

Robust and Accurate Visual Echo Cancelation in a Full-duplex Projector-camera System

Miao Liao, Mingxuan Sun, Ruigang Yang
Center for Visualization and Virtual Environment
The University of Kentucky, Lexington, KY 40506, USA
{mliao3,msun2,ryang}@cs.uky.edu

Zhengyou Zhang
Microsoft Research
Redmond, WA 98052-6399, USA
zhang@microsoft.com

Abstract

We developed a comprehensive set of techniques to address the “visual echo” problem in a full-duplex projector-camera system. A calibration procedure records the geometric and photometric transfer between the projector and camera in a look-up table. With the calibration information, the predicted camera view of the projected image is compared against the captured camera image to find echo pixels. Only non-echo pixels are used for display, therefore achieving the goal of suppressing visual echo. Compared to previous techniques, our approach’s main advantages are two-fold. First, it accurately handles full color images with no assumption about the surface reflectance or the photometric response of the projector or camera. Secondly, it is robust to geometric registration errors and quantization effect. It is particularly effective for high-frequency contents such as texts and hand drawings. We demonstrate the effectiveness of our approach with a variety of real images in a full-duplex projector-camera system.

1. Introduction

During the last few years, driven by the diminishing cost and size of digital light projectors and cameras, we have seen a proliferation of research projects in using them for a variety of applications. The combination of projectors and cameras in a common space provides both the input and the output capabilities that enable a new paradigm for human-computer interactions, from motion tracking (e.g. [16, 11]), immersive self-calibrating displays (e.g., [15, 12, 22, 3, 13, 7]), to remote collaboration tools that bring geographically distributed people to share virtually the same physical space (e.g. [14, 6, 19]).

In a typical remote collaboration setup, two or more projector-camera pairs are “cross-wired”, as shown in Figure 1 (top), to form a full-duplex system for two-way communication. Images from the projector are mixed with real objects (such as papers with writings) to create a shared

space. Therefore the camera will capture an image in which the projected image is also embedded. If we simply send the image for display on the other end, there could be a feedback loop that will distort the projected image. As shown in Figure 1 (middle), the captured image is directly projected back. After a few frames, some part of the image becomes saturated and some part of the real writing has ghosting effect. This is analogous to the audio echo effect in a full-duplex phone system.

While the idea of using projectors and cameras as shown in Figure 1 for remote collaborations can be traced back to 1993 [21], there has been little effort reported to address this “visual echo” problem explicitly. We have so far found only two related papers [18, 23]. Both papers focus mainly on finding non-echo pixels (due to real writings or sketches) from a relatively simple projected image with a small number of color variations.

To suppress visual echo boils down to a classification problem. That is, one needs to identify whether or not a pixel in the camera image is a projected pixel reflected directly off the screen, i.e., an “echo”. If so, this pixel should be removed to cut off the feedback loop. Given an image to be projected (i.e., the framebuffer content), one needs to first predict accurately its appearance in the camera’s view and then compare it with the actual camera image to make the identification on a pixel-by-pixel basis.

We here present a comprehensive set of techniques to address the visual echo cancelation problem in general. Drawn upon the extensive research on projector calibration, we developed a simple and accurate procedure to find a complete photometric and geometric mapping between projector and camera pixels. It is a per-pixel mapping that allows arbitrary color transfer between the projector and the camera. From the mapping, we designed a robust classification scheme to find out echo pixels in the presence of quantization errors that are unavoidable when warping images from projector to camera, a problem mentioned in [23]. Figure 1 (bottom) shows the camera image before and after classification.

