Next steps to immersion in Virtual Reality

Agenda

- Introduction
- HW Evolution
- Depth Reconstruction
- Body Immersion
- Redesign of UI

Introduction

Karol Gasiński

- http://www.linkedin.com/
- kuktus@gmail.com

Graphics Software Engineer @ Intel
In transition to:
Senior Software Engineer @ Apple

KHRONOS Group member
Contributes to OpenGL Specification

- http://www.khronos.org/

Founder of WGK - this conference ☺
Speaker at plenty of events:

- GameDev Shool, Digital Frontier, GameDay, WGK, IGK, Gdańsk University of Technology, AGH,…

OpenGL

Introduction

Projects:

- True Color graphics library for DOS
- Multitasking Operating System Microkernel in ASM
- Multiplatform game engine (Android/BlackBerry/iOS/Windows)

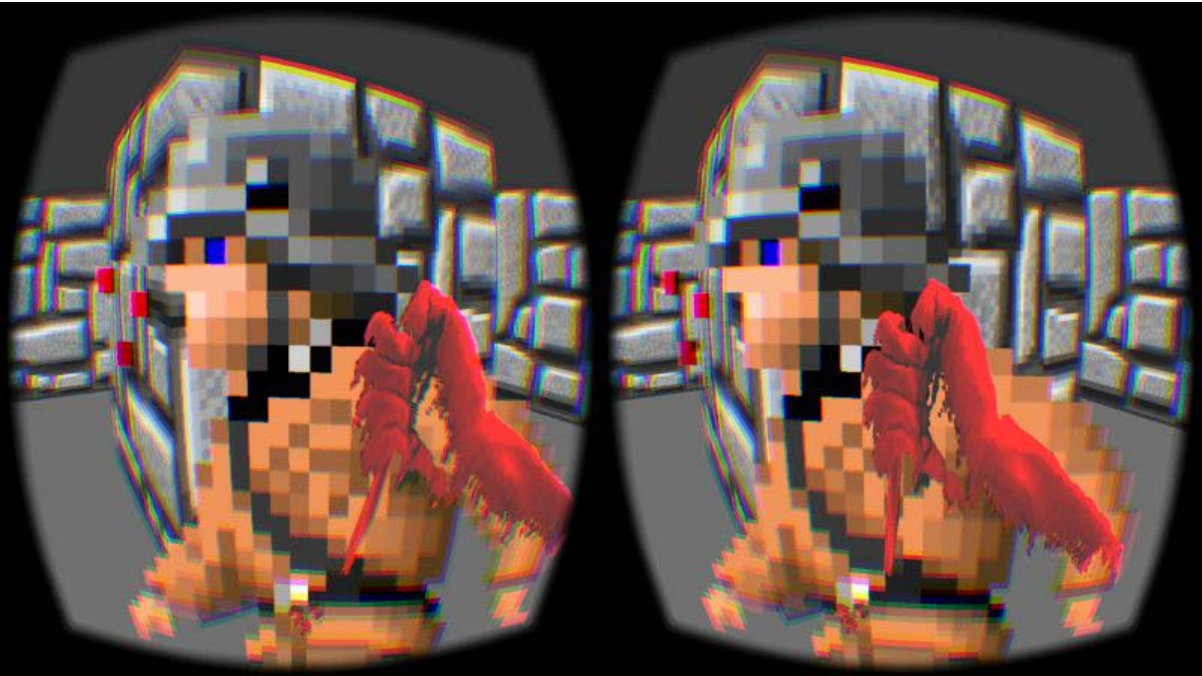
Worked on mobile versions of:

- Medieval Total War
- Pro Evolution Soccer
- Silent Hill
- Invaders Must Die!
- Other Indie games...

Currently developing:

- WolfensteinVR ☺

Introduction



WolfensteinVR is an experimental game developed to research how we can increase user immersion in Virtual Reality (and have fun ☺).

