The Couch, the Sofa, and Everything in between: Discussion on the Use Case Scenarios for Light Field Video Streaming Services

Peter A. Kara*, Aron Cserkaszky†, Attila Barsi†, Maria G. Martini*
*WMN Research Group, Kingston University, London, UK
Email: {p.kara, m.martini}@kingston.ac.uk
†Holografika, Budapest, Hungary
Email: {a.cserkaszky, a.barsi}@holografika.com

***Abstract*—The newest light field displays visualizing captured and rendered 3D contents not only bring their entrance to the consumer market closer, but also raise numerous research questions regarding the quality of this novel form of multimedia. Unlike their conventional 2D and stereoscopic 3D predecessors, glasses-free visualization can enable prolonged periods of 3D multimedia consumption, in the form of local playback and video streaming, where the perceived content depends on the user's viewing angle. In this paper, we discuss the potential use case scenarios of streaming services for such displays and other general cases of visualization for entertainment purposes. The paper also details the considerations regarding visual quality, as certain requirements may vary for different viewing conditions.**

***Keywords*—Light field display; Light field video; Quality of Experience; Perceived quality; Viewing conditions; Horizontal motion parallax**

I. Introduction

3D is dead. We've been hearing this statement quite frequently at the time of this paper. To be precise, the properly phrased statement that appeared in several forms of media is "3D TV is dead". The most notable recent push in this direction was that the two major stereoscopic 3D TV manufacturers announced that they will drop support for the following development and production period, meaning that none of their new TV sets will be 3D-capable. One of the companies' director of new product development particularly stated that "3D capability was never really universally embraced in the industry for home use, and it's just not a key buying factor when selecting a new TV." [1]

If we accept that 3D TV is dead, we must also accept that is has never been truly alive in the first place. Indeed, it was never really universally embraced in the industry for home use, and thus users in their homes could never really embrace it. Technological constrains and the limited available content never allowed the potential added value to properly surface.

Similarly to 3D cinema, the commercialized stereoscopic 3D televisions require additional equipment – special glasses – for usage. The viewing of the visual content is also position-independent, which means that the user sees the same video frame from each and every position of observation – just like in case of conventional 2D technologies – as one channel of information is allocated to each eye of the observer.

This does not apply to virtual reality (VR) head gears, where orientation defines the visualized content. However, VR gears at the time of this paper come with several issues, mostly considerations regarding health. In case of the well-known Oculus Rift [2], there is a long list of potential symptoms that can occur due to usage, including eye strain, dizziness, impaired balance, nausea, visual abnormalities and many more. Most of these symptoms are looked at as temporary side effects, but medical judgment on this topic is not confident yet due to the lack of long-term research studies.

Light field displays, unlike any of the previously mentioned technologies, provide a natural 3D visual experience without relying on an additional viewing device; it can be fully observed by the naked eye. Of course there are other glasses-free, autostereoscopic 3D technologies as well, e.g., the multiview displays, where content is visualized in a small field of view, that can be seen from limited locations (sweet points or sweet spots), and "repeats" for several positions so multiple users can observe the content at the same time. However, light field displays utilize the entire field of view, that can theoretically – and actually even practically – reach up to 180 degrees horizontally [3].

This makes light field displays a potential user device in the home environments. However, even though there are already such displays that can be purchased, it has not entered the consumer market yet, and there is still a lot of research to be done. What is definitely a question worth considering regarding such displays and services of the future, is the actual usage. How will we consume streaming video services on light field displays? How will the parameters of current scenarios transform?

In this paper, we discuss the use case scenarios of future light field video streaming services and how the viewing conditions affect the perceived quality of visualization. The discussion primarily focuses on viewing conditions in home environments, but other use cases are considered as well for light field visualization in general.

The remainder of the paper is organized as follows. Section II provides a brief overview of the technological history and the related work in this area of research. The discussions are separated into Section III and IV, regarding use case scenarios and visualization quality, respectively. The paper is concluded in Section V.

