Urashima Tarō Mirror: An Interactive Portrait of the Fictional Self in Time

Sho Fujimura
cy20250@shibaura-it.ac.jp
Shibaura Institute of Technology
Tokyo, Japan

Yuta Tsuchida
cy22260@shibaura-it.ac.jp
Shibaura Institute of Technology
Tokyo, Japan

Kaito Makino
cy20238@shibaura-it.ac.jp
Shibaura Institute of Technology
Tokyo, Japan

Daisuke Uriu
uriu@shibaura-it.ac.jp
Shibaura Institute of Technology
Tokyo, Japan

Figure 1: Experiencing the Urashima Tarō Mirror: The portraits displayed appear older as you approach the mirror (left), and younger as you move away (right). Although these three images are shown separately, the experience feels continuous, allowing users to appreciate the gradual transformation of their face without interruption.

Abstract

This article presents the Urashima Tarō Mirror (UTM), an interactive installation that allows users to visually experience self-transformation associated with aging through AI-generated portraits. UTM generates a series of portraits ranging from childhood to old age based on the user's facial photograph, and the displayed image changes—appearing younger or older—depending on the user's distance from the mirror. Through this mirror-like interface, users can experience in a short moment the physical changes that would normally take years to occur. This experience evokes awareness of one's own transformation and aging, offering an experience that encourages reflection on identity and self-acceptance.

CCS Concepts

• **Human-centered computing → Human computer interaction (HCI)**.

Keywords

Portrait, Photography, Mirror, Interaction Design, Generative AI

TEI '26, March 08–11, 2026, Chicago, IL, USA

ACM ISBN 979-8-4007-1868-7/2026/03
https://doi.org/10.1145/3731459.3779341

ACM Reference Format:
Sho Fujimura, Yuta Tsuchida, Kaito Makino, and Daisuke Uriu. 2026. Urashima Tarō Mirror: An Interactive Portrait of the Fictional Self in Time. In *Twentieth International Conference on Tangible, Embedded, and Embodied Interaction (TEI '26), March 08–11, 2026, Chicago, IL, USA.* ACM, New York, NY, USA, 6 pages. https://doi.org/10.1145/3731459.3779341

1 Introduction

Looking at yourself in the mirror is part of everyday life, while seeing a portrait of yourself is usually reserved for special occasions—such as passport applications, CV profiles, or sharing selfies on social media. A mirror always reflects your actual, present appearance, whereas a portrait often presents an idealized version. Before photography was invented, portraits were painted by artists. Today, anyone can take a selfie with a smartphone. Unlike mirrors, which simply reflect your current face, portraits can be edited and styled to suit your preferences. This everyday understanding of mirrors and portraits is sometimes reinterpreted or transformed in fictional stories. Now, with AI integrated into our daily lives, we can even create new fictional experiences through collaboration with AI.

We propose the Urashima Tarō Mirror (UTM), a device that allows users to browse and experience AI-generated portraits of themselves, ranging from childhood to old age. As shown in Figure 1, the user stands in front of the UTM, which captures an image of their face and displays a corresponding portrait. When the user moves closer to the UTM, progressively older versions of their face appear. As the user steps back, the portraits gradually become younger. Although we call the UTM a "mirror," it offers an unreal, imaginative experience—much like the story of *Urashima Tarō* [10], in which the main character suddenly becomes old at the end.

Technically, the experience with the UTM is supported by AI-generated images, image morphing techniques, and an interactive design using a mirror-like interface. The system generates a sequence of portraits—ranging from childhood to old age—based on the user's current facial photograph. These portraits are integrated into a continuous morphing animation. A display with a built-in camera detects the distance between the user and the device and adjusts the displayed portrait accordingly. The closer the user stands, the older the portrait becomes; the farther they move, the younger it appears. This synthesis creates an entirely new way of encountering one's own portraits—an experience that is largely uncontrollable yet deeply reflective, and at times even unsettling.

