

Real Avatar Production - Raspberry Pi Zero W Based Low-Cost Full Body 3D Scan System Kit for VRM format

Yukihiro IWAYAMA
EIDEN Inc., Tokyo, Japan

<https://doi.org/10.15221/19.101>

Abstract

We would like to introduce Raspberry Pi Zero W based 360-degree full body shooting / 3D scanning system kit in order to create "Real Avatar". "Real Avatar" is a word for photo-real 3D Humanoid Avatar. System configuration, flow, operation and resulting 3D scanned model are shown in this paper. In order to create mesh and texture, we use Photogrammetry software, RealityCapture, on Microsoft Azure virtual machine being configured with NVIDIA GPU. We also briefly introduce VRM format which is 3D Humanoid Avatar format that has the potential to be widely used for various VR/AR apps and platforms.

Keywords: 3D body scanning, Raspberry Pi Zero W, Photogrammetry

1. Introduction

We would like to introduce Raspberry Pi Zero W [1] based 360-degree full body shooting / scanning system. Raspberry Pi based body scanning systems have been implemented and commercialized in the past and present [2]. We have explored the possibility of further cost down because we believe in the expanding opportunities of "Real Avatar" production with the advent of social VR/AR platforms.

2. Purpose

Our motivation for this research and business development comes from our vision that is "We can meet anyone, anytime, anywhere all over the world in the near future by using Real Avatar". Real Avatar, representing identity / 3D Model, is important asset in XR, digital three-dimensional extended reality space. We value the importance of communication with physicality and a sense of distance. We believe in the future where we can reconstruct the whole world digitally and we want to have our own Real Avatar.

3. Overview

3.1. Real Avatar Production Overview



Fig. 1. A human subject surrounded by 88 units of Raspberry Pi Zero W + Camera Module and processed in photogrammetry software, RealityCapture [3]. The number of camera units vary from 80 to 112.

It takes about three weeks to source parts, install and set programs, build rigs and make adjustment.

3.2. 3D Scan Studio Overview in Asakusa, Tokyo



Fig. 2. Asakusa studio and posing human figures

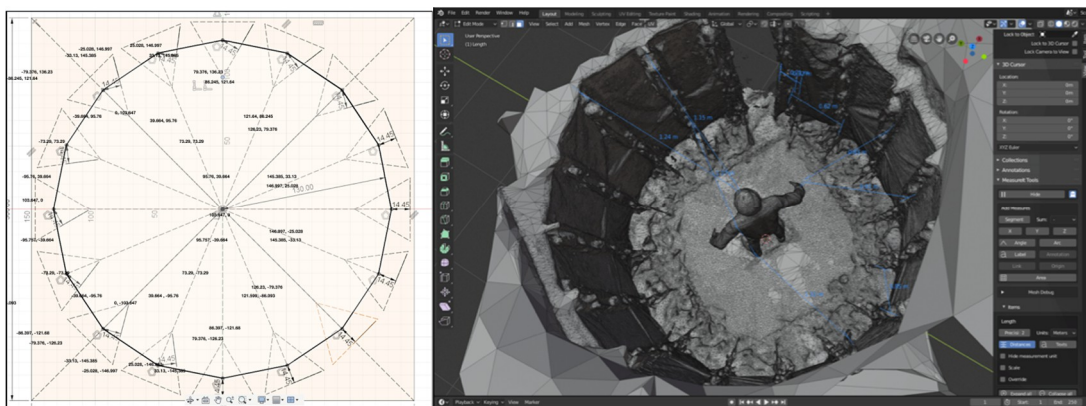


Fig. 3. Stand layout, 2D sketch and 3D model view from the above

We put more camera units mounted on the stand aiming for subject's face position than the ones mounted on the other equally spaced surrounding stands. The upper level and lower level of the cameras are angled appropriately for better coverage of the subject.