II. History and Related Work

The first recorded works regarding the human three-dimensional vision were done by the famous Greek thinkers Aristotle and Euclid. They were both concerned with the fact that the human eyes see two separate views of the world, yet only one image manifests itself in the mind. This question remained open until Arabic scholar Ibn al-Haytham (circa 1000 A.D.) described the basic merging process of the separate eye images and discovered some depth cues used by the brain to see the world in 3D.

In modern times, Sir Charles Wheatstone created the first 3D images by drawing simple objects from two perspectives and constructed the stereoscope that displayed these to the left and right eye [4]. Later, the design of the device was improved by Sir David Brewster and was transformed into a portable form, quite similar to today's VR glasses. With the technological (and social) advances of photography, the stereoscopic photographs and stereoscopes became widespread sources of entertainment in middle-class Victorian households. According to estimates, 70% of all civil war documentary photographs were shot in 3D [5].

The Lumières brothers made the first stereoscopic short movie in 1903 for a stereoscope viewer. Then from the 1920s onward, 3D cinema was brought to wide audiences by the anaglyph method, where the viewer wears color-filtered glasses (typically red and cyan) and these filters select the left and right eye image from what is displayed on the screen. This evolved into cinema systems with polarization filters, which differentiate the two images based on the left and right circular polarization of the light. Another concept is the active shutter system, that displays the left eye image while the glasses block the right eye, and this alternates through the movie. Currently, the most sophisticated visualization technology with widespread adoption for this use case is the Dolby3D [6]. This cinema system displays the right image on discrete wavelengths r, g, b for the primary colors, and the left one on the r', g', b' wavelengths, and then the interference filter in the glasses select the appropriate image for each eye.

Virtual Reality (VR) has made a lot of promises – as well as advancements – in recent times in home entertainment use case scenarios. In the ideal case, the VR user puts on a glasses-like head-mounted display that completely replaces the natural view of the eyes with a virtual 3D scene, hence the name. The early stereoscopes can be considered as the first VR glasses, but its first interactive use was the Sword of Damocles in 1965 [7], which was the first VR headset where the primitive, wireframe-based, virtual scene changed according to the viewing direction of the user. This led to the first large-scale application: the Aspen Movie Map in 1978 [8], where an entire town was photographed from all streets and the viewer could virtually drive through it. In present day, VR technology is widely available to anyone with cheap cardboard VR glasses and a smartphone as display inside it [9]. The main drawbacks of this technology currently are the known short-term medical issues and the unknown long-term ones. Prolonged use of VR systems induce problems due to the disconnect of our senses from the real world. This solution – similar to other stereoscopic displays – also suffers from the vergence-accommodation conflict [10]. These medical issues include general discomfort, headache, nausea, vomiting, fatigue, drowsiness, disorientation and even seizures [11].

3D content tried to conquer home entertainment also through the television. The first credible approach for the 3D TV was the WOWvx from Phillips, which also pioneered the content format "2D-plus-Depth" [12]. The format – with the appropriate software – allows the extension of the available 2D content with a (semi-) automatic process to contain a depth map, so in the resulting image every pixel has a color (rgb) and also a depth (z) value. This depth information allows the correct rendering of the content to the 3D TV. Then other large manufacturers like Samsung, Panasonic, LG and Sony also released 3D TVs aimed at the high-end of the consumer market, but as of 2017, all these approaches have since been discontinued due to low consumer demand.

Light field displays are the logical next step in the evolution of 3D displays. The light field was first described by Lippmann in 1908 [13] as a way to capture light coming from all positions and directions from a scene. The inverse of this is the aim of the light field display: to show the light coming from an object from all positions and directions and to give a natural 3D view. It is important to note here that while all displays are effectively light field displays – albeit with very limited directional views – the term is generally used when the angular resolution of the display is high and the viewer experiences continuous motion parallax. Such light field TV displays exist already today [3] but currently they lack adoption.

