Content Adaptive Embedding of Complementary Patterns for Nonintrusive Direct-Projected Augmented Reality

Hanhoon Park, Moon-Hyun Lee, Byung-Kuk Seo, Yoonjong Jin, and Jong-Il Park

Department of Electrical and Computer Engineering, Hanyang University, Seoul, Korea
{hanuni, fly4moon, nwseoweb, lakobe8}mr.hanyang.ac.kr,
jipark@hanyang.ac.kr

Abstract. In direct-projected augmented reality, the visual patterns for compensation may distract users despite users would not be interested in the compensation process. The distraction becomes more serious for dynamic projection surface in which compensation and display should be done simultaneously. Recently, a complementary pattern-based method of efficiently hiding the compensation process from users' view has been proposed. However, the method faced the tradeoff between the pattern imperceptibility and compensation accuracy. In this paper, we embed locally different strength of pattern images into different channels of the projector input images (AR images) after analyzing their spatial variation and color distribution. It is demonstrated that our content adaptive approach can significantly improve the imperceptibility of the patterns and produce better compensation results by comparing it with the previous approach through a variety of experiments and subjective evaluation.

Keywords: Nonintrusive direct-projected augmented reality, complementary pattern embedding, content adaptive.

1 Introduction

Direct-projected augmented reality (DirectAR) indicates a technique that displays (=projects) AR images on surfaces with arbitrary shape or color under arbitrary lighting environment without image distortion using a projector-camera pair. It needs the process of measuring the scene geometry and radiometry, and transforming the augmented reality images in advance, called usually compensation [1]. It is achieved by projecting code patterns using a projector, capturing the pattern images using a camera, and analyzing the camera images. In this process, the patterns are usually strongly perceptible and thus may visually intrusive to users. Moreover, in dynamic environments (e.g. camera, projector is moving, or geometry, photometry of surface is changing), the process of estimating the geometry and photometry of surface using visual patterns is not applicable any more [2]. To resolve these problems, there have been researches corresponding to projecting patterns invisibly or at high speed. We call them nonintrusive DirectAR.

Yasumuro et al. used an additional infrared projector to project near-infrared patterns which are invisible to human eyes [3]. Their system can augment medical

R. Shumaker (Ed.): Virtual Reality, HCII 2007, LNCS 4563, pp. 132 – 141, 2007.

images on dynamic human body without distortion and distraction. However, the frame rate is cut down by half because the system uses at least two projectors.

Raskar et al. have used special engineered digital light projector that is able to turn light on and off at a very high rate (over 1000Hz) [4]. This projector projects image bit-plane by bit-plane. Two of the 24 bit-planes are reserved to insert a light pattern and its complement. Because the switching is so fast, human eye is unable to distinguish between the bit-plane showing light pattern and the next one showing its complement.

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 projector input images (indicates augmented reality images in this paper) 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 projector input image values are approximated to the nearest values [5]. The adjustment can causes contrast reduction as explained in [6]. Moreover, sophisticated control of camera shuttering for detecting the short-term patterns is required.

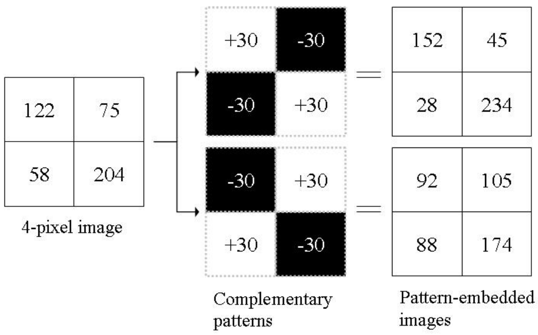
There have been other types of nonintrusive DirectAR, which are not fully vision-based ones because they use physical sensors to measure the projected code patterns but use visual patterns and try to reduce pattern perceptibility. Summet and Sukthankar proposed a hand-held projected display whose location is tracked using optical sensors in conjunction with temporally-coded patterns of projected light and tried to minimize the extent of distracting temporal codes to small regions [7]. They projected a small pattern over located sensors instead of projecting a full-screen localization pattern, but left the rest of the projection area free for user display purpose. Lee et al. used the concept of frequency shift keyed (FSK) modulation to reduce pattern perceptibility [8]. That is, they selected either a bright red or a medium gray that, when rendered by their modified projector (removing the color wheel) appears to be a similar intensity of gray to a human observer, but are manifested as a 180Hz signal and a 360Hz signal respectively. The optical sensor can discriminate the signals. By using these two colors, they could hide code patterns in what appeared to be mostly gray squares.

