Transactions Letters

Simultaneous Geometric and Radiometric Adaptation to Dynamic Surfaces With a Mobile Projector-Camera System

Hanhoon Park, Moon-Hyun Lee, Byung-Kuk Seo, Jong-Il Park, Moon-Sik Jeong, Tae-Suh Park, Yongbeom Lee, and Sang-Ryong Kim, *Member, IEEE*

***Abstract*—Existing geometric and radiometric compensation methods for direct-projected augmented reality focus on static projection surfaces rather than dynamic surfaces (with varying geometry in time). We aim at providing an effective framework for projecting a sequence of augmented reality images onto dynamic surfaces without geometric and radiometric distortion. We present our design of a special pattern image for simultaneous geometric and radiometric compensation and evaluate two different techniques for embedding the pattern image into augmented reality images. The validity of the proposed method is examined through a variety of experiments with a mobile projector-camera system.**

***Index Terms*—Direct-projected augmented reality (DirectAR), dynamic adaptation, pattern embedding, spatial encoding, temporal encoding.**

I. Introduction

GEOMETRIC compensation and radiometric compensation are two crucial aspects of direct-projected augmented reality (DirectAR) systems [1]. Unfortunately, most existing geometric and radiometric compensation methods are applicable only to static projection surfaces despite the fact that dynamic projection surfaces are also common in our everyday environment. In this paper, we focus on dynamic projection surfaces and propose a geometric and radiometric adaptation method for use on dynamic projection surfaces.

Most DirectAR systems use structured light techniques[1] [1]–[3] to compensate for geometric and radiometric distortion. That is, after projecting a known pattern image of geometry and color using a projector, the geometric and radiometric distortion is measured and corrected based on the pattern image captured by a camera. For this to work on a dynamic projection surface, the pattern image would have to be projected continuously onto the surface. However, projecting the pattern image along with AR images may interfere with user perception of the AR images. Therefore, we propose a pattern embedding method as an efficient way of making pattern images imperceptible.

We evaluate two different encoding methods for embedding imperceptible multiple pattern images: temporal encoding and spatial encoding. For spatial encoding, we designed a special-purpose pattern image which can be used for simultaneous geometric and radiometric compensation. Advantages and disadvantages of the two encoding approaches are discussed later.

To demonstrate the feasibility of the proposed method, we created a mobile projector-camera system by taking a CarPC and rigidly mounted a low-cost camera and projector onto it.

Due to space limitations, we have omitted an explanation of the DirectAR here; please refer to [1] and [4] for a detailed discussion of DirectAR.

A. Related Works

There have been some articles associated with geometric adaptation to dynamic surface in the literature. Yasumuro *et al.* used an additional infrared projector to project near-infrared patterns invisible to the human eye [5]. Although their system can augment medical images on a dynamic human body without distortion and distraction, the frame rate is cut down by half because the system uses at least two projectors. Cotting *et al.* measured the mirror flip (on/off) sequences of a digital light processing (DLP) projector for RGB values using a photo transistor and a digital oscilloscope and imperceptibly embedded arbitrary binary patterns into AR images by adjusting the mirror flips to be aligned to the binary pattern for a very short period of time in such a way that the original AR image values were approximated to the nearest values [6]. However, this method requires sophisticated control of camera shuttering for detecting short-term patterns. Yang and Welch proposed a pattern-less method of using underlying features in AR images to estimate the geometric relationship between projector and camera [7], which resolved the problem with the intrusiveness of pattern image; however, texture-less AR images made it difficult to compute the geometric relationship. Moreover, most feature extraction and matching methods are vulnerable to noise or geometric transformation.

To our knowledge, there is only one study associated with radiometric adaptation to dynamic surfaces. Fujii *et al.* proposed a coaxial projector-camera system with real-time photometric adaptation and fixed geometric correspondence for dynamic environments [8]. Their system was free from geometric distortion but lacked a solution for geometric compensation in the event

Manuscript received September 7, 2006; revised January 17, 2007. This work was supported by in part by a grant from the Samsung Advanced Institute of Technology. This paper was recommended by Associate Editor D. Schonfeld.

