A Drink in Mars: an Approach to Distributed Reality

P. Perez & E. Gonzalez-Sosa & R. Kachach & J. J. Ruiz & A. Villegas

Nokia Bell Labs

Abstract

We have developed A Drink in Mars application as a proof of concept of Distributed Reality, a particularisation of Mixed Reality which combines a reality transmitted from a remote place (e.g. live immersive video stream from Mars) with user interaction with the local reality (e.g. drink their favourite beverage). The application shows acceptable immersion and local interactivity. It runs on Samsung GearVR and needs no special green room for chroma keying, thus being suitable to test different use cases.

1. Introduction

In the continuum of Mixed Reality (MR), applications move from Augmented Reality (AR), where a real environment is augmented with virtual objects, to Augmented Virtuality (AV), where the virtual environment is augmented with real elements [Mil09]. In both cases, there is a main environment (either virtual or real) which is augmented with elements from the other.

We are starting to explore the concept of Distributed Reality (DR), an extension of MR where the "virtual" environment, which we call "remote" reality, is an actual multimedia representation of a real place in real time, streamed to the application, where it merges with the "local" reality. As both environments are real, it is not a traditional MR approach where one main reality is augmented with elements extracted from the other, but the blend of two equally important realities into a new "distributed" one.

We have developed A Drink in Mars as a proof-of-concept DR experience. Wearing a VR headset, the user can drink their favourite beverage while immersed in a completely remote reality: the surface of Mars.

2. Mixing local and remote realities

A Drink in Mars has been developed using Oculus GearVR Framework for Android. It runs on a Samsung Galaxy S8 phone with a Samsung Gear VR headset. The VR scene model is simple: the remote reality is projected into a sphere surrounding the user, while the local reality is projected into a plane in front of the user at position $(0,0,-z)$, and attached to the virtual camera rig (Fig. 1).

Local reality is obtained with an AV approach: introducing real-world elements in the remote reality by using the phone camera and object segmentation. This technique provides a good sense of self-perception and it allows for interaction with local objects without needing to remove the headset [MBMSB15]. To increase the visual field, we have attached a fish-eye lens on top of the camera. The optical system (camera + fish-eye lens) is calibrated offline using OpenCV, and the calibration parameters are used to correct the radial and tangential distortion.

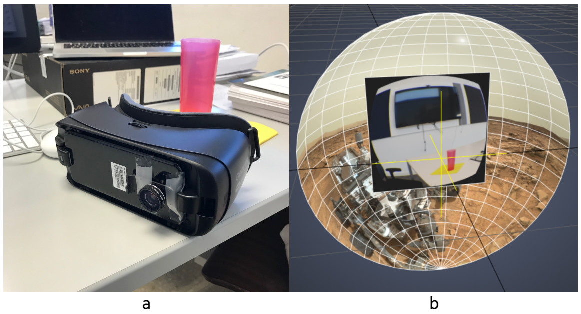

Figure 1: Experimental setup: a) headset with camera and attached fish-eye lense, b) graphical model, with local reality plane in front of remote reality spherical projection.

With respect to deciding which local objects are merged with the remote reality, the traditional approach has used chroma keying over a green background. However, this imposes a severe restriction on the local reality, which must be basically a green room. To overcome it, we have based our solution in YCbCr skin detection [LHL*17]. In particular, we do an alpha blending based on red chrominance:

$$\alpha(Cr) = \begin{cases} 1 & \text{if } Cr \geq C_1 \\ (Cr - C_0)/(C_1 - C_0) & \text{if } C_0 \leq Cr < C_1 \\ 0 & \text{if } Cr \leq C_0 \end{cases} \quad (1)$$

where, for $-128 \leq Cr \leq 127$, we use $C_0 = 8$ and $C_1 = 16$. Those

DOI: 10.2312/egp.20181015

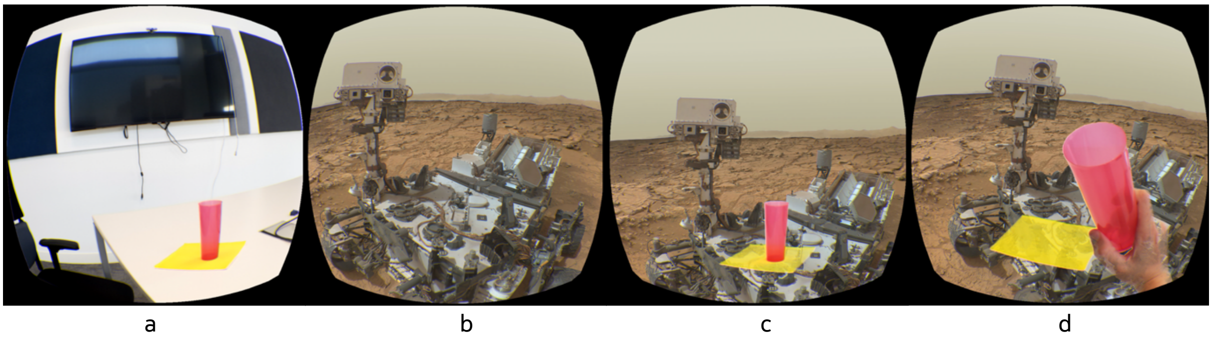

Figure 2: Left eye view of A Drink in Mars application: a) local reality (meeting room); b) remote reality (Mars); c) reality blending; d) user interaction.

