[Poster] Indirect Augmented Reality Considering Real-World Illumination Change

Fumio Okura* Takayuki Akaguma Tomokazu Sato* Naokazu Yokoya*

Nara Institute of Science and Technology (NAIST)

Abstract

Indirect augmented reality (IAR) utilizes pre-captured omnidirectional images and offline superimposition of virtual objects for achieving high-quality geometric and photometric registration. Meanwhile, IAR causes inconsistency between the real world and the pre-captured image. This paper describes the first-ever study focusing on the temporal inconsistency issue in IAR. We propose a novel IAR system which reflects real-world illumination change by selecting an appropriate image from a set of images pre-captured under various illumination. Results of a public experiment show that the proposed system can improve the realism in IAR.

Index Terms: H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities

1 Introduction

In IAR [4], as shown in Figure 1, virtual objects (C) are first superimposed on pre-captured image (B) of the real world in the offline process. An appropriate part of the pre-rendered augmented image is presented in the online process to a user through a mobile device based on roughly estimated orientation of the device. Unlike conventional AR methods that overlay virtual objects on images captured in real-time, IAR achieves high quality registration between pre-captured images (B) and virtual objects (C) with low computational cost on mobile devices.

However, IAR essentially causes spatial and temporal inconsistency between the pre-captured image (B) and the real world (A). IAR is intended to be used only at a place where the images are captured. Thus, spatial inconsistency appears when used at other places, and the user may find differences of viewpoints. The original IAR study [4] investigated the perceptual effects by the difference of the viewpoints between the user and the captured point. In case that the appearance of the real world is changed from those when the omnidirectional images were captured, the temporal inconsistency becomes noticeable. An experiment using an IAR system [1] reported that the temporal inconsistency caused by illumination change was perceived larger than the spatial one. The temporal inconsistency is a significant problem occurring in IAR nevertheless it has less been investigated. This paper focuses on the temporal inconsistency in IAR, particularly caused by temporal change of real-world illumination, and describes a novel IAR system considering the appearance change of real-world illumination.

2 IAR Considering Illumination Change

The proposed IAR system is divided into online and offline processes along with usual IAR systems. Unlike the ordinary IAR, in the offline process, omnidirectional images are captured under various illumination conditions at the same position. Virtual objects are then superimposed onto omnidirectional images and stored into a database with corresponding sub-image histograms. In the online process, a pre-generated augmented image, whose illumination condition is the most similar to that of the real scene acquired by a mobile device, is selected from the database based on a histogram similarity. Finally, the selected image is transformed to a planar perspective image, which is presented onto the mobile device.

*e-mail: {fumio-o,tomoka-s,yokoya}@is.naist.jp

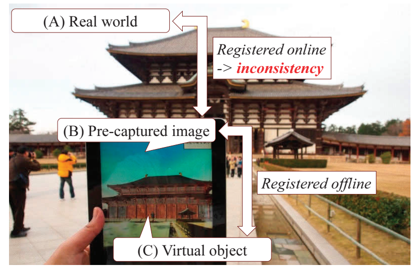

Figure 1: Registration problems in IAR. Inconsistency between the real world (A) and pre-captured images (B) appears due to illumination change (the real world is cloudy; while the IAR scene is sunny).

2.1 Offline Process

This process generates a database consisting of augmented omnidirectional images, from which dynamic objects are removed, and corresponding sub-image histograms. The histograms are used in online process for efficiently determining a similar image in respect to illumination condition. Although real-time pixel-wise matching methods for mobile and omnidirectional images [3] are available, sensor-based orientation, which does not achieve a pixel accuracy, is used in most IAR implementations owing to the light computational cost. Histogram-based similarity is robust against the orientation error, as well as efficient because it is not necessary to store whole pixel data into the database on a mobile device.

Multiple omnidirectional videos, each of whose duration is about several tens of seconds, are first captured at the fixed position under various illumination conditions. Omnidirectional images without dynamic objects are generated from the input videos using a panoramic photography approach for crowded environments [2]. Then, sub-image (i.e., block regions in the image) histograms are calculated from each omnidirectional image, in which dynamic objects are removed. Virtual objects are superimposed on the omnidirectional images without dynamic objects.

2.2 Online Process

An augmented omnidirectional image, whose illumination condition is considered to be the most similar to that of the real world, is selected based on the differences between sub-image histograms of each pre-captured image and real-world images captured by the mobile device. The selected image is presented on a display of the mobile device as the form of planar perspective image depending on the orientation of the device as with the usual IAR systems [4].

Since users can freely change their view direction during the IAR experience, we should consider the difference of illumination conditions for various viewing directions. Thus, assuming the real-world illumination does not vary during an IAR experience, differences between the i-th pre-captured image and the mobile images captured in various orientation θ are calculated, and then unified into a difference $D(i)$ of a whole direction. A pre-captured image with the minimum $D(i)$ is selected as the most similar one.

$$D(i) = \sum_{c \in R,G,B} \min_{s_c} \sum_{\theta} \mathit{diff}_c(i, \theta, s_c). \tag{1}$$

(a) Sunny (with small clouds), 12 am (b) Sunny, 4 pm (c) Cloudy, 12 am

Figure 2: Augmented omnidirectional images under various illumination as well as their perspective transformation for a given orientation.

A difference $diff_c(i, \theta, s_c)$ for an orientation θ and a color channel $c \in R, G, B$ is calculated using sub-image histograms corresponding to θ and the histogram of the mobile image. It is based on the sum of squared differences between the histograms, which are normalized by the number of pixels used for the difference calculation. The change of the luminance and the white balance between pre-captured and mobile images are compensated so as to minimize the difference with changing a scale factor s_c for each color channel.

