EMOVAC: Designing a Mixed Reality Experience for 5G

Ben Nicholson
University of Southern California
Mobile & Environmental Media Lab
Los Angeles, California
benjamin.nicholson@usc.edu

Fidelia Lam
University of Southern California
Mobile & Environmental Media Lab
Los Angeles, California
fidelial@usc.edu

Dave Warhol
Realtime Associates
2301 Rosecrans Ave, Suite 4150
El Segundo, California
davew@rtassoc.com

Scott S. Fisher
University of Southern California
Mobile & Environmental Media Lab
Los Angeles, California
scott.fisher@usc.edu

***Abstract*—Research on the future of gaming suggests that over the next 5-7 years there will be increasingly blended digital and physical worlds in the games and entertainment industry. That future is predicated on the low latency high bandwidth convolution of 5G, IOT and Edge Computing. This paper is a summary and postmortem examination of the main design and technology issues encountered in the development of a mixed reality experience designed to demonstrate the capabilities of 5G network technologies.**

***Keywords—component, formatting, style, styling, insert** (key words)*

I. Introduction

Research on the future of gaming suggests that over the next 5-7 years there will be increasingly blended digital and physical worlds in the games and entertainment industry. That future is predicated on the low latency high bandwidth convolution of 5G, IOT, and Edge Computing. To demonstrate and evaluate this mixed reality scenario, the University of Southern California's Mobile & Environmental Media Lab (MEML) collaborated with its industry partners in December of 2018 to launch EMOVAC, an immersive mixed reality experience that takes place in the comic book world of Batman.

Utilizing Warner Bros. and DC Comics' Batman property, along with Intel's network and computing power, and AT&T and Ericsson's 5G wireless data technology, MEML created an immersive mixed reality experience on the USC campus, set in the world of Batman. Over the course of four days, students and community members entered EMOVAC, hidden in an unused storefront at the USC University Village, and encountered their greatest fears.

This paper is a summary of what transpired within EMOVAC, as well as a design and technology postmortem that our lab intends to leverage in developing even more impactful mixed reality experiences in the future.

II. Research Objectives

A. Background

After decades of research and development, the spectrum of mixed reality technologies (AR, VR, and MxR) is rapidly becoming a new medium for content delivery and engagement with near-term commercial application across a wide array of disciplines ranging from immersive entertainment to spatial computing. Yet the effective and efficient deployment of these technologies is still uncertain and "best practices" have yet to be established. A 2017 paper by Orlosky, Kiyokawa, and Takemura [4] reviewed state of the art virtual and augmented reality communications technologies, outlining necessary next steps in 5G network design for emerging virtual and augmented reality application demands. Some of the challenges of designing for 5G mixed reality experiences include issues of latency, bandwidth, quality of service, colocation, and haptic response. Furthermore, a net publication by Qualcomm [5] noted the technological developments necessary and in motion for ubiquitous adaptation of mixed reality experiences, including capacity for low latency wireless connectivity and immersion. Though most mixed reality research has focused on technological and engineering developments, early research in site-specific mixed reality design emphasized the importance of balanced interconnected processes towards creating rich mixed reality experiences for users [1, 3]. A key remaining challenge to full deployment of mixed reality technologies is the ability to robustly support a very large number of mobile users in both indoor and outdoor unstructured environments. The rollout of ubiquitous connectivity through the 5G infrastructure will enable a critical next step in the evolution of this disruptive new medium and will support an entirely new form of context-aware user-centric experience that will become the next

This project was supported, in part, by a grant from Intel Corporation.

mobile computing platform. EMOVAC takes these experiential and technological mixed reality design considerations and concerns and both applies and tests them to evaluate the current state of mixed reality in combination with 5G and to develop a set of mixed reality project practices for future MEML projects.

