

AR Memory Viewer: Recreating Memorable Scenes through AR Superimposition of Past Photos

Shuya Tonooka^{1,*}, Taishi Iriyama¹ and Takashi Komuro¹

¹ Saitama University, Saitama, Japan

Abstract

In this paper, we propose *AR Memory Viewer*, which recreates personal memory scenes through AR superimposition of past photos. Deep learning-based feature point matching enables the accurate alignment of personal memory photos with the real-world scene for AR superimposition, even when there are significant visual differences between the past photo and the real-world scene. This enables *AR Memory Viewer* to provide an experience in which users can view past scenes such as landscapes in different seasons or moments when a pet was present through their device. We conducted a user study with the prototype system and confirmed its effectiveness as a new method for recalling personal memories.

Keywords

Mobile augmented reality, Reminiscence support, Feature point matching

1. Introduction

Memories are the recollections of events and emotions we have experienced in the past, enriching our lives. Past photos play an important role in evoking or recalling old memories. While the spread of mobile devices has increased opportunities to take photos, it has been pointed out that many of these photos remain unused, left behind in folders [1]. Therefore, there is a growing need for methods that go beyond simply viewing photos, aiming to derive richer experiences and greater value from them.

Some studies have effectively utilized past photos taken at the same location as the user's current position to recreate past scenes through AR superimposition [2, 3, 4]. In this field of research, accurately aligning past photos with the real-world scene remains a major challenge, and various approaches have been proposed to address this challenge.

One approach is to utilize GPS to enable AR superimposition of past photos corresponding to the user's location information [3]. However, this approach that utilizes GPS can only handle past photos with geotags, and GPS often suffers from decreased accuracy in indoor environments. Another approach is to perform feature point matching based on salient visual features at the location, enabling accurate alignment within the area where such features are visible [4]. However, conventional feature point matching methods may fail when salient visual features at the location become unrecognizable due to changes in weather or lighting conditions. To address this issue, a method has been proposed in which multiple reference images capturing the same location under different weather conditions are prepared in advance, allowing robust detection of visual features under varying environmental conditions [5].

✉APMAR'25: The 17th Asia-Pacific Workshop on Mixed and Augmented Reality, Sep. 26-27, 2025, Busan, South Korea

* You should use this document as the template for preparing your publication. We recommend using the latest version of the CEURART style.

* Corresponding author.

✉ s.tonoka.644@ms.saitama-u.ac.jp (S. Tonooka); iriyama@mail.saitama-u.ac.jp (T. Iriyama); komuro@mail.saitama-u.ac.jp (T. Komuro)



© 2025 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

In this paper, we propose *AR Memory Viewer*, which utilizes deep learning-based feature point matching to accurately align and present past photos as AR superimpositions, even when there are significant visual differences between the past photo and the real-world scene. As shown in Figure 1, *AR Memory Viewer* enables a new memory-recalling experience by allowing users to view personal past scenes through their device.



Using AR Memory Viewer

Camera image

Selected past photo

User's perspective

Figure 1: *AR Memory Viewer* allows a user to view past scenes through the screen of a user's mobile device. Feature point matching is performed between the camera image and a past photo, and the past photo is aligned with the camera image and displayed as an AR superimposition on the screen of the mobile device.

2. AR Memory Viewer

2.1. Core processing

An overview of the core processing of *AR Memory Viewer* is shown in Figure 2. Feature point matching is performed between the camera image and all past photos in the folder, and the past photo with the highest number of matched feature points is selected. A projective transformation is performed using the matched feature points from the selected photo, generating an image in which the past photo is aligned with the camera image. A transformation is applied to the generated image to simulate a magic lens from the user's point of view. The result is then displayed on the screen of the mobile device.