Recently, Park et al. proposed a method that embeds complementary patterns into AR images based on a simple pixel operation without using sensors or special-purpose projectors [2]. In their method, code pattern images and their complementary (inverse) images are embedded into consecutive AR images by increasing and decreasing the pixel values of the AR image, or vice versa. If the frame rate of the sequence reaches the critical fusion frequency (75Hz according to Watson [9]), the odd and even frames are visually integrated over time, so that the embedded pattern image would be imperceptible to the users. The embedded pattern images are detectable using the camera synchronized with the projector. However, their method needs to deal with the tradeoff between the strength (visibility) of the embedded patterns and the accuracy of estimating the geometry and photometry of surface [2]. Bimber et al. attempted to fade patterns in and out within a short sequence of subsequent projector input images [6]. However, the frame rate is severely cut down. As a more effective method, we propose a content-adaptive pattern embedding method of minimizing the visibility of patterns while maintaining the compensation accuracy. We locally adjust the pattern strength and the embedding channel after analyzing the AR images based on the characteristics of human vision system. The advantage of using the characteristics of human vision system was also experimentally verified by other researchers [10].

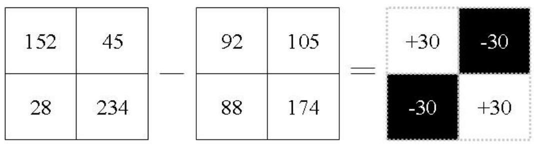
With our approach, the patterns with the same strength can be more imperceptible by changing the embedding channel. And the strong pattern images minimize the compensation error in the region with high spatial variation and the compensation error due to weak pattern images in the region with low spatial variation is unnoticeable. It is demonstrated that our content adaptive method can significantly improve the imperceptibility of the embedded patterns and produce better compensation results by comparing it with the previous methods through a variety of experiments and subjective evaluation.

2 Embedding and Detecting Complementary Patterns

In this section, we briefly explain the method of embedding and detecting complementary patterns which indicate a pair of a pattern and its inverse. They are generated by adding and subtracting a color (or brightness) of successive AR images as shown in Fig. 1-(a) and Fig. 2-(a).



(a) pattern embedding



(b) pattern detection

Fig. 1. Example of embedding and detecting complementary patterns into a 4-pixel gray AR image

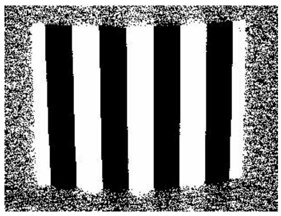
If complementary patterns are alternately projected at high speed, only the AR images will be perceived to the human eye because the human vision system integrates the amount of light seen over time in order to acquire an image as shown in Fig. 2-(c).

The detection of embedded pattern images is easily done by subtracting the odd and even frames of the camera image sequences and binarizing the result (see Fig. 1-(b) and Fig. 2-(d)) once the camera is synchronized with the projector [2]. When the resulting images are noisy, the median filtering may be effectively applied to the images. In ill-conditioned environment, simple median filtering suffices. More sophisticated algorithms may be required under ill-conditioned experimental environments.