B. Objectives

Before EMOVAC came into being, MEML and its industry collaborators worked to consider how emerging 5G technology and VR/AR technologies could be integrated in a proof-of-concept testbed to create visceral, compelling experiences in worlds we always wished we could inhabit. From these early conversations, an ambitious plan was put in motion that would combine the construction of physical sets, live performers, original sound design, professional voice actors, 3D modeling, motion capture animation, a VR/AR headset, and a 5G network. The aim was to discover:

- What are the challenges for experience designers to develop mixed reality with physical world interaction connected to a virtual or augmented world?
- How does it feel to not only be co-present with characters from a popular media franchise, but to play a role in their stories?
- Where is the boundary between immersion and distraction when using the latest VR/AR technologies?
- Can 5G technology enable graphically demanding, real-time VR/AR experiences? And what are the related technical requirements to achieve compelling context-aware user engagement?
- What best practices can be developed to optimize the production of mixed reality experiences?

We explored the answers to these questions by bringing participants into EMOVAC and observing their responses.

III. EMOVAC Experience Narrative

Over the course of several months, a small team of writers, artists, and interactive experience designers developed the backstory and narrative structure for the EMOVAC experience. After several iterations and consultations with Batman experts at DC Entertainment, the storyline can be briefly summarized as follows:

Jonathan Crane (aka Scarecrow) was a graduate student in psychology at Gotham State University. Along with nine other graduate students, he took part in a psychological study of emotions. The study, funded by the Wayne Foundation, used a new machine, an early computer, to induce an endocrine cascade that created an emotional response in its subjects. Affectionately dubbed EMOVAC (Emotion Modeling Operational Variable Automatic Computer), while it made users feel low levels of happiness or irritation, by some technological quirk it was vastly more effective at generating fear than expected. Of the ten original subjects, five ended up in Arkham Asylum. Many years later (today, in fact), Jonathan Crane sees 5G technology as a chance to spread fear much farther than his Fear Gas ever could. He remembers EMOVAC and is able to reconstruct the actual mechanical device, though this time the machine delivers its images across a digital network. However, Crane is worried about testing his improvised machine on himself; he coaxes participants into joining his 'scientific study' to test the effects of his 5G EMOVAC.

In the leadup week before the experience, participants are invited to make an appointment to meet with Dr. Crane. Once a participant has scheduled their appointment and enters the physical EMOVAC space, Dr. Crane asks them to complete a task - loading fear modules into EMOVAC. But even as the player is being subjected to mind-altering hallucinations, the Bat Cave monitoring systems registers the activity of EMOVAC because it uses immense amounts of power, alerting Batman.

Batman contacts the participant and tells them he has hacked into the system and they should play along with Scarecrow until he can get there. By the time the participant inserts the third of three modules, Batman discovers the location of EMOVAC and crashes into the space, confronts Scarecrow, and subdues him, stopping the experiment. The participant, by keeping Scarecrow occupied while Batman traces the location of the EMOVAC, saves Gotham City and the world from Scarecrow's nefarious plans.

IV. EMOVAC Experience Flow

On the day of the experience, once a participant had checked in with the receptionist, they would be instructed to follow the lab assistant, who led them through a disturbingly long, dark hallway and through a mysterious door. On the other side of the door was the brightly lit EMOVAC room with a VR/AR headset on a small rolling table in the room's center; the lab assistant would help the participant don the headset and provide them with simple instructions on how to proceed in the experiment.

Once the lab assistant had rolled the table to the corner of the room, Dr. Crane began to address the participant through speakers next to a projection screen on the north wall of the space; in the headset, the participant could see a virtual (AR) image of Dr. Crane overlaid onto the projection screen. From there, participants were asked to insert modules, physical boxes with small handles, into the EMOVAC interface one at a time. As the participant inserted each module, a new fear experience would be delivered through the headset as Dr. Crane would mock the participant for their fear (and Batman would send secret transmissions through the headset to encourage the participant to keep Dr. Crane distracted while Batman located EMOVAC). After the third fear sequence, Batman arrived and battled with Scarecrow, eventually succeeding and destroying EMOVAC.

After Batman had thanked the participant, the lab assistant helped them to remove the headset and apologized for what appeared to be a harrowing experience (they could not see or hear what the participant was experiencing through the headset). The lab assistant then led the participant back to the exit and released them.

