A User Study on Volumetric Video for Broadcasting, Metaverse Production, and RT3D Assets.

Gareth W. Young
School of Computer Science and Statistics
Trinity College Dublin
Dublin, Ireland.
YoungGa@TCD.ie

For this year's IBC, we've been exploring volumetric video for broadcasting, metaverse production, and creating real-time 3D assets. The objective of the user study on the Accelerator has been to examine the impact of our focus on creating, delivering, and applying Digital Humans on the viewing experiences of different audiences. To achieve this goal, I've been speaking to a variety of viewers to ascertain the potential of this technology to enhance media consumption practices.

The cohort with whom we engaged displayed a variety of technical expertise using current technologies in media entertainment but also self-described as being very good at using digital technology for consuming media and were very familiar with the different technologies used in the media entertainment domain.

- **22 participants; 16 male, 6 female**

- **Mean age of 36.13 (SD = 7.5)**

- **Ethnicity was mainly White (n = 17) and Asian (n = 5)**

- **Located in Ireland (n = 20) or Europe (n = 2)**

- **Educated to Irish National Framework of Qualification level 10 (n = 9), 9 (n = 9), and 8 (n = 5)**

How would you describe your ability to use digital technology for consuming media entertainment?
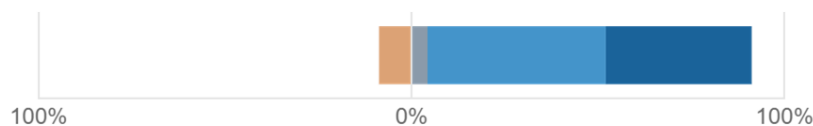
How would you rate your current knowledge of the different technologies used in media entertainment? (For example, do you watch content on your TV, laptop, tablet, and mobile phone?)
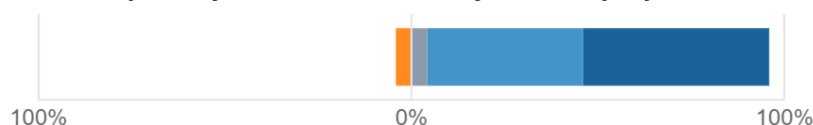
How would you rate your current knowledge of the different technologies used in media entertainment? (For example, do you watch content on your TV, laptop, tablet, and mobile phone?)
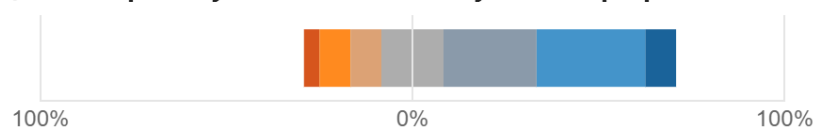
■ Bad ■ Very Poor ■ Poor ■ Average ■ Good ■ Very Good ■ Excellent

Q1. What did you like the most and the least about the content you just watched? Which did you prefer? Why?

Overall, the participants considered the 3D content more realistic; however, the 2D images were noted as being of higher quality and more stylized. At the same time, some participants couldn't discern between 2D and 3D content. Still, both formats were equally exciting and eye-catching, even for those who were "just not that into sports."

To begin with, it was felt that the 2D graphical overlays added more informational context to what was being presented, but both 2D and 3D were considered equally adequate for this task. At times, however, the 2D content dominated the screen to a level that threatened to overshadow the presenters, so there is a need for a carefully thought out scale and how the size of the virtual content can conflict with the audience's attention or impinge upon what they need to be focussed on. Moreover, sometimes the 2D assets looked as if they were stretched or squeezed, and that noticeably limited the way the camera could move around the studio.

While the 2D content was described as "smooth" and "polished," the 3D models provided a subjectively different experience because the audience felt as if the players were a part of the studio environment to a greater degree — even though the quality of the models was perceivably lower, and had some noticeable artifacts. This factor was particularly evident when the camera moved around the studio because the 3D models were noticeably different at their edges from the rest of the studio. However, it was also noted that the 3D content didn't become distorted when the camera panned around the studio. They looked consistent from multiple angles, especially when the two volumetric characters were presented side by side. Still, some 3D objects, like the player floating in the air, were thought to be less successful in communicating presence in the studio. The comparison in the final example was needed for some of the audience to appreciate how 3D gives a better representation of players on screen, and one participant commented that it could be improved if we could mix the 2D infographics with the 3D models.