VR Headsets

Resolution	Dev-Kit 1	DK2	CV1 (predicted)
Display	1280x800	1920x1080	2560x1440
Per Eye	640x800	960x1080	1280x1440
Field Of View	90	110-130	130
Refresh Rate	60Hz	60Hz, 75Hz	90Hz

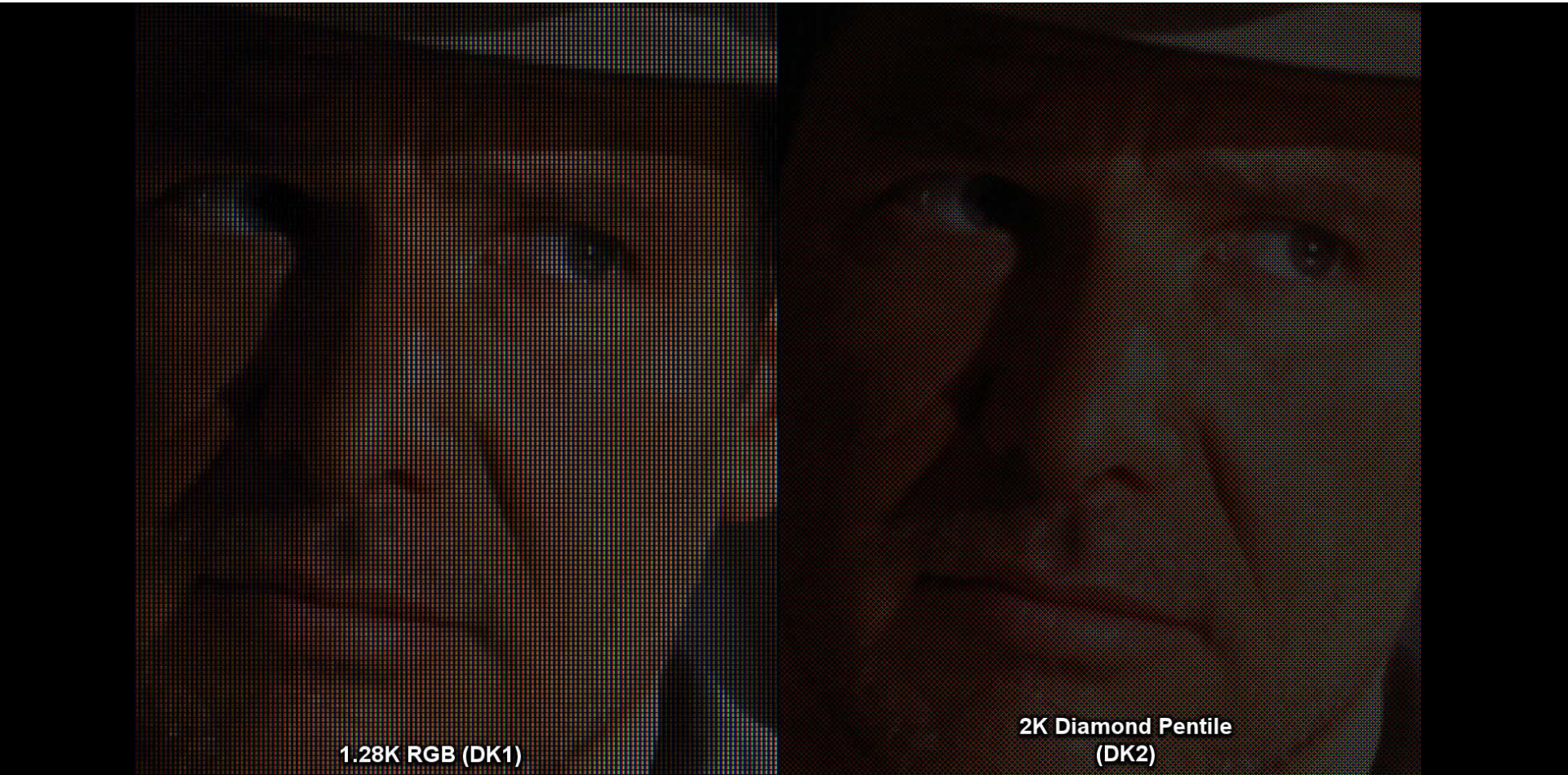
Oculus Rift (DK1)

