ICG Fina Project-Human Computer Interface

學號：D12944008　　　　　　　　　　　姓名：劉劭禹

Introduction

Augmented Reality(AR) enables users to personalize their surroundings by integrating digital content directly into their visual fields. For instance, users could embed virtual applications—such as web browsers and video call—into physical spaces or onto real-world objects. By seamlessly blending continuous digital information with the physical world, AR offers a rich and immersive experience that enhances interaction between virtual and real realms

In contrast, Diminished Reality(DR)—introduced by Steve Mann as part of his *Mediated Reality* framework[1]—addresses the issue of information overloaded by deliberately removing virtual and real-world content from users' perceived environment. By digital erasing or creating a "see through" effect on specific objects, DR produces a visual illusion that certain undesired or distracting elements have vanished.

While AR and DR differ in many aspects, they share a common goal: modulating user attention and altering the visual field. In this work, I aim towards explore how AR/DR influence users' perception, the considerations involved in their application, and their potential usages.

For this purpose, I conducted a user study with 8 participants(3 males, 5 females; age: *M*=27.8, *SD*=3.06). The study consisted of two trials and a semi-structure interview. Each trial took around 20 minutes. Participants were divided into pairs and asked to collaboratively complete the task: assembling 9 target blocks selected from a total 28 items.

Implementation

The prototype was developed on the Meta Quest 2 platform to simulate an AR experience within a VR environment, and it was built using Unity 2022.3 LTS[2] on a Windows platform, leveraging the robustness of Photon Fusion 2.0[3] for network connectivity. To present users within the VR environment, I integrated the Meta Avatar SDK[4].

To allocate users' attention, numerous approaches have augmented objects with various static and dynamic visual effects. In terms of attention capture, Bottomley et al.[5] noted that red is perceived as the most visually dominant color. Bortolotti et al.[6] and Singh[7] further demonstrated that red can stimulate brain activity and evoke emotions such as excitement and alertness. Additionally, Sutton et al. [8] mentioned that flickering light generates a distinct sensory response due to its intermittent presentation. Based on these findings—and considering the cognitive impacts of both color and motion—I designed the visual augmentation feature as a red flicker to emphasize the target blocks and effectively draw users' attention.

Conversely, to reduce users' attention, I referred to Cheng et al. [9], who compared and summarized various visual suppression techniques. Based on their findings, I implemented the visual diminishment feature as a semi-transparent effect, aiming to lessen the visual prominence of non-target blocks and minimize distraction.

Fig.1 Left: The appearance-maintained object. Middle: The appearance-augmented object. Right: The appearance-diminished object.

User Study & Result

To explore users' perception and the potential applications of visual augmentation and diminishment, I conducted a user study involving 8 participants(3 males, 5 females; age: *M*=27.8, *SD*=3.06). The study consisted of two trials and a semi-structured interview.

A within-subjects design was employed, in which each pair of participants performed two trials under different conditions. The order of conditions was fully counterbalanced to minimize potential order effects. The two experimental conditions were as follows:

- C1: Both the worker and the instructor had access to visual augmentation and visual diminishment.
- C2: Neither the worker nor the instructor had access to any visual augmentation/visual diminishment.

Each trial took around 20 minutes. Participants were divided into pairs and instructed to collaboratively complete a construction task: assembling 9 target blocks selected from a pool of 28 items. Within each pair, one participant assumed the role of the **worker**, responsible for locating and assembling the blocks, while the other acted as the **instructor**, who was given a design reference and tasked with identifying the correct target blocks and guiding the assembly process.

To investigate participants' attention allocation and their collaboration, participants were asked to complete the Social Presence questionnaire[10] with a 7-point Likert scale and a semi-structured interview after finishing the trials. I coded experiment session transcriptions to identify key findings and cross-referenced the analysis with the questionnaire results(Fig.3).