III. Discussion Regarding Use Case Scenarios

Light field displays enable users to observe the displayed content from any position inside the given field of view, in a position-dependent way. This novel sensation of natural 3D experience can motivate users to frequently change the position of observation, in order to maximize the visualization potential of the display – simply to try to make the most out of it. A practical example for this type of user behavior would be an exhibition, where objects are visualized in a still, animated or interactive manner. Interactivity refers to the option of enhancing the system with motion sensors that read a limited set of hand gestures (e.g., rotating, zooming in etc.) and the visualization changes accordingly. Even with the option of rotating, the user is still likely to change position during observation.

Usage can be rather similar in professional environments (e.g., medical applications, designing etc.); however, if interaction is an option, the users are more likely to operate and observe the 3D visualization from a given position or from a very limited space of movement. One reason for this is that interaction is more time-efficient than movement, and certain usages may be time-critical (e.g., in the limited preparation time of an urgent surgical procedure). Use cases in professional environments may prefer the "walking-around" approach, especially if there are several multiple human observers, who all need to extract certain visual information (e.g., presenting the design of a new building to a committee).

Still out of the home environment, but closer to the use case scenario of home video playback, is the cinematic experience on light field cinema systems. In such an environment, the observer is allocated to a fixed position during the visualization of the entire video content. Of course such systems may change the fundamental concepts of movie theaters – such as the fixed seats – but let us consider a scenario where it remains. In this case, the user has very limited observation points and can only see the screen from viewing angles originating from these points.

In the home environment, one of the greatest potential for light field displays beside video streaming is gaming. Multiplayer video games have always been a popular choice by users; whether we're talking about competitive or co-operative game modes, playing together with others is definitely an added value, even though the single player gaming experience is not inferior to the multiplayer one. At the time of this paper, multiplayer games are often played over the Internet, but playing with others in the same room, on the same screen, is still popular. Split screen gaming refers to a visualization method during which the screen is divided in a given way (e.g., vertically, horizontally, 2-by-2 etc.), and each section is allocated to a specific player. Before the spread of the world-wide network connectivity, this was one of the default ways to play together real-time, but allocation by time was and is also possible (i.e., turn-based games). In such a scenario, only one single display, computer or gaming console, and copy of the game is needed. Communication between players is also more natural, and the presence itself can enhance the experience. However, it comes with certain disadvantages. One is that only a portion of the screen is allocated to a user, which in case of four simultaneous local players means that a person only uses fourth of the total screen size and resolution. A rather game-dependent or game-mode-dependent problem is the so-called "screen-peeking", also known as "screen-cheating". In competitive games, seeing the user interface of the opposing player provides advantage, which can negatively affect the overall gaming experience. Light field systems can enable the angle-based separation of the content, which means that the visual user interface of a given player is allocated to a specific area of the field of view. This solves both issues, as all users use the same full screen size and resolution to play, and "screen-peeking" is not possible without moving from one section of the field of view to another. An extent of horizontal movement along the viewing angle is of course expected in certain games, especially for the ones using motion sensors.

Video viewing, regardless of media source, can be considered as the most essential use case scenario of large displays in home environments. While monitors, laptop screens and mobile handhelds enjoy a continuously and rapidly increasing popularity in terms of video watching, televisions (or television-like displays) in general can offer screens of grandiose sizes and higher resolutions – especially with the on-going spread of Ultra-High Definition (UHD) contents. Certain home video entertainment systems are also typically enhanced with additional high-quality audio equipments, which of course could be utilized for desktop computers and laptops as well.

In this discussion, we intend to highlight the use case scenario of light field video streaming. Light field displays can come in different "shapes and sizes"; some approach them more as monitors, while others focus more on the practical realizations as televisions. Of course the potential of this display technology covers a much larger collection of fields – e.g., head-up display (HUD) for the automotive industry – but in the scope of this paper, light field displays are narrowed down to the television-like ones.