Through this experience, users can visually recognize their own "change" by viewing portraits that represent different stages of life. Although humans are constantly changing, such transformations, especially those related to aging, are often too gradual to be consciously perceived. UTM offers an opportunity to make otherwise imperceptible changes visible through AI-generated portraits, experienced within a interactive interface. While it usually takes decades to experience such changes in appearance, UTM allows users to witness the passage of time within a short interactive experience. It functions as a "mirror that visualizes the passage of time," serving as a means of self-management and identity affirmation. Furthermore, confronting one's changing appearance becomes a ritual-like experience of "accepting one's own transformation." Through interactive control based on forward and backward bodily movement, the UTM system allows users to finely adjust the degree and speed of transformation, enabling an exploratory experience that goes beyond passive portrait viewing.

In the following sections, we present related work and the design details of the UTM, along with its technical implementation. We also report on participants' experiences with the UTM during a public exhibition and offer a brief discussion of the findings derived from this event. Throughout this article, we aim to explore the types of experiences, thoughts, and emotions that emerge in users as they interactively engage with their own transformation through portraits generated by the UTM.

1.1 GenAI Usage Disclosure

Generative AI tools were used in two distinct ways in this work. First, AI-based image generation (e.g., StyleGAN) was used as part of the research to generate portrait images for the proposed system, UTM. Second, ChatGPT (OpenAI, GPT-5 and later) was used to assist in improving the clarity and grammar of the English manuscript. All creative concepts and system integration of UTM were entirely developed by the authors.

2 Related Work

A wide range of experiences have been created by leveraging AI-based image generation technologies such as facial synthesis, portrait transformation, and morphing. For example, Shimizu et al. [8] designed Morphing Identity, in which the faces of two users gradually blend into each other. Pat et al. [7] developed a system called Future You, which generates a future portrait of the user and allows them to engage in a dialogue with their future self through chat. Martins et al. [5] proposed an interactive installation that generates and projects new portraits by combining the facial features of viewers, allowing them to experience the ambiguity of authenticity at the boundary between the real and the artificial. Several researchers [2, 6] have also proposed methods specifically focused on age progression in portraits, which is the central theme of this study.

Regarding the experience of viewing a portrait, Yoshida et al. [11] reported that applying artificial changes—such as smiles or sad expressions—to a user's real-time facial image can evoke emotional responses. Focusing on the experience of viewing portraits of the deceased, Uriu and Odom [9] proposed interaction designs such as gazing into a mirror or lighting a candle to create reflective moments. As a study focusing on the experience of viewing one's future portrait, Hershfield et al. [2] examined whether encountering an aged version of oneself in virtual reality would promote future-oriented decision-making, and found that participants who interacted with their future selves were more likely to choose delayed rewards over immediate ones.

This study creates a novel experience that allows users to perceive the age progression of their own portrait both visually and physically, through a combination of AI-generated imagery and an interaction design based on bodily movement—approaching or stepping away from the mirror. While the individual technologies used have been previously proposed, the interactive portrait frame of the UTM affords a unique user experience through its integrated and embodied combination of these elements.

3 Design

3.1 Experience

The UTM system enables each user to visually and physically experience their own "temporal transformation." UTM displays portraits generated by AI based on the user's facial photograph. When a user approaches the installation, the displayed portrait gradually changes into an elderly version, and when the user moves away from the installation, the portrait rejuvenates, shifting toward a childhood appearance. The UTM experience can serve as a means for users to reflect on their self-management and identity. Normally, recognizing changes in personal appearance requires several decades, but UTM compresses this temporal scale into a short interactive experience. This compressed experience allows users to observe the process of aging and rejuvenation in real time and to reflect on how they have changed and how they may continue to change in the future.

Such embodied temporal experiences offer a unique value that existing age-filter applications do not provide. Image-based age filters produce similar visual transformations, but the transformations are applied automatically whenever the user's face is captured, making the transformation a purely "screen-based" one with limited interactivity. In contrast, the UTM system generates temporal progression and regression through the user's physical movement. This bodily operation presents age transformation not as a simple visual effect but as an exploratory experience in which each user actively engages with their changing portraits.