Oculus HD

Oculus Crystal Cove

Oculus DK2

VR Headsets



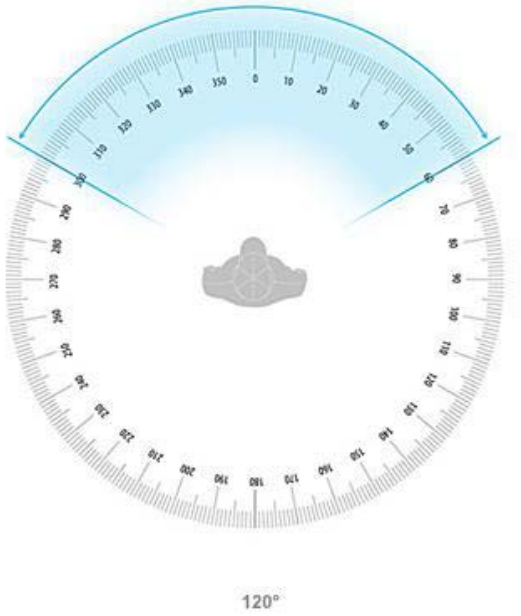
Field Of View

FOV (*pol. pole widzenia*)

Field of view in both, horizontal and vertical directions

(by FOV we mean horizontal FOV)

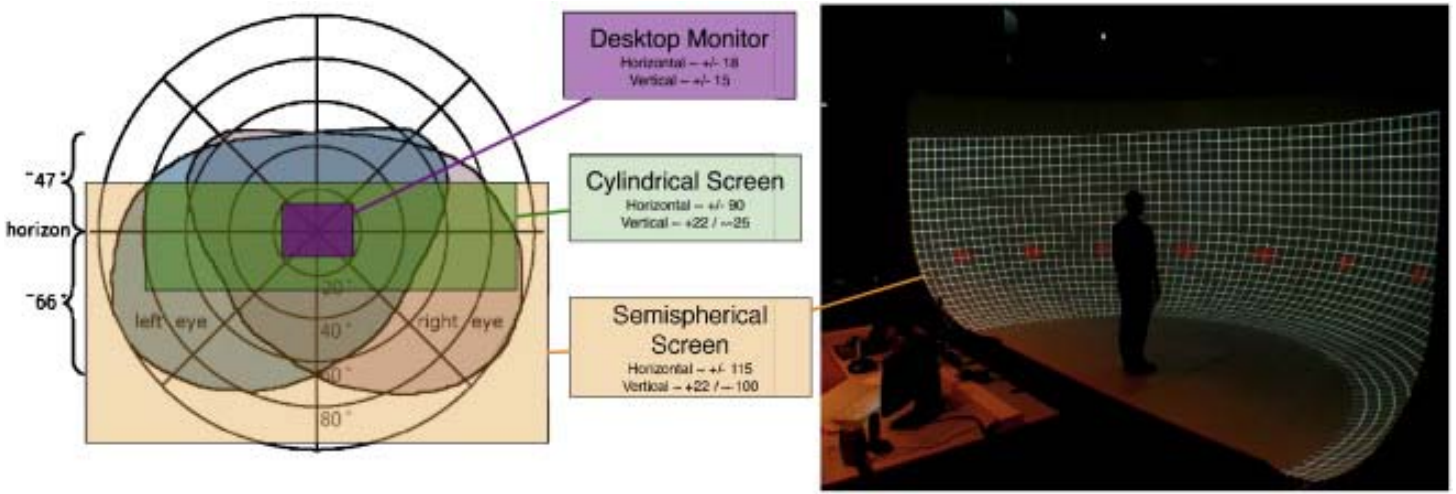
FOV (horizontal)
~150º Per eye
~120º Stereoscopic, center covered by both eyes
~180º total