Currently, we offer 360-degree shooting and 3D model creation service for \$70 in Asakusa, Tokyo. There is one dedicated staff in 15m² space. See <https://www.real-avatar.com> (in Japanese).

4. System Kit

4.1. Hardware parts



Fig. 4. parts

We bought various parts from amazon, DIY stores and electronic parts distributors.

Raspberry Pi Zero W, camera module v2, micro SD card 8GB, case, smartphone holder, bolt, clamp, bar light, stand 2m, USB cable, USB hub, white cloth, jigsaw mat, cable tie, and business-use wireless router.

4.2. System Configuration and Flow

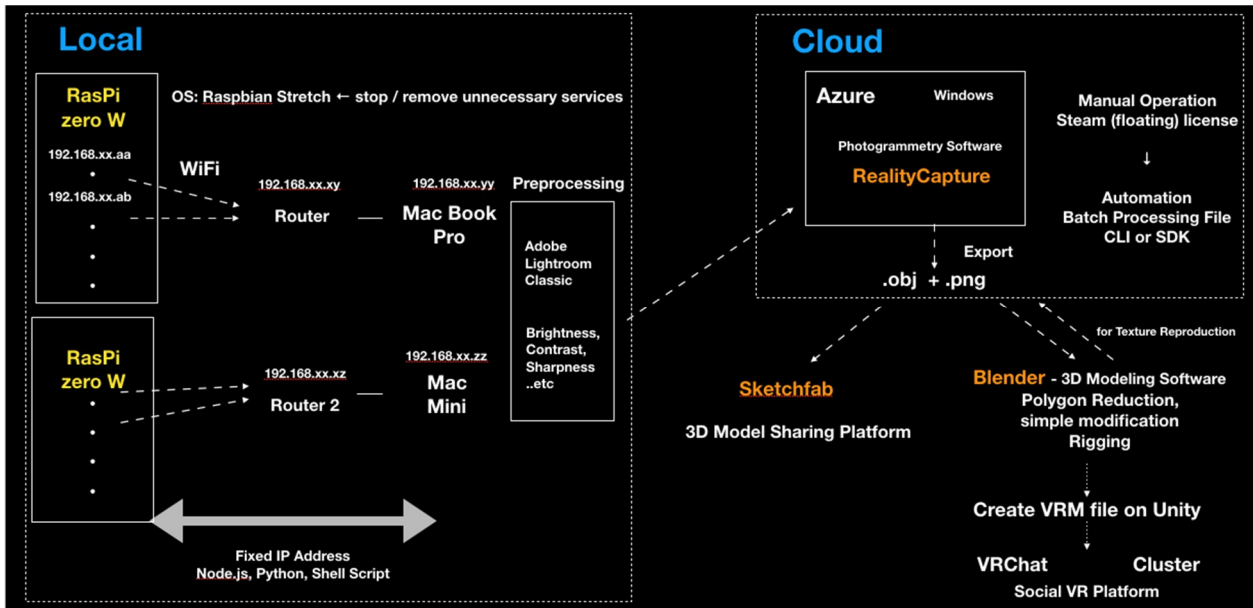


Fig. 5. System configuration and flow

We use script programs for taking and retrieving images. It is open source Node.js based programs [4, 5]. Fundamentally, it uses standard command “raspistill” of Raspberry Pi camera module.

Some modifications have been made to the original code for simplification and image file management. Simple Python scripts and shell scripts are also used for checking the status and stopping Raspberry Pi.

OS: Raspbian Stretch is installed on each micro SD card and unnecessary services are stopped and removed. Fixed IP address is assigned for each Raspberry Pi, Mac Book Pro, Mac mini and routers.

Currently, the lighting condition in the studio is not perfect so that we use Adobe Lightroom Classic [6] for increasing gain, sharpness and contrast to some degree.

We use RealityCapture, photogrammetry software, to create 3D Model on Microsoft Azure [7] virtual machine, Standard NC12_Promo, 12vcpu, 112GiB memory which has NVIDIA GPU [8] Tesla K80.

