Spatial Mobile Memories: Recording and Sharing Everyday Moments using Mobile Augmented Reality

Rishi Vanukuru
Ellen Yi-Luen Do
rishi.vanukuru@colorado.edu
ellen.do@colorado.edu
ATLAS Institute, University of Colorado Boulder
Boulder, Colorado, USA

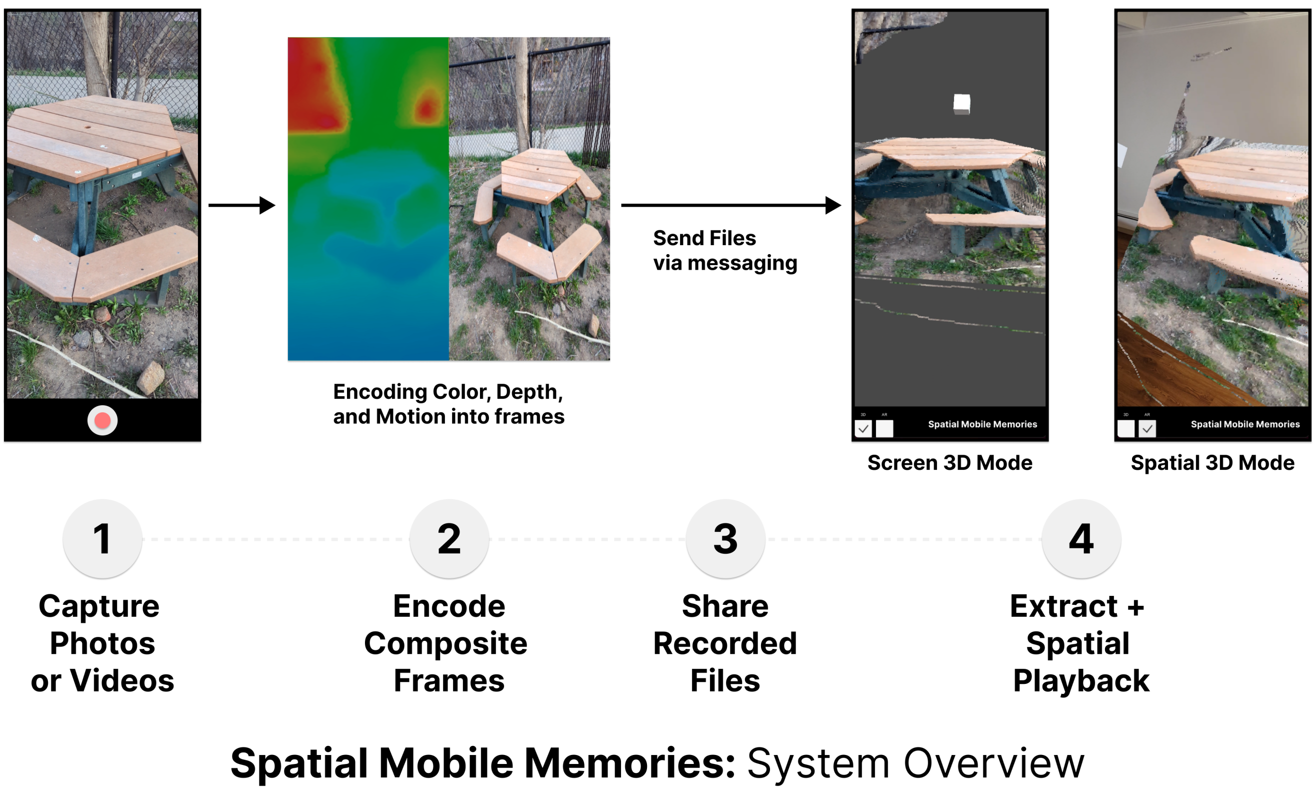

Figure 1: Spatial Mobile Memories - System Overview: (1) Recording photographs or videos using the application. (2) Encoding color, depth, and movement information for each frame, and converting into images or video files. (3) Sharing files over personal communication channels. (4) Viewing spatial recordings via application, on screen or in AR.

ABSTRACT

In this demonstration, we present a system for creating and viewing spatial memories using mobile devices. The concept of interacting with three-dimensional recordings of people and places has been explored using depth cameras and Augmented and Virtual Reality headsets, but these methods require specialized hardware that has not yet reached wider availability. Recent developments in mobile camera technology have presented new capabilities which we leverage in our application for Spatial Mobile Memories. The application captures visual and depth information, together with the orientation and movement of the mobile device. This is then combined into a single standard image or video file. Users can share these files via personal messaging channels, and replay them in 2D and 3D across devices using the same application. The goal of this project is to develop open and widely-accessible systems for the creation

IMX '24, June 12–14, 2024, Stockholm, Sweden

ACM ISBN 979-8-4007-0503-8/24/06
https://doi.org/10.1145/3639701.3661087

and viewing of spatial memories that leverage the ubiquity and familiarity of mobile devices.