	Collaboration		Allocating Attention	
	C1	C2	C1	C2
P1	7	1	5.34	5
P2	7	2	7	5.17
P3	7	6	6	4.67
P4	7	3	5	4
P5	7	7	6.67	5
P6	4	4	4.83	5
P7	7	7	7	5
P8	7	5	7	5

Fig.2 Left: The collaboration score of participants from the semi-structured interview. Right: The average of allocating attention from Social Presence

Questionnaire.

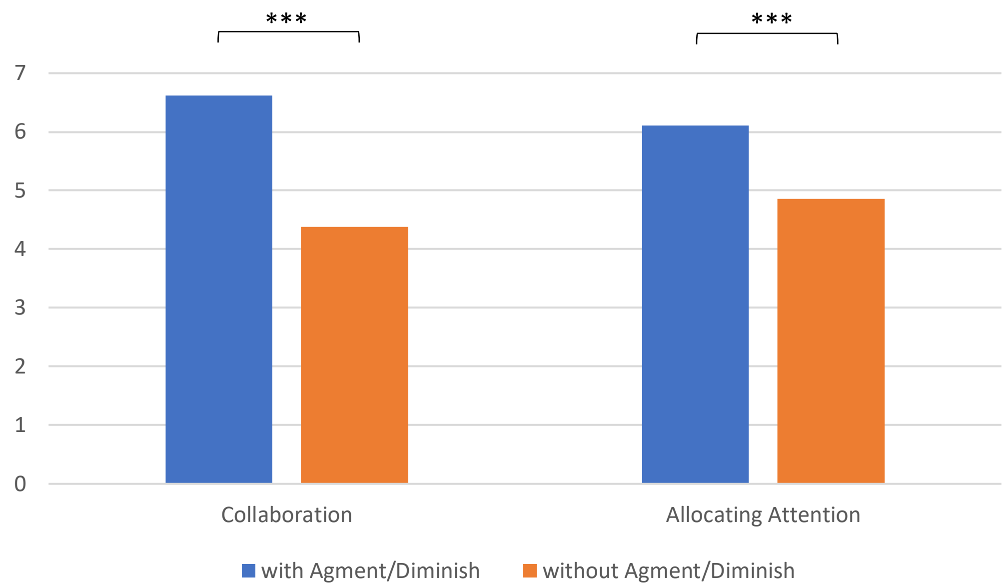

Fig.3 Left: Participant scoring of collaboration of visual augmentation and diminishment by the two conditions from the semi-structured interview. Right: Participants' attention allocation average in different conditions from Social Presence Questionnaire.

(*) indicates significant difference (p<0.05)

Participants rated their collaborative experience under both conditions using a 7-point Likert scale. In the quantitative data, the augmented condition(C1; *M*=6.63, *SD*=1.06) yields a significantly higher results(*p*=0.035) compared to the non-augmented condition(C2; *M*=6.10, *SD*=4.85). This finding aligns with the interview response, where 6 out of 8 participants mentioned that visual augmentation was essential for guiding and pointing, suggesting that instructors especially benefited from having access to it. On the contrary, the use of visual diminishment was more selective. While none of the workers actively use the feature, all instructors reported using it as a filtering tool to declutter the environment or distinguish target blocks from irrelevant ones.

According to the questionnaire results, participants demonstrated significantly higher attention allocation scores under the augmented condition(C1; *M*=6.10, *SD*=0.94) compared to the non-augmented condition(C2; *M*=4.85, *SD*=0.37), *p*=0.003. This indicates that visual augmentation and diminishment exerted a strong and reliable positive effect on participants' attentional performance during the task. Most participants reported being more alert due to the red color and the blink effect, aligning with findings by Bortolotti[6] and Singh[7], who emphasized the stimulating properties of red in capturing attention. Furthermore, the semi-transparent effect of visual diminishment effectively reduces the cognitive load by minimizing the visual prominence of irrelevant objects, allowing participants to more easily focus on appearance-maintained or appearance-augmented objects.

