Potential volumetric capture applications for GroupM clients

A Comprehensive Guide to Utilising Volumetric Capture

Company	**University**	**Student**
GroupM	Fontys	Jan Carlo Hendriks
Charlie Freestone	Eric Slaats	SN: 2366487

Table of contents

Version History

Date	Version	Author	Comment
2023-05-09	0.0	J.C. Hendriks	Initial document
2023-06-27	1.0	J.C. Hendriks	Updated abstract and conclusion with latest feedback.
2023-07-12	2.0	J.C. Hendriks	Added more in depth information and comparison matrix of my prototypes.
2023-08-09	3.0	J.C. Hendriks	Changed title, implemented latest feedback.

Abstract

This study aims to gain a comprehensive understanding of volumetric capture technology, its offerings, relevance, working principles, different techniques and different quality output. The goal is to make an informed decision on how to utilise the technology. The methods employed include a literature study, prototype development, and a comparison chart. Through this research, it is evident that volumetric capture is a highly relevant technology, predicted to become even more so as advancements continue. Volumetric capture enables the creation of digital assets that can be viewed from any angle in various environments, such as Web3D, augmented reality (AR), and virtual reality (VR), providing a realistic and immersive 3D experience. The technology allows for the scalability, duplication, and transformation of captured subjects, offering viewers a more engaging experience. The market potential for volumetric video is substantial, with estimations indicating it will approach a value of $5 billion by 2026 (Arcturus, 2022). Recommendations based on the findings suggest that volumetric capture should be taken seriously and closely monitored for further developments. As Extended Reality technologies like VR, AR, and the Metaverse gain prominence, volumetric capture will play a significant role. Investing in a high-quality volumetric capture studio can provide long-term value and production-ready digital assets, making it a valuable consideration for media companies like GroupM.

Research strategy

Main question	What is Volumetric Capture?
Sub Questions	● What is the offering of Volumetric Capture? ● Why is Volumetric Capture relevant? ● What are its benefits and considerations? ● How does Volumetric Capture work? ● What are the techniques for achieving Volumetric Capture? ● How can Volumetric Capture be categorised? ● How do the types of Volumetric Capture compare?
Methods	● Literature study ● Prototypes ● Comparison Chart

Volumetric Capture Explained

Volumetric capture refers to the process of recording a physical space, object, person, or event in a way that gives the impression of occupying three-dimensional space. It involves capturing the actual dimensions and shape of the subject, allowing viewers to observe it from various angles. In a volumetric captured image, viewers have the ability to "be the director" by freely looking and moving anywhere within the image, creating a heightened sense of presence.

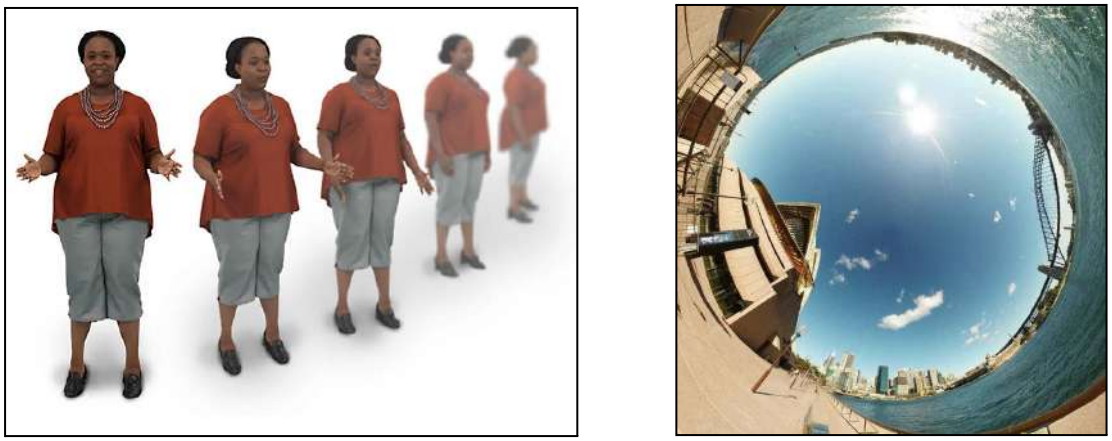
The result of volumetric capture is an image, video or hologram, which represents a genuine 3D space or performance. These volumetric assets can be viewed from any angle in a Web3D environment, augmented reality (AR) or virtual reality (VR), offering a highly realistic, immersive, and interactive 3D experience. Volumetric images and holograms combine the visual quality of traditional images with the unique immersion achieved through spatialized content.