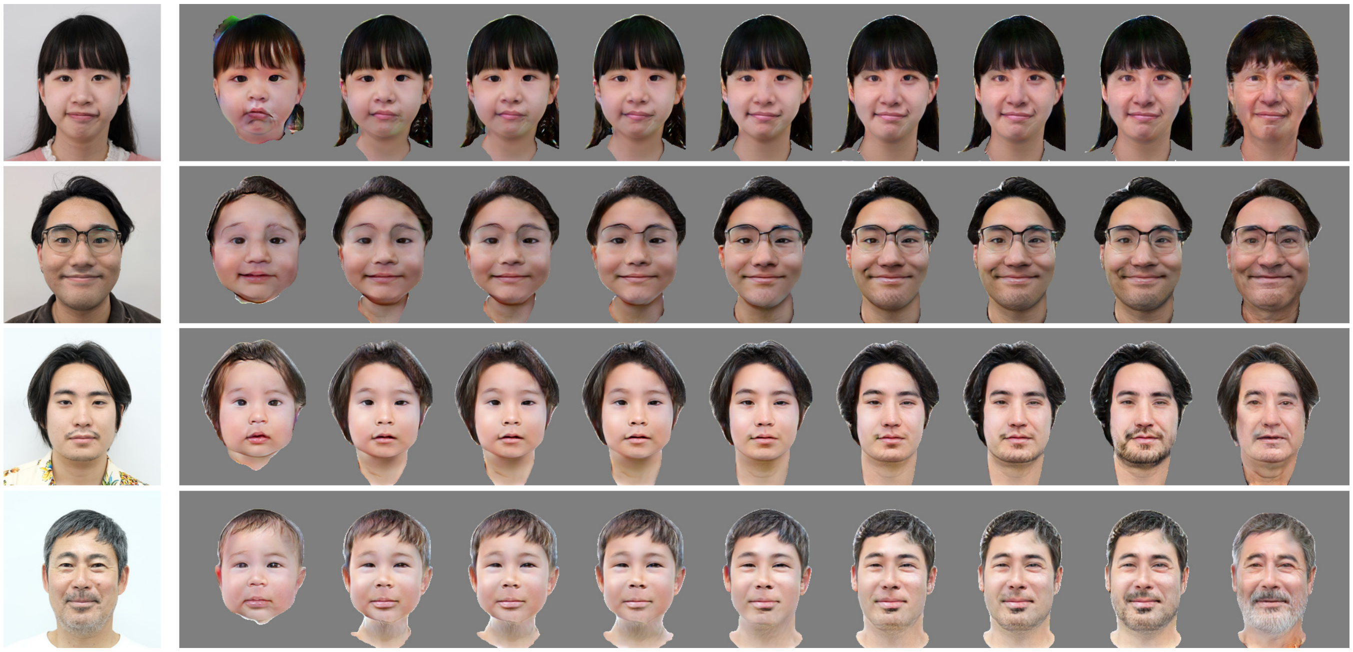

Figure 2: Generated portraits of four individuals. Left: the original input image. Right: nine representative AI-generated portraits selected from a set of 100 age-progressed images produced from a single input image, arranged from younger (left) to older (right).

The detailed experience with UTM consists of two phases: taking a photograph and interactively viewing the generated portraits. The experience begins when the user photographs their own face (Phase 1, Figure3-right). After several tens of seconds of processing, the user stands in front of the UTM installation and views a series of AI-generated portraits (Phase 2, Figure3-right). The user may retake the photograph at any time, and UTM generates a new set of portraits based on the updated image. The generated portraits vary according to visual features such as the user's facial expression, head angle, hairstyle, and makeup.

Once the portrait-generation process is completed, the UTM system displays portraits generated from the user's face photograph. Through forward and backward body movements, the user controls the temporal progression and regression by changing the displayed portraits. When the user approaches the display, the portrait gradually transitions into an elderly appearance, and when the user moves away from the display, the portrait rejuvenates and shifts toward a childhood appearance.