Light field displays support the concept of continuous field of view, which means that the visual content can be observed differently from each and every viewing angle inside the field of view. If and when such displays designed for home entertainment enter the consumer market, this novel visual sensation in video streaming would definitely encourage users to change positions within the field of view during the playback of content; e.g., one could start watching a video on the sofa, but later move to the couch or any other furniture with the purpose of sitting, from which the screen can be observed in a different horizontal angle. Observation height might also change during video watching in case of full-parallax displays, which means that the viewing angle determines the observable content both horizontally and vertically. It is of course important to note that at the time of this paper, there exists no full-parallax television-like light field display.

Movement during video playback can also be encouraged by the video content itself. Movies and films can be creatively directed in a way that they tell the story in a different context from different viewing angles. An easy-to-imagine example would be a crime drama where the subtle clues during the film seen from the left would differ from what can be seen from the right, suggesting alternate conclusions. Another example would be a tragicomedy or a melodrama, where the transmitted emotions would be angle-dependent; e.g., a drop of tear running down a person's face would be visible from a given direction only.

However, unless light field displays fundamentally revolutionize the way television is watched, it is expected that significant user motion during video playback would become non-existent or at least very limited. Apart from the special prior examples, it is assumed that content providers would optimize for center view, and would try to make the video enjoyable from any given viewing angle. After all, immersion during video watching on light field displays do not need to depend on the movement of the human observer. In fact, movement goes against the "sit back and relax" concept of the general video playback use case scenario.

Thus in this discussion, we point out that in the early phase of light field video streaming services and video playback in general, users are expected to have some movement inside the field of view of the display; however, in a long-term vision of such services, the traditional ways of video watching shall become the common use case scenario. On one hand, movement affects the user experience on different levels, but on the other hand, it comes with objective visual quality implications and considerations as well.

IV. Discussion Regarding Visualization Quality

In the previous section, every use case scenario was introduced with considerations regarding user movement. Movement during 3D light field visualization evokes the motion parallax effect, which is typically horizontal, but can be vertical or full parallax as well. However, no physical relocation is required in order to experience the parallax effect. First of all, the head of a living, breathing human being always moves with some minimal yet measurable extent, even if the person appears to be still during observation. Second, and more importantly, the sole movement of the eyes – realignment of orientation – alone results in a perceived parallax effect. In fact, one single eye with this natural orientation change is enough to experience it, even if its location in space if completely fixed. This single eye can be theoretically replaced with a pinhole camera in this viewing scenario. If we eliminate motion in space, and also exclude any change in orientation, then the parallax effect cannot be perceived. However, motion and orientation change of the human eyes cannot be fully avoided, thus the parallax effect during observation is always present.

The reason why user movement was particularly considered in the use case scenarios is that motion can result in more perceptual sensitivity towards the parallax effect. This research topic is currently being investigated through a series of subjective quality assessment experiments. Because of this phenomenon, any disturbance in the parallax effect can potentially reduce the overall user experience more if the observer is moving. Disturbances originate from insufficient source content angular resolution, which is derived from the number of views visualized in the given field of view. Low angular resolutions can result in the appearance of discrete source image views and the crosstalk effect.

The levels of the users' sensitivity towards angular resolution reduction contribute to the quality requirements of future services, as the satisfaction of the users is the ultimate goal. Combined with the ways of multimedia consumption in specific use cases, those scenarios with anticipated user movement necessitate higher angular resolutions than those where static observers are expected. In case of light field video streaming, assuming more or less static viewers, this means that a reasonable extent of lower angular resolution can be subjectively tolerated. As angular resolution is a major parameter of the total data size that needs to be transmitted over the network, the load put on the network capacity by light field streaming services may be significantly lessened by angular resolution reductions.