Understanding The Difference Between Volumetric Capture & 360-Degree Images

When comparing volumetric capture and 360-degree images, it's important to note the fundamental difference between the two. While both offer unique experiences, they differ significantly in terms of depth and immersion.

A 360-degree image is captured using cameras that capture the entire 360-degree field of view. When viewers watch a 360-degree image, they have the ability to look in any direction within that sphere. However, the image lacks depth, resembling the experience of being inside a snow globe.

On the other hand, volumetric capture goes beyond the limitations of 360-degree image by providing depth and spatial awareness. With volumetric capture, viewers can perceive the three-dimensional nature of the captured content. It allows users to interact with the image as if they were a director, granting them control over the perspective and the ability to explore the scene by adjusting the viewing distance.



Images: Volumetric Capture (left), 360-degree (right)

Relevance

Volumetric capture allows for versatile use of recorded content, as captured individuals can be scaled, duplicated, and digitally transformed. This digital representation enables various applications in the enterprise sector, including employee training, education, customer service, corporate communications, product inspection, marketing, advertising, and brand recognition. The potential use cases are extensive.

One significant advantage of volumetric video is the immersive experience it offers to viewers. For example, rather than physically attending a concert or play, volumetric video brings the performance directly to the audience by creating photoreal holograms of the actors or musicians. This enables viewers to virtually sit next to them. Similarly, in video conferences, instead of interacting with a grid of talking heads on a screen, volumetric video technology allows individuals to stand in front of a full-body digital avatar of the person they are conversing with. They can even move around the avatar while the other person views a photorealistic digital representation. This immersive experience can be applied to learning scenarios as well, such as practising dance moves or receiving feedback from a photorealistic instructor while perfecting a golf swing.

The potential for volumetric video is vast, and significant financial investments are being made in this technology. By 2026, the volumetric video market is projected to approach the $5 billion mark (Arcturus, 2022), with the potential for even higher growth. These estimations highlight the substantial opportunities and market demand surrounding volumetric capture and its applications.

Benefits

Realism
Volumetric capture can create highly realistic 3D models that capture the true depth and detail of the subject.

Interactivity
These models can be used in virtual and augmented reality applications to create immersive and interactive experiences for users.

Flexibility
Volumetric capture can be used to create models of a wide range of subjects, including people, objects, and environments.

Efficiency
Compared to traditional 3D modelling techniques, volumetric capture can be faster and more efficient, especially when dealing with complex shapes or moving subjects.

Considerations

Costs
Volumetric capture can be expensive, particularly when using advanced cameras and sensors.

Storage and processing
The resulting 3D models can be large and require significant processing power to work with, which can be a challenge for some systems.

relevant data input, computing power taken care of by team at GroupM

Limited field of view
Depending on the technology used, volumetric capture may have a limited field of view, which can make it challenging to capture large or complex environments.

Lighting and environmental conditions
The accuracy and quality of the resulting 3D model can be affected by lighting and other environmental conditions during the capture process.

Workings of Volumetric capture

Volumetric capture is a process that involves the use of multiple cameras and sensors to record a subject from various angles, creating a three-dimensional representation of the subject rather than a flat image.

