ARphy: Managing Photo Collections Using Physical Objects in AR

Han Joo Chae
Seoul National University
Seoul, Republic of Korea
hanjoo@hcil.snu.ac.kr

Gwanmo Park
Seoul National University
Seoul, Republic of Korea
gmpark@hcil.snu.ac.kr

Youli Chang
Seoul National University
Seoul, Republic of Korea
ylchang@hcil.snu.ac.kr

Jinwook Seo
Seoul National University
Seoul, Republic of Korea
jseo@snu.ac.kr

Minji Kim
Seoul National University
Seoul, Republic of Korea
mjkim@hcil.snu.ac.kr

CHI '20 Extended Abstracts, April 25–30, 2020, Honolulu, HI, USA.

ACM ISBN 978-1-4503-6819-3/20/04.
http://dx.doi.org/10.1145/3334480.3382885

Abstract
ARphy is a tangible interface that extends current ways of organizing photo collections by enabling people to interact with digital photos using physical objects in Augmented Reality. ARphy contextually connects photos with real objects and utilizes physical affordances so that people can add more meanings to their collections and interact with them naturally. We also created an ARphy Interaction Design Toolkit, which can add ARphy-compatible interactions to any object, so that people can register their own things for organizing collections. We developed a prototype using seven everyday objects and evaluated ARphy through a qualitative user study. Our findings indicate that ARphy is intuitive, immersive, and enjoyable and has the potential for selectively managing collections using photos and objects that have personal meanings.

Author Keywords
Photo interface; tangible interaction; augmented reality.

CCS Concepts
•**Human-centered computing** → **Mixed** / **augmented reality; User interface design;** *User studies;*

Introduction and Background
Photo collections are commonly used as a means to remember precious moments. While such groups of photos

Figure 1: An example collection created with ARphy using photos and souvenirs. A Blue Jays fan wanted to commemorate his first visit to an MLB game.

are themselves valuable, contextual meanings could add more value to each photo collection. For example, a bunch of photo frames decorating a living room along with other related things like gifts and souvenirs have always been an effective way for family members and friends to share and recall their memories. However, most current photographic experiences are separated into two different worlds: digital and physical. Although merging them is already technologically possible, digital photos are generally disconnected from printed ones as well as other related objects; we easily experience intensive scrolling on a smartphone searching for relevant photos related to real-world objects or events.

Moreover, photos only in digital form lack the tangibility and manipulability of physical photos—aspects many people consider important [1]—and they are often perceived as less valuable than physical photos because they are not persistently encountered [7]. Many efforts have been made to address these issues by materializing digital photos or providing ways to connect physical objects with digital photos. For instance, Odom et al. [6] randomly printed photos once in a while to promote anticipation and re-visitation of the past, and Nunes et al. [5] used physical memorabilia to trigger memories and enrich storytelling.

We believe AR technology can extend current ways of managing photo collections by allowing people to spatially organize and tangibly manage them along with related physical objects in situ. However, many photo applications in AR still use traditional grid-style user interfaces similar to those on mobile devices and existing AR-specific approaches mostly focus on spatial placement of photos with only providing fragmented examples [2, 4, 8]. Tighter contextual integration between digital photos and physical objects as well as utilization of physical affordances and tangible interactions could further enrich personal photo experiences.

In this paper, we present *ARphy*, a tangible AR interface that allows people to create and manage their own photo collections by spatially and temporally connecting digital photos with real objects (Figure 1). In addition, ARphy takes advantage of tangible interaction and affordances for more natural usages. For example, people can attach travel photos to a souvenir, arrange photos by date on a wall calendar, or delete photos by putting them into a trash can. We also create an ARphy Interaction Design Toolkit (AIDT), which can add ARphy-compatible interactions to physical objects so that users can register their own things to ARphy, and develop a proof-of-concept prototype with seven everyday objects. Lastly, we evaluate ARphy through a qualitative user study and discuss our major findings on managing photo collections along with real objects in AR.

ARphy: A Tangible Photo Interface

ARphy is a tangible AR interface that allows people to create photo collections along with real objects. ARphy is not intended to replace existing photo applications but to provide people more options for managing photo collections depending on their personal scenarios and needs.

To track and communicate with real objects, ARphy adopts the concept of *smart objects* from cooperative augmentation [4], whereby smart objects describe themselves and their sensor events. Since the cooperative augmentation was designed for projector-camera systems, we enhanced it to fit portable AR devices (e.g., HMDs), as in Figure 2. By abstracting each module into a service, we enable ARphy to be used with any smart object on various AR devices.