H. Park, M.-H. Lee, B.-K. Seo, and J.-I. Park are with the Department of Electrical and Computer Engineering, Hanyang University, Seoul 133791, Korea (e-mail: jipark@hanyang.ac.kr).

M.-S. Jeong, T.-S. Park, Y. Lee, and S.-R. Kim are with the Advanced Systems Research Laboratory, Samsung Advanced Institute of Technology, Suwon 440600, Korea.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCSVT.2007.903322

[1][Online]. Available: http://www.cs.unc.edu/~raskar/Projector/projbib.html

that the projector and camera were positioned arbitrarily. Furthermore, the method of adaptation chosen gives only an approximation of the results that can be achieved with full radiometric compensation.

B. Contribution of This Paper

There is little research on geometric or radiometric adaptation to dynamic surfaces and what exists focuses on either geometric or radiometric adaptations. To the best of our knowledge, this paper is the first attempt to simultaneously address geometric and radiometric adaptation to dynamic surfaces.

We propose a single pattern based geometric and radiometric compensation method, called *spatial encoding*, for mobile projector-camera systems and have designed a special pattern image which can be used simultaneously for both geometric and radiometric compensation for dynamic surfaces.

This paper proposes a new pattern embedding technique for making embedded pattern images imperceptible. Our method is inspired by the complementary patterns that are used mainly for 3-D video acquisition [9], [10]. The main difference lies in the fact that the pattern images in this paper are modulated into AR images by embedding (adding/subtracting) a certain value (brightness or color) to the pixels of the AR image.

II. Embedding and Detecting Imperceptible Pattern Images

A. Principle of Embedding and Detecting Imperceptible Pattern Images

Given a sequence of AR images, it is possible to embed (= unnoticeably insert) a pattern image by adding a certain value (brightness or color) to the pixels of odd frames of the sequence and subtracting the value from the pixels of even frames of the sequence, or vice versa. Letting $I_{\mathrm{odd}}(x,y)'$ and $I_{\mathrm{even}}(x,y)'$ be pattern-embedded odd and even frames of AR images, then

$$I_{\mathrm{odd}}(x,y)' = I_{\mathrm{odd}}(x,y) + \frac{P(x,y)}{2}$$
$$I_{\mathrm{even}}(x,y)' = I_{\mathrm{even}}(x,y) - \frac{P(x,y)}{2}$$

where

$$P(x,y) = \alpha \quad \text{or} \quad 0. \tag{1}$$

If the complementary AR images are alternately projected with a high refresh rate, only the AR images will be perceived by the human eye because the human vision system integrates the amount of light seen over time in order to acquire an image.

Due to the limited dynamic range of an image, embedding a pattern image (adding or subtracting a value) results in underflow or overflow in some dark or bright pixels. Thus, if the value is k, the AR images are compressed in luminance such that their luminance ranges from k to 255-k. The visual degradation due to contrast reduction can be minimized using a content-based pattern embedding method [11].

Note that the degree of the imperceptibility varies according to the magnitude or kind of variation of the pixels [11]. In general, a large variation makes the pattern image robust to noise but quite easily perceptible, while a low variation makes the pattern image completely imperceptible but weak to noise. Humans are sensitive to variations of brightness but less sensitive to those of color [12]. It is well known that humans are least sensitive to variations in the Q channel of the YIQ color space. With these facts in mind, the desirable variation for maximizing imperceptibility must be determined experimentally.

The pattern images are internally and simply extracted from the "difference image" between the odd and even frames of a camera image sequences as

$$P(x,y) = C_{\mathrm{odd}}(x,y) - C_{\mathrm{even}}(x,y). \tag{2}$$

When the resulting images are noisy, median filtering may be effectively applied. In our experimental environment, simple median filtering suffices. More sophisticated algorithms may be employed in ill-conditioned experimental environments.