Once the video capture is completed, the scene is processed to generate a sequence of 3D models. Computer algorithms are employed to stitch together the captured views, combining them into a cohesive model. These models undergo further processing, such as unwrapping the meshes and generating textures. Finally, the entire volumetric image is compressed into a data file that is ready for viewing.

The resulting image can be viewed from any angle, providing a realistic sense of depth, colour, and lighting. It is compatible with various platforms, including mobile devices.

Example of volumetric video being used by ANAYI

Techniques for achieving Volumetric Capture

Converting real life subjects into 3D assets has been around for many years and it knows many techniques. Below is a list of them.

Traditional 3D Modelling
Manually recreates the shapes, textures, and details of the image in the 3D environment.
Software: *Blender, Maya, 3ds Max*

Photogrammetry
It involves capturing multiple photographs of an object or scene from different angles and then using specialised software to reconstruct a 3D model. The software analyses the images and calculates the shape and texture information based on the differences between the photos.
Software: *RealityCapture, Agisoft Metashape, Meshroom, , Autodesk ReCap*

Depth-sensing cameras
A depth camera also known as a depth sensor, is a type of camera engineered to analyse a scene to determine depth and distance of objects within the environment, from this image the camera can create a 3D map of the current scene. Types of depth-sensing cameras:

1. Structured Light: Observes distortions in projected pattern
2. Stereo Vision: Compares features in two stereo images
3. LiDAR (Light Detection and Ranging): Measures transit time of reflected light from an object.
4. ToF (Time-of-Flight): Measures transit time of reflected light from an object

NeRFs
Neural radiance fields (NeRFs) is a technique that generates 3D representations of an object or scene from 2D images by using advanced machine learning.
Software: *Luma AI, Nerf Studio*

Volumetric Studios
These work by using multiple cameras from different angles that are all filming at the same time. Then, computer algorithms stitch the views from these angles together to create volumetric images. A human performance can be filmed from all angles creating a three dimensional (3D) video, allowing the user to view any point of the performance from any angle.
Studios: *Volumetric Studio Kawasaki, Microsoft's Mixed Reality Capture Studios, 4DR Studios, Dimension Studio, Evercoast*

Categorising Volumetric Capture by Quality

Volumetric capture can be created in various ways, using different technologies such as those previously discussed. For creatives, the technologies are not as important as the results and creative possibilities that they can offer. To make it easier for GroupM in the decision tree at the moment of choosing an appropriate form of volumetric capture, I divided the available volumetric capture options into three categories: high-end capture, sensor-based capture, and mobile device capture.

High-End Capture

These are done in professional recording studios with systems that include various cameras positioned around the subjects, that capture the performance from every possible angle and convert these shots into holograms. Some studios, like Canon's Volumetric Studio Kawasaki or Microsoft's Mixed Reality Capture Studios, work with more than 100 cameras to capture the holograms. The results are incredibly realistic and the best out there.

Pros	Cons
Creates incredibly realistic holograms with outstanding quality.	Expensive due to technology and operation costs involved.
Offers good support for clients.	Physical presence at the studio may be required

Studios: Volumetric Studio Kawasaki, Microsoft's Mixed Reality Capture Studios, 4DR Studios, Dimension Studio, Evercoast

Sensor-based Capture

These are depth-sensor cameras, which have depth and motion sensing technology at their core, which project a known pattern onto the space and capture back the distortion caused by objects in the scene, which then can be computed into a volumetric video capture.

Pros	Cons
Affordable and there is a supportive community around this form of volumetric video capture.	Quality is less realistic than high-end studios.

Cameras: Xbox Kinect, Azure Kinect, Intel RealSense

Mobile-based Capture

When iPhone X was released, the built-in TrueDepth camera was utilised by developers as a mobile 3D scanner, opening space for what is today a form of volumetric video capture enhanced with a LiDAR sensor in newest iPhones. There are a few apps out there that allow any user to capture a hologram with their phones – no previous knowledge necessary. This opens the doors of 3D communication, as people are now able to record themselves in 3D and share their holograms with others in AR.

