White Paper

A 3D Volumetric VOD Capture and Stream Solution for Public Cloud

How We Implemented an End to End Solution Using AWS

Authors: Anna Qingfeng Li, William Cheung, Dusty Robbins, Tomasz Frankowski, Tian Jun, Chimay Tyagi, Brandon Gavino, Gang Shen – Intel

Contributors: Ben Nunez, Sebastian Marino - Evercoast

Introduction

Content and communications are evolving from 2D to 3D; volumetric technologies are at the forefront of that transition. A wide range of 3D usages are emerging from healthcare, e-commerce, education, telepresence and extensively in the media & entertainment industry; broadcasting, journalism, music, sports, live entertainment and 'live action' storytelling. To date, creators have operated within the confines of a camera lens, a two-hundred-year-old technology that gives readers a single-view of 2D representation. Volumetric Video allows us to give viewers the opportunity to experience content as they'd like to see it. Recent availability of low-cost and effective computer-vision sensors, advancements in machine learning, and the international rollout of 5G mobile and 10G broadband networks are converging to make end-end 3D volumetric video experiences achievable at scale.

Contents

At IBC 2019, Intel implemented a volumetric VOD tiling and delivery framework that delivers high quality volumetric video content to mobile devices in AR environments over 5G bandwidth [1]. In collaboration with Evercoast, Intel is extending the pipeline with capture and processing capabilities over on-premise and public cloud infrastructure respectively.

When people think of volumetric content, they typically think of point cloud video, requiring Gbps of bandwidth to stream. Any streaming solution that requires such a high network bandwidth is not scalable & deployable economically. Evercoast has successfully built a scalable commercial capture solution with commodity depth cameras Realsense D415[6] and a scalable cloud-based processing service. Augmented with Intel tiling, open sourced 3D compression, and streaming technologies, we can deliver a scalable commercial solution for volumetric usage.

1. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
2. Configurations: [Listed above in the Implementation and Test Setup section, 08/25/2020].
3. For more information go to www.intel.com/benchmarks

Volumetric Video Usages: A New Dimension for Video

Volumetric video adds an incredible capability for content creators in 2D and 3D consumption. The methods to capture and stream this content is similar, but now provides an extra dimension (z) that's part of the source capture, giving full depth and free viewpoint control. These benefits include:

CAPTURE

- Camera Control: Control can be a "virtual camera" handled by the creative director, and/or end user, to simulate the movements of a physical camera.
- 3D Worlds: Captured 3D characters can be placed in any 3D scene or backdrop.
- 3D Visual Effects: The capture sequence can be augmented with visual effects, such as lighting and shading.

OUTPUT

- Standard 2D Video: The 3D content can be encoded as a 2D video stream viewable on any device and platform.
- 3D World on a 2D Screen: Users can interact with the 3D content seamlessly on devices today.
- Mobile AR: Volumetric video can be placed into any real world AR scene viewed through a mobile camera view.
- VR/AR Headsets: Volumetric video easily integrates into Apple, Facebook, and Google VR/AR solutions today.
- Volumetric Displays: Volumetric display like the Looking Glass allow viewing in real 3D on our walls or desks.