V. Design Findings

After the intense development of EMOVAC, MEML was excited to see what participants would think of the experience and how the project would address its aims. As an experimental collaboration between entities with varied interests, the experience itself was certainly not flawless, but ultimately led to learnings that should be useful in the ideation and execution of new mixed reality experiences. Some key findings that MEML intends to further develop in future projects include:

- As a foundational premise, mixed reality (combining virtual/augmented reality experiences with elements of the physical world) appears to be a compelling medium for immersive story telling; through EMOVAC, we engaged participants at every step along the way, from discovery to recruitment, from intake to experience, from conclusion to debrief. By allowing the world of EMOVAC to exist outside of the headset alone, participants felt like they were not merely trying a tech demo but stepping into a total world. The ability to coordinate an interdisciplinary team and create such experiences, while resource intensive and challenging to facilitate, shows great potential to impact participants' lives in ways that differ from more singular media forms; MEML intends to continue assembling interdisciplinary teams to develop mixed reality experiences.

- The opportunity to share physical space with characters and worlds that were previously only accessible through "the screen" is exhilarating for participants; the shock of Scarecrow finally emerging from the screen and walking towards the participant yielded responses from joy to horror and everything in-between. The additional surprise of Batman appearing behind the participant and proceeding to protect them from the evil of Scarecrow by engaging in battle captivated participants, who followed the action and reacted bodily and verbally whenever the fight seemed like it might crash into them. While the utilization of character models in EMOVAC was limited, we felt that these moments were a true delight for participants and believe, for future projects, that additional presence (and enhanced interactivity) will lead to greater emotional investment from participants.

- One major challenge we encountered in the development of EMOVAC is that VR/AR technology is still in its early commercial stages and, though current devices show tremendous potential, there are shortcomings that break immersion and remind participants of the novelty of the apparatus. Two major issues that diminished the impact of EMOVAC were the low-resolution pass-through camera in the headset and lack of true occlusion tracking for hands and other objects surrounding the participant.

  The first issue, which relates to the "real world" image provided to participants by the cameras on the front the headset, caused a discrepancy between the crispness of the virtual overlay images and the fuzziness of the space actually surrounding the player, a bizarre inversion where the computer-generated images appeared more "present" than the real location the participant occupied. A solution would be for headset manufacturers to invest in higher-resolution pass-through capture, though this sort of decision is subject to market imperatives (i.e., whether the increased cost of this functionality would drive corresponding sales). Alternatively, experience designers could render virtual content at a lower resolution to match that of the pass-through cameras, though we imagine that this would result in an overall lower-quality experience.

  The second issue, in which it is difficult to determine which physical objects in a space should be placed in front of or behind virtual objects (particularly as a participant moves around), prevents the execution of a number of a experiences, the most vital of which may be the presence of multiple participants. Because current occlusion solutions only allow for real-time tracking of objects like hands or objects that have tracking devices attached to them (and defined spatial dimensions), it is impractical to produce experiences where large number of objects are free to move around the space. The inability to provide immersive experiences for multiple participants severely limits the potential impact of said experiences (as they must be solitary), discouraging designers from experimenting with VR/AR experiences that are highly social. Fortunately, software solutions are in development that can better recognize the relative location of various virtual and physical objects in a space; we hope that continued developments in this area will lead to increased possibilities for immersion and, ideally, social co-presence.

- While there is still development to be done, we are excited about the potential for 5G technology to enable real-time, immersive mixed reality experiences. A major limitation of VR/AR technology is its expense, with the need for powerful graphical computation on-site to render images convincingly on a headset in real-time. If this computation can be offloaded to consolidated processing resources off-site, then transmitted to the local headset with the low latency that 5G is capable of, then VR/AR experiences might only require a high-resolution viewing device (available in most people's pockets in their smartphones) and a network capable of delivering high-resolution visual content in real-time.

If VR/AR could be primarily delivered through smartphones (with a simple headset that allows a person to affix their phone in front of their eyes), the mobility and accessibility of mixed reality experiences would take a tremendous step forward, allowing not only for easier development of those experiences, but also the potential for a much greater audience (which, in turn, would encourage more development, improving learning and the ability to innovate).