Volumetric capture techniques vary depending on the operating system. The new versions of Apple Iphones PRO come with a LiDAR sensor integrated. While high end android phones include a ToF sensor. Both operating systems allow the use of NerF applications.

Pros	Cons
Most accessible form of volumetric capture out there.	Single-camera capture may only generate one-sided holograms or back view generated by machine-learning, which isn't as good as reality.
Allows for user-generated content for AR and 3D communication.	

Apps: Record 3D (LiDAR), Luma AI (NerF), Polycam (LiDAR, NerF), Vologrames (LiDAR)

Comparing the categories

In the context of providing GroupM with additional knowledge on the available resources, I decided to test and compare the different types of volumetric capture techniques available.

Criteria

Criteria to take into consideration for producing volumetric capture with different devices for future projects:

Criteria	Description
Set up complexity	How difficult is it to get the hardware running and getting results? This is relevant to take into account in the planning for future projects. Set up complexity can be divided into easy, medium and difficult. The easier the better.
Output type	Does it capture static images or video? This helps take into account the sort of asset attainable. Output type can be divided in image and video. Video being the best.
Output quality	How realistic is the resulting capture? This helps in determining which technology is suitable for production. Output quality can be divided into low, medium, high. High being the best and low the worst.
Polishing	How much polishing does the result need in order to be production ready? This is relevant in order to take into consideration in planning for future projects. No polishing being the desired outcome.
Scanning size	Can it scan small, medium or big sized subjects? This helps take into account the sort of asset possible. The more sizes possible the better.
Pricing	How much does it cost? This helps take into account the moment of budgeting. The lower the price the better
Live streaming	Can it live-stream? This is relevant to take into consideration for aligning with the project objectives. Feature being present the best.
Support	Is there a support page or a community available to fix any problem encountered during production? This is relevant to know in case unforeseeable problems emerge during production. Any type of support is desirable.

Devices

The following devices were selected for study based on their availability so I could test the different categories of volumetric capture and compare them:

Mobile-based

- Polycam
- Volograms
- Luma AI

Sensor-based

- Kinect v2
- Kinect Azure

High-end

- Evercoast

Although each device is different and there are many, they all use the same technique based on their category. In the following page a more thorough analysis.

Mobile-based

- **Polycam**: easy setup, just download the app. Only produces static 3D images. The output quality is medium, the subject needs to stay still. The resulting assets need to be further polished after being captured. You can scan small or human size subjects. There is a free and a pro tier. It doesn't provide streaming services, but it has discord server support.
- **Volograms**: easy setup, just download the app. Produces 3D video holograms taking the whole body, but the back part of the subject is filled with random textures. Static images can be obtained from the video. The output quality is medium, the subject can perform movement. The resulting assets do not need to be further polished after being captured. You can scan small or human size subjects. There is a free and a pro tier. It doesn't provide streaming services, but it has discord server support.
- **Luma AI**: easy setup, just upload any video captured with your phone. Only produces static 3D images. The output quality is medium, the subject needs to stay still. The resulting assets do not need to be further polished after being captured. You can scan small or human size subjects. There is a free and a pro tier. It doesn't provide streaming services, but it has discord server support.

Polycam

Volograms

Luma AI

Sensor-based

- **Kinect v2**: reasonably easy setup, requires the installation of the SDK and other software. It can produce static full body 3D images and 3D video but not full body. The output quality is between small and medium, it depends on the handyness of the person scanning, in which case the subject needs to stay still. The resulting assets do not need to be further polished after being captured. You can primarily scan human size subjects. Buying the hardware can only buy second hand, and it would cost less than €100. It doesn't provide streaming services, and any support has been discontinued since the device is officially deprecated.
- **Kinect Azure**: in comparison to its predecessor, this one is not that easy to get running. Reason being because this device is targeted towards developers who have an understanding of coding. To get it running you will need to install the SDK and use the Command Prompt to execute functions, or write your own program in Visual Studio. Once setup is completed, capturing 3D images or 3D video works really easy. The quality is medium, better than its predecessor. The resulting assets do not necessarily need to be further polished after being captured, this depends on the ability to make a good capture. Its current price tag amounts $399.00. It does not offer streaming services out of the box, although a program for that could be developed. It offers support by Contact form, Stack Overflow, Twitter and the MICROSOFT Q&A Azure community.