USAGES

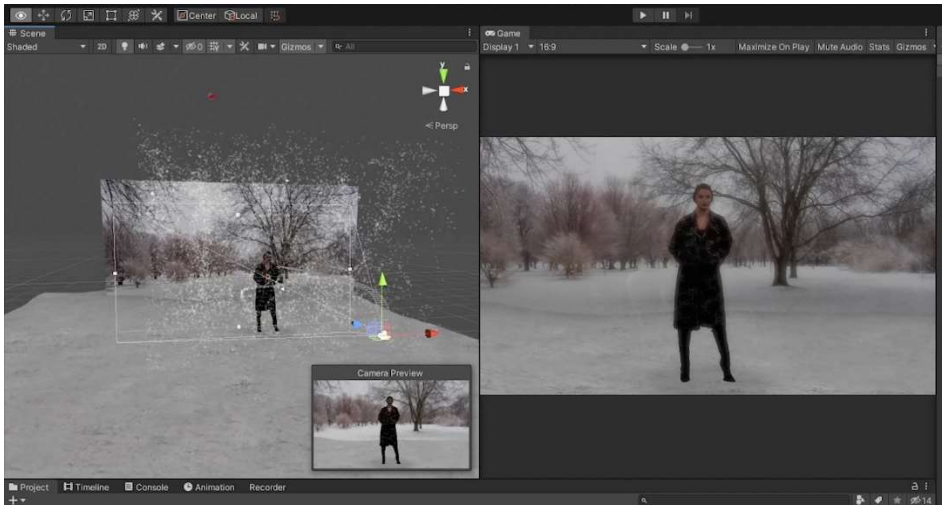
The following are existing or in exploration examples of multi-platform usages for volumetric content:

- Film: A huge, storied movie studio capturing their entire film costume archive (1,000s of pieces) using live actors.
- Music / Festivals: Major venue operators using volumetric to place artists in scenes beyond their backyards.
- Sports: Use of volumetric capture at team training facilities and stadiums and for fan content.
- News / Journalism: Reporters and journalists can integrate and interact with scenes on air and for production.
- Conference Keynotes + Talks: Capture executives for a virtual 3D "stage" for announcements & product release.
- Influencers: Volumetric fits well for social media scalable content and provides great usages for marketing.
- E-sports: eSports teams looking at volumetric for live broadcast.

For examples of the Evercoast's current VOD content, please refer to www.vimeo.com/volumetric

1. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
2. Configurations: [Listed above in the Implementation and Test Setup section, 08/25/2020].
3. For more information go to www.intel.com/benchmarks

Evercoast's post-processed / on-demand output: high-resolution textured meshes



Evercoast's Volumetric Video in a Blended Scene

Volumetric VOD Streaming Solution Implementation on the Public Cloud

Intel approaches volumetric VOD streaming with priority in scalability and deployability on existing content delivery technology. Consequently, Intel's solution chooses to stream the actual 3D assets over any CDN to the client devices [1] for local rendering. Unlike remote rendering approach which requires higher compute requirements per user in the edge cloud, this solution has much higher scalability for better TCO and lower motion-to-photon latency for improved user experience. Volumetric Streaming

1. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
2. Configurations: [Listed above in the Implementation and Test Setup section, 08/25/2020].
3. For more information go to www.intel.com/benchmarks

has many challenges with respect to bandwidth and latency even on a local network. Our volumetric VOD framework uses Intel Tiling SDK and standard compression to reduce the network bandwidth consumption by more than 15x to below 50Mbps, well within 5G connectivity range.

The tiling SDK will break a 3D asset into 3D tiles frame by frame, and then we apply Draco [4] compression to the 3D tiles and FFmpeg[5] compression to the texture content per frame. This tiled and compressed content can then be sent to a NGINX server for distribution. From the NGINX server, client devices trying to access the content will send its viewport/view information to the NGINX server. The NGINX server performs the visibility test to determine which tiles are entailed in the viewport and then pushes them to the client device.

The client device then receives the set of 3D tiles from the CDN, runs the Draco 3D asset decode and runs the media decoder on the client device for texture decoding. This forms the 3D content that will be rendered by the 3D GFX engine on the client device itself. We can implement the render both through openGL/DirectX directly or leverage render engines like Unity to generate this final image. Because this final image is rendered on the device itself, the VR/AR motion to photon latency, is only limited by the device's rendering capability. The assets we are testing use 2K-4K color textures and 20-30K vertices per human subject per frame, which most mobile devices are more than capable of rendering today.