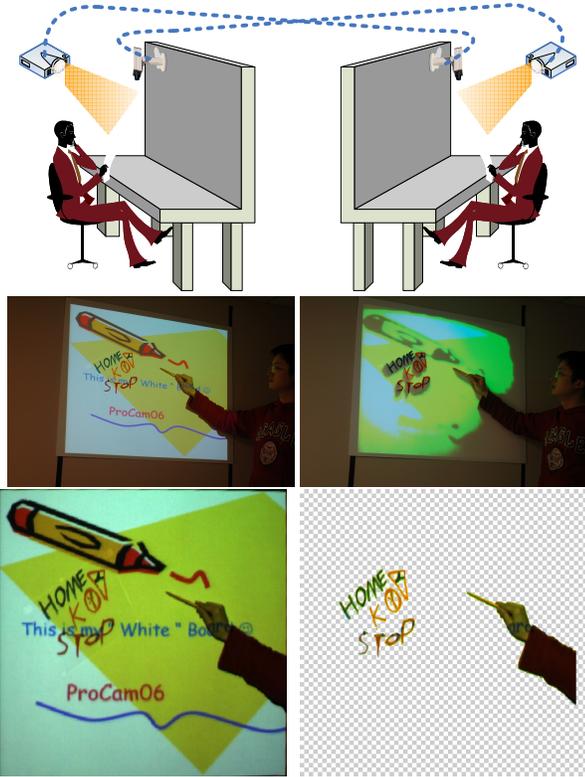


Figure 1. The problem of visual echo in a full duplex projector-camera system. (Top) two pairs of projector and camera are connected for remote collaboration. (Middle row) In the left, a user is annotating a presentation slide. A camera is capturing his annotations as well as the projected slide. The image is displayed directly, causing a visual echo. The right image shows the effect after seven frames. (Bottom row) To suppress visual echo, we present techniques to segment the captured image (left) into echo and non-echo pixels. The right image shows the non-echo pixels (i.e., the user’s writing and gesture). The light-colored checkerboard is used to illustrate the transparent background.

2. Background

The use of projectors and cameras for augmented reality applications has long been proposed. But it was until recently that the problem of “visual echoing” received attention. In the Telegraffiti system, a user’s sketch is captured by a camera and projected on a piece of paper in a remote site [19]. To avoid visual echoing in two-way communication, they increase the ambient lighting so that the real sketch (in black) is significantly darker than the projected sketch [18]. But the amount of lighting necessary is tricky to adjust. In their technical report [20], the visual echo problem is formally studied. The proposed solution is to adjust gain properly. It is demonstrated that it works quite well for extracting sketches from a clean background, but has difficulties with color or complex patterns.

The term *visual echo cancelation* is introduced in [23],

in which the basic steps to address such a problem are also articulated, which include geometric and photometric calibrations of the projector-camera pair and image difference. To find the color transfer, they proposed the use of a large 5D look-up-table, i.e., for each pixel (x, y) in the project, iterate through its full range (r, g, b) and record the observed color in the camera. The location dependency is necessary since commodity projectors typically have poor color uniformity. While this approach in theory can deal with *any* color transfer, it requires a prohibitively large table (e.g., $256^3 \times 1024 \times 768$). To make it practical, their implementation quantized the projector’s color space into $9 \times 9 \times 9$ bins and divided the image into 32×32 blocks. While it may be adequate for the driving application in [23]—to capture whiteboard contents, extending it to a more general scene is difficult. We show in this paper that by linearizing the camera’s spectral response, the 5D look-up table can be factorized into several independent tables that are orders smaller.

Geometric calibration of projectors and cameras has been very well studied [15, 12, 3, 13]) under the context of constructing tiled projection-based display that are *seamless*. Many of these techniques are able to register images with sub-pixel accuracy. Interested readers are referred to a recent survey of camera-based calibration techniques [1]. On the other hand, photometric calibration has received much less attention. Most projector-based displays simply blend linearly the pixels in the overlap regions. Previous methods mainly attack the intensity or color variations within or between projectors [17, 10, 9]. To deal with visual echo, we need more accurate chromatic estimation between the projector color space and camera color space.

The issue of predicting a projected image from the camera’s perspective has been studied under the context of shadow removal [2, 8] and adaptive projection on colored-surfaces [5]. In [2], the projected images are pre-captured by the cameras in the absence of any other object. In other words, the predicted images are captured directly. Therefore this method cannot deal with dynamic images not known a priori. The approach in [8] is a step forward: it estimates the photometric transfer function so it can process dynamic contents. They assume that the three color-channels in the projector can be mapped to the corresponding channels in the camera via independent intensity transfer functions. This assumption is valid for projectors and cameras that are color balanced and have narrow-band spectral responses. However, the typical spectral responses of cameras and projectors are wide band and have large overlaps [5]. The photometric model in [5] is probably the most general one to measure the color transfer between a projector and a camera. They have achieved some of the most accurate predictions, but they require the projector and camera to be co-axial to avoid the geometric alignment issue. This requirement does not scale to a multi-projector setup. Our

photometric model is similar to that in [5], but instead of trying to solve the color transfer matrix numerically, we use a look-up-table based approach to deal with projectors with non-linear responses.