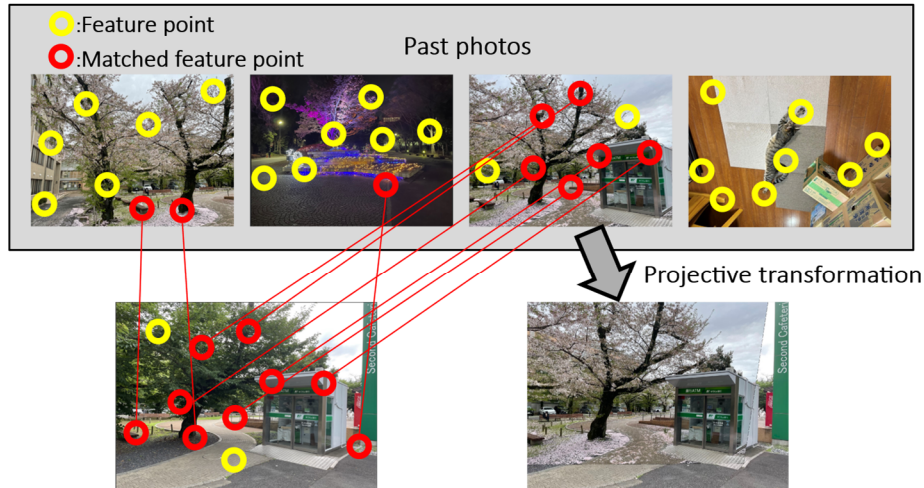


Figure 2: Overview of the core processing

2.2. Implementation details

The prototype system used a Surface Pro 7 as the mobile device, and a PC equipped with an NVIDIA GeForce GTX 1070 GPU as the server for performing deep learning-based feature point matching. SuperPoint [6] was used for feature point extraction, with a maximum of 4,096 feature points per image. SuperGlue [7] was used for matching, configured with the outdoor version of the model trained on outdoor datasets. The camera image captured by the mobile device is sent to the server via TCP communication. Feature point matching is performed on the server, and

the aligned image is sent to the mobile device. The image displayed on the mobile device was transformed under the assumption that the distance from the user to the device is 50 cm, and the distance from the device to the real-world scene is at infinity. After receiving the image, the mobile device performs alignment using AKAZE [8], a computationally efficient feature matching method, enabling accurate AR superimposition even when the camera is moved.

The processing results of the prototype system are shown in Figure 3. We verified the operation using scenes from different seasons for outdoor locations and scenes with a pet for indoor locations. Sufficient matching between the selected past photo and the camera image was achieved, enabling a magic lens with accurate alignment of the past photo to the real-world scene.

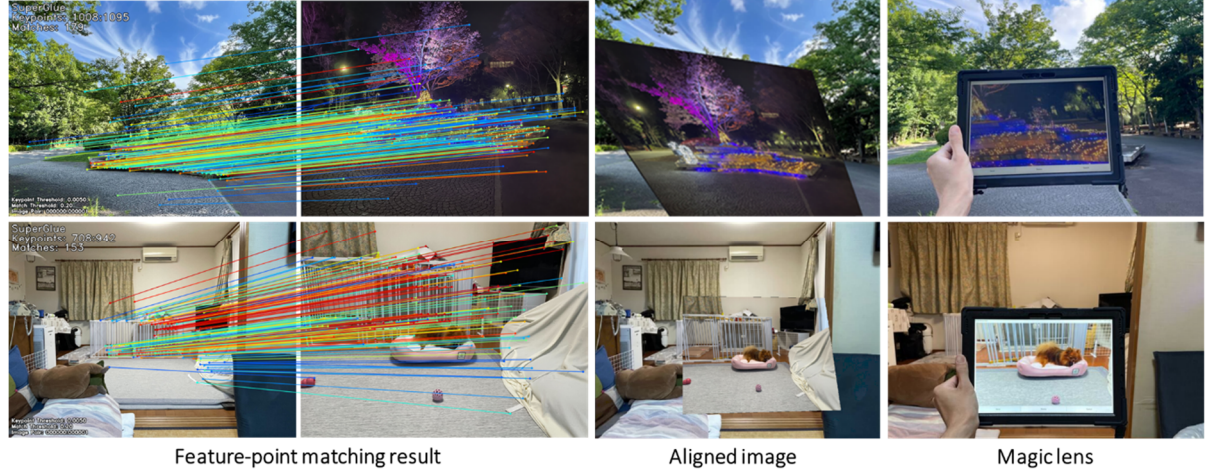


Figure 3: Processing results of the prototype system

3. User Study

3.1. Experimental design

We conducted a comparative experiment to investigate the effectiveness of the proposed method. As a baseline method, we employed a manual selection approach in which participants chose a past photo similar to the real-world scene from a folder and presented it as an AR superimposition at the center of the screen. The set of past photos in the folder consisted of 10 images: one photo used for scene reproduction and nine randomly selected photos from Flickr using the keywords “indoor place” and “outdoor place”.