This interaction, in which "moving closer results in aging and moving away results in becoming younger," was inspired by the seek-bar interaction commonly used in video playback interfaces. Moving the slider forward advances the video's timeline, while moving it backward rewinds it. In UTM, we explored whether this idea of "controlling time through spatial manipulation" could be translated into a bodily form of interaction.

Furthermore, based on the cognitive metaphor that interprets the user's forward and backward bodily movements as the progression and regression of time, we designed a mapping between distance and age. Boroditsky [1] stated that abstract domains such as time appear to be structured through metaphorical mappings from more concrete and experiential domains such as space. In Boroditsky's paper, two types of temporal metaphors are introduced. One of them is the Ego-moving metaphor, in which the "self" or observer moves forward along the time axis toward the future. The interaction implemented in this study was also inspired by this Ego-moving understanding of time. Based on this understanding, the system is designed so that approaching the screen reveals the future self (an aged appearance), while moving away reveals the past self (a younger appearance).

3.2 Technical Processing

The technical processing of the UTM system involves three main phases: generating portrait images, creating a morphing video, and implementing an interactive interface.

Initially, this system generates a series of portrait images ranging from childhood to old age based on a single input photo. It employs the method proposed by [6], trained on the FFHQ dataset [4], which realistically simulates age progression from 0 to 70 years. In addition to conventional changes in skin texture, such as wrinkles, this method also captures structural changes in the face, including variations in head shape, enabling more lifelike representations of aging.

As shown in Figure 2, the system generates a set of 100 age-progressed portraits from a single input image. These portraits are arranged in chronological order, illustrating the gradual transformation of the face over time. For clarity, the figure shows nine representative portraits selected from the full set, shown alongside the original input photo.


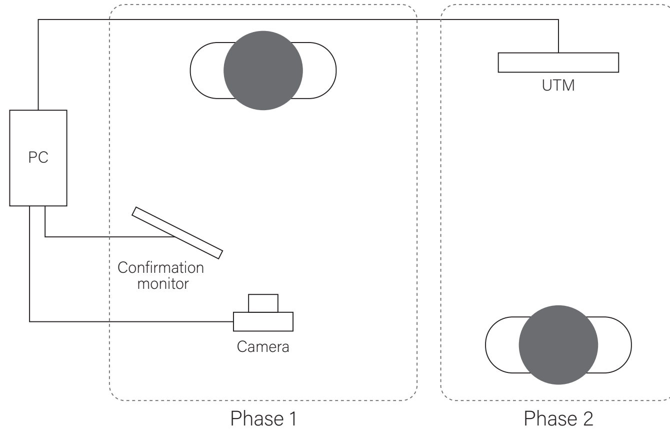

Figure 3: The UTM exhibition at ENDEX Japan 2025. Left: a participant experiencing UTM. Right: the UTM installation setup.

Secondly, the generated portraits are combined into a morphing video that presents a continuous and smooth transformation of the face over time. The system creates seamless transitions between age stages, enabling viewers to experience the aging process as an uninterrupted visual flow.

Finally, an interactive interface allows users to engage with the generated portraits in a dynamic way. A small camera is embedded in the frame of the display monitor, which displays the portraits and detects the face of the user standing in front of it. The system is implemented in Python and uses Google MediaPipe for face detection. By measuring the width of the detected face in the camera feed, the system estimates the user's distance: a larger face indicates closer proximity, while a smaller face suggests the user is farther away. Based on this estimated distance, the system controls the playback position of the morphing video—starting from the beginning when the user is farther away and progressing toward the end as the user moves closer. In this way, the video dynamically responds to the user's movement, enabling an interactive and embodied viewing experience.

