Proceedings of the 2nd International Conference on Financial Technology and Business Analysis
DOI: 10.54254/2754-1169/65/20231669

Using XR to Transform Today's Short-form Video Content Consumption

Youyang Li[1,a,*]

[1]*Bejing National Day School, Haidian, Beijing,100039, China*
a. liyouyang0605@163.com
**corresponding author*

Abstract: The purpose of this research is to investigate how Extended Reality (XR) technologies can help address the effects of consuming short form content, on social media platforms. With the rise in popularity of video apps it's important to understand their impact on users, younger generations. While these platforms cater to peoples entertainment and news needs during their time they can have effects such as shorter attention spans, changes in cultural appreciation and even social disconnection. This research proposes an approach by combining Extended Reality technology with short form content to enhance user engagement while minimizing negative consequences. The proposed system includes experiences created by users that can be accessed individually or by viewers through an interactive feature called "The Hall". Additionally the system includes features that encourage world connections and a transparent content recommendation mechanism that prioritizes user privacy. By prioritizing physical health well as privacy this Extended Reality based application has the potential to offer an effective solution to the challenges posed by current 2D short form content platforms. The findings are based on a survey conducted among 30 high school students in Beijing, which indicates the feasibility of this XR based application to transform short-form content consumption.

Keywords: mixed-reality, AR, VR, content consumption, short-form content

1. Introduction

Short video apps are going viral globally. It's a more intense form of video-based social media, with a typical video length of between 10 to 60 seconds [1]. Due to their "tic" interaction mode, short attention span and algorithm users of such apps are very likely to addict to such contents, and content creators must find ways to make their videos eye-catching through preposterous content [2]. This leads to a negative trend in people's mental health, causing problems such as addtion, short attention-span, and loneliness especially for the younger generations [2-4]. Moreover, the preposterous contents lead to a decrease in aesthetic appreciation and the interest-based algorithms created information-cocoons that splits social opinions [3,5]. The current medium for such content consumption is mostly flat screens with 2D content and visual and sound feedback, without other kinds of interaction, causing low efficiency, negative impacts such as addiction due to interaction mode, and destructive cultural implications. Virtual Reality (VR) and other Mixed Reality (MR) technologies can provide a more immersive visual and sound experience and more delicate sensational feedback, maximizing the contents' efficiency and power. The research is carried out via an online questionnaire, where 30

DOI: 10.54254/2754-1169/65/20231669

high-school students in Beijing are asked about their motivations, preferences, and evaluations of using short video apps [5]. This research aims to discuss a XR-based application towards reducing the negative impact of short-form content consumption in the near future.

2. Background

2.1. Short-form Video Content Platforms

Short-form videos typically refer to a video that is 10-60 seconds long. Today, the major short-form video platforms include TikTok, Instagram Stories, Snapchat, etc. [1] These platforms use a "swiping" interaction mode to renew the information feed, while the content recommendation model is based on interest-based internal algorithms [2]. These platforms have taken the world by storm and become a significant content resource for people. In the case of TikTok, a product released in 2016, there are over 1.8 billion monthly active users in 2023, and 38.5% of TikTok users are between the ages of 18 and 24. The average time spent on TikTok per user per day is about 1.5 hours, resulting in an average session length of 10.85 minutes for the app [6].

2.2. The Needs Being Fulfilled by Short-form Content and Related Platforms

It fulfills people's fragmented time with news and entertainment, fitting their busy and sometimes dull urban life. Such short-form video content fills in the gaps in people's daily life and becomes a chance for new expression. The convenient interaction mode and the personalized algorithms made the user experience satisfying. According to a survey among 30 high-school students in Beijing, 83% agree that "feeling bored" is a reason for short-form content use, and 75% agree that "finding relaxation" is a reason for short-form content use [5]. Short-video content platforms also work well on "eye-opening", as over 60% agree that short-form video content lets them know something they have never heard of before.

2.3. Negative Effects of Short-form Video Content and Related Platforms

Adverse effects of such short-form content have been suggested acccording to the results of the questionnaire and common observations [5].