Finally it should be noted that it is possible to avoid the visual echo problem in design. In [14], the cameras and projectors are synchronized, so the camera takes an image only when the projector is off or showing a specific pattern. This effectively interleaves the operation of the projector and camera in a time-sequential manner. To avoid visual flicking, DLP projectors have to be used to provide a fast enough switching rate ($> 1000\text{Hz}$). This approach is probably the most robust way to avoid visual echoes. But it requires modifications in hardware and it is usually difficult to synchronize a large network of projectors and camera in practice.

3. Our Methods

Our processing pipeline for visual echo cancelation starts with a calibration procedure to find out the geometric and photometric mapping between projector and camera pixels. For the scope of our paper, we assume that the projector, the camera, and the display surface are fixed. Therefore the calibration only needs to be done once per setup. At runtime, a classification procedure determines for each pixel in the camera image if it is a visual echo or from a real object, given the corresponding projector image. In the next few sections, we present in details how we perform these tasks.

3.1. Geometric Calibration

A straightforward way to find out the geometric mapping between the projector and the camera is to turn on one projector pixel at a time, and record its position in the camera’s frame. This approach is valid for any type of projector-camera configuration, even if the display surface is not planar. Furthermore, it can also deal with the lens distortions in both the projector and the camera. In [12], the structure-light technique is adopted to reduce the running time from $O(n)$ to $O(\log(n))$ where n is the number of projector pixels. It is also reported that the projector pixels can be sub-sampled on a regular grid to reduce the number of pixels to record. Those that are not sampled can be approximated by linear interpolation. Usually, this direct-mapping approach generates the most accurate mapping within a single pixel.

If the display surface is planar and the lens distortions can be ignored, the mapping can be expressed simply as a 3×3 perspective transformation, or a *homography*. In order to estimate the homography, one needs to project a minimum of only four points.

We have implemented both approaches for geometric calibration.

3.2. Photometric Calibration

Given a pixel in the projector space, we know its corresponding position in the camera space through geometric calibration described above. The task of photometric calibration is to predict the corresponding color from the camera’s perspective.

We first describe the photometric model we used. For a single point M on the display surface, it is illuminated by a point light source—a projector pixel. For the sake of discussion, let us for now assume that the projector and the camera have just one channel each. Let I be the pixel value to be projected and P be the projector brightness, then we have

$$P = h(I), \quad (1)$$

where h is the non-linear response function of the projector. Typically, the response curve is “gamma” like.

The projector brightness is then modulated by the spectral response $s(\lambda)$ of the projector where λ is the wavelength. Considering the effect of ambient illumination, the irradiance at M is written as

$$D(\lambda) = f(\lambda) + P \cdot s(\lambda), \quad (2)$$

where $f(\lambda)$ is the ambient light. We found that the inclusion of the ambient light term is important since most commodity projectors leak a substantial amount of light even when a blank image is projected.

Let $r(\lambda)$ be the spectral reflectance of M in the viewing direction of the camera. M ’s radiance in this direction is therefore:

$$L = (f(\lambda) + P \cdot s(\lambda))r(\lambda). \quad (3)$$

If $t(\lambda)$ is the spectral response for the camera, then the irradiance detected by the camera sensor is:

$$\begin{aligned} Q &= \int L \cdot t(\lambda) d\lambda \\ &= \int f(\lambda)r(\lambda)t(\lambda) d\lambda + P \int s(\lambda)r(\lambda)t(\lambda) d\lambda \end{aligned} \quad (4)$$