To simplify the deployment of the Intel tiling and compression framework, Intel has dockerized its solution to simplify integration with the Evercoast volumetric video creation pipeline on public cloud. Intel's solution will run on any generic Linux server configuration. For the content distribution, the reference solution uses the standard open source NGINX server, but it can be replaced by any commercial CDN solution. On the client side, a reference client application is provided to demonstrate how the intel tiling SDK, Draco 3D decompression, and texture decoding can be integrated with any Unity or iOS client application framework. We currently utilize ARkit to demonstrate volumetric video in an AR environment, however, we can also easily port this solution via Unity to other client AR environments or devices.

Volumetric VOD Capture and Stream End to End Architecture

Evercoast and Intel's components provided here are modular and can be provided as a separate or integrated pipeline. Evercoast's end to end solution has many benefits particularly in ease of use, scalability, cost and fast deployment as well as best in class quality, which will be discussed in further details in the next section.

1. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
2. Configurations: [Listed above in the Implementation and Test Setup section, 08/25/2020].
3. For more information go to www.intel.com/benchmarks

Evercoast Volumetric Video Creation

The features and benefits of the Evercoast creation solution include:

- **Easy Operation:** Evercoast deliver easy to use software with hardware deployment options that can fit in a conference room or even on a desk.
- **Scalable:** From 2 cameras on a laptop or next unit of computing (NUC) to 20+ cameras on a single high-powered PC, configurations and options can be specialized to meet customer demand.
- **Cost-Effective:** Use of low-cost off-the-shelf hardware positions Evercoast at a 10x installation cost advantage, with an even greater cost advantage for rendering operations.
- **Live Broadcast Streaming:** The only system to provide real-time live broadcast streaming of volumetric content in multiple formats from volumetric domains to textured meshes.
- **Fast Cloud-based Turnaround:** Thousands of AWS instances are simultaneously deployed for rendering. The post-production process is frame-parallel such that a core-per-frame model renders content, regardless of duration, in about an hour from start to finish.
- **Truly Portable:** Evercoast's system is portable in a practical manner. It has been specifically designed, including customized cases that adhere to both size and weight limits, so that it can be transported as checked luggage on commercial flight.
- **Best in Class Aesthetic:** Evercoast's post-process rendering is a close analog to high-cost apex quality volumetric systems. When considering final deliverable assets, with geometric and texture compression, the differences are difficult to perceive.

The picture below is a typical 20 camera Evercoast capture system layout [2]

20 camera Evercoast Mavericks System deployed in a 12' x 12' conference room in Los Angeles, CA

Cloudbreak on AWS: Seamless Post-production on the Cloud

Recorded capture is uploaded to Cloudbreak for 3D post-processing and streaming. Evercoast's powerful, fast, high-resolution volumetric rendering in the cloud offers a fully automated, easy-to-use web interface to launch renders using farms of servers in the cloud to crunch through pipeline jobs. Cloudbreak dynamically manages a render farm across thousands of cores on AWS

1. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
2. Configurations: [Listed above in the Implementation and Test Setup section, 08/25/2020].
3. For more information go to www.intel.com/benchmarks

coupled with a WebRTC streaming platform to deliver rendered assets. Evercoast Cloudbreak on AWS offers the highest level of scalability with endless compute power, storage, and bandwidth to render captured content in minutes. Cloudbreak offers pay-as-you-go rates as well as block pricing and is priced low to enable creativity and iteration with volumetric content while providing an advanced post-processing 3D pipeline that generates high-quality renders.

Cloudbreak Edge: Post-production on the Edge for On-prem and Network Edge

Evercoast Cloudbreak Edge enables network operators, data center providers, and companies with in-house server farms to run Evercoast's volumetric rendering pipeline for internal use or as a service. Edge computing can also enable real-time rendering and the use of micro compute devices on-prem. Cloudbreak Edge offers reduced pricing, passing on savings from eliminating public cloud fees to leverage existing capital investments.