2.3.1. Attention

Approximately 75% felt that short-form videos shortened their attention span, combined with the addictive "swiping" or "tic" interaction mode for content refreshing, making them hard to focus on reading, studying, appreciating art, and natural sceneries, etc. [5].

2.3.2. Cultural Appreciation

Such platforms' contents are primarily shallow, sometimes preposterous, and even gross. Exposure to such content could cause a negative trend in people's cultural appreciation, especially for the younger generations. In the survey mentioned above, nearly 80% agree that short-form videos have a good or bad impact on their aesthetics [5].

2.3.3. Social

The physical limitation of 2D screens also blocks real-world human connections, as people have to look "down" on their mobile devices, meaning they won't be able to see others, causing a distant atmosphere in some public spaces. Meanwhile, as people are paying too much time on individualized content consumption, the time avalible for face-to-face interation is significantly reduced.

DOI: 10.54254/2754-1169/65/20231669

2.3.4. Information Cocoons

Moreover, the interest-based personalized algorithms used by short-form content platforms caused the split of public value, as what is being recommended to individual users are the fields they have been interested in and the accepted points. As time moves, people are likely to be trapped in an "information cocoon," having a misbelief in/her long accepted value [2-3]. In addition, there is no effective fake news filtration system, and people are easy to be manipulated as pawns in political clashes.

3. XR Mediums and Short-Form Content

Today's platforms' harm mainly comes from the lack of efficiency of flat mediums. With multi-sensational immersiveness and diverse interaction methods, Extended Reality (XR) mediums can make information transmission more effective and efficient. There will be no more need and acceptance of low-quality content. Combining virtual and physical worlds through sensors and image synthesizing, XR mediums can play an essential role in bringing back real-world human connection.

3.1. Existing XR Applications for Content Consuming

As XR mediums can bring people a more immersive experience, many developers have been trying to build content-consuming experiences based on existing XR hardware products. According to their approaches, it can be divide into two main categories, VR experience of existing media platforms with a developed content-creating, and VR Media Players without content creation ecosystem.

3.1.1. VR Experience of Existing Media Platforms with a Developed Content-Creating

The use of Virtual Reality experiences on existing media platforms like Youtube that allow content creation offers benefits in terms of mainly providing an immersive experience for point of view (POV) videos. For example users can virtually participate in a car test drive or enjoy a roller coaster ride immersing themselves in situations that may otherwise be inaccessible physically. However despite its nature this approach has some limitations. One notable drawback is the lack of experiences and engaging feedback which can hinder user involvement and interest over time. Additionally it's important to note that such VR applications cannot replace short form content platforms like TikTok, due to their usage scenarios and unique elements that attract their user base.

3.1.2. VR Media Players Without Content Creation Ecosystem

There are many VR media players on the market that don't have a content creation ecosystem. These platforms mainly focus on converting existing 2D content into an experience as seen in platforms, like SkyBox VR Player and Simple VR [7]. The advantage of these platforms is the heightened immersion they bring to consuming 2D media offering users a new level of engagement. However there are limitations to consider well. One major drawback is the lack of interactivity, which might eventually lessen user interest over time. Additionally since the original content wasn't specifically designed for VR the 3D experience might come across as forced or artificial thus diminishing the user experience.

3.2. Experiences- An Alternative to Today's Short-Form Content

This project is about to fulfill people's using scenarios for short-form content, preserving the advantages of existing applications and eliminating the harms. To achieve that, an alternative form of today's short-form content is suggested, called "Experiences."

DOI: 10.54254/2754-1169/65/20231669

Experiences are immersive user-generated content involving as many sensations as the user/content creator's hardware supports. It can be launched quickly during the fragmented time, including various content recommended to users based on a simple, transparent system. It will be less addictive for the innovative interaction mode. For the immersiveness of XR mediums, the quality and efficiency of information of these XR experiences will increase compared to current 2D mediums, as people would not be able to be immersed in noisy, raffish content in such experiences. Experiences can be entered in multi-viewer or single-viewer mode, and the multi-viewer method connects people.