thresholds are less strict than what it is normally used for skin detection, but this way the solution provides also a simple methodology to introduce elements in the local reality: red or yellow objects with high saturation will be displayed as local, while blue and neutral tones will be filtered out as background. With this method, it is possible to deploy the system in a normal meeting room in the office, provided that the walls and furniture have black, white, grey and blue tones, instead of needing a full green room. Besides, the simplicity of the technique allows the processing to introduce a delay of less than a frame.

The remote reality texture is obtained by using a modified ExoPlayer, which plays in real-time a 360° video stream. As there is no actual live streaming coming from Mars, we have simulated it from a 360° panorama generated from pictures from the NASA Mars Science Laboratory rover Curiosity. The system has also been tested with a live video stream coming from a Rico Theta V camera, encapsulated in MPEG-DASH and transported over the Internet.

Fig. 2 shows some screens captured from the application (the left eye view). The local reality (Fig. 2.a) is a meeting room, as described before, where the elements that need to be included in the DR are red or yellow (glass, napkin). The background (Fig. 2.b) is the simulated video from Mars. Fig. 2.c shows the blending of both elements, while Fig. 2.d shows the interaction of the user, taking the glass to drink.

Fig. 2.d shows some holes in the segmentation of the hand. Besides, as the video is monoscopic and not completely aligned to the eye position, self-perception is not perfect either. However, informal testing carried out in the laboratory shows that the perception is good enough to interact with the local objects and that, after a couple of minutes of usage, users are able to pick up the glass in a natural way, while keeping a high sense of immersion in the remote reality.

3. Escaping to local

There are situations where the user may need to go back to a fully-local mode (that we call FLM), i.e., setting $\alpha(Cr) = 1 \; \forall Cr$, so that they can interact with the reality without removing the headset. This can be forced by using the headset touchpad, but we also offer an implicit user control based on head motion detection.

The application monitors two Android motion sensors: linear accelerometer and step detector. When the acceleration in any of the axes exceeds a threshold, the application switches to FLM for 5 seconds; while when the system detects that the user is walking (step detector is triggered), it switches to FLM for 15 seconds. The rationale for that is allowing the user to become active (stand up, walk...) in FLM, do whatever they have to do in the room, and then move back to the DR when they sit, relax, and enjoy their drink in Mars.

4. Conclusions

We have implemented a simple system, yet very effective, to test the concept of Distributed Reality. Unlike other MR approaches, DR merges two realities: for instance having a drink (local reality) while visiting Mars (remote reality). With its limitations, our system provides an effective framework to test DR applications, as it includes a very flexible AV-based method to introduce local elements into the remote reality.

In our future work we will use A Drink in Mars to explore the DR concept more in depth. On the one hand, it will be the baseline platform to test new DR use cases, such as bidirectional communication or remote car driving, where there is more interactivity between both realities. On the other, it will be the benchmark for further improvements in the implementation: deep-learning based object segmentation, 3D local camera, implementation in a high-end VR headset, etc.

References

[LHL*17] LIN J.-W., HAN P.-H., LEE J.-Y., CHEN Y.-S., CHANG T.-W., CHEN K.-W., HUNG Y.-P.: Visualizing the keyboard in virtual reality for enhancing immersive experience. In *ACM SIGGRAPH 2017 Posters* (New York, NY, USA, 2017), SIGGRAPH '17, ACM, pp. 35:1–35:2. doi:10.1145/3102163.3102175. 1

[MBMSB15] MCGILL M., BOLAND D., MURRAY-SMITH R., BREWSTER S.: A dose of reality: Overcoming usability challenges in VR head-mounted displays. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (2015), ACM, pp. 2143–2152. doi:10.1145/2702123.2702382. 1

[Mil09] MILGRAM P.: A taxonomy of (real and virtual world) display and control interactions. In *Proceedings of the 16th ACM Symposium on Virtual Reality Software and Technology* (New York, NY, USA, 2009), VRST '09, ACM, pp. 10–10. doi:10.1145/1643928.1643932. 1