Field Of View

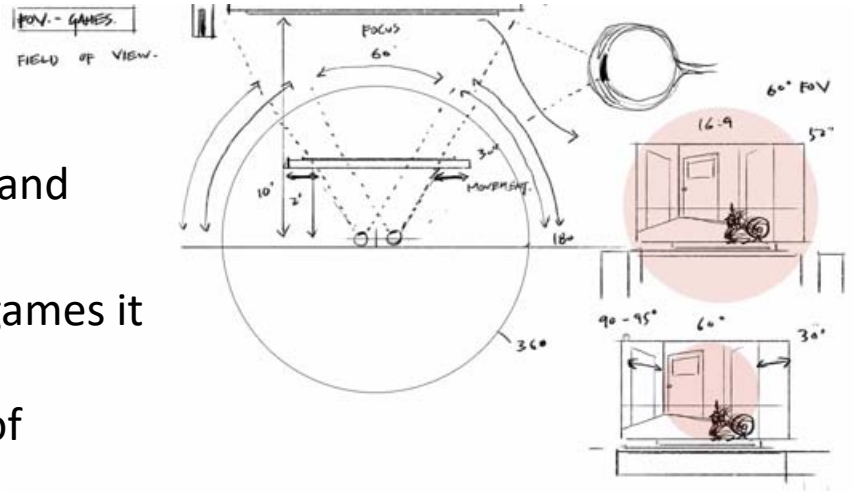
Vertical FOV

- Nominal line of sight is 15° below horizon line
- We focus only on area of +/- 15° from that line (from horizon line to -30 degree)
- ~100° total



Field Of View

Adjusting FOV



- FOV depends from display aspect ratio, size and distance from viewer.
- For PC standard is 90º FOV and for console games it is set to 60º. This ensures that in both cases displayed content will cover the same area of Human FOV.

More explanation at:

- http://www.youtube.com/watch?v=blZUao2jTGA

Gesture Recognition

Currently available devices on the market:

Kinect for XBox 360

Creative Senz3D

Real Sense 3D*

Kinect v2.0 for XBox ONE

MYO Armband*

Gesture Recognition – Depth Sensors

Creative Senz3D and Real Sense 3D are currently best choice for interactive gesture recognition

	Kinect for XBOX 360	Creative Senz3D	Kinect 2.0 for XBOX One	Real Sense 3D
Sensor Type	Pattern	TOF	TOF	TOF (?)
Resolution	320x240	320x240	512x424	640x480
Field of View	H57.5 V43.5	H57.5 V43.5	H70.0 V60.0	(?)
Refresh Rate	30fps	60fps	30fps (?)	60fps (?)
Range	0.8m – 4.0m (0.4m-3.6m)	0.15m-1.0m	0.8m – 4.0m	0.15m-1.0m (?)
Accuracy	~1.5mm	1mm	~1.5mm	1mm (?)
Streams	Color, Depth	Color, Depth, IR, Vector Field	Color, Depth, IR	Color, Depth, IR, Vector Field
Color Resolution	640x480 (24bpp 30fps)	1280x720 (32bpp 30fps)	1920x1080 (16bpp 30fps)	1920x1080 (32bpp 30fps)

Gesture Recognition – Depth Sensors

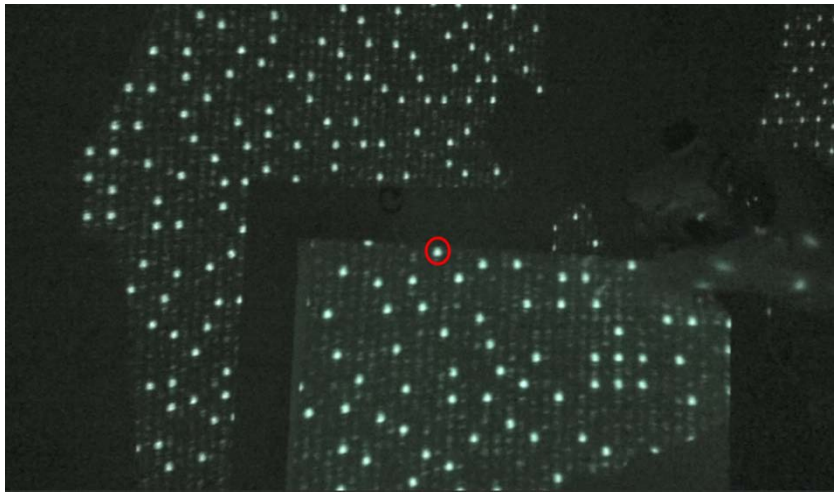
Depth sensors FOV is different from Oculus HMD FOV. This needs to be taken into notice during depth reconstruction.