B. Method of Embedding and Detecting Imperceptible Patten Images

Multiple pattern images are necessary for geometric or radiometric compensation [1]; however, it is not easy to come up with a method of embedding more than one pattern image into a single pair of odd and even frames of AR images. We propose two different ways of embedding multiple pattern images: spatial encoding and temporal encoding.

1) Temporal Encoding: Temporal encoding is most commonly used for projection surfaces that do not change dynamically [3]. In temporal encoding, the pattern images are separately embedded into successive AR images. That is, the information for geometric and radiometric compensation is multiplexed in time across the multiple pattern embedding AR images as shown in Fig. 1(a). Fig. 2 shows examples of embedding pattern images for geometric and radiometric compensation using temporal encoding. In the case of pattern embedding for geometric compensation, the embedding (addition/subtraction) can be alternatively performed in neighboring stripes to improve the imperceptibility of a pattern image [see Fig. 2(a)].

Detection of the embedded patterns can be achieved by repeatedly applying the single pattern detection method explained in Section II-A to each pair of odd and even frame of camera image sequences. The pattern image when $t = t_0$ is computed as

$$P(x,y)|_{t=t_0} = C_{\mathrm{odd}}(x,y)|_{t=t_0} - C_{\mathrm{even}}(x,y)|_{t=t_0}. \tag{3}$$

2) Spatial Encoding: In spatial encoding, pattern images are embedded together into a pair of odd and even AR images as shown in Fig. 1(b). For this purpose, we have designed a special-purpose pattern image in such a way that its pixels are identified easily for geometric compensation and include the changes in RGB channels for radiometric compensation. Notice that neighbor pixels usually have the same intensity and can thus be united to a large pixel, which we call a *united pixel*. Therefore, each of the neighbor pixels can embed its allotted information for geometric and radiometric compensation (for example, in a united pixel consisting of four pixels, four neighbor pixels have a chessboard-like layout and experience the variation in different channels). All of this information can be used together for geometric and radiometric compensation in the location of the united pixel.

In general, camera resolution is lower than projector resolution. Thus, the mapping between a projector image pixel and a camera image pixel is not 1-to-1, but n-to-1. Actually, a united

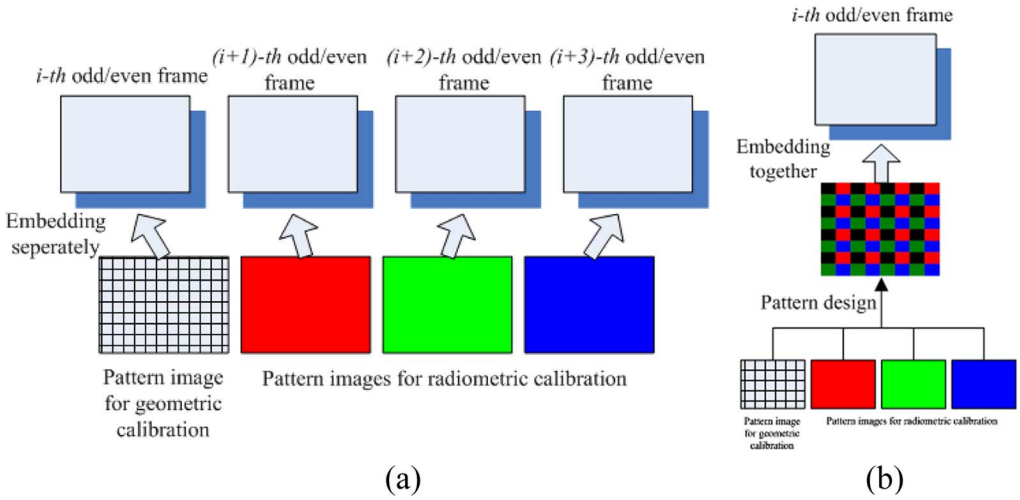


Fig. 1. Two different methods of embedding imperceptible pattern images. In temporal encoding, the pattern images for geometric and radiometric compensation are multiplexed in time. In spatial encoding, a special pattern in which the information for geometric and radiometric compensation is spatially included is designed first and is embedded into a single pair of odd and even AR images. (a) Temporal encoding. (b) Spatial encoding.