The audience preferred the 3D assets, although it was noted that some model details needed refinement. The audience liked the realism of 3D as the 2D assets looked more like a picture but without any background context, whereas with 3D, they could "feel the presence," which was a much more enjoyable experience. Therefore, the 3D players felt qualitatively better, were more immersive, and provided a sense of depth in the studio. Integrating the 3D assets improved the viewing experience, even with the quality of the assets still requiring work. The negative factors that were noted could be improved by creating shadows of the models on the floor and creating a more realistic presence on the screen. It was also suggested that some movement of the 3D components could further enhance this experience, and the addition of shadow would also help to better position these assets within the studio.

"I liked when the two female athletes were facing wherever the camera went. It was fun, but on a serious note, while it felt much more natural, it was also somewhat uncanny."

"I liked the additional sense of life and presence within the studio."

"It brought the players they were talking about to life rather than just discussing the sports. I think integrating the animated effects into the 3D would add to them."

"Overall, I preferred the 3D as it's more natural looking and awe-inspiring."

"I preferred the 3D version to the 2D version overall. I liked the 3D content because it felt more immersive. In particular, I liked when the video zooms in on the couch."

"I least liked that the 3D content was blurry around the edges of the sports players."

"I liked viewing the players from multiple perspectives/angles."

"I preferred the 3D because it felt more immersive."

"The screen sometimes felt a little busy, especially in the first 2D video."

"I liked how well they incorporated the 2D transparent images into their broadcast. Even though the 2D visuals lacked depth, they were well integrated."

"I struggled to see the difference between the two groups of images."

"I liked the 2D cutouts because they were presented much better."

"The 2D was very slick. The written content added that little bit of extra info for people that are just getting up to speed. I loved the map of Ireland and when the county was highlighted. I also liked the 3D content because of its similarity with the physical world."

"If treated and captured well, I would prefer the 3D models (more so if they were volumetric videos — this will look like video games with a very high degree of immersion)."

"I liked the 3D characters compared to the 2D characters in the scene. I liked the 3D characters that stayed at a fixed point without turning toward the camera."

Q2. Did you notice the difference between the sports players displayed while viewing the videos? Could you describe the differences?

The nuanced differences between the 3D and 2D content were sometimes difficult for the audience to discern on the screen. For the participants who could differentiate, using soft shadows and camera angles made a difference in their viewing pleasure. This factor was particularly impactful when there was camera movement, as the 2D images felt "flat" at certain angles, whereas the 3D models felt like part of the studio.

It was noted that the appearance of the 2D models did not change when the camera altered angles. In contrast, the 3D effect was noticeable when the camera movements created a parallax when there was an overlap between the players. This parallax gave depth to the viewing experience. It also provided more dynamic viewing options for the camera, especially as the view moved up and down on the jib and arced around the players.

The 2D assets were described as "simply pictures" or "cardboard cutouts," and that the 3D assets were "more like real people in the studio." The audience commented that the 3D versions looked more integrated with the environment and reacted better to the camera's movement, where the camera movements could be much more significant. Again, it was suggested that extra details, such as reflections and shadows, would help to enhance the production value of the 3D assets.

"2D felt like cardboard cutouts planted in the scene."

"While watching the 2D players, you cannot see the depth."

"2D versions looked like cutout people made out of cardboard, while 3D had the depth."

"The 2D cutouts were clearer and high-resolution, but we can see that it's a 2D plane. However, the 3D version is less high-resolution, and the lack of proper illumination and shadows trigger an uncanny feeling due to the increase in realism compared to 2D."

"When viewing the 3D players, you can see the tops of their heads, which have a depth to their bodies.

"The 3D sports players looked more realistic."

"The 2D only allowed specific camera movements to pan in and out with the fixed images, but the 2D was more impactful. The 3D allowed slight camera rotation but didn't add to the experience as there is no other animation or movement, not even a blink, wink, or smile."

Q3. Which videos had the most influential displays and substantially impacted your viewing experience?