Gesture Recognition – Depth Sensors

Kinect 360 (as well as first Kinect for Windows) has too low quality, and it's HW was designed for far range depth scanning (0.8m - 4.0m).
It's also not portable and thus, not practical as integrated sensor.

It uses IR patterns to calculate scene depth as well as to recognise depth discontinuities.

It's real resolution is about 80 x 60 points which is used as a base for interpolated depth of 320x240.



Gesture Recognition – Depth Sensors

Kinect for Xbox ONE (also known as Kinect v2.0 for Windows) has different depth sensor, it uses Time Of Flight depth camera which samples depth probing space one probe at a time.

This means that all depth probes in 514x424 resolution depth buffer provides meaningful data. It also has FullHD color camera which provides much more color samples to pick from.

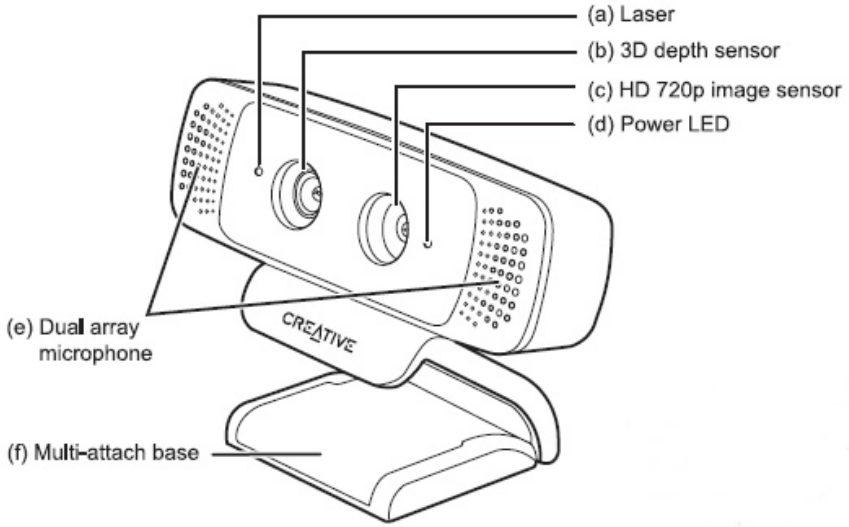
Unfortunately it’s size allows it to be used only as static sensor.

Gesture Recognition – Depth Sensors

Creative Senz 3D thanks to it’s small size can be used as portable depth camera. It also uses TOF sensor but is designed for close range depth scanning and can provide frames with 60Hz frequency.

It’s a perfect fit for HMD mounted hands scanner, and it also provides additional data like vector field buffer or ID buffer.

(Intel Perceptual Computing SDK that is used to access the data is somewhat non trivial to use :/)