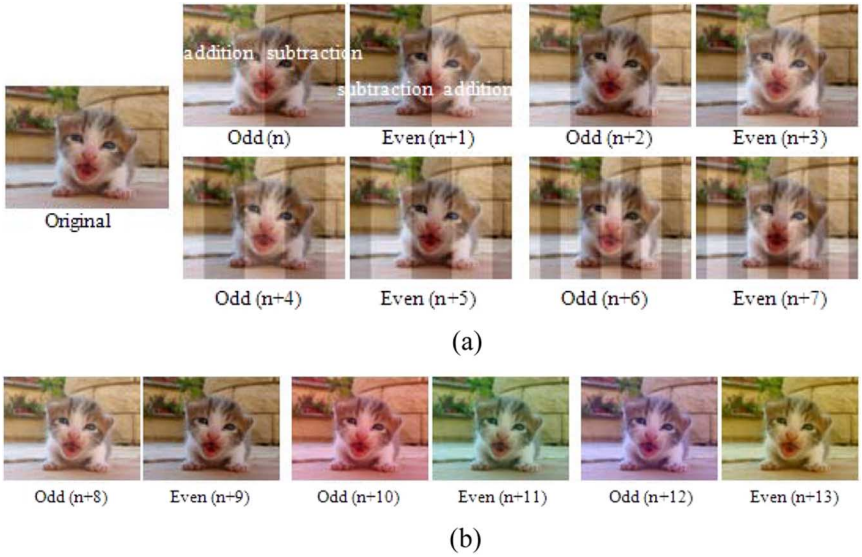


Fig. 2. Examples of embedding pattern images using temporal encoding. The numerals in parentheses indicate the frame number. (a) Intensity values are added and subtracted alternatively. (b) R, G, B color values are added and subtracted alternatively. (a) Pattern image embedding for geometric compensation. (b) Pattern image embedding for radiometric compensation.

pixel consists of four n-pixel blocks in which the pixels experience the same variation as

$$I_{\text{odd}}(2nx+i,2ny+j)' = I_{\text{odd}}(2nx+i,2ny+j) + \frac{P(2nx+i,2ny+j)}{2}$$
$$I_{\text{even}}(2nx+i,2ny+j)' = I_{\text{even}}(2nx+i,2ny+j) - \frac{P(2nx+i,2ny+j)}{2}$$

where (4), shown at the bottom of the page. Addition and subtraction for pattern embedding can be alternately performed in the neighbor united pixels to improve the imperceptibility of a pattern image (see Fig. 3).

In spatial encoding, detection of the embedded patterns becomes slightly different. Letting C_{odd} and C_{even} be the camera image frames, the pattern image for radiometric compensation is obtained as

$$P^{\text{rad}}(x,y) = \text{sgn}(D(x,y))D(x,y) + \text{sgn}(-D(x,y))\{-D(x,y)\}$$

$$\begin{cases} P(2nx+i,2ny+j)\big|_{\text{red}} = \alpha, & i=0,\ldots,n-1 \quad \text{and} \quad j=n,\ldots,2n-1 \\ P(2nx+i,2ny+j)\big|_{\text{green}} = \alpha, & i=n,\ldots,2n-1 \quad \text{and} \quad j=0,\ldots,n-1 \\ P(2nx+i,2ny+j)\big|_{\text{blue}} = \alpha & i,j=n,\ldots,2n-1, \\ P(2nx+i,2ny+j) = 0 & \text{otherwise.} \end{cases} \tag{4}$$