ARphy visualizes digital photos as 3D objects (i.e., 2D photos as thin cuboids and 360-degree panoramic photos as spheres) so that users can freely manipulate them in 3D space and connect them to any smart object. Two differ-

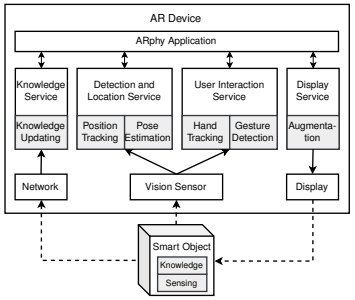

Figure 2: An overview of ARphy's architecture. Smart objects transmit their knowledge to an AR device over the network, and the *knowledge service* keeps the information received from each object up-to-date. The *detection and location service* and *user interaction service* utilize vision sensors to track objects and detect user gestures, respectively. While the current design uses hand gestures on AR HMDs like the HoloLens, it can easily be extended to other input types, such as external controllers and touch gestures. The *display service* augments the scene with visual artifacts based on the geometry of each object and the environment.

ent directions can be used to connect or disconnect photos and objects depending on the source—either a photo or object—of interaction. For instance, photos can be sources of interaction ($P \rightarrow O$) and dragged into or out of objects to connect or disconnect, respectively. In contrast, objects can also be sources of interaction ($O \rightarrow P$) and tangibly moved into or out of photos. These directions should be carefully selected considering the purpose of the interaction as well as the physical affordance of the object. For instance, ($P \rightarrow O$) might feel natural when a user wants to attach photos to an object (Figure 3a), and ($O \rightarrow P$) might be useful when a user wants an object to change its appearance based on where one puts it (Figure 3d).

Once the connection is made, users can interact with photos and objects based on the post-connection interaction configured for each object. ARphy currently provides four different post-connection interactions that were designed to reflect common real-life photo managing behaviors.

Attaching
Attaching allows users to place photos anywhere near objects. This basic interaction is crucial for ARphy because it not only reflects common photo managing practices, such as clipping printed photos on a picture board or placing them near related objects like souvenirs, but also enables users to add spatial contexts to their collections. For instance, one might arrange photos around an Eiffel tower statue based on the relative location from which each photo was taken. Once a photo is attached to an object, it always stays in that position and orientation relative to the object. Detaching simply disconnects photos from objects so that they become completely independent from each other.

Embedding
Inspired by printing photos on the surface of an object (e.g., mugs and t-shirts), embedding allows users to texturize the outer surfaces of objects with photos. The appearance of a photo is determined by the shape of the surface onto which the photo is embedded. For instance, photos embedded on a music box will rotate following the movement of the music box (Figure 3a), and those on an elastic fabric will be scaled when the fabric is stretched (Figure 3f).

Stacking
Stacking enable users to naturally organize digital photos in physical containers like boxes along with other objects (Figure 3b). For a more realistic experience, we added gravitational effect inside the containers so that users can not only stack up the photos but also physically shake the containers to move those inside; they will roll and slide accordingly.

Automatic Arrangement
Instead of manually setting the position and appearance of each photo, objects can be configured to arrange photos automatically; when users insert photos to the objects, the objects decide the appearance of the inserted photos based on their layouts and ordering mechanisms. While various layouts and metadata can be used, ARphy currently supports calendar and map layouts using date and location data, respectively. The calendar layout is intended to be used with physical calendars. When users specify a calendar area with year and month, thumbnails of representative photos for each date are overlaid on top of the corresponding date numbers on the calendar so that users can have a monthly overview of important photos. Similarly, the map layout can be used with printed maps to arrange photos using GPS location.

Action Triggering
Objects can act as *physical icons* and trigger functions that are related to their affordances. As drag-and-drop interactions are already ubiquitous, it is not too difficult to expect certain actions to be triggered when physical objects and

(a) Music box **(b)** Box container **(c)** Photo book **(d)** Mug **(e)** Calendar **(f)** Elastic fabric **(g)** Trash can

Figure 3: Examples of ARphy using seven real objects. The first row shows the original objects without any visual augmentation and the second and third rows demonstrate how photos can be connected to and interacted with each object. (a) Photos attached around a music box and embedded on its body and virtual flag can be viewed as it rotates with music playing. (b) Photos can be stacked inside a box container and viewed by looking inside, grabbing individual photos, or shaking the box. (c) Photos in a photo book can be viewed by turning each page. (d) A mug's face is texturized when it is tangibly moved to a photo. (e) A calendar arranges photos by date and displays thumbnails overlaid on the dates of the calendar. More details can be viewed when thumbnails are clicked. Custom visualizations such as a year ago and monthly/annual highlights are also shown on the left of the calendar. (f) A photo embedded on an elastic fabric can be scaled by pulling each side of the fabric and making it stretch. (g) Photos can be deleted by dragging them into a trash can.

digital content collide with each other. For instance, interactions like touching a photo with an eraser or putting a photo into a trash can could delete the photo (Figure 3g).