Reference

[1] Mann, Steve, et al. "All reality: Virtual, augmented, mixed (x), mediated (x, y), and multimediated reality." *arXiv preprint arXiv:1804.08386* (2018).
[2] Unity. 2024. Unity Real-Time Development Platform | 3D, 2D, VR & AR Engine — unity.com. https://unity.com/. Manuscript submitted to ACM
[3] Photon Engine. 2024. Fusion 2 Introduction | Photon Engine — doc.photonengine.com. https://doc.photonengine.com/fusion/current/fusion-intro.
[4] Meta Developers. 2024. Overview of Meta Avatars SDK. https://developers.meta.com/horizon/documentation/unity/meta-avatars-overview.
[5] Bottomley, Paul A., and John R. Doyle. "The interactive effects of colors and products on perceptions of brand logo appropriateness." Marketing Theory 6.1 (2006): 63-83.
[6] Bortolotti, Alessandro, et al. "The promise of color in marketing: use, applications, tips and neuromarketing." Cultura e Scienza del Colore-Color Culture and Science 15.01 (2023): 76-85.
[7] Singh, M., R. M. Singari, and M. Bholey. "A review study of cognitive design research on colors from a visual psychological perspective." International Journal of Experimental Research and Review 30 (2023): 75-86.
[8] Sutton, Jonathan, et al. "Flicker Augmentations: Rapid Brightness Modulation

for Real-World Visual Guidance using Augmented Reality." Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems. 2024.
[9] Cheng, Yi Fei, et al. "Towards understanding diminished reality." Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems. 2022.
[10] Lowenthal, Patrick R. "Social presence." Social computing: Concepts, methodologies, tools, and applications. IGI global, 2010. 129-136.

ICG Fina Project-Image Metamorphosis

學號：D12944008 姓名：劉劭禹

Introduction

Inspired by the technique described in F*eature-Based Image Metamorphosis*[1] by the Beier et al., the process of matching a photo with a portrait was divided into 2 phases:

- Warping: Modifying the shapes of the source images and destination images to achieve better geometric alignment.
- Cross-dissolving: Interpolating pixel values between the warped images to create a smooth and natural transition throughout the morphing process.

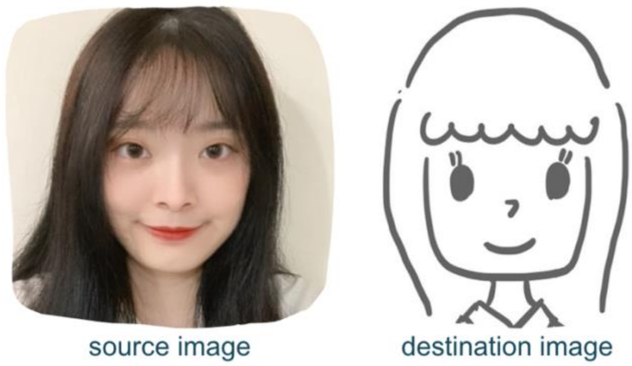

Fig.1 Left: The source image(photo). Right: The destination image(portrait).

Implementation

According to the approach presented in the paper, I manually selected key feature points on both images. These points were paired and connected as feature lines, forming a one-to-one correspondence between the source image and the destination image(Fig.2).

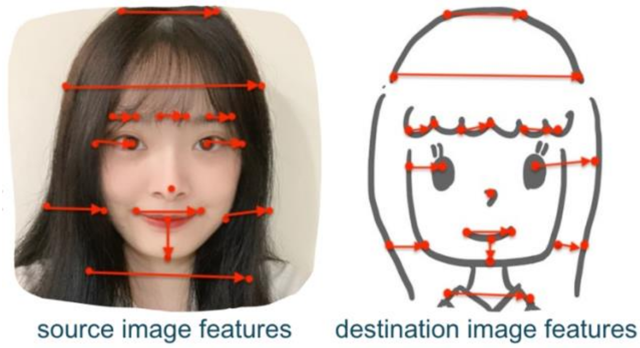

Fig.2 Left: The source image with feature lines. Right: The destination image with feature lines.