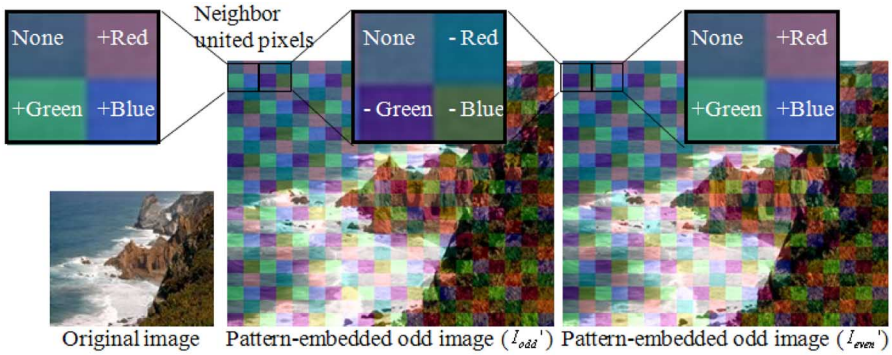


Fig. 3. Examples of embedding pattern images using spatial encoding. Each neighbor block in a united pixel experiences different changes in R, G, B channels. The changes are offset in a pair of the pattern embedded odd and even frames.

CarPC Spec.
CPU – Intel Pentium 4 2.8GHz
Memory – RAM 1GB
VGA – On-Board
Mainboard – LV-670
Interface – USB 2.0 / IEEE 1394A
OS – Windows XP Embedded
Size – 178mm×169mm×50.5mm
Weight – 1.8 Kg

Fig. 4. Overview of our mobile projector-camera system.

where

$$\mathrm{sgn}(x) = \begin{cases} 1 & x > 0 \\ 0 & x \le 0 \end{cases}$$
$$D(x,y) = C_{\mathrm{odd}}(x,y) - C_{\mathrm{even}}(x,y). \tag{5}$$

The pattern image for geometric compensation is obtained as

$$P^{\mathrm{geo}}(x,y) = \begin{cases} 1, & AD(x,y) > \varepsilon \\ 0, & AD(x,y) < \varepsilon \end{cases}$$

where

$$AD(x,y) = |C_{\mathrm{odd}}(x,y) - C_{\mathrm{even}}(x,y)|. \tag{6}$$

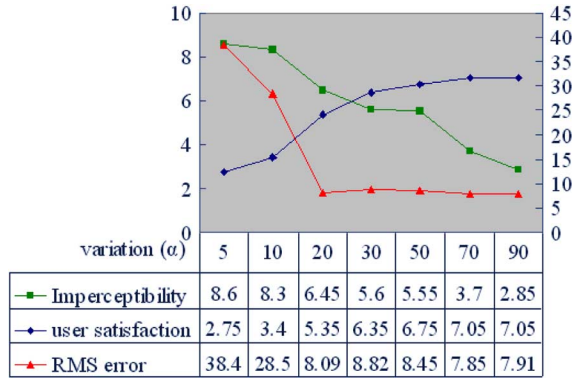
III. Experimental Results and Discussion

To demonstrate the validity of our pattern embedding technique, we created a mobile projector-camera system which consists of a CarPC, a projector (HP MP-2210),[2] a camera (Prosilica EC 655C),[3] and a keypad as shown in Fig. 4.

The projector and camera were synchronized to capture the image projected by the projector without frame loss at a frame rate of 60 Hz. A vertical synchronization signal of a VGA input signal connected to a projector was provided to an external trigger port of a camera as an external input signal. An input image with a resolution of 640 × 480 pixels was used. The projection surface was nonplanar and color-textured. The P matrices and nonlinear response functions of the projector and camera were computed in advance to allow the system to work in real-time. The algorithms of [1] and [13] were used for geometric and radiometric compensation.