(a) pattern embedded AR images

(b) camera images of projection of (a)

(c) image seen to the users



(d) extracted pattern from the images of (b)

Fig. 2. Example of embedding and detecting of code pattern for geometric compensation

3 Content-Adaptive Pattern Embedding and Detecting

In this section, we explain the content-adaptive method of embedding and detecting imperceptible pattern images. We locally embed patterns based on the analysis on the color distribution, spatial variation of AR images (= projector input images) as shown in Fig. 3. We divide the AR images into small $n \times n$ ($n < w, h$) blocks. w and h indicate the width and height of the AR images. For each block, we embed patterns with the strength proportional to the spatial variation into different channel depending on the color contribution.

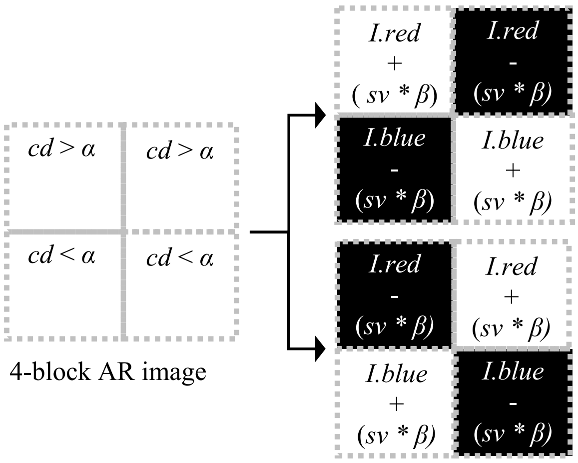


Fig. 3. Content-adaptive pattern embedding for a 4-block AR image. α and β indicate user-defined constants. According to the color distribution (*cd* in Eq. (2)), the pattern images are embedded into different channels. According to the spatial variation (*sv* in Eq. (1)), the pattern images are embedded with different strengths.

The spatial variation is computed using a 2-D derivative filter as follows:

$$sv = \sqrt{\{F_h * I\}^2 + \{F_v * I\}^2} \tag{1}$$

where * denotes spatial convolution, F_h and F_v denote horizontal and vertical derivative filters, respectively, and I denotes projector input image. We use the filter kernel $\begin{bmatrix} -1 & 8 & -8 & 1 \end{bmatrix}$ for derivative calculation.

The color distribution is analyzed in the YIQ color space. The transformation of RGB color space to YIQ color space is defined as

$$y = 0.299r + 0.587g + 0.114b,$$
$$i = 0.596r + 0.274g + 0.322b,$$
$$q = 0.211r + 0.523g + 0.312b.$$

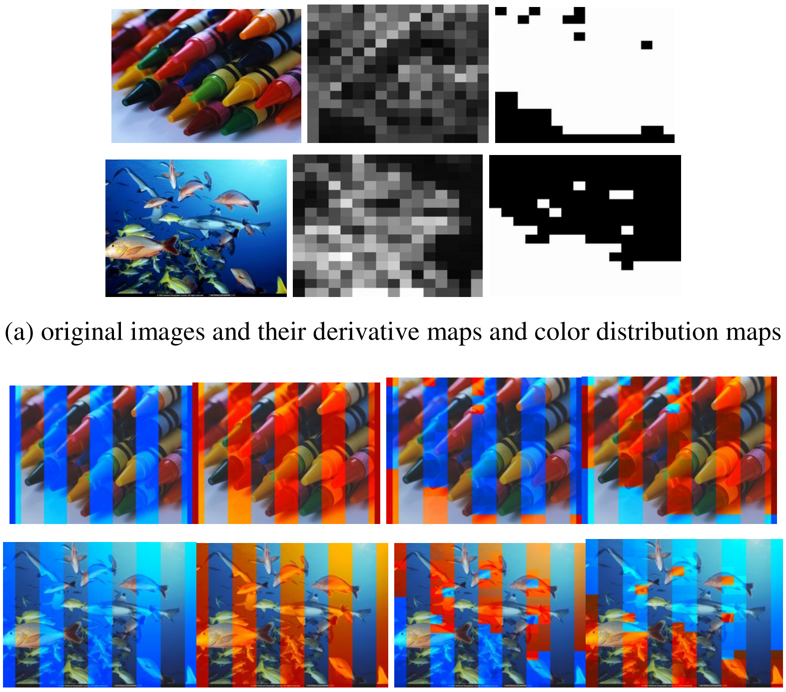
Here, i value is dominated by r value while q value is dominated by g value. Therefore, we embed pattern images into the Q-channel in the region with a large value of r while into the I-channel otherwise. Specifically, the color distribution value is computed as