3.3 Limitation

The quality of generated images in our current system is limited by the FFHQ dataset [4]. For example, the authors and their colleagues—many of whom are Asian—sometimes find the results unsatisfactory, as the model appears to be trained primarily on Western-centric facial features. While users can compare generated portraits with their real childhood photos, it is, of course, impossible to verify the accuracy of the aged images, since future appearances are inherently unknown.

In this system, the transformed facial images are presented as static portraits. The current prototype faces technical and computational limitations that make real-time tracking and transformation difficult to perform reliably. Nevertheless, using static portraits is effective for clearly illustrating the correspondence between forward–backward bodily movement and age-related changes.

Despite these limitations, we believe that the experience of witnessing the transformation of one's own portrait over a lifetime offers a uniquely reflective and engaging form of interaction. The core experience of the UTM lies in the process of encountering how one's face changes across time—fictional, yet meaningful.

4 Participants' Experience at ENDEX Japan 2025

4.1 Settings

We exhibited UTM at ENDEX Japan 2025 held in Tokyo [3]. This section reports our observations and findings obtained through the exhibition. A total of 31 participants (30 Japanese participants and 1 participant from another Asian country), who provided written consent for research participation, experienced the system, and their behaviors were observed during and after the interaction.

As noted, the UTM system integrates two key elements: the visualization of personal transformation through AI-generated portraits and an interactive experience driven by forward and backward body movements. To further refine the design of UTM and enhance our understanding of its experiential qualities, we conducted these observations to explore how users perceive both the generated portraits and the interactive experience.

As shown on the right side of Figure 3, the exhibition booth was divided into two sections: a photo-capturing area and (Phase 1) an interactive portrait-viewing area (Phase 2) . Visitors first had their portraits taken on-site. Capturing the image at the venue allowed the system to generate portraits that reflected the participants' appearance and condition on that particular day.

After the photo was taken by the camera, a preview image was displayed on a confirmation monitor. Once the participant approved the image, the UTM system began generating the portrait images. While waiting for the generation process to complete, participants were asked to provide their consent for the use of the UTM-generated images and video recordings of their experience for research purposes.

4.2 Common Behaviors

Before participants used the UTM, we provided them with an explanation of its concept and operation. The participants were informed that the UTM generates and presents a series of portraits of themselves from childhood to old age using generative AI based on their current portrait photograph, and that as they move closer to the monitor, the portrait appears older, while moving away makes it appear younger. Participants were also instructed that moving slowly would make it easier to perceive the transitions between portraits.

We observed two main patterns of interaction. The first group tended to take one step at a time, pausing to carefully observe the portrait after each change. Participants in this group appeared to focus on examining how the AI-generated portraits transformed in relation to their distance from the display. The second group moved continuously toward and away from the screen without stopping, engaging more with the process of transformation itself. Both styles of interaction were observed with roughly equal frequency.

We observed that most participants spent the longest time viewing portraits corresponding either to the participants' current ages or to the oldest stage (old age). This tendency may be explained by the increased resemblance of the generated portraits as the images approached the participants' actual appearances. When the portrait reached an age close to each participant's real age, many participants verbally or facially expressed recognition and interest, saying things like “This looks like me” or “I'm around here.” In contrast, only a few participants showed particular interest in the younger portraits. Many participants commented that the childhood portraits “did not look like me,” although two participants mentioned that the images “somehow resembled” their younger selves. As discussed in the technical section, this limitation can be attributed to the generative model being primarily trained on Western-centric facial datasets, resulting in a noticeable sense of discrepancy for some participants.

Additionally, many participants took photographs of the generated portraits. Some recorded first-person-view videos using their smartphones, while others took two-shot photographs with their portraits at different life stages—childhood, current age, and old age.

4.3 Undesirable Behaviors

We also observed that certain features of the AI-generated images elicited a variety of unexpected or undesirable responses from the participants. For example, as shown in Figure 4 there were cases where glasses were added to portraits from participants later years, even though the participants were not wearing them during the actual photo shoot. In addition, when portraits were generated based on participants who had facial hair, the resulting images also showed facial hair even in the childhood stages.