Kinect v2

Kinect Azure

High-end

- **Evercoast***: it has the most complicated setup since it requires a whole room with lots of cameras, and also specialised personnel to operate the system. It can output incredible high quality 3D images and 3D video. It can scan everything that fits into the room- - small objects, humans and more. Price ranges depending on how it is used. Renting 1 day shoot at a Partner Studios amounts $7,500. While building your own studio amounts to USD $39,000.

Evercoast

Comparison Matrix

Criteria	Types					
	Mobile Device			Sensor-based		High-end
	Polycam	Volograms	Luma AI	Kinect v2	Kinect Azure	Evercoast*
Setup complexity	easy	easy	easy	easy	medium	high
Output type	static	video	static	static/video	static/video	static/video
Output quality	medium	medium	medium	low	medium	high
Polishing	Yes	No	No	No	No	No
Scanning size	small, human size	small, human size	small, human size	human size	human size	small, human size, big
Pricing	Free	Free	Free	< EUR €100	USD $399.00	> USD $7,500
Live streaming	No	No	No	No	No	Yes
Support	Discord server	Discord server	Discord server	Deprecated	Many	Reviews

Legend: Does not suffice criteria, merely suffices criteria, excels criteria

** During the project the Evercoast Volumetric Studio was not available for me to test, yet I found it useful to compare it with the other technologies to get the whole picture. The information regarding evercoast can be found on their website.*

Comparison Results

There are many ways for making volumetric captures, each with their up and down sides. Depending on the situation and the requirements of a media project, an appropriate technique might be chosen.

For quick and low cost solutions, mobile devices prove to be a viable solution, most apps will only produce static assets, but Vologrames takes it a step further by creating 3D holograms. The only downside is that these 3D holograms do not record the full body, only what is in front of the camera. Another thing to take into consideration is static 3D asset generation requires the subject to stay still, which might make it difficult to record faces or living creatures. This option would mainly be preferable for amateur use or for capturing of small 3D assets to combine with other bigger 3D environments. Think of video games, or social media marketing where low resolution is acceptable.

Depth sensor cameras, such as Kinect v2 or Kinect Azure might take less time to use since the camera is meant for capturing humans. In terms of quality, the Kinect Azure creates better results since it is also the most modern version of the Kinect. The Kinect Azure also lets you capture 3D video out of the box once properly set up. The down side of both is that some degree of technical knowledge is required. The Kinect v2 requires installing the SDK and other software in order to obtain results, but a regular user might be able to accomplish that. The Kinect Azure on the other hand requires the use of the Command Prompt, which might be frightening for the regular user. The Azure Kinect is intended for developers. It is also possible to arrange multiple Kinect Azure cameras around a subject in order to get a full body hologram. All and all, depending on its intended use, a depth sensor based solution might be helpful where installation doesn’t need to be repeated, but where gear is planned to be used many more times. A direct use might be 3D video calls or teleconferences.

Lastly, high end volumetric capture studios provide the best quality overall. It can capture full body and in 3D video, and of course just static subjects as well. Out of all the options, this technique has the highest capture range, being able to capture a whole body at once just like you would shoot a simple photograph, needs to reduce the production time by a lot. The downside is obviously its cost. It is by far the most expensive solution out there and it is not for any entry level hobbyist. It also requires specialised personnel to operate the machine. Apart from the cost, it is the most versatile solution, as it can create both static and video assets with the same high quality all at once. It also has the broadest capture range, as it can capture small, medium and big subjects, everything that fits in the volumetric capture studio. This option would be more suitable for professional production, where all kinds of assets need to be created at a fast pace for entities that have the financial capacity. Think of media campaigns and branding.