$$cd = \operatorname{sgn}\left(\frac{\sum_N r/N}{\sum_N g/N} - 1 \right) \quad \textit{where } N\text{: the number of pixels,} \tag{2}$$

$$\operatorname{sgn}(x) = \begin{cases} 1 & x > 0 \\ 0 & x \le 0 \end{cases}.$$

The detection of embedded pattern images is not different from that of Section 2. Figure 5 shows the examples of detecting content-adaptively embedded patterns.

Although the patterns were embedded into locally different channels with locally different strengths, the pattern images were clearly extracted. It implies that the content-adaptive method does not influence the compensation accuracy.



(a) original images and their derivative maps and color distribution maps

(b) non-adaptive (c) adaptive

Fig. 4. Content-adaptive pattern embedding for a real AR image. The derivative map has continuous values ranging from 0 to 1 while the color distribution map has two values, i.e. 0 and 1.

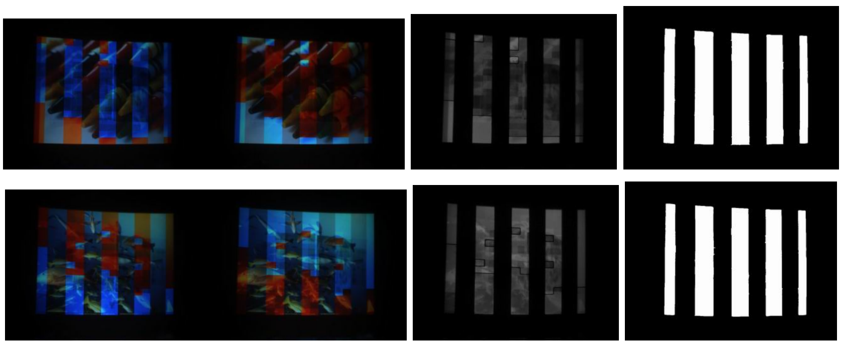


Fig. 5. Projection of adaptively pattern-embedded AR images, their difference images, and detected patterns. Even for the image having low spatial frequency and contrast all over, the weakly embedded pattern is also clearly detected.

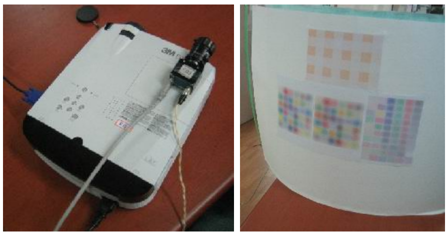
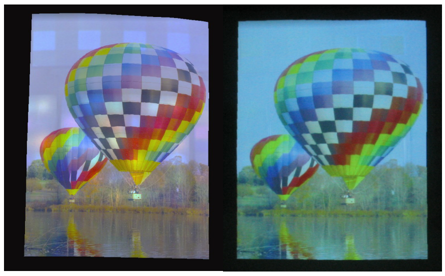


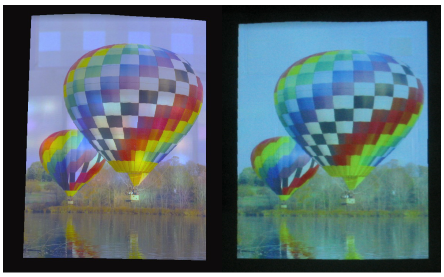
Fig. 6. Experimental setup

(a) original image and its projection (not compensated)