For a fixed setup, the integrations remain constant, therefore equation 4 can be simply written as

$$Q = A + P \cdot v, \quad (5)$$

where:

$$v = \int s(\lambda)r(\lambda)t(\lambda) d\lambda, \text{ and } A = \int f(\lambda)r(\lambda)t(\lambda) d\lambda \quad (6)$$

Finally the measured irradiance Q is converted to a pixel value C via a camera response function $g()$ that is typically non-linear too. So the entire transform from a projector pixel value I to a camera pixel value C is

$$C = g(Q) = g(A + h(I) \cdot v). \quad (7)$$

Now let us consider the case that the projector and the camera have three color channels (R, G, B). Following the similar analysis outline above, we can expand and rewrite equation 7 to:

$$\mathbf{Q} = \begin{bmatrix} g_R^{-1}(C_R) \\ g_G^{-1}(C_G) \\ g_B^{-1}(C_B) \end{bmatrix} = \mathbf{A} + \mathbf{VP}, \quad (8)$$

where :

$$\mathbf{A} = \begin{bmatrix} A_R \\ A_G \\ A_B \end{bmatrix}, \mathbf{V} = \begin{bmatrix} v_R^R & v_G^R & v_B^R \\ v_R^G & v_G^G & v_B^G \\ v_R^B & v_G^B & v_B^B \end{bmatrix}, \mathbf{P} = \begin{bmatrix} h(I_R) \\ h(I_G) \\ h(I_B) \end{bmatrix}.$$

The vector \mathbf{A} is the contribution due to the environmental light, including the black level of the projector. The matrix \mathbf{V} , typically referred to as the *color mixing matrix*, is the interaction of the spectral responses of the projector and camera. The superscript and subscript of v indicate the response from the corresponding projector or camera color channel.

Note that because of the non-linear transforms ($h(), g()$), the color transfer from the projector to the camera is not linear. In [23], a brute-force approach is used to record the entire transform as a look-up table, resulting in a prohibitively large table. However, if we find out the camera response curve $g()$, we can recover the irradiance value from a pixel value. Then we can decompose equation 8 as:

$$\mathbf{Q} = (\mathbf{A} + \mathbf{VP}_1) + (\mathbf{A} + \mathbf{VP}_2) + (\mathbf{A} + \mathbf{VP}_3) - 2\mathbf{A}, \quad (9)$$

where:

$$\mathbf{P}_1 = \begin{bmatrix} h(I_R) \\ 0 \\ 0 \end{bmatrix}, \mathbf{P}_2 = \begin{bmatrix} 0 \\ h(I_G) \\ 0 \end{bmatrix}, \mathbf{P}_3 = \begin{bmatrix} 0 \\ 0 \\ h(I_B) \end{bmatrix}.$$

Now we can use four separate look-up tables, three for RGB and one for ambient, to record the color transfer function because the resulting irradiance values are linear and additive. For a projector pixel value $I(r, g, b)$, its predicted color could be obtained by subtracting the ambient contribution twice from the sum of the responses of all three channels.

Our photometric model is similar to that in [5], but we did not solve the color transfer matrix \mathbf{V} numerically. Instead we use a look-up-table based approach to avoid the step for projector calibration (i.e., finding out the function $h()$). It is also worth mentioning that some previous approaches [18, 8] to predict camera images treat the three RGB channels independently, i.e. red maps only to red, green maps only to green, and blue maps only to blue. This is equivalent to assuming the color mixing matrix \mathbf{V} is diagonal.

Our photometric calibration procedure is summarized in Algorithm 1. After calibration, for each pixel (x, y) in the

Algorithm 1 Photometric Calibration

- 1: Estimate the camera response curves ($g_R(), g_G(), g_B()$) and their inverse as in [4];
 - 2: project the scene with a blank image (i.e., color value $\leftarrow [0 \ 0 \ 0]$)
 - 3: capture the projected image A_0
 - 4: Initialize the ambient look-up table A such that $A[x, y] = \text{linearized}(A_0[x, y])$
 - 5: **for all** color channel $i \in \{R, G, B\}$ **do**
 - 6: **for all** intensity value $j \in [0 \cdots 255]$ **do**
 - 7: project the pure-color image for j in i and capture the camera view C
 - 8: Initialize the color look-up table LUT such that $LUT[x, y][i][j] = \text{linearized}(C[x, y])$
 - 9: **end for**
 - 10: **end for**
 - 11: **function** *linearize*($[R, G, B]$)
 - 12: return $[g_R^{-1}(R), g_G^{-1}(G), g_B^{-1}(B)]$
-