3 Public Experiment Using Historical Application

We have developed a practical historical application set in the Todaiji temple, a UNESCO World Heritage site in Nara, Japan, and have performed a public experiment in cooperation with Todaiji, to confirm the effect of illumination consideration on IAR.

In this application, appearances of three buildings at the time of the foundation of Todaiji were superimposed on pre-captured omnidirectional images.[1] Nine omnidirectional images were captured under several time of day and weather conditions at a fixed position using an omnidirectional camera, Ladybug3 (Point Grey Research Inc.). Figure 2 shows examples of augmented omnidirectional image, on which a virtual buildings were superimposed using 3ds Max (Autodesk, Inc.) with offline manual editing (i.e., adjusting light sources, and designating occlusion masks). The perspective images in Figure 2 are examples of IAR images displayed onto mobile devices. This application has three display modes for the comparison:

1. Proposed system: Display an IAR scene with the smallest difference $D(i)$.
2. Ordinary IAR: Display an IAR scene randomly selected.
3. Worst case: Display an IAR scene with the largest $D(i)$.

The difference $D(i)$ was calculated using mobile images of a whole circumference that were captured for approximately every 30° just before each trial. The IAR scenes were displayed onto mobile devices, iPad2 and iPad 4th generation (Apple Inc.) with full-screen mode except for a question sentence as well as small buttons for answering scores and switching the display modes. The physical location of participants during the experiment was within a few meters from the position where the omnidirectional images were captured. The experiment was carried out in the whole daytime under either sunny or cloudy weather (depending on the participant).

Total 87 public participants experienced the application. The number of participants for each generation were 1, 6, 6, 11, 19, 25, and 19 for their 10's to 70's, respectively. 56 participants were male, and remains were female. Most participants were not familiar with AR/MR fields. The participants first experienced the application with switching among these modes, and then scored each of them through the question about the realism of IAR scenes: "Does the virtual buildings look as if it is really in the real world?" on a scale of 1 (absolutely no) to 7 (absolutely yes). The order in which the display modes were presented was randomized, and participants were allowed to switch and watch among them as many times as they wished even while answering the question. Free commentary was corrected after scoring all images.

[1] These models were created with reference to the miniature model of Todaiji with architectural validation by Yumiko Fukuda, Hiroshima Institute of Technology, and patina expression technology validation by Takeaki Nakajima, Hiroshima City University.

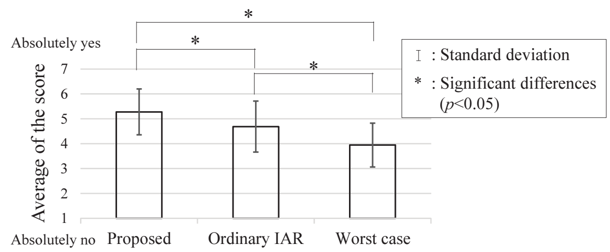

Figure 3: Realism scores in a public experiment.

Figure 3 shows the average score for each display modes. The score for the proposed system is the best for the realism of IAR scenes. The scores are significantly different by an one-way repeated measures ANOVA with a post-hoc paired t-test with Bonferroni correction ($p < 0.05$). This result indicates that considering the illumination change of the real world contributes the improvement of the realism of IAR scenes.

This result clearly shows realism improvement of IAR scenes; however, several participants pointed out that the *aesthetics* of IAR scenes affected their scores in addition to the consistency between the IAR scenes and the real world. A few participants actually scored better for IAR scenes using sunny pre-captured images than cloudy ones, even though they experienced the application under cloudy weather. This issue may bring a large benefit to IAR applications when it will be investigated further; i.e., depending on the application and/or the user, presenting *aesthetic* IAR scenes can improve the realism in an IAR experience.

4 Conclusions and Future Work

This paper has proposed an IAR system considering the real-world illumination changes, which displays an IAR scene similar to the real-world scene from among those captured under various time of day and weather conditions. The experiment indicated that our prototype system contributed to improve the realism of IAR scenes; while the aesthetics of the scene can also affect the realism. In future work, we plan to develop a more efficient method to prepare the pre-captured images, which does not require multiple capture under various illumination by employing a relighting technique.

Acknowledgements

This research was supported by JSPS KAKENHI (A) 23240024, Grant-in-Aid for JSPS Fellows 25-7448, and the "Ambient Intelligence" project funded by MEXT. We thank members of NAIST-Keio joint research group. We also thank Todaiji for giving us the opportunity to conduct public experiments.

References

[1] T. Akaguma, F. Okura, T. Sato, and N. Yokoya. Mobile AR using pre-captured omnidirectional images. In *Proc. ACM SIGGRAPH Asia'13 Symp. MGIA*, pages 26:1–26:4, 2013.

[2] I. Arai, M. Hori, N. Kawai, et al. Pano UMECHIKA: A crowded underground city panoramic view system. In *Proc. DCAI'10*, pages 173–180, 2010.

[3] T. Langlotz, C. Degendorfer, A. Mulloni, et al. Robust detection and tracking of annotations for outdoor augmented reality browsing. *Computers & Graphics*, 35(4):831–840, 2011.

[4] J. Wither, Y.-T. Tsai, and R. Azuma. Indirect augmented reality. *Computers & Graphics*, 35(4):810–822, 2011.