3.3. The Hall- An Alternative to Today's Swiping Interaction Mode

As one of the significant factors of harm to current 2D short-form content platforms is their addictive swiping interaction mode, a new interaction mode of switching between contents called "The Hall" is suggested. As the name suggests, it is a hall in the virtual world, with doors-like windows in the hall displaying the preview of the contents and acting as the way to enter the contents. As mixed reality can combine the real with the virtual world through sensors and graphic processing, The Hall has a translucent white wall, which displays the real-time environment of the physical space. For various considerations, the primary interaction method in The Hall will be eye control through eye-tracking. When the user enters the application, contents appear in different doors, with each door displaying a single piece of content from a specific field different from the ones in other doors. Three doors will be displayed in a single field of view, corresponding to three content categories. When wanting to browse other categories, the user can stare at the "field of view switching buttons" for two seconds, and three new doors will appear. To get into an experience, the user stares at the corresponding door for two seconds. After watching one experience, the user will enter a new Hall with two doors. The door on the left leads to the original Hall, where the user can switch to different categories of contents. The door on the right leads to another experience within the same category. Entering each door requires staring at it for two seconds.

3.4. Forming Connections

One of the major problems of existing fragmented time content consuming platforms is its effect of splitting people apart, physically and mentally. Improving this situation became an essential objective of this project.

Experiences will be recommended to each user based on location, meaning that users are more likely to share standard content with nearby people.

Experiences can be entered in both single-viewer and multi-viewer modes. Users in a multi-viewer experience can see nearby people sharing that experience through avatars, providing chances for self-expression. Avatars with advertisements will be restricted, and the user feedback system helps. After finishing a multi-player experience, the translucent hall would label the location of other people who enjoyed the same experience. The user can stare at that tagged person for three seconds, and basic information about that person (user provided) will be shown on the screen, including their comments on the last experience. Suppose one user is interested in another person. In that case, they can strike up a conversation by having a direct talk in a physical space or starting a 20 lines-maximum chat within the application.

The Hall effectively uses XR mediums' ability to combine virtual and physical worlds, reminding the users of real-world environments. Moreover, as the display device is placed in front of the users' eyes, it changes the situation where people have to look down on their phones, eliminating the possibility of natural eyesight connections. Now, the application users can see other people in the physical world from a natural angle of view, reminding people of real-world human connections.

DOI: 10.54254/2754-1169/65/20231669

Personalized algorithms have been splitting people's values. This problem can be solved with this application's simple recommendation system, which randomly provides contents of different fields. The user can choose what to watch depending on what is displayed in different doors of the hall. This is a simple way of breaking knowledge-related barriers by exposing users to other content areas.

3.5. Content Creator Side

As this application is based on a UGC (user-generated content) mode, it is essential to consider the experience of content creators. For the content creators, the essential hardware requirement would be a 360-degree camera with a microphone. Odor and haptic feedback devices are optional. The editing can be entirely done within the application, including setting viewer's positions and routes, timeline (of multiple sensations) editing, music editing, and hiding easter eggs in the experience that can be obtained by the user by staring at it. Collecting such easter eggs gives the users medals that can be shown on their profiles.

4. Moral Concerns

4.1. Mental Health

To avoid addictions and negative results of such addictions such as loneliness and anxiety, the daily use time of this product is limited to 2.5 hours. There will be notifications when the use time is achieved one h, 1.5h, and two h. Each experience has an optimal time limit of 30 min, following by a 5 minute breaking time. This can ensure the focus on content and general well-being of the users, according to the well-known Pomodoro technique [8]. There will be guidance for content producers about how to make their immersive story-telling more effective.

4.2. Physical Safety

When safety-related activities (driving, walking, etc.) are detected through sensors (cameras, LiDARs, GPS, etc.), the app won't open. When an unsafe environment (crosswalk, etc.) is seen, the app won't run. Having eye-control as the main interaction mode ensures that the user won't hurt others/themselves through body movements. The experience includes irritative contents, which might cause the viewer to make instinctive extensive body movements or noises (screaming, etc.) and disrupt other people in the environment, even causing danger, won't be accepted by the platform, and recommended to the users.