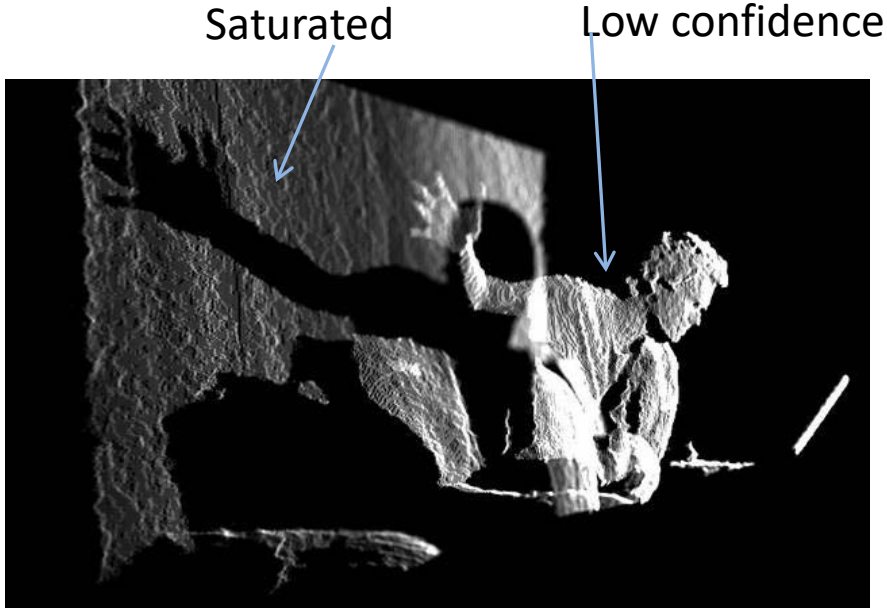
Player hands image reconstruction

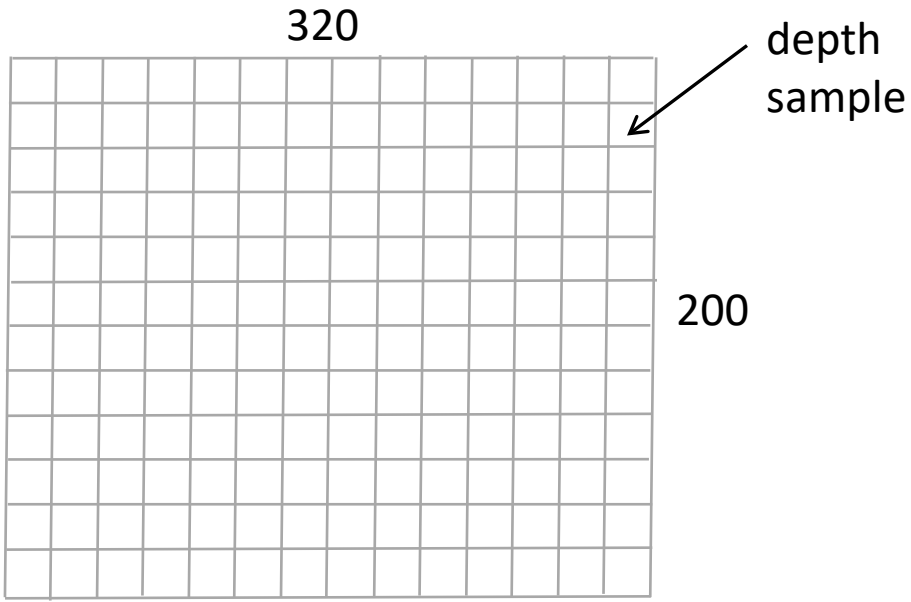
First prototype setup was based on Oculus DK1 and Creative Sezn 3D. Depth sensor was pointing straight in the looking direction of player. It resulted in great image quality, but required player to hold his hands up in front of the face to see them which was unpractical.

Player hands image reconstruction

Depth buffer:

Measured data is returned with micrometer (10^{-6}) precision as uint16.