5. Operation overview

5.1. Taking Images

We take images of 88 camera modules. Each camera module has a 8 megapixel image sensor.

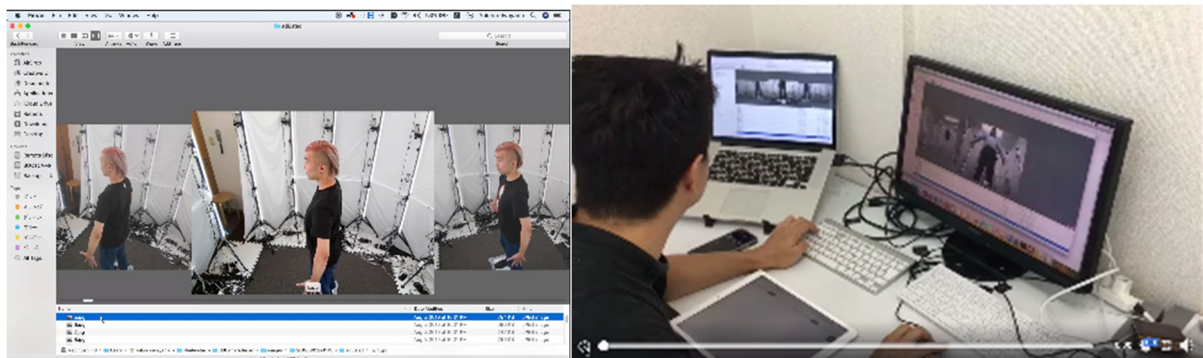


Fig. 6. Checking images

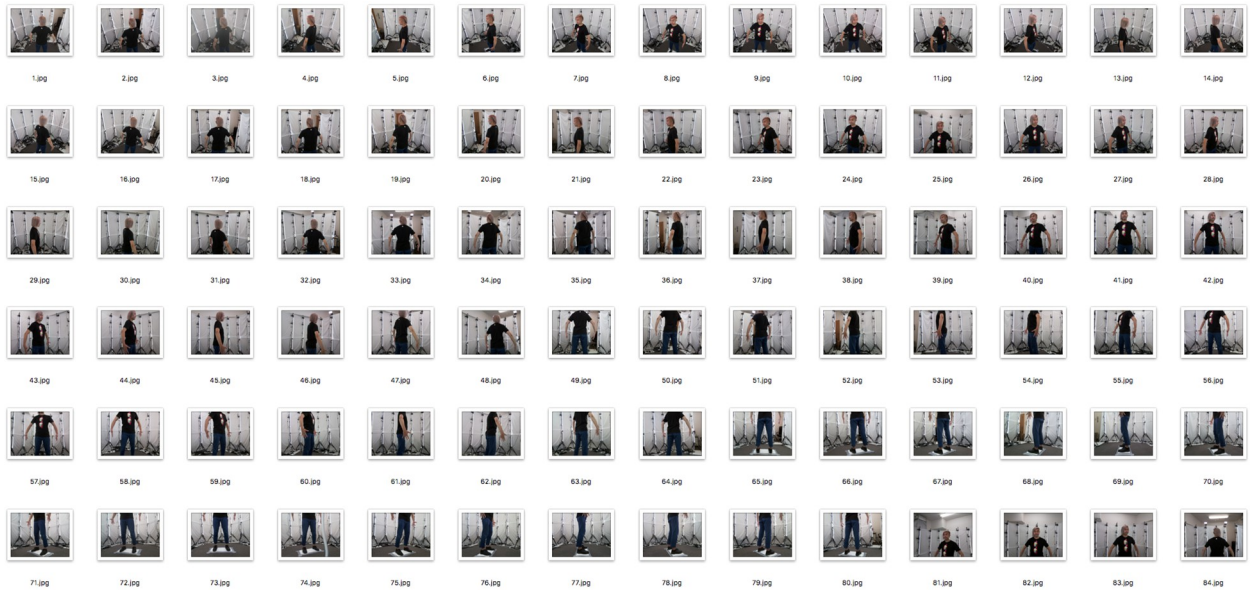


Fig. 7. Taken images

5.2. RealityCapture Operation

Import, Alignment, Reconstruction, unwrap and texture creation. It is standard operation.