[2][Online]. Available: http://www.hp.com

[3][Online]. Available: http://www.prosilica.com



variation (α)	5	10	20	30	50	70	90
Imperceptibility	8.6	8.3	6.45	5.6	5.55	3.7	2.85
user satisfaction	2.75	3.4	5.35	6.35	6.75	7.05	7.05
RMS error	38.4	28.5	8.09	8.82	8.45	7.85	7.91

Fig. 5. Tradeoff relationship between imperceptibility of embedded pattern images and level of user satisfaction (or RMS error) with compensation results according to the magnitude of the variation of the AR image pixels for pattern embedding.

To find the desirable strength for embedded pattern images, we asked twenty university students to complete a questionnaire based on their satisfaction of geometric and radiometric compensation results and the perceptibility of different strengths of the embedded pattern image. We divided the degree of user satisfaction and imperceptibility into ten levels. Fig. 5 shows the trade-off relationship between user satisfaction and imperceptibility of embedded pattern images corresponding to the strength of embedded pattern image. In the subjective evaluation, the desirable variation was 60 ($\alpha = 30$). Actually, because it was difficult to detect weakly embedded pattern images with the standard commercial camera, the aspect of imperceptibility ended up being sacrificed to guarantee reasonable compensation results.

Fig. 6 shows the geometric and radiometric compensation results according to the strength of embedded pattern images. The radiometric root-mean-square (RMS) error of the compensation results is given in Fig. 5. The radiometric RMS error was affected by the geometric RMS error and thus we evaluated the system with only the radiometric RMS error. Hereafter, RMS error means radiometric RMS error. The RMS error of an estimated image $\hat{I}$ was computed as

$$\mathrm{error} = \sqrt{\frac{\{\bar{I}_R - \hat{I}_R\}^2 + \{\bar{I}_G - \hat{I}_G\}^2 + \{\bar{I}_B - \hat{I}_B\}^2}{3}}$$

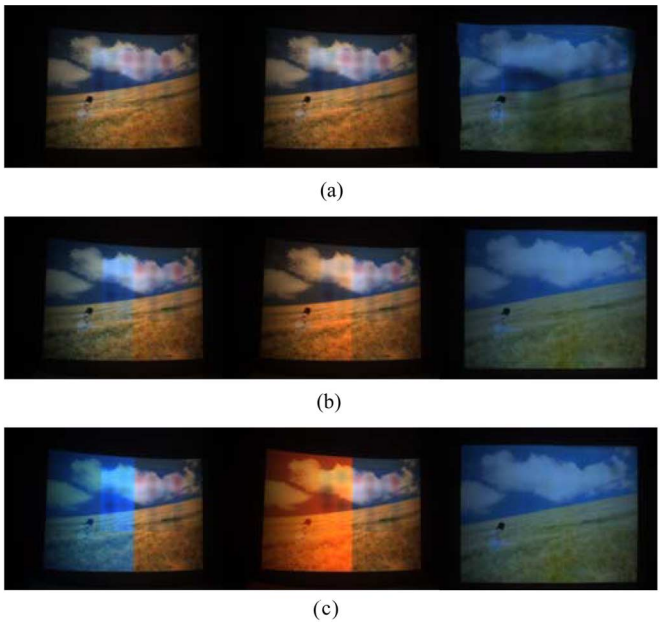


Fig. 6. Geometric and radiometric compensation results according to the strength of embedded pattern images. The geometric and radiometric distortion disappeared in the compensated images when the strength exceeded a certain value. Left two images: first two frames (odd and even) of AR image sequences which embed the gray code pattern image in their Q channel. Right image: compensated image. (a) Strength: 10 ($\alpha = 5$). (b) Strength: 60 ($\alpha = 30$). (c) Strength: 180 ($\alpha = 90$).

where

$$\hat{I} : \text{target image.} \tag{7}$$

The RMS error was almost unchanged when the strength of embedded pattern images was greater than 60. Therefore, a variation of 60 was optimal in our system. The reason the RMS error was not small as expected is that radiometric compensation generally results in the decrease of brightness of the compensated images [14]. Actually, the RMS error in chromaticity was 4.7 on average when the strength of embedded pattern images was greater than 60. The RMS error in chromaticity was computed by transforming the target image and the estimated image to HSV color space and differencing them in the hue channel.