VI. Best Practices for Future MEML Projects

Because of the relatively ad-hoc nature of projects like EMOVAC, it can be difficult to extract best practices that would be applicable to mixed reality endeavors in general. However, there are three takeaways we feel can be applied to the future mixed reality efforts of our own lab that, if kept in mind, we believe will lead to more efficient and impactful outcomes:

- Invest in struggle: though perhaps counterintuitive as a "best practice", MEML must keep in mind that mixed reality is an experimental endeavor and that, in order to make these experiences more efficient and impactful, we need to take risks and engage with challenging projects if we hope to someday arrive at a more standardized form of making. Now is the time for MEML, as an early adopter of mixed reality technologies, to do the hard and messy work of trying and, in some cases, failing, but always learning and sharing our learnings. This may be a difficult premise to valorize (particularly when it comes to determining how our projects might be financed), but we must encourage an audacity of effort, not conservatism, if MEML is going deliver on its promise of intelligent integration between the physical and the digital. Whether it be utilizing as-yes-unproven new technologies or experimenting with untested social and physical contexts, we will try to not let the unknown act as an obstacle; we have found that audiences have a tendency to appreciate ambitious but flawed experiences more than mediocre but polished ones.

- The physical is our friend: we often have a desire to over-invest in the capabilities of new digital graphic technologies to deliver innovative experiences, yet MEML has found that we tremendous tools already available in the physical world to facilitate mixed reality. Whether it be live actors, physical spaces, social media, or, most vital of all, sound, we aim to use what is already there to immerse an audience in the world we are building. While the novelty of VR/AR is certainly a draw, we must remember that for many these experiences are so new and novel that, without a grounding in media forms they are more familiar with consumers, mixed reality can feel more like a stylized tech demo than an immersive experience. We will limit the use of VR/AR to when it is most necessary for what it accomplishes best: the introduction of fascinating visual content in the context of a broader world. Otherwise, we will defer to that same broader world for all of its capabilities and conveniences to not only bring our experiences to life, but to bring the life of the real world to our experiences.

- Project management: while true of any complicated project, we have found that project management is particularly vital for our mixed reality projects, where different disciplinary capabilities and emotional/professional investments can make it difficult for the team to all move in the same direction. Project management for our mixed reality experiences requires a somewhat dispassionate stance towards the technologies being deployed and an extreme of empathy for those involved in the production of the experience; amidst the particularities of each of the project's various elements, a total experience needs to be developed and no isolated interest can overtake the whole. This means that MEML attempts to adopt an emphatically practical approach from day one to set a precedent for how the team will work together, an understanding that not every effort will make "the final cut" but that the team will do its best to honor each effort and communicate to avoid wasted work. There is a fine line between ensuring execution and stifling creativity, so a deep knowledge of the personal motivations and needs of each team member is vital so that our project manager can both help direct the team, but also absorb their ideas and needs to constantly reconstitute the collaborative process in real-time. This is easier said than done and impossible to do without mistakes (and, ultimately, only works if the whole team participates in the project management process) but an established way of working based on mutual respect and clarity of mission helps the excitement of experimentation and innovation see the light of day (at the end of the day).

VII. Technical Findings

The technical goals of this project were to demonstrate the potential for delivering AR, VR, and mixed reality experiences over a 5G network. Our approach was to take a single Unity project containing the user experience and adapt it to run over the 5G network. This involved turning the Unity project into two Unity projects. The first Unity project is a Thin Client that runs on a computer connected to an HTC Vive pro, either cabled or over WiGig technology, to collect video pass-through data, 6DOF information, and the LeapMotion occlusion data, and to transmit it over the 5G network to the second Unity project running on a remote Edge Server. The Edge Server collects the video pass-through data, the 6DOF data, and the LeapMotion occlusion data, combines it with a rich gameplay graphic environment, renders the appropriate game graphics state, and passes the resulting stereo frame buffers back to the Thin Client for display to the users. The objective is to show the low latency and high graphic fidelity possible in an edge computing environment, where a co-located computer is doing most of the graphics experience work, allowing for the lightest possible graphics display. An early 2018 report by researchers at Microsoft posit edge computing technique "collaborative acceleration", which leverages the specific hardware capabilities of a mixed reality devices and edge computing nodes to partition an application's workflow for optimized efficiency [2]. EMOVAC takes this technique and provides an early glimpse into the system design challenges of mixed reality that utilizes edge offloading. We also created a back-up 5G demonstration of pulling all the experience assets over the 5G network to demonstrate how graphically rich experiences can be facilitated to a local computing solution.