camera, there is a 256×3 look-up table, each table cell, indexed by a color channel and an intensity value, holds a *linearized* RGB tuple. In addition, there is a table to record the ambient lighting term. Thus, if a projector pixel with color $I(r, g, b)$ is warped to a pixel location (x, y) in the camera's view, its predicted color could be obtained by:

$$Q(r, g, b) = LUT[x, y][R][r] + LUT[x, y][G][g] + LUT[x, y][B][b] - 2\mathbf{A}[x, y] \quad (10)$$

3.3. Online Visual Echo Removal

After the calibration, we are ready to segment camera images to remove visual echo. Given a projector image and its corresponding camera image, the following three steps are applied:

1. *Geometric warp* the projector image is warped into the camera's reference frame, either through a direct mapping or a homography;
2. *Color Transfer* for every pixel in the warped projector image I_p , its appearance in the camera is predicted by equation 10. The captured camera image is also transferred by the recovered camera response curve $g()$ to obtain an "irradiance" image I_c .
3. *Image Classification* The color difference is computed between I_p and I_c , i.e.,

$$e = \|I_p(x, y) - I_c(x, y)\|. \quad (11)$$

If e is smaller than a threshold, than it is a visual echo.

The output from the above procedure is a segmented image in which all the "echo" pixels have been removed. That

image can be sent for display on the remote site without causing any visual echo.

Feature-based Post-processing In practice we found that there are quite some false negatives in the identified echo pixels. This is primarily due to the error in the geometric calibration and quantization. Here we introduce a novel procedure to increase the classification robustness.

For a pixel $I_c(x, y)$ in the camera image, if it is an echo, it should appear somewhere near $I_p(x, y)$ if not exactly at $I_p(x, y)$. So we could search around $I_p(x, y)$ in a small neighborhood, typically less than 3×3 , to find if there is a corresponding matching pixel. While matching a single pixel is usually not robust enough to differentiate the echo pixels from non-echo ones, we match over a small (3×3) template window $T_c(x, y)$. For $T_c(x, y)$ and $T_p(u, v)$ where (u, v) is a small neighborhood around (x, y) , if the sum of absolute intensity differences between them is less than a certain threshold, we classify $I_c(x, y)$ as an echo. We have found that this method *dramatically* reduces the false negative rate for echo pixels.

To further suppress noise, we also apply a 3×3 median filter to the non-echo images.

Projector-Camera Image Synchronization Another practical issue we discovered in building our projector-camera pairs is image synchronization. The display and the capture routine are in two different threads. Both the projector and camera have certain delays in processing the imagery, and these delays vary due to many factors. Therefore, for dynamic contents, we need to have a means to establish frame correspondences between projector images and camera images. To address this, every time we project an image, we draw a four-digit bar code in a small corner of the image and store this image with its ID in a circular buffer. When the camera captures an image, we first decode the bar code on the image, and then based on the ID, retrieve the corresponding projector image in the buffer. Note that since we know the projector-camera mapping from calibration, there is no need to search for the bar code in the camera image, yielding a very efficient algorithm. Furthermore the bar code can be made very small to avoid causing distractions.

4. Experimental Results

We have set up a full duplex system. At each end, we use an XGA projector and a video camera, both pointing to a white matte surface. The projector-camera pair is connected to a computer equipped with a 3GHz CPU, 1G RAM, and a Geforce 6800 graphics card. Both PCs are connected via LAN to transfer image data. The whiteboard covers approximately a 500-pixel-squared area in the camera. Therefore

we set the color look-up table to be $512 \times 512 \times 256 \times 3$, with a (r, g, b) tuple in each table cell.

Since our display surface is flat, we estimated the homography between the projector and the camera. During our experiments we found that the surface is not very rigid and the projector image is not stable (the projector in the ceiling is connected to the PC via a 30-foot long VGA cable). All these resulted in the inaccuracy in our geometric calibration. Nevertheless, our feature-based post-processing successfully overcame this limitation.

