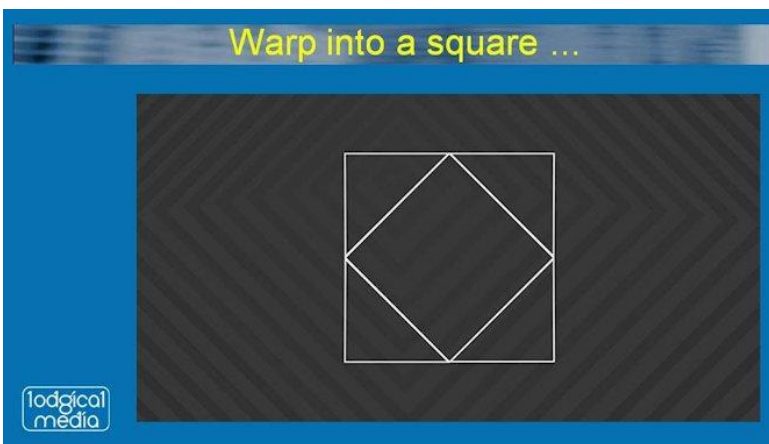
•

49



- it looks something like this.
- As the peripherals contain lower resolution they can be squeezed more than the crucial ROI, which remains unaffected.
- From this it is a small step to warping it into a square...

50



- As before the square represents a convenient pixel array for storage or compression prior to streaming.

51



- So for our image above (a view from the cockpit of a stunt aircraft), where the viewer is looking centrally through the window (his ROI), this is what the streamed video might look like.

52



- A broken down Oculus Rift showing the system specification.
- The lenses are hybrid Fresnel lenses.
- The weak point of today's HMDs is the image resolution (especially sad, as we can *capture* much higher resolutions). Much work is going into this challenging task. Oculus have said that it might take a decade to get to 8k resolution.

53

### The Roller Coaster Ride

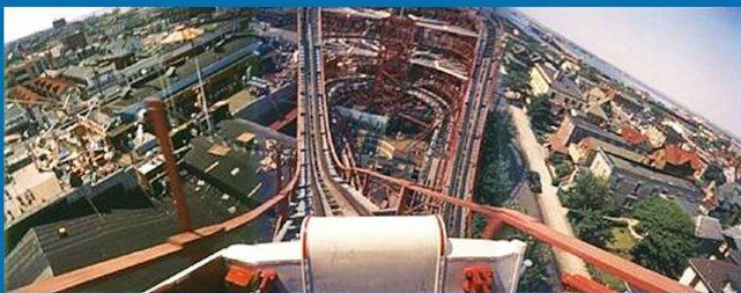


1odgical media

- The roller coaster ride is a VR classic but if many people have motion sickness, why aren't more of them sick. Or, why aren't lots of people sick on real roller coasters?
- The answer lies in the banking.

54

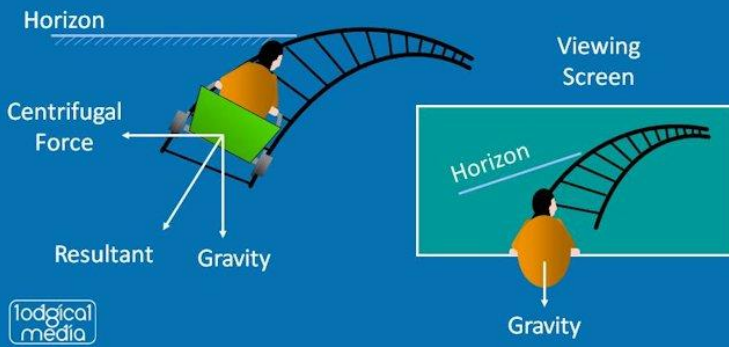
### Cinerama Roller Coaster Ride 1952



- A roller coaster ride Rockaway Playland's 'Atom Smasher' was used in the 1952 Cinerama film 'This is Cinerama'.

55

### The Roller Coaster



1odgical media

- Roller coasters are designed so that the banking ensures that there are few dangerous sideways forces on the riders. Left-hand diagram.
- The fact that forces always act vertically on the body helps us to create realistic VR rides. As the viewer's reference is the screen, the image can be tilted to ensure that the apparent *resultant* force on the immersed subject corresponds to ordinary gravity.