One of the primary research tasks in this field of investigation is to determine intervals of angular resolution in which service provisioning could be acceptable for the users in the use case scenario of light field video streaming in home environments. By enhancing the understanding of user satisfaction regarding this novel perceptual phenomenon, future services could be cost-effectively optimized and the construction of service protocols would be supported.

Source content spatial resolution visualized on light field displays is also worth investigating, as its reduction does not result in perceptual effects analogous to conventional displays; light rays hit irregular positions, thus content pixellation on low spatial resolutions is replaced by blur. Also, the perceived differences between higher resolutions can be lower on certain displays [14]. However, the blur due to insufficiently low angular resolutions might actually compensate some visual phenomena of low spatial resolution, i.e., the perception of discrete source image borders between views. This research topic is included in another one of our ongoing investigations, which is – similarly to the other experiments – key to the construction of our visual-experience-centric light field video transmission protocol for future services.

V. Conclusion

In this paper, we presented a discussion on the potential use case scenarios of light field displays, with particular focus on those in home environments, such as video streaming services. Instead of envisioning fundamentally new home use cases, we detailed the adoption of light field visualization in existing ones. We conclude that service usage shall have a major impact on the requirements of such future services, as it will affect the perception of quality. Our primary concern is regarding angular resolution, which determines the smoothness of the parallax effect, and thus the glasses-free 3D experience of the user.

Acknowledgment

The work in this paper was funded from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 643072, Network QoE-Net. The work in this paper was also funded from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 676401, Network ETN-FPI.

References

[1] D. Katzmaier, "Shambling corpse of 3D TV finally falls down dead", https://www.cnet.com/news/shambling-corpse-of-3d-tv-finally-falls-down-dead/ (retrieved March 2017).

[2] Health and safety documentation of the Oculus Rift, https://static.oculus.com/documents/310-30023-01_Rift_HealthSafety_English.pdf (retrieved March 2017).

[3] Holovizio 80WLT light field display, http://www.holografika.com/Documents/HoloVizio_80WLT-emailsize.pdf (retrieved March 2017).

[4] C. Wheatstone, "Contributions to the physiology of vision.–part the first. on some remarkable, and hitherto unobserved, phenomena of binocular vision," *Philosophical transactions of the Royal Society of London*, pp. 371–394, 1838.

[5] The Center for Civil War Photography, http://www.civilwar.org/photos/3d-photography-special/ (retrieved March 2017).

[6] D. Auld, "Passive eyewear stereoscopic viewing system with frequency selective emitter," Aug. 25 2011, US Patent App. 12/913,229.

[7] I. E. Sutherland, "The ultimate display," *Multimedia: From Wagner to virtual reality*, 1965.

[8] R. Mohl, "Cognitive space in the interactive movie map: an investigation of spatial learning in virtual environments," Ph.D. dissertation, 1982.

[9] Google Cardboard, https://vr.google.com/cardboard/ (retrieved March 2017).

[10] J. Kim, D. Kane, and M. S. Banks, "The rate of change of vergence–accommodation conflict affects visual discomfort," *Vision research*, vol. 105, pp. 159–165, 2014.

[11] E. M. Kolasinski, "Simulator sickness in virtual environments." DTIC Document, Tech. Rep., 1995.

[12] A. Redert, R.-P. Berretty, C. Varekamp, O. Willemsen, J. Swillens, and H. Driessen, "Philips 3D solutions: From content creation to visualization," in *3D Data Processing, Visualization, and Transmission, Third International Symposium on*. IEEE, 2006, pp. 429–431.

[13] G. Lippmann, "Epreuves reversibles. Photographies integrales," *Comptes-Rendus Academie des Sciences*, vol. 146, pp. 446–451, 1908.

[14] P. A. Kara, P. T. Kovács, M. G. Martini, A. Barsi, K. Lackner, and T. Balogh, "Viva la resolution: The perceivable differences between image resolutions for light field displays," in *5th ISCA/DEGA Workshop on Perceptual Quality of Systems (PQS)*, 2016, pp. 107–111.