In the portraits representing old age, some participants' generated images included features such as eyeglasses that were not present in the original photographs. Such additions were generally accepted as one possible version of the participant's future appearance and did not appear particularly unnatural. In contrast, when facial hair appeared even in the childhood portraits generated from photographs of participants with beards, the result seemed highly unnatural in the context of observing how each participant's appearance had changed over time.

As illustrated by these examples, the outcomes produced by the particular quirks of the AI model used in this study were not fully aligned with the intended experience of “observing one's own transformation” that we sought to provide through UTM. We regard this as an aspect that requires further refinement in future iterations.

Figure 4: Undesirable images generated by AI. Left: An old-age portrait; glasses added despite a participant not wearing them. Right: A childhood portrait; beard appears despite participant being a child.

4.4 Reflections on the Limitation of Interaction Design

At the exhibition, participants were instructed to move slowly in order to better perceive the changes in the portraits. However, in our observations, participants exhibited two main patterns of interaction: one group took a step forward and then paused to carefully observe the transformed portrait, while the other group continuously moved toward or away from the screen without stopping. Notably, the fact that participants diverged into these two interaction styles despite receiving identical instructions suggests that the way the instructions were communicated may have influenced the participants' behavior.

Another constraint of the current interaction design is that the generated portraits do not include explicit age labels. Some UTM participants also asked questions regarding the specific age represented by the displayed portraits. Since the system does not assign precise age labels to the generated portraits, the authors cannot definitively determine the exact age of each portrait. Consequently, participants had to infer the age of the portraits using visual cues such as wrinkles and other facial features. The generated portraits cover an approximate age range from 0 to 70 years. According to the interaction design, the portrait corresponding to age 0 (childhood) is displayed when the participant is at the farthest distance from the display, and the portrait corresponding to age 70 (old age) is displayed when the participant is at the closest distance. Therefore, the distance between the participant and the display also provides an additional cue for estimating the age of the portrait. For example, when the participant is positioned at the midpoint between the farthest and closest distances, the displayed portrait corresponds approximately to age 35.

4.5 Discussion

We noticed that participants tended to spend the most time viewing the portraits representing their current age. This tendency cannot be explained solely by the higher visual resemblance of the generated portraits at that stage. When viewing portraits close to the participants' actual age, they could directly compare the generated image with their real-time appearance. In contrast, the portraits depicting old age provided an opportunity for the participants to imagine how their appearance might change in the future. In this sense, the UTM may not only enable the viewing of generated portraits but also offer the participants a moment for self-reflection through the viewing. Here, self-reflection refers to the process in which participants contemplate their own temporal changes—thinking about their aging, growth, and personal identity—through the visual experience of generated portraits. Participants tended to deepen their understanding of themselves and their perception of time by comparing their present appearance with the generated portraits representing past and past and future appearances.

This tendency may also be partially attributed to design-related factors of the system. Most participants were in their 30s to 60s, and the optimal viewing distance from the monitor coincidentally correspond to the position associated with their actual age within the system's spatial mapping. In the current configuration, moving closer to the display corresponds to aging, while moving away corresponds to rejuvenation. Reversing this mapping could provide further insight into which age range participants engage with most strongly. However, the generative model used in the system tends to produce more accurate and realistic depictions at older ages. If the mapping were reversed, the portraits would appear less similar as participants approach the monitor. From this perspective, the current design—where "approaching the screen leads to greater resemblance"—is consistent with both the system's design intention and the experiential structure of "seeing oneself as if in a mirror."

5 Conclusion and Future work

In this paper, we introduced the Urashima Taro Mirror (UTM), an interactive installation that enables users to engage with AI-generated portraits representing different stages of their lives. Through embodied interaction, users can experience the visual process of aging and rejuvenation—compressing decades of change into a short encounter. UTM functions as a "mirror that visualizes the passage of time," inviting users to reflect on their personal growth and to rediscover their identity through their changing appearance.