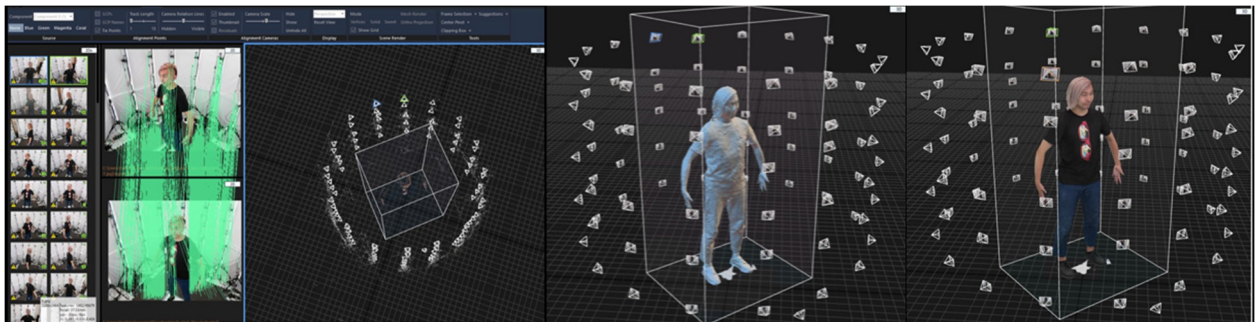


Fig. 8. Operation in RealityCapture

We export resulting .obj file and .png file and then we use Blender, 3D modeling software, to cut the bottom stand and unnecessary meshes between fingers, crotch, .etc.

6. Results / 3D Model



Fig. 9. Resulting 3D model. See in full at <https://skfb.ly/6MPZS>

We often use Sketchfab [9] for quick view and conveniently “Upload to Sketchfab” function is in RealityCapture. Overall quality of 3D model is good enough for current casual VR/AR apps.

7. Application

7.1. Rigging



Fig. 10. AR business card, Walking in the house, Dancing in the shrine and Real-time maneuvering

Once we rig the 3D Model, we can make various animations which are provided by third parties though you need to have some basic skills to achieve it.

We use mixamo [10] for simple rigging and then modify it in 3D modeling software Blender as needed.

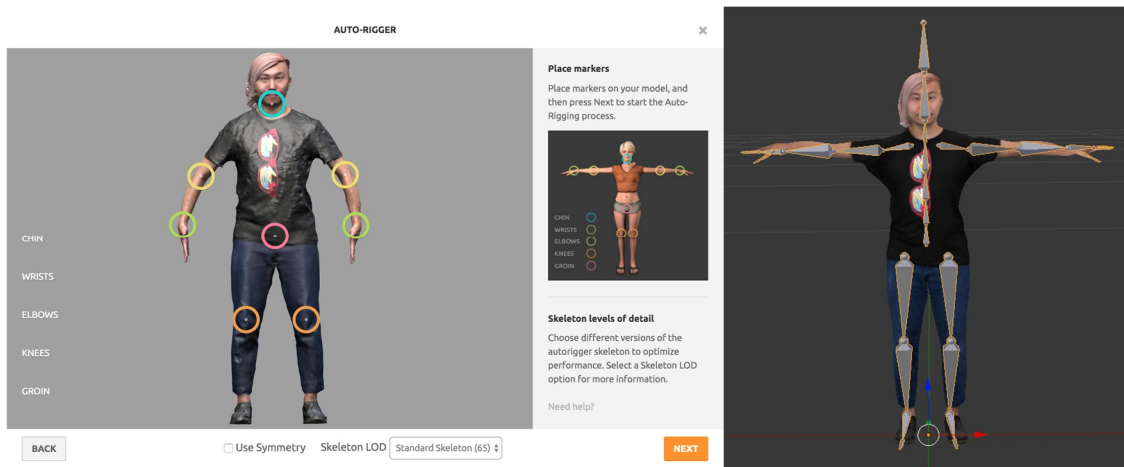


Fig. 11. Using mixamo and adjusting the rigging in Blender.