Each destination feature point X maps to multiple transformed positions X’ in the source image, each derived from a different feature line. To determine the final mapped position X, I calculated a weighted average of all corresponding X’ values. The weighting was based on two main criteria:
- Line length: Longer lines were assigned more influence.
- Proximity: Lines closer to point X received higher weight.

Based on the above criteria, the weight of each line was calculated using the formula:

$$weight = \left(\frac{length^p}{(a + dist)}\right)^b$$

Result

To visualize the warping process, I generated a metamorphosis animation illustrating the gradual transformation from the source image to the destination image. The animation was divided into 20 times steps(t=0 to 20).
- t=0, the frame: 0% destination image + 100% source image (=original source image)
- t=1, the frame: 5% destination image + 95% source image
- t=2, the frame: 10% destination image + 90% source image

…
- t=20, the frame: 100% destination image + 0% source image (=original destination image)

All frame were compiled sequentially into both a GIF and a MP4 file to present the full animation.

Problems

Since the time step parameter(t) were set to only 20, the resulting morphing(Fig.3) seems not smooth enough. Several intermediated frames displays black edges artifacts, likely due to insufficient interpolation.
The problem could be solved by increasing the number of time step parameter(t), which would produce smoother transitions and reduce or eliminate edge artifacts. However, this improvement would also result in higher computational cost and longer rendering time, maybe there exists other better function to practice the metamorphosis.

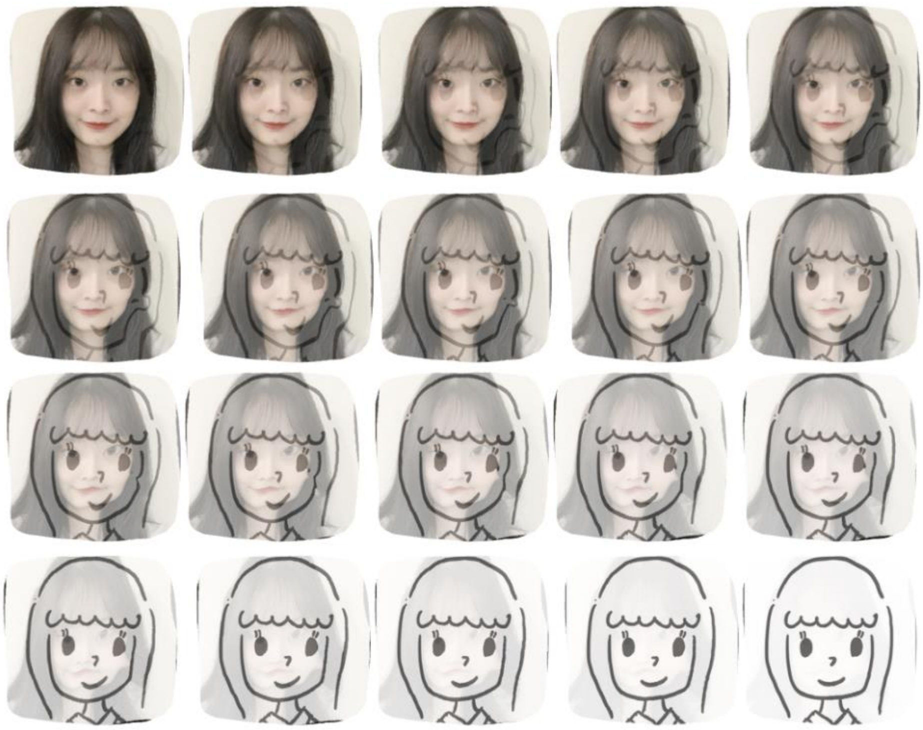

Fig.3 The intermediate frames are arranged sequentially, progressing from left to right and then from top to bottom.

Reference

[1] Beier, Thaddeus, and Shawn Neely. "Feature-based image metamorphosis." Seminal Graphics Papers: Pushing the Boundaries, Volume 2. 2023. 529-536.