For some participants, 2D and 3D impacted their viewing experiences similarly. While the set, text, and video wall graphics enhanced the viewing experience with the 2D AR graphics, the 3D assets looked more natural on the studio floor, so the same 2D features combined with the 3D assets would provide increased production value and facilitate more dramatic camera movements. 2D images are currently commonly utilized on sports shows — so they were as expected and didn't have any additional impact other than providing some visual clues for the content. In the 2D video, using more dynamic poses from play, although frozen on the screen, gives the viewer a feel for the intensity or athleticism required in the game, as opposed to the stationary poses of the 3D player with arms crossed, etc. However, the 3D models provided greater immersion and interest-capturing potential since they drew attention due to camera position and studio layout changes. As a result, when the 3D models were utilized, the audience's attention was diverted to the players to a greater degree. The 3D content could have had a more substantial impact; as the camera passes the players, the participants could see the depth, giving the viewer a sense that the image was a person, more than just a 2D image. With a 3D object, the camera operator could use more studio space by looking back on the 3D image.

"I liked the 2D set more because they seemed more finished."

"It's easier to pick better images to use as 2D cutouts on the show; therefore, the result is more lively and integrated more interestingly and engagingly."

"The lack of movement in the 3D detracts from the impact."

"In the 3D version, with some refinement and character motion, the viewer will experience a rich visual experience far exceeding the 2D version."

"The 2D was much more engaging by adding information like graphs and maps."

"I think 3D video had the most influential display, particularly when the camera zooms in on 3D objects such as the couch or player."

"It was easier to imagine the space in the 3D video."

Q4. Do the signage and graphics onscreen compel you to watch a show? Why?

The cohort was divided on how influential the onscreen graphics were in compelling them to watch a show. While the graphical elements in the 2D video were captivating, the viewers would also have liked to see a similar treatment in the 3D sample. However, these graphics were only considered vital if they provided additional information, summaries of gameplays, or some additional knowledge of what was being presented or discussed on screen. Moreover, if the show presented had a new and exciting use of graphics technology, some participants would watch a specific show to see what technology the other broadcasters use.

"Absolutely yeah."

"Yes, they enhance the viewing experience."

"Yes. I believe the signage and graphics provide nice details in addition to the presenter(s)."

"Maybe yes because it acts as a support for the immersion and visualization of the news."

"It might work as an eye-catcher for the first few minutes, but ultimately it depends on how interesting the show is."

"No. But some extra infographics look good."

"In this specific case, no. "

The type of content seemed to play a big part in audience impressions on this point. While the 3D content adds to the visual experience, too much information would be distracting. Therefore, the quality of the signage and type of visuals used could make the information clearer and more readily understood, for example, real-time election or sports outcomes, to give the big picture on what results are in or pending, what numbers are up or down, what are the hotspots or areas of interest and are there patterns that are otherwise hard to detect or grasp.

Q5. Would you watch a particular show because of the onscreen images or stats?

Generally speaking, the audience said they would watch a show that used advanced onscreen graphics — the reason behind this varied. Some participants highlighted that the onscreen graphics helped them to summarise the main stats and share these findings across multiple platforms. Others pointed out that it would undoubtedly improve the production values of channels that haven't embraced 3D technology. These advantages were suggested for cross-platform synergies, such as the team position across TV and social media during the midseason. Ultimately, the visual quality of the graphics is irrelevant if it is not backed up with discerning presenters and good coverage.

"Yes, I find it's best to display stats and ensure they're easily understood."

"Yes, I think it's a great addition to the broadcast."

"Yes, this treatment has worked well in other countries news, sports, and entertainment TV."

"Between two similar shows, I would be more attracted to the one with onscreen images, videos, or stats."

"It looked better than just the usual pundits chatting away and would help get someone like me to watch more."

Q6. Do you think this video format could be applied to specific TV Shows? If so, what other TV shows would it be suited?

Using 3D content was envisioned to be advantageous for lots of different types of shows. Examples given by the cohort included reports, sports commentary, nature documentaries, history and culture programming, news channels, current affairs, kids' TV, quiz shows, cooking shows, talk shows, toy shows, fitness programs, science and technology shows, as well as music concerts.

"Anything where 2D images are currently being used to present information."

"There are numerous use cases."

"Pretty much all live content would benefit from including some rich media content. Provided it appropriate and helps the user/viewer experience."

"Additional pop-up written content can help engage the viewer, such as gardening shows or home renovation shows."

"Massive scope for climate programs — making essential information points more accurate, proximate, and believable."

"The news and sports but also talk shows and possibly documentaries."