We are currently researching and developing other auto-rigging system to eliminate manual operation.

7.2. VRM

"VRM" is a file format for using 3D humanoid avatars (and models) in VR applications. VRM is based on glTF2.0 [11]. It is intended to be used in cross platform so we can use the same avatar across all supported applications.

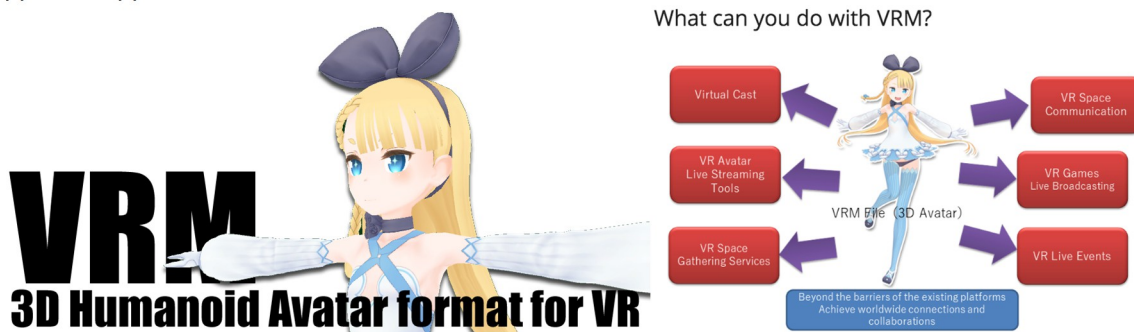


Fig. 12. VRM and What can you do with VRM? [12]

This VRM format is mainly developed and proposed in Japanese companies which later incorporated VRM consortium [13]. There are many services and platforms which have already adopted VRM such as VRoidHub [14], VirtualCast [15], Cluster [16], etc. It is easy to create VRM file if you have humanoid 3D model files (.fbx + .png .etc) by using UniVRM and also we can implement VRM uploading and importing program in VR/AR apps and services. Though it is an early phase, we believe VRM format can be widely adopted in other VR/AR platforms and services such as VRChat [17], AltspaceVR[18], .etc globally or it could be linked to Facebook Horizon [19] in the future.

8. Operation Detail

Posing

“A Pose”, the subject should pay attention to upright posture, open underarms, open fingers, foot position and orientation, eye sight direction and facial expression.

Shooting

Count of five and shoot on operator’s command. Remain still for 1 second in case a few cameras react milliseconds later. Even this rare delay doesn’t affect the quality in vast majority of the case.

Retrieving images

It takes about 2 to 3 minutes through Wifi 2.4GHz and routers. We could reduce the time through 5GHz using such as PLANEX GW-450D2 5GHz driver [20]. Wired Ethernet cable connection surely enables transferring images a lot faster but our original idea was to reduce cables and other parts as much as possible in order to build the system lightly.

Checking images

The subject may had been closing the eyes during the shot. In that case, we take another round.

Importing images in Adobe Lightroom Classic. Applying pre-set adjustment and Exporting

This process is done mainly because the lighting condition in the Asakusa studio is not perfect and the color of the face is a little dark in the original taken photo set. However, if we overdo it, it affects feature detection and reconstruction adversely in photogrammetry software process.

Compressing the files, transferring it to virtual machine on Microsoft Azure.

This should be done with script program through SSH connection but currently we manually transfer the compressed file through online storage.