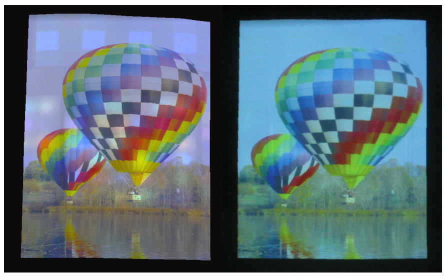
(b) non-adaptive



(c) adaptive to color distribution



(d) adatptive to spatial variation

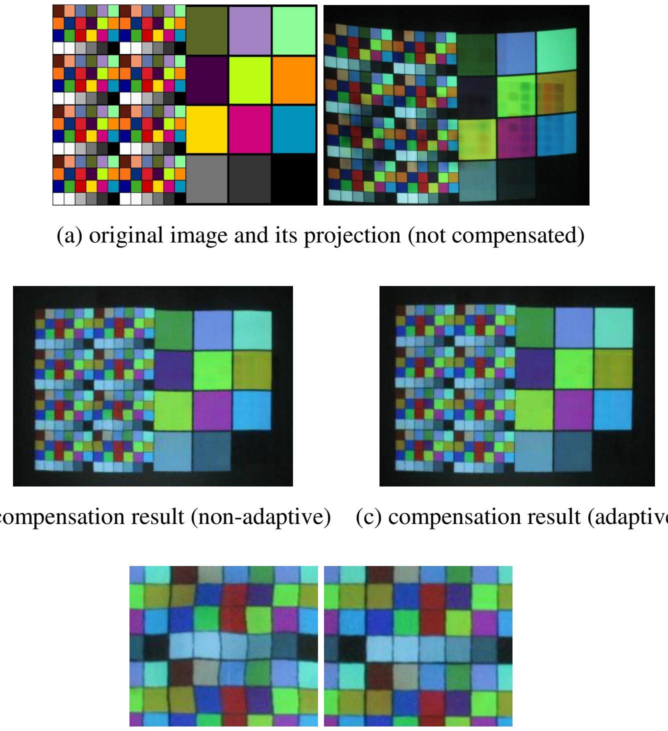


(d) adaptive to both

Fig. 7. Comparison of compensation results using the non-adaptive embedding and adaptive embedding (for real image). Left images: modified projector input image, right images: compensated projection.

4 Experimental Results and Discussion

We created a projector-camera system which consists of a projector (3M DX60) and a camera (Prosilica EC 655C) as shown in Fig. 6. The camera was synchronized with the projector and it can capture the images projected by the projector without frame loss. A vertical synchronization (shortly, v-sync) 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. Input images with a resolution of 1024 by 768 pixels were used and the projection surface was non-planar and color-textured (see Fig. 6). The P matrices and nonlinear response functions of the projector and camera were computed in advance. Therefore, the system could work in real-time at a frame rate of 60Hz (= the refresh rate of the projector).



(a) original image and its projection (not compensated)

(b) compensation result (non-adaptive) (c) compensation result (adaptive)

(d) enlargement of a region with high texture in (b) and (c)

Fig. 8. Comparison of compensation results using the non-adaptive embedding and adaptive embedding (for synthetic image). The pattern strength was intentionally reduced to see the superiority of the adaptive embedding to the non-adaptive one.

Figure 7 and 8 show the compensation results for a real and synthetic image using content-adaptive embedding method and non-adaptive embedding method on the same experimental environment. The compensation results showed little difference.

Rather, the content-adaptive one outperformed the non-adaptive one in the region with high spatial variation as we see in Fig. 8[1]. Actually, the pattern strength was intentionally reduced for compensating the synthetic image. Thus, the overall compensation results were worse than those of the real image. However, the adaptive embedding method could reduce the compensation error in the region with high spatial variation by increasing the pattern strength.

To confirm the usability of the content-adaptive pattern embedding, we asked fourteen university students to complete a questionnaire based on their satisfaction of compensation results and the perceptibility of the embedded patterns. We divided the degree of user satisfaction and imperceptibility into ten levels. Figure 9 shows the results. It seems that users thought that the accuracies of both are similar but they suffered from the intrusiveness of patterns when using non-adaptive one. On the whole, they preferred adaptive one to non-adaptive one.