Figure 2 shows our visual echo cancelation results on various projected background. Notice the busy background both in color and contents. Figure 3 shows the differences between the camera images and images predicted by different photometric calibration methods. We can see that our method achieved more accurate prediction with the least pixel error. Figure 3 also shows the robustness of our post-processing method in the presence of misalignment.

5. Conclusion

We have demonstrated a comprehensive pipeline for visual echo cancelation in a full-duplex projector camera system. Both ends conduct a geometric and photometric calibration off-line. With the recorded calibration information, at each end, the camera view image predicted from projector image is compared against the real captured image to find echo pixels at run-time. Only non-echo pixels are transferred in network and used for display, therefore achieving the goal of suppressing visual echo.

Compared to some previous methods, our color look-up-table model obtained in the photometric calibration shows a significant advantage in initial echo-pixel classification. The following feature-based post-processing method further reduces the false negative rate for classifying each pixel, which leads to dramatically improved segmentation result.

Looking into the future there are some places for improvement. We would like to better set up our projectors to reduce the misalignment. A further step is to improve the performance for video-rate update. Our current un-optimized code takes about 500ms to update a frame. Most of the time is spent on the feature search. We hope to dramatically speed up this process through software optimization or even using the computational power in the graphics card.

Acknowledgement This work is supported in part by University of Kentucky Research Foundation and US National Science Foundation award IIS-0448185.

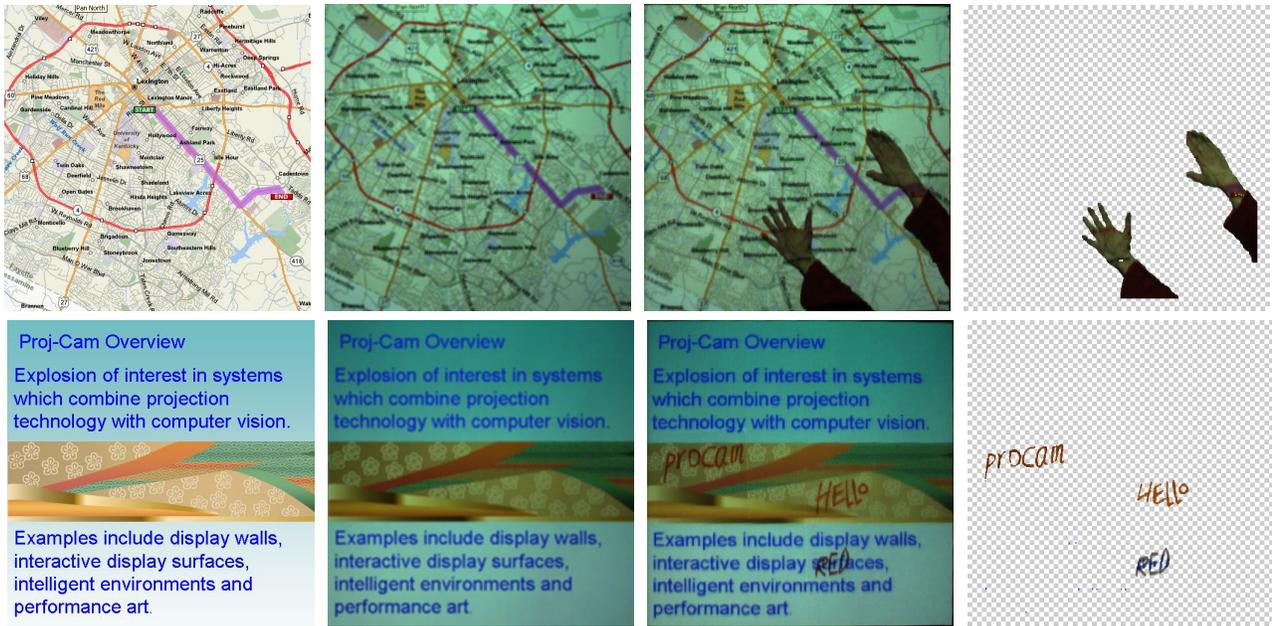


Figure 2. Visual echo cancellation results produced by our method. From left to right, (1st column) original projector images; (2nd column) predicted images by our photometric calibration method; (3rd column) images captured by the camera; (4th column) images after visual echo cancellation; only non-echo pixels are shown.