Currently, the system employs forward and backward movement as the primary mode of interaction. While effective, this design also imposes constraints. UTM displays only static age-transformed portraits and does not respond to users' facial movements or expressions; age change is triggered solely by forward–backward movement. Introducing additional interactive behaviors, such as subtle portrait motion or responsiveness to user movement, may enrich the exploratory experience at the core of UTM.

In future work, we plan to refine the interaction design and examine how participants interpret and experience their temporal changes through UTM.

Acknowledgments

This work was partially supported by Tomoetogyo, Co., Ltd. We gratefully acknowledge Ruri Maeda, Haru Yasuda, and Syohei Yasuda for providing portrait photographs and for their assistance with image generation using UTM and the recording of supplemental videos.

References

[1] Lera Boroditsky. 2000. Metaphoric structuring: Understanding time through spatial metaphors. *Cognition* 75, 1 (2000), 1–28.

[2] Hal E Hershfield, Daniel G Goldstein, William F Sharpe, Jesse Fox, Leo Yeykelis, Laura L Carstensen, and Jeremy N Bailenson. 2011. Increasing saving behavior through age-progressed renderings of the future self. *Journal of marketing research* 48, SPL (2011), S23–S37.

[3] ENDEX Japan. 2025. *ENDEX Japan 2025*. https://www.tokyohakuzen.co.jp/endex-exhibition/en/ Accessed: 2025-12-12.

[4] Tero Karras, Samuli Laine, and Timo Aila. 2019. A Style-Based Generator Architecture for Generative Adversarial Networks. arXiv:1812.04948 [cs.NE] https://arxiv.org/abs/1812.04948

[5] Tiago Martins, João Correia, Sérgio Rebelo, João Bicker, and Penousal Machado. 2020. Portraits of no one: an interactive installation. In *International Conference on Computational Intelligence in Music, Sound, Art and Design (Part of EvoStar)*. Springer, 104–117.

[6] Roy Or-El, Soumyadip Sengupta, Ohad Fried, Eli Shechtman, and Ira Kemelmacher-Shlizerman. 2020. Lifespan Age Transformation Synthesis. In *Computer Vision – ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VI* (Glasgow, United Kingdom). Springer-Verlag, Berlin, Heidelberg, 739–755. https://doi.org/10.1007/978-3-030-58539-6_44

[7] Pat Pataranutaporn, Kavin Winson, Peggy Yin, Auttasak Lapapirojn, Pichayoot Ouppaphan, Monchai Lertsutthiwong, Pattie Maes, and Hal Hershfield. 2024. Future You: A Conversation with an AI-Generated Future Self Reduces Anxiety, Negative Emotions, and Increases Future Self-Continuity. arXiv:2405.12514 [cs.HC] https://arxiv.org/abs/2405.12514

[8] Kye Shimizu, Santa Naruse, Jun Nishida, and Shunichi Kasahara. 2023. Morphing Identity: Exploring Self-Other Identity Continuum through Interpersonal Facial Morphing Experience. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (Hamburg, Germany) *(CHI '23)*. Association for Computing Machinery, New York, NY, USA, Article 500, 15 pages. https://doi.org/10.1145/3544548.3580853

[9] Daisuke Uriu and William Odom. 2016. Designing for Domestic Memorialization and Remembrance: A Field Study of Fenestra in Japan. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 5945–5957. https://doi.org/10.1145/2858036.2858069

[10] Wikipedia. [n. d.]. *Urashima Tarō*. https://en.wikipedia.org/wiki/Urashima_Tar%C5%8D Accessed: 2025-6-19.

[11] Shigeo Yoshida, Tomohiro Tanikawa, Sho Sakurai, Michitaka Hirose, and Takuji Narumi. 2013. Manipulation of an emotional experience by real-time deformed facial feedback. In *Proceedings of the 4th Augmented Human International Conference* (Stuttgart, Germany) *(AH '13)*. Association for Computing Machinery, New York, NY, USA, 35–42. https://doi.org/10.1145/2459236.2459243