Next, we analyzed the performance of spatial encoding and temporal encoding in two scenarios: one on a static projection surface with a fixed system; the other on a static projection surface on which the system is moving continuously relative to the projection surface (as commonly happens in mobile DirectAR systems such as RFIG Lamps by Raskar *et al.* [15]). In temporal encoding, multiple stripe pattern images were embedded into successive AR images for geometric compensation. R, G, B color pattern images were also embedded into successive AR images. R, G, B color pattern images indicate the images in which r-, g-, and b-values were added and subtracted by a given level (30) across the whole area of AR images, respectively. In spatial encoding, the multiple pattern images for geometric and radiometric compensation were substituted with our designed pattern image. The designed pattern image was embedded into AR images with a strength of 60.

In the first scenario, the compensation results of temporal encoding were better than those of spatial encoding, as shown in Fig. 7. In the results, the RMS error caused by spatial encoding was (about 7) larger than that caused by temporal encoding. Spatial encoding was inferior because the camera image resolution was reduced by the spatial encoding mechanism (at least four pixels are merged into a united pixel) and because the camera image pixels which include the spatially different information were more vulnerable to the color variation of the projection surface or to edges in AR images. The remaining color distortion in the left image of Fig. 7(b) could be mitigated by reducing the target brightness as shown in the right image of Fig. 7(b).

Fig. 7. Compensation results when the projection surface is static and the system is fixed. Using (a) temporal encoding and (b) spatial encoding, with different target brightness.

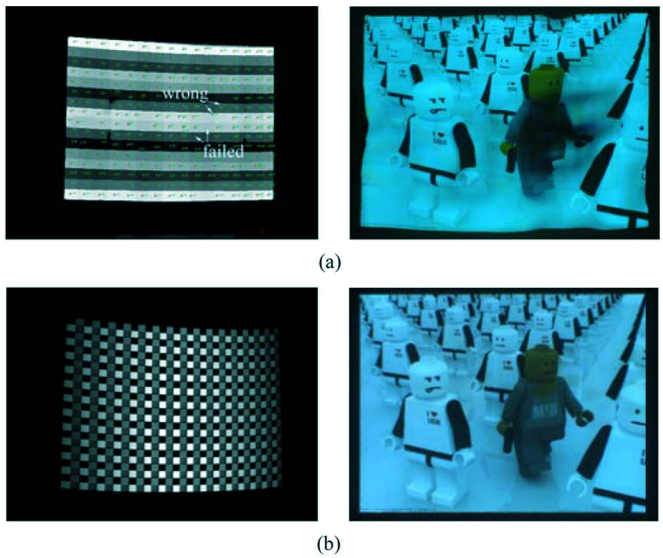


Fig. 8. Compensation results when the projection surface is static but the system is mobile. (a) Temporal encoding. Left image: computed gray codes for geometric compensation. Right image: compensation result. (b) Spatial encoding. Left image: detected and processed pattern image for geometric compensation. Right image: compensation result.

In the second scenario, the performance of spatial encoding was not different from the first scenario as shown in Fig. 8(b). However, temporal encoding was not available (image distortion was wrongly compensated) because the movement of the

system broke the temporal correlation between the embedded pattern images as shown in the left image of Fig. 8(a). Consequently, for mobile projector-camera systems, spatial encoding would be better than temporal encoding. However, it should be considered that spatial encoding is affected by the spatial frequency of projection surfaces and AR images.