A. Results

We configured the Thin Client and Edge Server architecture such that it ran at 30FPS over a local area network. When deployed over the 5G network, frame rate dropped to 10FPS which was not sufficient for a satisfactory 5G presentation. So, the deployed 5G demonstration component used the back-up approach of bringing assets over the 5G network. Upon analysis, the 5G network showed conflicting information about the source of latency. On one hand, the 5G network showed it was capable of moving 1GByte of data across the network in approximately one second, both through its bandwidth tools, and through a Windows file copy from the Edge Server to the Thin Client. On the other hand, when running the Client/Server experience over the network, it showed that we were using only 15% of its capacity. This is a clear indicator that when the source of the latency is discovered and minimized, 5G bandwidth is available to increase the frame rate many fold.

B. Complications

We encountered two main technical complications over the course of the project. The first was reliance on the Microsoft Streaming Toolkit. We had previous experience with the MST and were relying on it for communication between the Thin Client and Edge Server and to provide graphics compression from the Edge Server back to the Thin Client. As the MST is provided to the public, it primarily supports the Microsoft HoloLens. Ultimately it proved impractical to adapt the MST to the HTC Vive. The MST-supplied Unity Thin Client failed to connect to the MST-supplied Edge Server, which was a known issue from the start of the project. We tasked an Ericsson-supplied resource to debug this while we had a parallel task of gathering the Thin Client frame buffers and 6DOF information to send to the Edge Server. The external resource was unable to compile the MST Unity Thin Client, much less solve the network connectivity issue, which had us rely on a Microsoft-supplied non-Unity Thin Client we had used in our previous work. Here, we were unable to change the resolution of the frame buffers in the pipeline, requiring us to create our own communication layer between the Thin Client and Edge Server, which proved to be an ambitious task for the end of the project, and had us abandon the feature of compressing graphics from the Edge Server back to the thin client for purposes of this deployment.

The second complication was that, in an effort to test our custom network code in a 5G environment prior to deployment, the 5G equipment was not ready for the task. The 5G equipment had been upgraded prior to our testing in a manner that disabled the TCP/IP layer we use for communication. Had the equipment been working, we would have had a test case showing the 10FPS experience and would have had a chance to troubleshoot it, both from the architecture we built as well as in the 5G implementation, in advance of deployment.

C. Sources of Latency

Before and during the experience deployment, Realtime, Ericsson, and Intel performed numerous tests in the 5G environment and gathered volumes of information about the sources of latency. We cataloged two main sources which should be addressed, two optional technical tasks that may reduce latency, and an additional issue that didn't show an impact but should be explored and solved.

- Client/Server Handshaking. Our implementation uses handshaking between the Client and Server to announce the availability of data and to request data be sent. On a local area network the latency is much lower for these, but over the 5G, each transmission no matter how small creates a delay. If there is an inordinate amount of handshaking during an upload or download operation, this would impact the number of frames possible. This explanation is the most consistent with observed results. Ideally, the code can be architected to reduce or to remove the need for a handshake for all of the frame data being transmitted each way by having both the Client and Server continuously transmitting data when it is ready and using the most recently completed set of data when ready instead of requesting it be sent. This will result in a true 'streaming' architecture.

- Unknown 5G Packet Resends. The Ericsson/Intel engineers catalogued an interesting and unexplained phenomenon. The 5G network reported an unusually high amount of TCP packet retransmission requests (approximately 30%) in the communication from the Edge Server to the Thin Client when running the experience over the Client/Server architecture, but not when running their bandwidth tools or the Windows file transfer. This alone cannot explain the 10FPS result, but how a TCP/IP implementation could increase 5G transmission errors is an area for future exploration and resolution, because left as is, it will slow any implementation down 30%.

- Communication Parallelization. Our implementation opens a single TCP/IP connection to transmit all frame data sequentially. An alternative approach is to open multiple connections, each of which is responsible to carry a portion of the data. This may be an unnecessary complication, as the 5G network should be dedicating all of its bandwidth to a single connection: for example, the Windows file transfer of 1GB in one second uses a single TCP/IP connection. Action was deferred on this until the prior two methods were investigated and optimized.