The experiment was conducted in July 2025. The participants were 20 students from our university (5 females; mean age = 22.9, SD = 1.74). Four past photos taken on our university campus (two indoor locations taken on June 2025, and two outdoor locations taken on December 2023 and April 2025) were used, and participants experienced one of the two methods at each location. The conditions were counterbalanced so that each was experienced an equal number of times across the four locations. After guiding the participants to each location, we provided instructions on the appropriate camera angle when using *AR Memory Viewer*. We measured the time from launching each application to performing AR superimposition of past photos, as well as the time from the AR superimposition to exiting the application. After the experience, participants were asked to complete a questionnaire in which the “game” section of the GEQ [9] was adapted by replacing “game” with “application,” along with three custom questions. The GEQ has been used to evaluate AR experiences in the study by Lee et al. [3], and it is considered effective as an indicator for measuring user experience. Memories are deeply rooted in each user’s internal experiences, and it is difficult to directly measure the quality of recreating memorable scenes using quantitative scales. Therefore, we adopted this evaluation method.

3.2. Experimental results

The time required from launching the application to performing AR superimposition was 18.1 seconds on average (± 3.65) for the baseline method and 18.56 seconds on average (± 4.31) for the proposed method, showing no substantial difference between the two methods. Since only 10 past photos were included in the folder in this experiment, no difference was observed; however, the advantage of the proposed method is expected to become more evident as the number of photos increases. Regarding the time from the start of AR superimposition to exiting the application, the baseline method required 27.82 seconds on average (± 11.81), whereas the proposed method required 43.13 seconds on average (± 19.89), indicating that the proposed method enabled significantly longer application usage ($p < 0.001$).

Next, the results of the GEQ questionnaire are shown in Figure 4. A paired t-test revealed that the proposed method demonstrated superiority in all subscales except for Tension and Challenge. This indicates that, compared to the baseline method, the proposed method enhances the quality of the application experience.

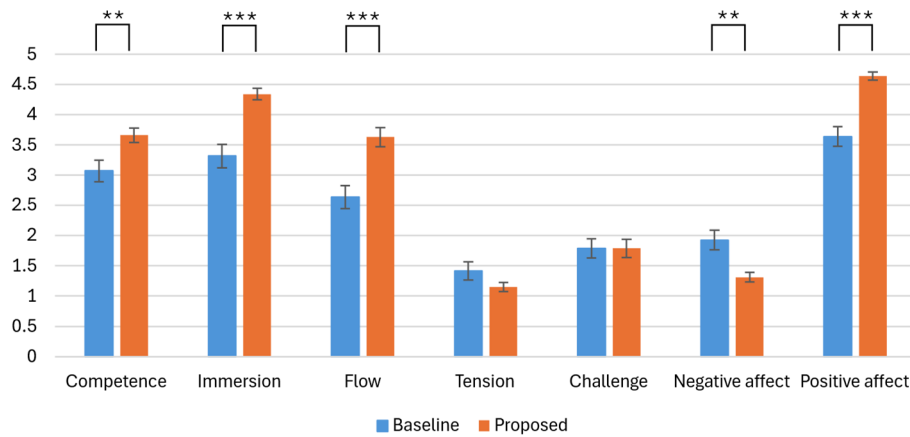


Figure 4: GEQ scores for each condition. Error bars denote SEM. Significant differences are marked with ** ($p < .01$), and *** ($p < .001$).

Finally, the results of the custom questions are shown in Figure 5. For Q1, ratings of 7 and 6 accounted for 80% of the responses, suggesting that the concept of the proposed method was sufficiently conveyed. For Q2, all responses were either 7 or 6, indicating that the proposed method may be effective as a new way of referring to personal memories. For Q3, the ratings were limited to 7, 6, and 5, implying that the burden of photo selection was reduced and that past photos were continuously aligned with the real-world scene at the correct position.