4.3. Privacy

Privacy has been a significant concern of this project, as it is a fundamental human right in this information era. The data processing and content recommendations are transparent, as explained below. Sensors will be used to collect and process the data of surrounding environments to keep the user from using the app in dangerous environments. All those safety-related visual and location analytics will be done entirely offline. The location based content recommendation through Ultra-wideband (UWB) technology is completely offline [9]. The users purely provide their identity information on the personal profile page. There are no unique algorithms included in the content recommendation process.

5. Conclusion

The information above discussed the possibility of applying the traits of XR mediums to transform today's short-form video content. This project solved current addiction problems caused by short-

DOI: 10.54254/2754-1169/65/20231669

form content platforms due to their interaction mode with a virtual Hall, where eye control slows down the information renewal speed and brings back the users' patience. The current 2D short-form videos are being transformed into immersive user-generated experiences that can be entered in single-viewer and multi-viewer modes. Details have been added to the hall and experiences to remind the users about building real-world human connections. The society-harming personalized algorithms are being replaced with a simple, translucent content recommendation system. User privacy is also concerned with having most of its analytics done offline. Shortly, XR hardware has the potential to take over smartphones and even become people's major computing hardware. They will be more portable so that people can wear them comfortbly as today's regular glasses like Apple's Vision Pro, or even as contact lenses, a possible future being realized by products like Mojo Lens by Mojo Vision [10-11]. This can lead to a natural and necessary improvement in society's content-consuming habit, leading to a better cultural appreciation and mental health condition of the public.

References

[1] *Apple Vision Pro. (n.d.). Apple. https://www.apple.com/apple-vision-pro/?afid=p238%7CERmWyX8s-dc_mtid_%5Btracker_id%5D_pcrid_78615288092096_pgrid_1257842600642414_&cid=wwa-us-kwbi-VisionPro-slid---productid--Brand-Avalanche-announceBrand-*

[2] *Aslam, S. (2023, February 27). TikTok by the Numbers (2023): Stats, Demographics & Fun Facts. Omnicore Agency.https://www.omnicoreagency.com/tiktok-statistics#:~:text=41%20percent%20of%20TikTok%20usersare%20aged%20between%2016*

[3] *Li, Y. (2022) [Unpublished raw data on The Usage of Short-form Content Platforms].*

[4] *Mandal, A. (2020, May). The Pomodoro Technique: An Effective Time Management Tool - NICHD Connection - Science@NICHD. Science.nichd.nih.gov. https://science.nichd.nih.gov/confluence/display/newsletter/2020/05/07/The+Pomodoro+Technique%3A+An+Effective+Time+Management+Tool*

[5] *Mojo Vision, The Micro-LED Company. (n.d.). Mojo Vision. http://www.mojo.vision/mojo-lens/*

[6] *Petrillo, S. (2021, December 13). What Makes TikTok so Addictive?: An Analysis of the Mechanisms Underlying the World's Latest Social Media Craze. Brown Undergraduate Journal of Public Health. https://sites.brown.edu/publichealthjournal/2021/12/13/tiktok/*

[7] *Ren, S., Liu, L., Yang, S., & Jiang, J. (2022). Investigating Information Cocoon Attitudes in Short-Form Video Applications. Communications in Computer and Information Science, 1655, 89–96.*

[8] *Stokel-Walker, C. (2022, February 21). TikTok Wants Longer Videos—Whether You Like It or Not. Wired. https://www.wired.com/story/tiktok-wants-longer-videos-like-not/*

[9] *Why Short Form Video Content is Taking Over Social Media. (n.d.). Spliceapp.com. http://spliceapp.com/blog/why-short-form-video-content-is-taking-over-social-media*

[10] *Williams, E. (2023, July 9). 11 BEST VR VIDEO PLAYERS IN 2023. FixThePhoto.com. https://fixthephoto.com/best-vr-video-player.html*

[11] *Zhou, Y., Choi Look Law, & Xia, J. (2012). Ultra low-power UWB-RFID system for precise location-aware applications. Wireless Communications and Networking Conference Workshops (WCNCW), 2012 IEEE. https://doi.org/10.1109/wcncw.2012.6215480*