- Frame Buffer Compression. Another point of optimization is to compress the composited video frames on the Server prior to sending to the Client. NVidia has made its real-time compression libraries available for developers. Compression is not immediate and may add latency, so the tradeoff has yet to be determined if the time it takes to compress and transmit smaller frames will be more responsive than sending uncompressed frames immediately.

- Unstable 5G Implementation. Sometimes the 5G network would get in a state where communication from the Thin Client to the Edge Server would only run at 5% of its capacity in a Windows file transfer, the only solution for which was to restart the 5G network. This was determined not to be a factor in the performance of the Client/Server architecture, because the experience still ran at 10FPS when it was demonstrated the 5G network was running at full bandwidth from the Client to the Server. While therefore addressing it is not necessarily a requirement for additional 5G Client/Server deployments, any such unexplained dramatic loss of performance could point to the source of a problem that contributed to another issue.

- Development Path Observations. In hindsight, a different sequence of development priorities could have supported identifying the problem spots further. By starting with the MST implementation in-house instead of tasking it to an unknown resource, we would have been able to determine its lack of viability and take a different development path earlier.

D. Network Protocol Correction

After the initial deployment in December 2018, the network architecture for this experience was rewritten using RTP/UDP protocol instead of TCP/IP. With this change, the 3D graphics software was able run at 90-120fps at under 20ms roundtrip with data rates of 1.2G up and 750M down in raw instead of compressed format. This change removed all network issues previously experienced and ran seamlessly and stably with the Vive Pro headset and WiGi adapter.

VIII. Lessons Learned

After decades of research and development, the spectrum of mixed reality technologies (AR, VR, and MxR) appears to be rapidly emerging as a new medium for content delivery and engagement with near term commercial application across a wide array of disciplines ranging from immersive entertainment to spatial computing. But a remaining challenge to our full deployment of these technologies is the ability to robustly support a very large number of mobile users in both indoor and outdoor unstructured environments. The rollout of ubiquitous connectivity through the 5G infrastructure will enable us to take a critical next step in the evolution of this disruptive new medium and will support an entirely new form of context-aware user-centric experience that may become the next mobile computing platform.

The project described in this paper developed a proof of concept mobile experience in which users interacted with a high-resolution virtual world spatially correspondent with an extensive indoor physical site. The PoC integrated a 5G connected environment, IoT-enabled physical props, a context-aware physical space, AR/VR display technology, and DC Comic content to create an immersive mixed reality experience. The project expanded our understanding of 5G and how it might enable context aware environments and experiences and demonstrates new forms of mobile digital content and user interaction that are only enabled when deployed using 5G.

Acknowledgment

This project was done in collaboration with industry partners from Intel, Ericsson, Warner Bros, DC Entertainment, and AT&T. The MEML team responsible for EMOVAC consists of faculty, graduate students, and undergraduate students from across disciplines but with a shared interest in emergent technology, mixed reality, and storytelling.

A more detailed project description and development credits are available on the project website here: https://www.emovac.org/

A full walkthrough of the experience can be seen here: https://vimeo.com/321577400

References

1. S. S. Fisher, “An Authoring Toolkit for Mixed Reality Experiences,” in *Entertainment Computing: Technologies and Application*, R. Nakatsu and J. Hoshino, Eds. Boston, MA: Springer US, 2003, pp. 487–494.*)*
2. K. Lebeck, E. Cuervo, and M. Philipose, “Collaborative Acceleration for Mixed Reality,” p. 6.
3. S. S. Fisher, “Recent developments in virtual experience design and production,” in *Stereoscopic Displays and Virtual Reality Systems II*, 1995, vol. 2409, pp. 296–302.
4. J. Orlosky, K. Kiyokawa, and H. Takemura, “Virtual and Augmented Reality on the 5G Highway,” *Journal of Information Processing*, vol. 25, pp. 133–141, Feb. 2017.
5. Qualcomm Technologies, Inc., “VR and AR pushing connectivity limits,” Oct. 2018. https://www.qualcomm.com/media/documents/files/vr-and-ar-pushing-connectivity-limits.pdf