A Note for Microsoft Azure

When we do 3D scanning in a continuous manner for multiple people, we think that cloud-based system, where you can create multiple virtual machines easily, would become necessary for speedy parallel production. However, needless to say, everything can be done in any PC with recommended GPU configuration, CUDA 2.0+ is required for RealityCapture.

Capturing Reality Inc., the provider of RealityCapture, has recently introduced PPI license. This license also supports using virtual machine setup.

Processing in RealityCapture

We change some of default parameters but in most cases, the default parameter settings are good enough. Notably in our studio, we use custom texel size setting in unwrap parameters and manually changing texel size to achieve better texture quality and utilization.

Modifying 3D model

We use Blender to cut the white stand at the bottom and make dent in the crotch part because unnecessary mesh is often created due to the shortage of images around the area. We also cut the unnecessarily meshes between fingers. Often times, we reduce the number of polygons appropriately to ex. 32,000 and reimport the modified mesh into the RealityCapture and re-texture it to achieve better texture file.

9. Challenges

9.1. Camera Focus Manual Adjustment

One of the most important things in photogrammetry is how sharp the images are. Raspberry Pi Camera module doesn’t have auto-focus function.

We needed to set focus manually by rotating the small camera part very delicately. The small amount of change such as 10-degree rotation shifts the point of focus rather a lot. Depending on distributors and timing, some of the units are already almost pan-focus though.

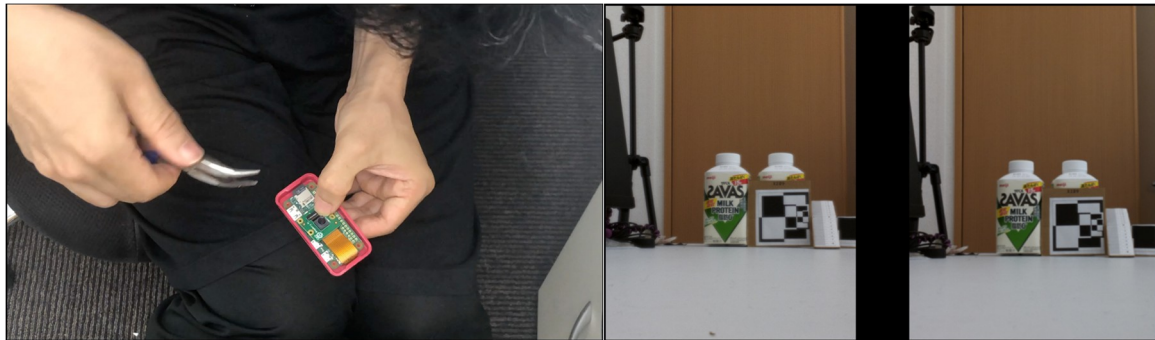


Fig. 13. Manual Focus Adjustment. See in full <https://youtu.be/bHqP4CBoJxg>

We may need to consider using different camera module which has auto-focus function with different circuit board rather than Raspberry Pi in the future.

9.2. Challenge in Feature Point Detection and Reconstruction

There are some difficulties to reconstruct things that has following features.

Shiny / reflective material (ex. Shine shoes), Thin or frill clothe, Smooth black material with no pattern, etc.

We below show the reconstruction result of shoes. The left is one of the images taken and the right is the one view of reconstructed 3D model.



Fig. 14. The surface mesh of the black shiny shoes isn't reconstructed properly

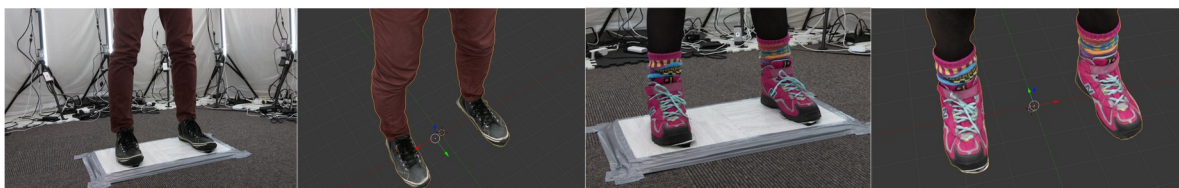


Fig. 15. The reconstruction example of shoes which are not shiny. A part of the slight artifact is due to the smaller number of images, which covered the area visibly, compared to the ones for the upper body.