- 3D volumetric asset management
- Browser-based interface
- Runs entirely on AWS
- Cost-efficient use of EC2
- Queuing system to control load
- Immediately stream rendered output
- Manage live streams from Mavericks 2.0
- Store, search and secure volumetric data

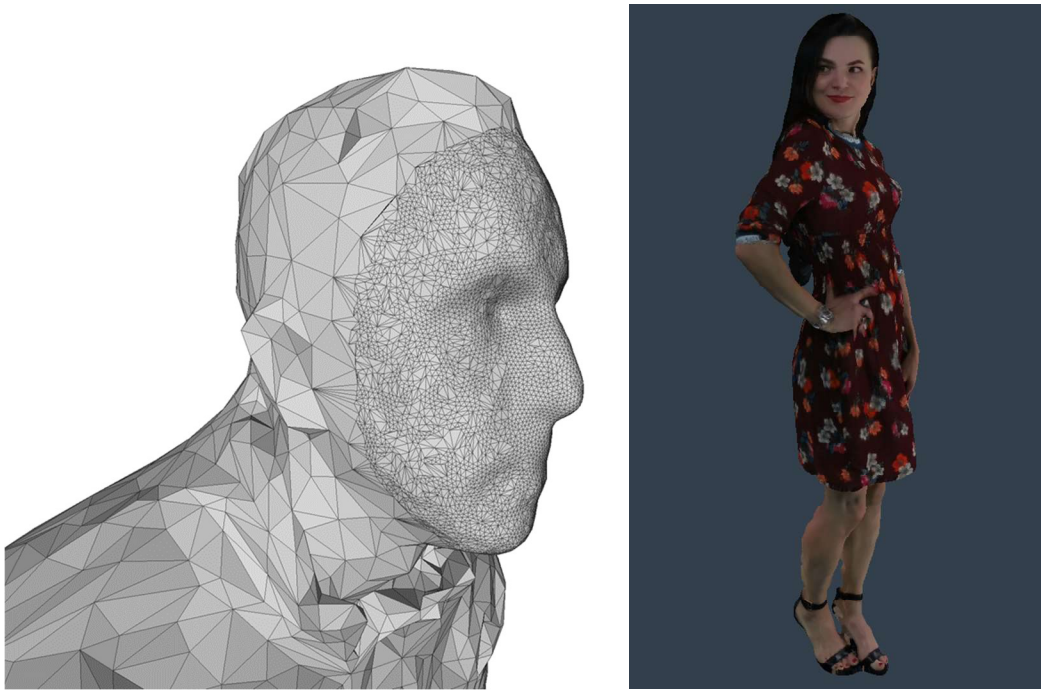
Advanced 3D Pipeline: Fully Rendered 3D Assets Loadable into any Rendering Framework or Application

Evercoast's industry-leading post-processing is comprised of 20+ proprietary applications that can logically be broken into four phases:

- 2D image processing (Color & Depth)
- Point cloud processing
- Mesh processing
- Texturing

Refined details in the pipeline allow for further compression possibilities by using things like face detection to optimize resolution in both the mesh and texture by preserving quality in important regions like the face while decimating other portions of the body that don't necessitate high levels of detail under most circumstances.

1. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
2. Configurations: [Listed above in the Implementation and Test Setup section, 08/25/2020].
3. For more information go to www.intel.com/benchmarks