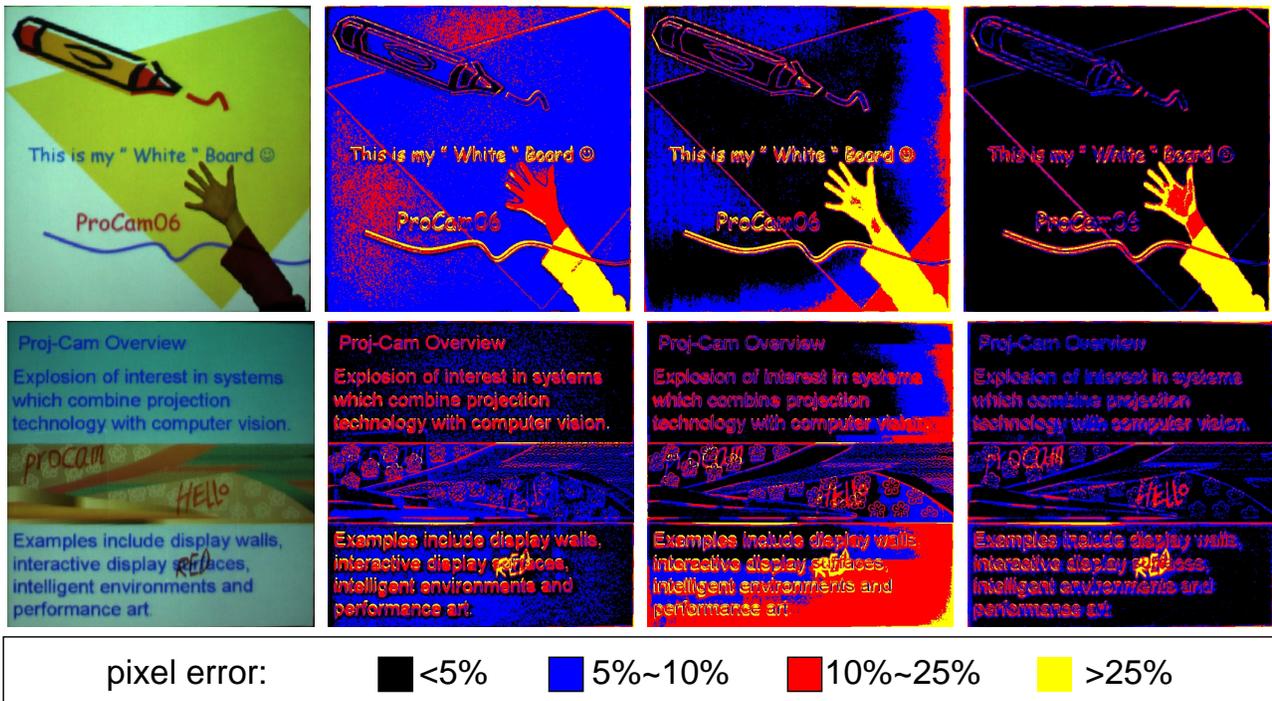


Figure 3. The pixel color error between the captured images and predicted images by several photometric calibration methods. From left to right, (1st column) images captured by the camera. (2nd column) color error from images predicted by three (RGB) independent intensity transfer functions—an assumption made in [18, 8]. (3rd column) color error from images predicted using the method in [23]. The projector's color space is quantized into $16 \times 16 \times 16$ bins and the image is divided into 32×32 blocks. (4th column) color error from our predicted images.