One way of overcoming this reconstruction challenge had been proposed and implemented in other 3D scanner products which incorporate projector. One example incorporates quick two-step shooting. One take is for geometry reconstruction while grid pattern is being projected for the subject and the other take is for texture creation [21]. Currently, we haven't implemented such projection function in our system but we want to implement it in the future.

9.3. Raspberry Pi Zero W is not perfectly stable

Raspberry Pi Zero W occasionally doesn't start on initial power supply or refuse to connect due to instable power supply or other reasons. We need to reboot or in some cases need to reinstall the OS and related programs from the backup OS image file. We have tasks to simplify the system and programs further.

9.4. Future Research and Development Area

It resides mainly in the area of real-time three-dimensional capturing and mapping of human motion.

- Mapping of Body Tracking data using such as Azure Kinect [22] to the Real Avatar
- Applying OpenPose [23] library to the image set and the Real Avatar

Point cloud based hologram created by using multiple Azure Kinect is very prospective.

There is another approach in which precise meshed Real Avatar which is created separately and rigged, and the Real Avatar is maneuvered according to body parts coordinates information being tracked real-time in image processing system. It is also possible of real-time facial expression mapping / transfer if the face is retopologized properly in such as Wrap3 [24] and the movement is captured by using such as Rokoko Face Capture [25].

10. Summary

We developed Raspberry Pi Zero W based low-cost full body 3D scan system and we are currently operating it as small business in Asakusa, Tokyo.

References

- [1] Raspberry Pi Zero W <https://www.raspberrypi.org/products/raspberry-pi-zero-w/>
- [2] Richard Garsthagen, "An Open Source, Low-Cost, Multi Camera Full-Body 3D Scanner", 5th International Conference on 3D Body Scanning Technologies, <http://dx.doi.org/10.15221/14.174>
- [3] RealityCapture <https://www.capturingreality.com/>
- [4] 3D Camera Server <https://github.com/ArthurGuy/3dCameraServer>
- [5] 3D Camera <https://github.com/ArthurGuy/3dCamera>
- [6] Adobe Lightroom Classic <https://www.adobe.com/products/photoshop-lightroom-classic.html>
- [7] Microsoft Azure <https://azure.microsoft.com>
- [8] NVIDIA GPUs for Azure <https://docs.microsoft.com/en-us/azure/virtual-machines/windows/sizes-gpu>
- [9] Sketchfab <https://sketchfab.com>
- [10] Adobe mixamo <https://www.mixamo.com/#/>
- [11] glTF2.0 <https://github.com/KhronosGroup/glTF/blob/master/README.md>
- [12] What can you do with VRM? <https://vrm.dev/en/>
- [13] VRoid Hub <https://hub.vroid.com/>
- [14] Virtual Cast <https://virtualcast.jp/>
- [15] Cluster <https://cluster.mu/>
- [16] VRChat <https://www.vrchat.net/>
- [17] AltspaceVR <https://altvr.com/>
- [18] Facebook Horizon <https://www.oculus.com/facebookhorizon/?>
- [19] VRM Consortium <https://vrm-consortium.org/>
- [20] PLANEX GW-450D2 <https://www.planex.co.jp/products/gw-450d2/>
- [21] https://www.3dbodyscanning.org/cap/papers/2014/14174_45garsthagen.pdf, Figure 10
- [22] Azure Kinect DK <https://azure.microsoft.com/en-us/services/kinect-dk/>
- [23] OpenPose <https://github.com/CMU-Perceptual-Computing-Lab/openpose>
- [24] R3DS WRAP <https://www.russian3dscanner.com/>
- [25] ROKOKO Face Capture <https://www.rokoko.com>