Examples of Evercoast's Mesh Output and Volumetric Video Final Output

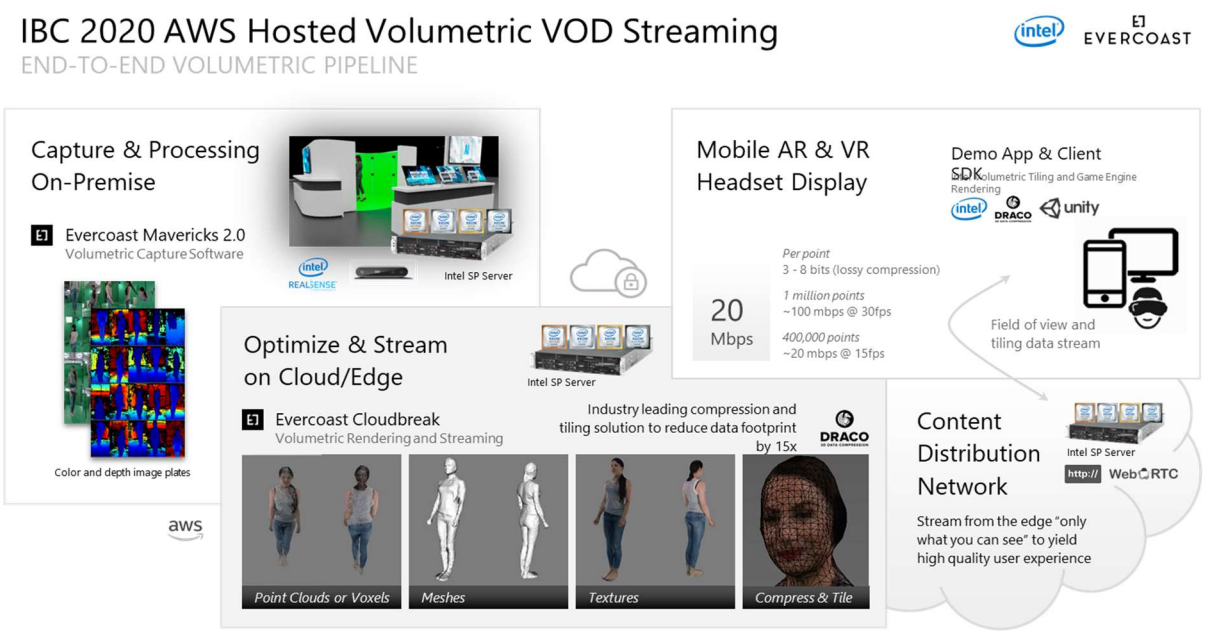
Implementation and Test Setup

Implementation

The Intel Volumetric VOD Streaming pipeline is integrated with the volumetric video capture and video generation from Evercoast as follows. Once a 3D asset becomes available from the Evercoast volumetric video generation on AWS, the Intel Tiling SDK takes the content and generates 3D tiles from the content frame by frame all on AWS. After the tiling SDK generates the tiles, we apply Draco [4] compression to the 3D tiles and FFmpeg [5] compression to the texture content per frame. This tiled and compressed content can then be sent to either a CDN distribution on a cloud server, a network edge, or a local network server for offices and enterprise usages. From that CDN network node, whether it be local or remote, client devices trying to access the content will send its viewport/view information to the network node. The network node performs the visibility test to determine which tiles are entailed in the viewport and then pushes them to the client device from the CDN. We have a windows client and an iOS client which leverages Unity for the final image render and display, which we currently render at 30 or 60FPS depending on the client device.

This VOD streaming solution has been tested on AWS and implemented to be integrated into Evercoast's AWS creation solution or other public cloud distribution solutions. The exact server configurations we used for both AWS and our local server as well as our client devices are outlined in detail in the test setup section.

1. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
2. Configurations: [Listed above in the Implementation and Test Setup section, 08/25/2020].
3. For more information go to www.intel.com/benchmarks



Public Cloud Deployed Volumetric VOD Streaming

Implementation and Testing Setup

Below are the system configurations we used to test our VOD streaming framework in both AWS and on a local server. The AWS instances are listed as below as well as the local server configurations we tested on and the clients we used. The measurements discussed in our results section below are focused on the Windows Client, but the iOS client should get similar results.