IV. Conclusion

In this paper, we aimed at providing an effective method for projecting a sequence of AR images without geometric and radiometric distortion onto dynamic surfaces. For this purpose, spatial and temporal encoding techniques of imperceptibly embedding multiple pattern images into AR images were proposed and evaluated through experiments with a mobile projector-camera system. A special pattern image for simultaneous geometric and radiometric compensation was designed and used in the spatial technique.

In our mobile projector-camera system, the adaptation based on the simple embedding mechanism could be done in real-time.

However, the CarPC may not be convenient to carry about so we are currently considering substituting CarPC with PDA.

In spatial encoding, we used a chessboard pattern image for geometric compensation. For improving the performance and flexibility, it would be good for future work to spatially embed multiple gray code pattern images for geometric compensation.

References

[1] H. Park, M.-H. Lee, S.-J. Kim, and J.-I. Park, "Surface-independent direct-projected augmented reality," in *Proc. ACCV'06*, 2006, pp. 892–901.

[2] O. Bimber and R. Raskar, *Spatial Augmented Reality*. New York: A K Peters, 2005.

[3] J. Salvi, J. Pages, and J. Batlle, "Pattern codification strategies in structured light systems," *Pattern Recognit.*, vol. 37, no. 4, pp. 827–849, 2004.

[4] H. Park and J.-I. Park, "Modern approaches to direct-projected augmented reality: A review," in *Proc. ISUVR'06*, 2006, pp. 15–18.

[5] Y. Yasumuro, M. Imura, Y. Manabe, O. Oshiro, and K. Chihara, "Projection-based augmented reality with automated shape scanning," in *Proc. SPIE EI'05*, 2005, pp. 555–562.

[6] D. Cotting, M. Naef, M. Gross, and H. Fuchs, "Embedding imperceptible patterns into projected images for simultaneous acquisition and display," in *Proc. ISMAR'04*, 2004, pp. 100–109.

[7] R. Yang and G. Welch, "Automatic projector display surface estimation using every-day imagery," presented at the WSCG'01, 2001 [Online]. Available: http://wscg.zcu.cz/wscg2001/Papers_2001/R51.pdf

[8] K. Fujii, M. D. Grossberg, and S. K. Nayar, "A projector-camera system with real-time photometric adaptation for dynamic environments," in *Proc. CVPR'05*, 2005, pp. 814–821.

[9] M. B. Vieira, L. Velho, A. Sa, and P. C. Carvalho, "A camera-projector system for real-time 3-D video," in *Proc. CVPR'05*, 2005, p. 96.

[10] M. Waschbüsch, S. Würmlin, D. Cotting, F. Sadlo, and M. Gross, "Scalable 3-D video of dynamic scenes," *Vis. Comp.*, pp. 629–638, 2005.

[11] H. Park, M.-H. Lee, B.-K. Seo, Y. Jin, and J.-I. Park, "Content adaptive embedding of complementary patterns for nonintrusive direct-projected augmented reality," in *Proc. HCII'07*, 2007, pp. 132–141.

[12] S. E. Palmer, *Vision Science*. Cambridge, MA: MIT Press, 1999.

[13] M. D. Grossberg, H. Peri, S. K. Nayar, and P. N. Belhumeur, "Making one object look like another: Controlling appearance using a projector-camera system," in *Proc. CVPR'04*, 2004, vol. 1, pp. 452–459.

[14] D. Wang, I. Sato, T. Okabe, and Y. Sato, "Radiometric compensation in a projector-camera system based on the properties of human vision system," in *Proc. PROCAMS'05*, 2005, p. 100.

[15] R. Raskar, P. Beardsley, J. van Baar, Y. Wang, P. Dietz, J. Lee, D. Leigh, and T. Willwacher, "RFIG Lamps: Interacting with a self-describing world via photosensing wireless tags and projectors," in *Proc. SIGGRAPH'04*, 2004, pp. 406–415.