56

### Mediated Coaster Experience



1odgical media

- So if you watch people on a VR coaster ride they don't move very much. But they think that they do!
- The visual looming cue is a powerful attention-grabber.
- The speed of the car along the track is matched by the rate of a mechanical ratcheting sound.
- Air is usually blown into the subject's face at the relevant speed too.

57

## Mediated Grand Prix Experience



- For non-banked rides like motor racing we must provide sideways forces in order to provide a sense of reality.
- This can be done with a motion platform. Great fun!
- The movements of the platform have to be carefully choreographed, however. It is not possible to produce a sustained sideways force for example. Jolts can be made but gravity remains the only constant force at the designer's disposal.

58

## The Latest Twist on the Coaster



- So now the latest twist is to use synchronised VR while riding a real roller coaster!
- The coaster is thus a motion platform.
- The clever idea is to take an old, 'tame' coaster ride and turn it into a new exciting proposition! A great business idea.

59

## Experiences of 360° Production

- 360° filming requires an entirely different 'spherical or cylindrical' approach
- How do you use the presenter?
- Where do you hide the crew and kit?
- Lights and mics must be concealed
- Camera technology is not mature
- Post production workflow laborious – stitching, working on unwrapped video, 360° audio editing

- Remember that 360° shooting often uses available light (a rig will be in shot) and often levels will vary around the scene. This is why HDR is particularly advantageous.
- Presenters report that an informal style works well with first person format. Being led around by a guide works well (eg a character from the mediated environment).

60

## The 'Pufferfish' Display



- How do we represent a spherical image for the purpose of editing and post-production?
- One way might be to use a spherical display like the Pufferfish.
- The image is inside out, however.

61

## A New Production Grammar

- Will you always want to be immersed?
- And for how long?
- Production grammar needs to be re-invented; cuts, zooms, moves may disorientate the viewer – but can they be avoided?
- Will close-ups be needed?
- How do you direct the viewer's attention?

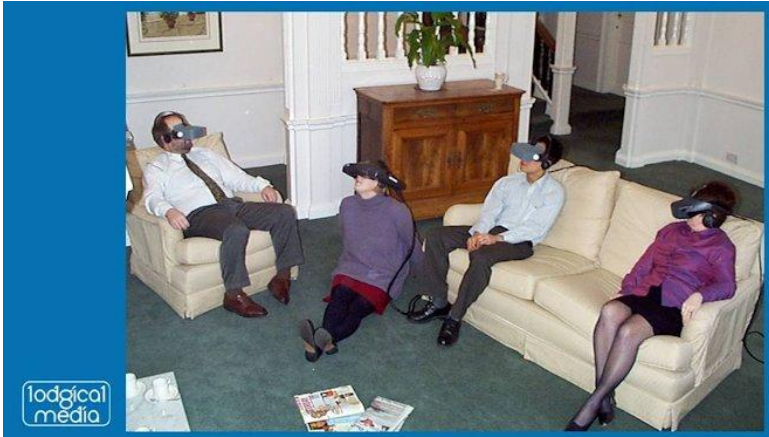
- Most people report that 360° video is best consumed in short productions of perhaps 5-10 minutes. Will there be a time when we might enjoy longer experiences – dramas, concerts, sporting events?
- Creative producers have found ways to introduce scene dissolves and ways to direct the viewer's attention.

62

## Summary

- My verbal summary of the main topics covered in this masterclass.
- It has only been possible to cover some fundamentals here. The world of immersive entertainment is rich with exciting experiences, strong emotions and a wealth of fascinating applications.
- We do not yet know if VR or 360° video will become mainstream entertainment. Perhaps a family evening in front of the TV might look something like this! ...

63



- (I shot this scene in 1998 even before the days of digital photography – the original is on 35mm slide!)

64



Proost !

[nick.lodge@iee.org](mailto:nick.lodge@iee.org)

+44 7900 394656

- Cheers!
- Find me at Lodgical Media

Phone: +44 7900 394656

Email: [nick.lodge@iee.org](mailto:nick.lodge@iee.org)

© 2016 Nicolas Lodge