AWS Servers	**Tiling server (Tiling, Draco, Docker etc)**	**Webserver (NGINX server)**
OS	Ubuntu Server 18.04 LTS	Ubuntu Server 18.04 LTS
Instance Type	c5.4xlarge (8 core, 16GB), 10G NIC	c5.xlarge (2 core, GB), 10G N8IC
EBS Volume (standard SSD)	8G OS /dev/sda1, 64G additional /dev/sdb	8G OS /dev/sda1, 64G additional /dev/sdb
Local Server	**Server System**	
OS	Ubuntu Server 18.04 LTS	
CPU, Memory	Intel i7-8665U, 8G Memory	
Storage	235GB	
Client Hardware	**Windows**	**iOS iPad Pro**
OS	Windows 10 64-bit, 8 GB Ram	iPadOS
CPU	Intel i&-8665U	A10X
GPU	Intel UHD Graphics 620	PowerVR 7XT

Volumetric Public Cloud Streaming Results

Bandwidth Savings from Tiling

To quantify the amount of bandwidth savings we get from our Tiling and Viewport based Tile Streaming, we will compare the measured bandwidth via Wireshark [7]. We will compare our solution, where 3D tiles are downloaded based on the client's viewport information and downloading all the tiles. This is summarized in the table below in absolute bandwidth savings including mesh and textures. By using tiling and viewport-based tile streaming, we saved overall bandwidth by 31%. In this white paper we do

1. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
2. Configurations: [Listed above in the Implementation and Test Setup section, 08/25/2020].
3. For more information go to www.intel.com/benchmarks

not cover what the compression we were able to get with Draco because its already been covered in our previous white paper[1], which is included in the reference section.

	Total packets	Total bits per Second	Megabit per second
Download all Tiles	96247	43179733	43.18
Viewport Tile Streaming	72151	29864590	29.86
% of BW Improvement			**30.84**

Furthermore, we isolated the vertices bandwidth data in Wireshark to illustrate true benefit of tiling. Because we current do not applying our tiling technique to the textures, the texture bandwidth stays the same whether tiling is implemented. However, texture tiling is a feature that we intend to add in the future. When we isolate the vertices bandwidth, we see that on average we get about a 40% in bandwidth savings.

	Total packets	Total bits per Second	Megabit per second
Download all Tiles	58347	20359769	20.36
Viewport Based Tiles	33024	12281237	12.28
% of BW Improvement			**39.68**

Latency and Jitter Deltas between AWS Server and Local Server

In the following section, we measured the latency and jitter of our local server configuration vs a generic AWS configuration. Latency is defined as how long it takes to receive tiles. Jitter is defined as how many tiles are missing vs how many tiles are expected. Latency was expected to be significantly different between the local server and the AWS server. Although the max and min jitter is very different between AWS and local server, the average jitter is reasonably similar between the AWS configuration and the local server. We believe this jitter can be further reduced by employing local cloud services such as CloudFront and other edge server deployments.

Latency and Jitter on AWS vs Local Server		
Server Config	**AWS Server**	**Local Server**
Min Latency	26.7 ms	0.16 ms
Max Latency	290 ms	145.6 ms
Average Latency	52.9 ms	2.7 ms
Min Jitter = Tiles Missing/Tiles Expected	0%	0%
Max Jitter = Tiles Missing/Tiles Expected	45.5%	21.4%
Avg Jitter = Tiles Missing/Tiles Expected	1.8%	0.9%

1. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
2. Configurations: [Listed above in the Implementation and Test Setup section, 08/25/2020].
3. For more information go to www.intel.com/benchmarks

Summary

Volumetric video is at the forefront of the 2D to 3D consumption shift, with AR devices and applications becoming more pervasive, volumetric video gives users a unique window into immersive experiences that allow full movement and interactions. In our work, we have shown that high quality, best in class volumetric video can be captured, post-processed and streamed as VOD content to common place client devices over 5G connectivity. We showed that this pipeline is fully deployable over public cloud networks. It is our belief that this pipeline can be further enhanced with local cloud servers and set the next stage for volumetric video AR interactions. This whitepaper summarizes how Intel and Evercoast were able to build a high-quality volumetric video system that is scalable and deployable to 5G devices. To learn more about the Evercoast solution, please refer to www.evercoast.com. To learn about Intel's Visual Cloud solutions, including white papers, blogs, case studies and videos at www.intel.com/visualcloud.

1. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
2. Configurations: [Listed above in the Implementation and Test Setup section, 08/25/2020].
3. For more information go to www.intel.com/benchmarks

Glossary

AR	Augmented Reality
CDN	Content Delivery Network
Cloud	Cloud server
CPU	Central Processing Unit
Edge	Edge server
FPS	Frames Per Second
GHz	Giga Hertz
GPU	Graphics Processing Unit
Mesh	A set of triangles that are connected by the common edges
Texture	Surface texture or color information applied on a 3D generated model
VOD	Video On Demand
Vertex	A point where two more or curves meet
Volumetric Video	Video technique that enables capture and playback in 3D space

References

1. AQ Li, W Cheung, R Kawiak, D Robbins-Intel, M Chen, P Quesada-Comcast, T Tran, J Juang, CY Jiang-HypeVR, Enabling High Quality Volumetric VOD Streaming Over Broadband and 5G https://www.intel.com/content/dam/www/public/us/en/documents/white-papers/v2-volumetric-vod-streaming-whitepaper.pdf, Accessed 08/25/2020
2. Ben Nunez, Sebastian Marino, www.evercoast.com, Accessed 08/25/2020
3. Ben Nunez, Sebastian Marino, www.vimeo.com/volumetric, Accessed 08/25/2020
4. Google Inc. and other Authors, Draco 3D Data Compression, https://google.github.io/draco/, Accessed 09 Sept 2019
5. FFmpeg video and texture compression, https://ffmpeg.org/, Accessed 08/25/2020
6. Intel Inc., RealSense Camera https://www.intelrealsense.com/, Accessed 08/25/2020
7. Wireshark, https://www.wireshark.org/, Accessed 08/25/2020

1. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
2. Configurations: [Listed above in the Implementation and Test Setup section, 08/25/2020].
3. For more information go to www.intel.com/benchmarks

Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors.

Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.

Performance results are based on testing as of 08/25/2020 and may not reflect all publicly available security updates. See configuration disclosure for details. No product or component can be absolutely secure.

Configurations: [Listed above in the Implementation and Test Setup section, 08/25/2020].

Intel technologies' features and benefits depend on system configuration and may require enabled hardware, software or service activation. Performance varies depending on system configuration. Check with your system manufacturer or retailer or learn more at [intel.com].

No license (express or implied, by estoppel or otherwise) to any intellectual property rights is granted by this document.

Intel disclaims all express and implied warranties, including without limitation, the implied warranties of merchantability, fitness for a particular purpose, and non-infringement, as well as any warranty arising from course of performance, course of dealing, or usage in trade.

This document contains information on products, services and/or processes in development. All information provided here is subject to change without notice. Contact your Intel representative to obtain the latest forecast, schedule, specifications and roadmaps.

The products and services described may contain defects or errors known as errata which may cause deviations from published specifications. Current characterized errata are available on request. No product or component can be absolutely secure.

Copies of documents which have an order number and are referenced in this document may be obtained by calling 1-800-548-4725 or by visiting www.intel.com/design/literature.htm.

Intel, the Intel logo, are trademarks of Intel Corporation or its subsidiaries in the U.S. and/or other countries.

*Other names and brands may be claimed as the property of others.

© Intel Corporation.

1. Software and workloads used in performance tests may have been optimized for performance only on Intel microprocessors. Performance tests, such as SYSmark and MobileMark, are measured using specific computer systems, components, software, operations and functions. Any change to any of those factors may cause the results to vary. You should consult other information and performance tests to assist you in fully evaluating your contemplated purchases, including the performance of that product when combined with other products.
2. Configurations: [Listed above in the Implementation and Test Setup section, 08/25/2020].
3. For more information go to www.intel.com/benchmarks