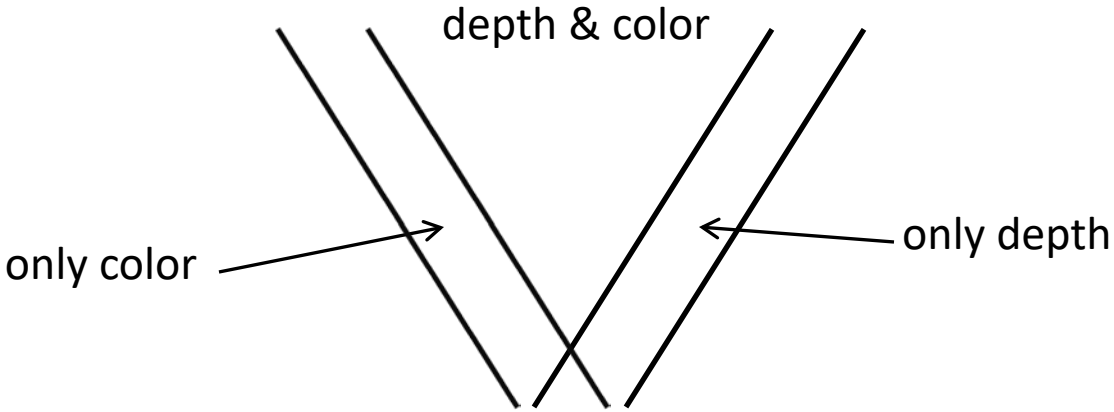




This data needs to be cleared from noise by using different filters.

Player hands image reconstruction

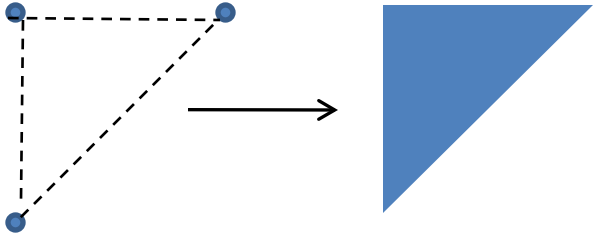
Color buffer:
Color stream is in 1280x720 resolution and 30Hz. Color sensor is also physically shifted by around 2cm from depth sensors which causes that it’s frustum is not covering depth sensor frustum.



To match color and depth samples we need Vector Field.

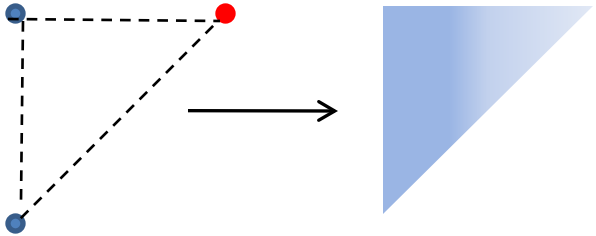
Player hands image reconstruction

Geometry shader samples depth buffer to create mesh.
Low confidence and saturated samples are rejected.

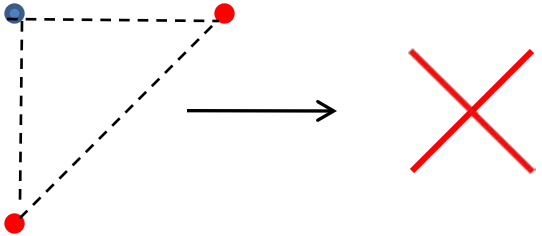


Color buffer is mapped using
Vector Field displacement.

All 3 samples are correct. We build mesh triangle.