ARphy Interaction Design Toolkit (AIDT)

In order for each smart object to be able to interact with photos, ARphy-compatible interactions must be predefined. AIDT is created for this specific purpose; it consists of Unity prefabs and C# scripts that provide a high-level interface for constructing photo interactions for ARphy in a Unity scene. At this point, AIDT is targeted for deployment on HoloLens and uses Vuforia Engine [9] to track smart objects. With AIDT, users can author interactions in two steps: (i) select an area or surface to add interactivity and (ii) add interactions to the area or surface including the direction. These steps can be repeated as many times as needed if users want to have multiple interactive areas and surfaces on a single object. While AIDT provides scripts of predefined interactions for users without programming knowledge to add basic ARphy interactions to their objects, it is also highly scalable and allows advanced users to build custom interactions by writing their own scripts.

Adding Post-Connection Interactions
Users first need to specify the part of the object to which they want to add interactivity. In Unity, this is done simply by selecting a GameObject that contains a 3D mesh. By default, basic primitive mesh shapes like a cube, sphere, cylinder, and plane can be used, but custom meshes can also be used. Then, users can drag-and-drop one of the scripts with predefined interactions to the selected GameObject. Currently available interactions are *attach*, *embed*, *stack*, *calendar*, and *action trigger*. After adding the script to the GameObject, users need to configure them including the direction of interaction. If the GameObject does not have colliders defined, AIDT generates one on the fly appropriate for the mesh and interaction.

Prototype and Example Applications
As a proof of concept, we selected seven representative everyday objects (i.e., a music box, box container, photo book, mug, calendar, fabric, and trash can) and mapped their physical attributes to basic photo interactions, as shown in Figure 3. All the interactions were designed and developed using AIDT considering each object's affordance, interaction methods, and corresponding visualizations.

User Evaluation

We conducted a qualitative user study to evaluate ARphy and investigate how people feel about using real objects and their physical affordances for managing photos in AR. Microsoft Photos [3], a basic photo application with a 2D-grid user interface, was used on the HoloLens as a baseline and compared with ARphy in photo managing interactions. Five example objects (i.e., the music box, photo book, box container, calendar, and trash can) were used with ARphy during the user study.

Study Design
We designed four tasks to compare basic photo interactions using ARphy and the Photos app. As a within-subjects design study, each participant used both interfaces, and the order of the interfaces was counterbalanced. After completing each task, the participants subjectively evaluated each interface using seven-point Likert scales on the following measures: ease of learning, ease of use, enjoyment, fatigue, intuitiveness, immersiveness, and satisfaction. Brief interviews were held in between tasks and after all the tasks were completed.

Task 1: We asked the participants to view the given collection using each interface and describe each photo. The purpose of the task was to let the participants see how ARphy's collections can be viewed and interacted with. When using ARphy, the participants connected the given photos to the music box and viewed them as it rotated with the music playing.

Task 2: The participants performed actual photo managing interactions (i.e., removing and adding photos from and to a collection, respectively) during this task. For ARphy, the box container was used to manage collections and the trash can for deleting individual photos.

Task 3: To demonstrate the photo book as another example of ARphy, we gave the participants the photo book with photos already connected and asked them to view the photos by turning each page.

Task 4: The last task was navigating photos that were automatically arranged by date. The calendar, which was used for ARphy, displayed thumbnails overlaid on the dates of the calendar, while the Photos app arranged the photos from newest to oldest so that users could scroll to navigate. We

Task 1	ARphy	Photos
Enjoyable	6.0	4.0
Immersive	5.5	4.5

Task 2	ARphy	Photos
Enjoyable	6.0	4.0
Intuitive	6.0	5.0
Immersive	5.5	4.0

Task 4	ARphy	Photos
Enjoyable	5.0	4.0
Satisfying	5.5	4.0

Table 1: Median scores of the subjective evaluation. Only the ones with the statistical significance are shown.

asked the participants to find the photos taken on a given date and describe them.

Results

We recruited 12 participants (7 men and 5 women) from a local university between the ages of 20 and 32 ($\mu = 24, \sigma = 4$), and they all had normal or corrected-to-normal vision.

We analyzed the participants' subjective evaluations of both ARphy and Microsoft Photos for each task using the Wilcoxon signed-rank test. Since there were no comparisons during Task 3, we only analyzed the results of Tasks 1, 2, and 4. For viewing albums in Task 1, ARphy was both more enjoyable ($Z = -2.54$, $p = 0.011$) and more immersive ($Z = -2.33$, $p = 0.020$) than Photos. For managing albums in Task 2, ARphy was more enjoyable ($Z = -2.72$, $p = 0.007$), intuitive ($Z = -2.39$, $p = 0.017$), and immersive ($Z = -1.97$, $p = 0.048$) than Photos. For navigating the photos by date in Task 4, ARphy was more enjoyable ($Z = -2.04$, $p = 0.041$) and satisfying ($Z = -2.85$, $p = 0.004$) than Photo. No statistically significant differences were found in the other measures and the median scores for the above analysis are reported in Table 1.