All and all, choosing the appropriate technique depends on its intended use and depends on the user to analyse the cost effectiveness required. Mobile devices and high end capture might be two of the most common appropriate solutions, as they could be easily combined. High-end provides overall the best quality and versatility for production.

Conclusion

By analysing the technology, its capabilities and its types, I have found that this technology is quite relevant, and it is also predicted to become even more so as it is only being improved over time.

Volumetric capture offers a unique feature which is to bring real life subjects to the digital space. These volumetric assets can be viewed from any angle in a Web3D environment, augmented reality (AR) or virtual reality (VR) and more, offering a highly realistic, immersive, and interactive 3D experience. Volumetric images and holograms combine the visual quality of traditional images with the unique immersion achieved through spatialized content.

Capture a person once and that video can be used in many different ways. The person can be scaled, duplicated, and even transformed since they are a digital capture of their real selves. At its core, volumetric video offers viewers a more immersive experience

The possibilities are endless, and there is serious money involved. By 2026, the volumetric video market is expected to approach the $5 billion mark (Arcturus, 2022), and those are likely conservative estimates.

All and all, choosing the appropriate technique for using volumetric capture depends on its intended use and depends on the user to analyse the cost effectiveness required. Mobile devices and high end capture might be two of the most common appropriate solutions, as they could be easily combined. High-end techniques provide overall the best quality and versatility for production.

Recommendations

Based on my knowledge, volumetric capture represents not only the next generation of 3D asset generation, but also a transformative force in our interaction with digital worlds. It is imperative to take it seriously and closely monitor its further developments. With the imminent rise of Extended Reality technologies like VR, AR, and the Metaverse, volumetric capture will assume a progressively significant role. While a standardised approach is yet to be solidified, media companies such as GroupM can gain considerable advantages by being pioneers in exploring the possibilities offered by this technology. Therefore, investing in its own high-quality volumetric capture studio would yield long-term value, as it enables the creation of versatile and production-ready digital assets.

References

Buy the Azure Kinect developer kit – Microsoft. (n.d.). Microsoft. Retrieved August 9, 2023, from https://www.microsoft.com/en-us/d/azure-kinect-dk/8pp5vxmd9nhq

Hackl, C. (2020, September 27). *What Is Volumetric Video And Why It Matters To The Enterprise*. Forbes. Retrieved August 9, 2023, from https://www.forbes.com/sites/cathyhackl/2020/09/27/what-is-volumetric-video--why-it-matters-to-the-enterprise/

Jaehnig, J. (2021, September 28). *What Is Volumetric Capture and How Does It Work?* MakeUseOf. Retrieved August 9, 2023, from https://www.makeuseof.com/what-is-volumetric-capture-and-how-does-it-work/

Kinect V2. (n.d.). Amazon.com. Retrieved August 9, 2023, from https://www.amazon.com/kinect-v2/s?k=kinect+v2

Pricing. (n.d.). Polycam. Retrieved August 9, 2023, from https://poly.cam/pricing

Types of Volumetric Capture in 2021. (2021, July 15). Sense of Space. Retrieved August 9, 2023, from https://www.senseofspace.io/blog/types-of-volumetric-capture-in-2021/

Volu - future of content creation. (n.d.). Vologrames. Retrieved August 9, 2023, from https://www.vologrames.com/volu

Volumetric capture. (n.d.). Wikipedia. Retrieved August 9, 2023, from https://en.wikipedia.org/wiki/Volumetric_capture

Volumetric Capture Production Services | Evercoast. (n.d.). Evercoast. Retrieved August 9, 2023, from https://evercoast.com/volumetric-capture-production-services/

What is volumetric video? A beginner's guide. (2022, April 5). Arcturus Studios. Retrieved May 10, 2023, from https://arcturus.studio/blog/2022/04/05/what-is-volumetric-video/