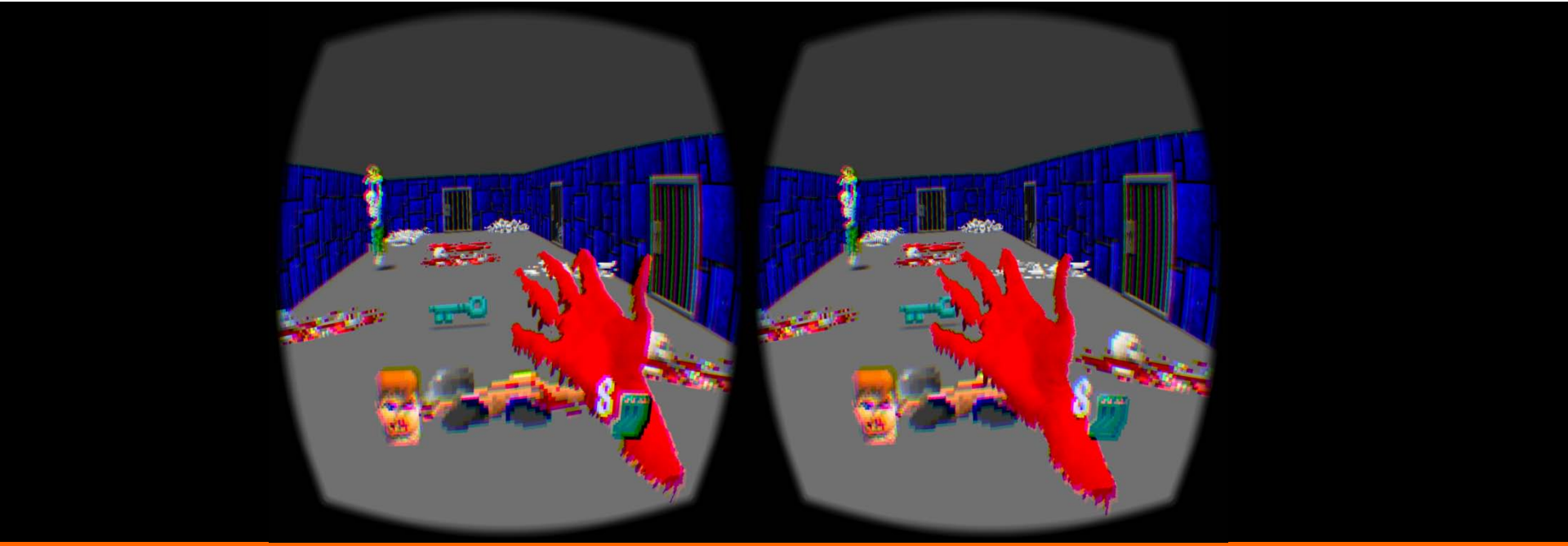
One sample is rejected, we blend-out.



Two or all samples are rejected, we discard triangle.

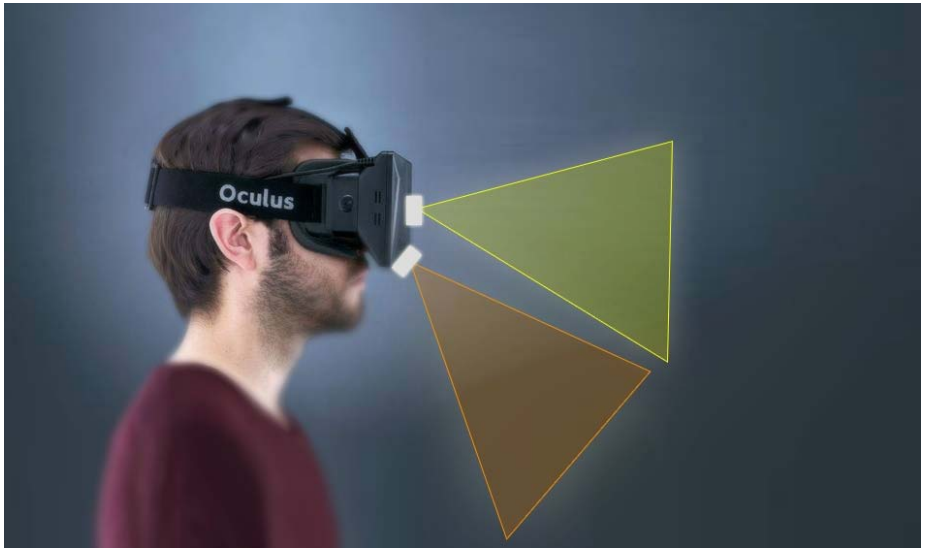
Player hands image reconstruction

Finally, in Fragment Shader we add filters to create desired effect.
For WolfenstainVR it is effect of blood and ghost like appearance.



Player hands image reconstruction

Camera relative position and orientation dictates quality of received data. Camera facing upfront will provide best data for reconstruction but has limited FOV for our use case. Camera facing down perfectly captures hands but introduces samples noise.



Conclusion

Any questions?