CCS CONCEPTS

• **Human-centered computing → Mixed / augmented reality**; **Mobile devices**.

KEYWORDS

Augmented Reality, Mobile Devices, Asynchronous Communication

ACM Reference Format:
Rishi Vanukuru and Ellen Yi-Luen Do. 2024. Spatial Mobile Memories: Recording and Sharing Everyday Moments using Mobile Augmented Reality. In *ACM International Conference on Interactive Media Experiences (IMX '24), June 12–14, 2024, Stockholm, Sweden.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3639701.3661087

1 SPATIAL MEMORIES TODAY

Recording and sharing everyday moments has become an essential part of our everyday lives. We record voice notes, photographs, and videos to help ourselves remember, and to share our lives with the ones we hold close. The growing ubiquity of mobile devices has played a significant role in enabling this, and more recent advancements in camera technology have given rise to the possibility of recording memories with elements of spatial perception and interaction.

The notion of "spatial", "holographic", or "volumetric" communication—being able to interact with three-dimensional representations of people and places—has existed in popular culture for many years. The Holoportation system developed by Orts-Escolano et al. [3] is a compelling example of an early end-to-end pipeline for such kinds of 3D communication. While Holoportation focused on real-time communication, the authors also discussed the idea of recording and playing back "living memories" using the same underlying technology. The system required multiple external depth cameras for capturing volumetric information, and Augmented Reality headsets to view the recorded holograms, making it difficult to scale outside of laboratory environments. One way to achieve wider accessibility is to leverage mobile devices and their ability to support AR experiences. A recent example of this approach is the Memento Player developed by Liu and Ritchie et al. [2]. While the memories in this system are still captured using multiple depth cameras (and AR glasses), they can be played back and experienced in AR using mobile phones. To address the capture challenge, many mobile-based AR systems have explored the use of on-device cameras and depth sensing technology to construct and share spatial content in real-time [1, 4, 5], and some of these techniques have found their way into commercial products as well. Leveraging the multiple cameras and on-board LiDAR sensors in newer-generation iPhones, Apple introduced spatial videos[1], which are stereoscopic videos that can be viewed in AR/VR headsets. Applications such as Wist[2] provide more three-dimensional content capture via iPhones, and enable users to view and walk around recordings of memories using mobile devices and AR/VR headsets.

All of these developments point to the growing potential of mobile devices for recording and sharing spatial memories. However, a few key issues still persist. Even if spatial memories can be *viewed* on mobile devices, many systems still require complex capture setups involving depth cameras and post-processing (such as Memento Player). When on-device capture is available (as is the case with Apple's spatial videos, or the Wist application), the recording capability is limited to specific devices, and the format of the recordings further limits the devices that can play the recordings. Addressing these issues, we present a system for recording and replaying ***Spatial Mobile Memories***.

2 SPATIAL *MOBILE* MEMORIES

We developed our application for Spatial Mobile Memories using Unity 2021.3[3], for both Android and iOS devices. The application (1) supports the capture of spatial photographs and videos across a wide range of mobile devices, (2) encodes visual and spatial information into a single PNG image or MP4 video file, and (3) plays back memories in 2D and 3D modes across devices. Each of these steps are visually explained in Figure 1, and described below.

2.1 Recording

The interface of recording a spatial mobile memory is similar to conventional camera applications, with options to take spatial photographs or videos. Once recording begins, the application collects and encodes depth information of the scene in front of the phone, as well as the position and movement of the phone itself.

2.1.1 Depth. We obtain scene color and depth information using the respective device AR frameworks—Google's ARCore API[4] and Apple's ARKit[5]. For each color pixel, we obtain a corresponding value of depth (either as a true depth value from LiDAR sensors, or an estimate when hardware depth sensors are unavailable). These depth pixel values are used to generate a colorized depth image, where the color of each pixel in the image encodes the depth value of that pixel. We then stitch the color image and the colorized depth image together to create a single composite frame.

2.1.2 Movement. The orientation and movement of the capturing device is crucial information for the spatial playback of content, and this information is not recorded by most conventional camera applications. Our application for spatial mobile memories keeps track of the device orientation with respect to its relative environment by leveraging the AR frameworks' tracking systems. This information is also encoded as color values in pixels of the composite frame.