Discussion

The results of the subjective evaluation indicate that compared to the conventional 2D-grid-based photo app, ARphy provided a more immersive, intuitive, and enjoyable experience in general while it was similarly easy to learn and use. Many participants expressed positive reactions about ARphy's interactions using physical objects with affordances and visual effects. P8 said, *"It was like putting things in a box to store"*; and P5 said, *"It was amusing to see the photos rotating with the orgel."* Also, many expressed their desire to use ARphy on special occasions, as P1 said, *"I want to use it when I want to remember."* Moreover, we discovered that ARphy could add new value to objects—as P1 mentioned, *"I will be curious about everything"*—thereby potentially increasing the user's attachment to the objects and collections. These findings align well with our motivation to support selectively managing collections with meaningful photos and objects.

Though our evaluation positively demonstrated ARphy for managing photo collections, it did not use the participants' personal photos and objects and social interactions were missing. To evaluate how those factors could affect ARphy's experience, personal and social scenarios, where people use their own photos and objects while communicating with family and friends, need to be further investigated.

Conclusion

In this paper, we presented ARphy, which enables users to contextually connect photos with real objects and utilize physical affordances to organize photo collections in AR. We also showed how AIDT can add ARphy-compatible photo interactions to objects and demonstrated seven examples designed and developed with AIDT. The user evaluation revealed that ARphy is intuitive, immersive, and enjoyable and has the potential for selectively managing collections using photos and objects with personal meanings.

Acknowledgments

This work was supported by Samsung Electronics Co., Ltd.

REFERENCES

[1] David Frohlich, Allan Kuchinsky, Celine Pering, Abbe Don, and Steven Ariss. 2002. Requirements for Photoware. In *Proceedings of the 2002 ACM Conference on Computer Supported Cooperative Work (CSCW '02)*. ACM, New York, NY, USA,

166–175. DOI: http://dx.doi.org/10.1145/587078.587102

[2] Asier Marzo and Oscar Ardaiz. 2013. CollARt: A Tool for Creating 3D Photo Collages Using Mobile Augmented Reality. In *Proceedings of the 21st ACM International Conference on Multimedia (MM '13)*. ACM, New York, NY, USA, 585–588. DOI: http://dx.doi.org/10.1145/2502081.2502154

[3] Microsoft. 2019. Microsoft Photos. (2019). Retrieved March 29, 2019 from https://www.microsoft.com/en-us/p/microsoft-photos/9wzdncrfjbh4?activetab=pivot:overviewtab.

[4] David Molyneaux, Hans Gellersen, and Joe Finney. 2013. Cooperative Augmentation of Mobile Smart Objects with Projected Displays. *ACM Trans. Interact. Intell. Syst.* 3, 2, Article 7 (Aug. 2013), 35 pages. DOI: http://dx.doi.org/10.1145/2499474.2499476

[5] Michael Nunes, Saul Greenberg, and Carman Neustaedter. 2008. Sharing Digital Photographs in the Home Through Physical Mementos, Souvenirs, and Keepsakes. In *Proceedings of the 7th ACM Conference on Designing Interactive Systems (DIS '08)*. ACM, New York, NY, USA, 250–260. DOI: http://dx.doi.org/10.1145/1394445.1394472

[6] William T. Odom, Abigail J. Sellen, Richard Banks, David S. Kirk, Tim Regan, Mark Selby, Jodi L. Forlizzi, and John Zimmerman. 2014. Designing for Slowness, Anticipation and Re-visitation: A Long Term Field Study of the Photobox. In *Proceedings of the 32Nd Annual ACM Conference on Human Factors in Computing Systems (CHI '14)*. ACM, New York, NY, USA, 1961–1970. DOI: http://dx.doi.org/10.1145/2556288.2557178

[7] Daniela Petrelli and Steve Whittaker. 2010. Family memories in the home: contrasting physical and digital mementos. *Personal and Ubiquitous Computing* 14, 2 (01 Feb 2010), 153–169. DOI: http://dx.doi.org/10.1007/s00779-009-0279-7

[8] Kari Rainio, Petri Honkamaa, and Kaisa Spilling. 2015. Presenting Historical Photos using Augmented Reality. *ARea* 15 (2015), 11–12.

[9] Vuforia. 2019. Vuforia Engine Developer Portal. (2019). Retrieved March 28, 2019 from https://developer.vuforia.com/.