Q7. Does using enhanced production tools, such as 3D, add to the production value of a show? If so, can you describe or explain how it does?

Enhanced production tools were considered advantageous for delivering valuable viewing experiences to the audience. 3D assets were described as "attention getters" that put the audience in the middle of the game and can capture the viewer's imagination. The participants interpreted the term "production value" to mean the quality of content based on how it was "produced," for example, actors, cameras, video quality, and graphics quality. In this context, the use of 3D has the potential to add production value. However, from the examples given, it seemed as though there would be a more significant effort (and perhaps also cost) required to increase the quality of 3D content because any errors in the models (for example, the 3D capture quality, blending into the environment, etc.) would be more evident when compared to current 2D technology due to significant variances in their quality.

"To allow audiences to use their imagination, to know what it's like, to have a virtual taste."

"I think so because people are somehow integrated with the environment; this is much better than seeing a billboard."

"I think showing the 3D model (or volumetric video of a sports moment) is much more powerful than describing the event verbally."

"In my opinion, yes. Adding more visuals of better quality always results in more production value because it removes ambiguity in the viewer's interpretation of what is being said, explained, or discussed."
"
It could add production value by helping to reduce the VFX budget."

"Yes, it can make the show look slick and modern, but we shouldn't lose sight of the fact that they are there to help tell a story or convey a message."

"Perhaps? It offers a whole new perspective. For now, shows aren't using 3D to its full potential; more like a gimmick."

The viewer may assume a show must have a high budget, with a lot of time spent working on it by a highly skilled team and expensive equipment used throughout without 3D technology. The 3D assets looked more natural on a studio floor than the 2D images, which were distorted depending on camera positioning and looked like cardboard cutouts. However, the 3D assets usually take longer to create and cost more, regardless of how they can enhance the show's look. Depending on how the 3D asset is used, it could add value as it would give the viewer at home a more rounded view of the 3D item and aid the user in understanding the scale of the model used. With a 2D image, it can be hard to understand size and scale without some reference. At the same time, it offers the opportunity to present complex content, such as statistics, to the user in a compelling way. 3D could also bring a certain amount of interactivity between the host and the 3D model, especially if it could be picked up or moved. This interaction potential could add entertainment and give the host a new way of telling a story.

Q8. If these types of 3D assets were available to download for personal use, could you think of using them for anything?

The potential use of 3D assets outside the studio also generated some interest for our participants. Examples of this extended to independent video creation, providing additional educational or informative content to supplement a primary performance, interactive demos for research, and marketing content across different platforms. Other use cases included access to the collectibles market, allowing fans to trade 3D avatars of their favorite players as NFTs. Suppose there were attractive 3D assets made available. In that case, audiences could start using them in social media or experimenting with them as we collectively work towards a shared metaverse and cocreated virtual environments.

"A mobile app that allows people to take selfies next to their favorite sports stars or celebrities and share photos on social media."

"I would love to take a photo next to some of these athletes in AR
"I can place them in other virtual environments, for example."

"There are some excellent 3D assets stores online. I see massive potential for a volumetric video or 3D scanned asset store, similar to Getty Images, where 3D items could be made available to aid their VR environment with assets."

"Maybe 3D structure or reconstruction of home, furniture, decoration, etc."

"Presentations, music videos."

"Community Education — climate, health, history."

"For personal use, they could make for a great art project."

"Viewing clothes from different angles."

Q9. Would you like to create a personal 3D avatar? What would you use it for and why?

About two-thirds of our participants could see innate value in having their 3D avatar and gave some compelling use case examples, such as delivering presentations and virtual meetings, using them on social VR platforms for enhancing reality, adding themselves to different football teams, and replacing photographs to show personal development and changes in health, gaming, dating apps, and online social media.

"Presentations — to add some additional interest and make the experience feel more immersive."

"It would be helpful for VR platforms where avatars are usually very cartoon-like rather than representing a realistic version of the user."

"Yes. You can use it for fun."

"Yes, I would make a game with me as the character; it would be fun."

"I'm not sure the technology is up to speed on the delivery we would require."

"I have an avatar for Instagram, Snapchat, and other social media platforms already, so using one for volumetric videos seems normal."

It was also noted that it is now possible to create volumetric videos with mobile phones and that this technology can help people feel more connected to the next generation of digital technology.