As a result, each recorded frame, whether as an individual photograph, or a series of frames in a video, consists of information of the color and depth of the recorded scene, and the position and orientation of the capturing device. The application saves these photographs and videos as standard PNG image files or MP4 video files in the devices' media gallery.

[1]"Apple introduces spatial video capture on iPhone 15 Pro", from apple.com/newsroom
[2]Wist: Step Inside Your Memories, from wistlabs.com
[3]Unity: unity.com
[4]Google AR Core developers.google.com/ar
[5]ARKit Framework developer.apple.com

2.2 Sharing

In developing this application, we decided against centrally recording any user information (interaction logs, user recordings). As a result, the application does not directly control file storage or manage sharing. As the spatial videos and photographs are saved in standard file formats directly in the media gallery, users can share these files via any personal communication channel, and once received, these files can be played back using our application.

2.3 Playback

In playback mode, users can open specific images and videos from their media gallery in our application. The application extracts the color, depth, and movement information from the files, and presents multiple options for playback: (1) ***2D mode***, where users see a standard photograph or video on their screen, (2) ***Screen-based 3D mode***, where users see a point-cloud representation of the recording that they can inspect by interacting with the device screen, and (3) ***Spatial 3D mode***, where the recorded point cloud representation is displayed in the user's local environment in AR. In the case of a photograph, the 3D representation is anchored in space and projected from the relative orientation of capture of the photograph, while in the case of a spatial video, the 3D representation moves based on the recorded movement of the device. The user can independently walk around and inspect the representation from different viewpoints. At present, the resolution of the point-cloud representation is low, but we anticipate that the visual fidelity will significantly increase with improvements in the colorization and encoding process.

3 NEXT STEPS

Our application for Spatial Mobile Memories presents users with a pipeline to record spatial photographs and videos, and replay them in a range of formats across devices. In doing so, the application addresses many key challenges faced by existing systems for spatial recordings. In this demonstration, we will enable participants to view prerecorded spatial memories using our application, as well as record spatial photographs and videos of their own in the conference venue. We also aim to release the application for general use, and will create documentation for potential users who might want to try creating their own spatial recordings. We hope to invite users to share their comments and experience of using the application with us, and will use this feedback to further refine the design of the application. The overarching goal here is to develop open and widely-accessible systems for the creation and viewing of spatial memories that leverage the ubiquity and familiarity of mobile devices.

REFERENCES

[1] Yuan Li, Sang Won Lee, Doug A. Bowman, David Hicks, Wallace santos Lages, and Akshay Sharma. 2022. ARCritique: Supporting Remote Design Critique of Physical Artifacts through Collaborative Augmented Reality. In *Proceedings of the 2022 ACM Symposium on Spatial User Interaction (SUI '22)*. Association for Computing Machinery, New York, NY, USA, Article 10, 12 pages. https://doi.org/10.1145/3565970.3567700

[2] Yimeng Liu, Jacob Ritchie, Sven Kratz, Misha Sra, Brian A. Smith, Andrés Monroy-Hernández, and Rajan Vaish. 2023. Memento Player: Shared Multi-Perspective Playback of Volumetrically-Captured Moments in Augmented Reality. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI EA '23)*. Association for Computing Machinery, New York, NY, USA, Article 205, 9 pages. https://doi.org/10.1145/3544549.3585588

[3] Sergio Orts-Escolano, Christoph Rhemann, Sean Fanello, Wayne Chang, Adarsh Kowdle, Yury Degtyarev, David Kim, Philip L. Davidson, Sameh Khamis, Mingsong Dou, Vladimir Tankovich, Charles Loop, Qin Cai, Philip A. Chou, Sarah Mennicken, Julien Valentin, Vivek Pradeep, Shenlong Wang, Sing Bing Kang, Pushmeet Kohli, Yuliya Lutchyn, Cem Keskin, and Shahram Izadi. 2016. Holoportation: Virtual 3D Teleportation in Real-time. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 741–754. https://doi.org/10.1145/2984511.2984517

[4] Rishi Vanukuru, Suibi Che-Chuan Weng, Krithik Ranjan, Torin Hopkins, Amy Banic, Mark D. Gross, and Ellen Yi-Luen Do. 2023. DualStream: Spatially Sharing Selves and Surroundings using Mobile Devices and Augmented Reality. In *2023 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. 138–147. https://doi.org/10.1109/ISMAR59233.2023.00028

[5] Jacob Young, Tobias Langlotz, Steven Mills, and Holger Regenbrecht. 2020. Mobileportation: Nomadic Telepresence for Mobile Devices. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 4, 2, Article 65 (jun 2020), 16 pages. https://doi.org/10.1145/3397331