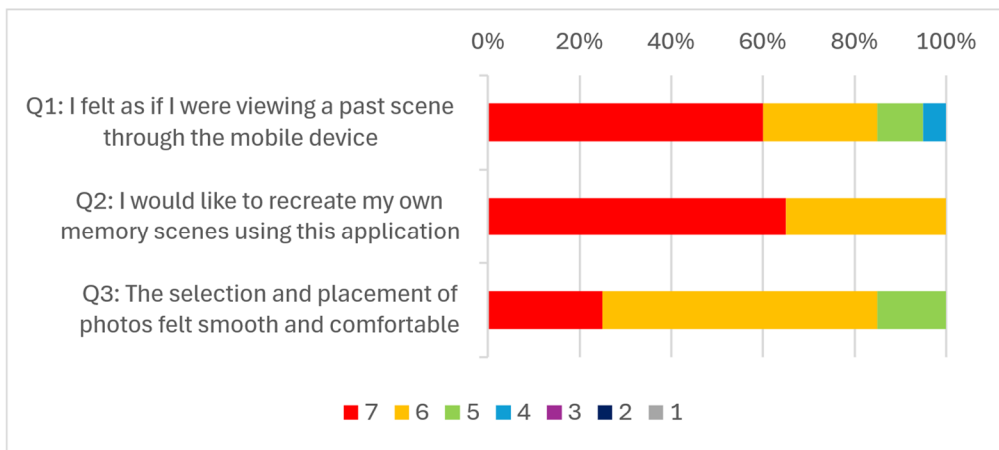


Figure 5: Results of the custom questions

4. Conclusion and future work

In this paper, we proposed *AR Memory Viewer*, which utilizes deep learning-based feature point matching to accurately align personal memory photos with the real-world scene for AR superimposition. *AR Memory Viewer* allows users to view past scenes through their device even when the appearance of past photos differs from the real-world scene, as long as the geometric structure is similar. The results of the evaluation experiment demonstrated that the concept of the prototype system was sufficiently conveyed and that it is effective as a new way of referring to personal memories.

In the prototype system, feature point matching with SuperGlue was difficult to perform solely on a mobile device, so communication with a desktop PC was used. However, LiteGlue [10], a lightweight version, has recently been proposed, and it may enable operation solely on a mobile device.

Declaration on Generative AI

During the preparation of this work, the authors used GPT-5 for grammar and language refinement. After using this tool, the authors reviewed and edited the content as needed and took full responsibility for the publication's content.

References

- [1] David McGookin. Reveal: Investigating Proactive Location-Based Reminiscing with Personal Digital Photo Repositories. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. pp. 1–14. <https://doi.org/10.1145/3290605.3300665>
- [2] Tommy Hasselman, Wei Hong Lo, Tobias Langlotz, and Stefanie Zollmann. ARephotography: Revisiting Historical Photographs using Augmented Reality. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23)*. pp. 1–7. <https://doi.org/10.1145/3544549.3585646>
- [3] Gun A. Lee, Andreas Dünser, Seungwon Kim, and Mark Billinghurst. CityViewAR: A mobile outdoor AR application for city visualization. In *2012 IEEE International Symposium on Mixed and Augmented Reality - Arts, Media, and Humanities (ISMAR-AMH '12)*. 57–64. <https://doi.org/10.1109/ISMAR-AMH.2012.6483989>
- [4] Marco Cavallo, Geoffrey Alan Rhodes, and Angus Graeme Forbes. Riverwalk: Incorporating Historical Photographs in Public Outdoor Augmented Reality Experiences. In *Adjunct Proceedings of 2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct '16)*. pp. 160–165. <https://doi.org/10.1109/ISMAR-Adjunct.2016.0068>
- [5] Silvia Blanco-Pons, Berta Carrión-Ruiz, Michelle Duong, Joshua Chartrand, Stephen Fai, and José Luis Lerma. Augmented Reality Markerless Multi-Image Outdoor Tracking System for the Historical Buildings on Parliament Hill. *Sustainability* 2019, Vol.11, No.16, Art. no. 4268. <https://doi.org/10.3390/su11164268>.
- [6] Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. SuperPoint: Self-Supervised Interest Point Detection and Description. In *Proceedings of 2018 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW '18)*. pp. 337–349 <https://doi.org/10.48550/arXiv.1712.07629>

- [7] Paul-Edouard Sarlin, Daniel DeTone, Tomasz Malisiewicz, and Andrew Rabinovich. SuperGlue: Learning Feature Matching With Graph Neural Networks. In *Proceedings of 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. 4937–4946. <https://doi.org/10.1109/CVPR42600.2020.00499>
- [8] Pablo F. Alcantarilla, Adrien Bartoli, and Andrew J. Davison. Fast explicit diffusion for accelerated features in nonlinear scale spaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)* 34, 7 1281–1298.
- [9] Wijnand A. Ijsselstein, Yvonne A. W. de Kort, and Karolien Poels. The game experience questionnaire. Technische Universiteit Eindhoven, Eindhoven, The Netherlands.
- [10] Philipp Lindenberger, Paul-Edouard Sarlin, and Marc Pollefeys. LightGlue: Local feature matching at light speed. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV '23)*. 17627–17638. <https://doi.org/10.48550/arXiv.2306.13643>