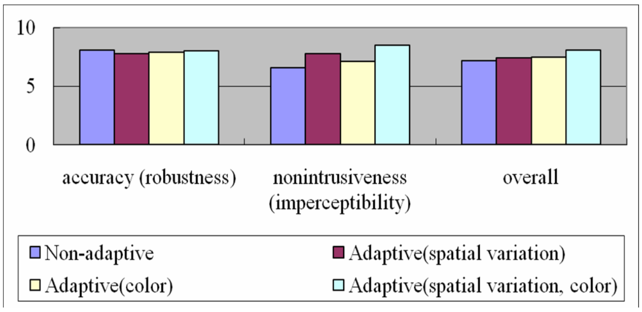


Fig. 9. Subjective evaluation on compensation results and pattern imperceptibility using the non-adaptive embedding and adaptive embedding. The overall rating was computed by averaging the two ratings for accuracy and nonintrusiveness.

5 Conclusion

In this paper, we proposed a content-adaptive method of embedding imperceptible pattern images and demonstrated that the method can significantly improve the imperceptibility of the embedded patterns on the one hand, produce better compensation results on the other hand, through a variety of experiments and subjective evaluation.

The block size for analyzing the color distribution and spatial variation would influence the imperceptibility of embedded pattern images. Therefore, it would also be valuable to divide AR images into different size of blocks depending on their spatial frequency and contrast, e.g. making a quad-tree.

Acknowledgments. This study was supported by a grant(02-PJ3-PG6-EV04-0003) of Ministry of Health and Welfare, Republic of Korea.

[1] Since the compensation error in the region with high spatial variation is usually well-perceptible, it is crucial to reduce the error.

References

1. Park, H., Lee, M.-H., Kim, S.-J., Park, J.-I.: Surface-Independent Direct-Projected Augmented Reality. In: Narayanan, P.J., Nayar, S.K., Shum, H.-Y. (eds.) ACCV 2006. LNCS, vol. 3852, pp. 892–901. Springer, Heidelberg (2006)
2. Park, H., Lee, M.-H., Seo, B.-K., Park, J.-I.: Undistorted Projection onto Dynamic Surface. In: Chang, L.-W., Lie, W.-N. (eds.) PSIVT 2006. LNCS, vol. 4319, pp. 582–590. Springer, Heidelberg (2006)
3. Yasumuro, Y., Imura, M., Manabe, Y., Oshiro, O., Chihara, K.: Projection-based Augmented Reality with Automated Shape Scanning. In: Proc. of SPIE EI'05, vol. 5664 pp. 555–562 (2005)
4. Raskar, R., Welch, G. Cutts, M., Lake, A., Stesin, L., Fuchs, H.: The Office of the Future: A Unified Approach to Image Based Modeling and Spatially Immersive Displays. In: Proc. of SIGGRAPH'98 (1998)
5. Cotting, D., Naef, M., Gross, M., Fuchs, H.: Embedding Imperceptible Patterns into Projected Images for Simultaneous Acquisition and Display. In: Proc. of ISMAR'04 pp. 100–109 (2004)
6. Bimber, O., Grundhöfer, A., Zollmann, S., Kolster, D.: Digital Illumination for Augmented Studios. J. of Virtual Reality and Broadcasting (2006)
7. Available at: http://gonzo.uni-weimar.de/ bimber/Pub/DIAS.pdf
8. Summet, J., Sukthankar, R.: Tracking Locations of Moving Hand-Held Displays Using Projected Light. In: Proc. of Pervasive'05, pp. 37–46 (2005)
9. Lee, J.C., Hudson, S.E., Summet, J.W., Dietz, P.H.: Moveable Interactive Projected Displays Using Projectors Based Tracking. In: Proc. of UIST'05, pp. 63–72 (2005)
10. Watson, A.B.: Temporal Sensitivity. Handbook of Perception and Human Performance (1986)
11. Wang, D., Sato, I., Okabe, T., Sato, Y.: Radiometric Compensation in a Projector-Camera System Based on the Properties of Human Vision System. In: Proc. of PROCAMS'05 (2005)