References

- [1] M. Brown, A. Majumder, and R. Yang. Camera-Based Calibration Techniques for Seamless Multi-projector Displays. *IEEE Transactions on Visualization and Computer Graphics*, 11(2):193–206, 2005. 2
- [2] T. Cham, J. Rehg, R. Sukthankar, and G. Sukthankar. Shadow elimination and occluder light suppression for multi-projector displays. *Proceedings of Computer Vision and Pattern Recognition*, 2003. 2
- [3] H. Chen, R. Sukthankar, and G. Wallace. Scalable Alignment of Large-Format Multi-Projector Displays Using Camera Homography Trees. In *Proceeding of IEEE Visualization 2002*, pages 339–346, 2002. 1, 2
- [4] P. E. Debevec and J. Malik. Recovering High Dynamic Range Radiance Maps from Photographs. *Proceedings of ACM Siggraph*, pages 369–378, 1997. 4
- [5] K. Fujii, M. Grossberg, and S. Nayar. A Projector-Camera System with Real-Time Photometric Adaptation for Dynamic Environments. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 814–821, 2005. 2, 3, 4
- [6] D. Hall, C. L. Gal, J. Martin, O. Chomat, T. Kapuscinski, and J. Crowley. Magicboard: A contribution to an intelligent office environment. In *Proc. of the International Symposium on Intelligent Robotic Systems*, 1999. 1
- [7] C. Jaynes, B. Seales, K. Calvert, Z. Fei, and J. Griffoen. The Metaverse - A Collection of inexpensive, self-configuring,immersive environments. In *Proceeding of 7th International Workshop on Immersive Projection Technology/Eurographics Workshop on virtual Environments*, 2003. 1
- [8] C. Jaynes, S. Webb, and R. Steele. Camera-based detection and removal of shadows from interactive multiprojector displays. *IEEE Transactions on Visualization and Computer Graphics*, 10(3):290–301, 2004. 2, 4, 6
- [9] A. Majumder, D. Jones, M. McCrory, M. E. Papka, and R. Stevens. Using a Camera to Capture and Correct Spatial Photometric Variation in Multi-Projector Displays. In *Proceeding of IEEE International Workshop on Projector-Camera Systems*, 2003. 2
- [10] A. Majumder and R. Stevens. Color Nonuniformity in Projection-Based Displays: Analysis and Solutions. *IEEE Transactions on Visualization and Computer Graphics*, 10(2), 2003. 2
- [11] C. Pinhanez. ”the everywhere displays projector: a device to create ubiquitous graphical interfaces. In *Ubicomp01, Proceedings of an International Conference on Ubiquitous Computing*, 2001. 1
- [12] R. Raskar, M. Brown, R. Yang, W. Chen, G. Welch, H. Towles, B. Seales, and H. Fuchs. Multi-projector displays using camera-based registration. In *Proceeding of IEEE Visualization 1999*, pages 161–168, 1999. 1, 2, 3
- [13] R. Raskar, J. van Baar, P. Beardsley, T. Willwacher, S. Rao, and C. Forlines. ilamps: Geometrically aware and self-configuring projectors. *ACM Transactions on Graphics (SIGGRAPH 2003)*, 22(3):809–818, 2003. 1, 2
- [14] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stessin, and H. Fuchs. The Office of the Future: A Unified Approach to Image-Based Modeling and Spatially Immersive Displays. *Computer Graphics*, 32(Annual Conference Series):179–188, 1998. 1, 3
- [15] R.Surati. *Scalable Self-Calibrating Display Technology for Seamless Large-Scale Displays*. PhD thesis, Department of Computer Science, Massachusetts Institute of Technology, 1998. 1, 2
- [16] Y. Sato, Y. Kobayashi, and H. Koike. Fast tracking of hands and fingertips in infrared images for augmented desk interface. In *Proc. of the 4th Intl. Conf. on Automatic Face and Gesture Recognition*, 2000. 1
- [17] M. C. Stone. Color and brightness appearance issues in tiled displays. *IEEE Comput. Graph. Appl.*, 21(5):58–66, 2001. 2
- [18] N. Takao, S. Baker, and J. Shi. Steady-state feedback analysis of tele-graffiti. In *Proceedings of the IEEE International Workshop on Projector-Camera Systems*, October 2003. 1, 2, 4, 6
- [19] N. Takao, J. Shi, and S. Baker. Telegraffiti: A camera-projector based remote sketching system with hand-based user interface and automatic session summarization. *International Journal of Computer Vision (IJCV)*, 53(2):115–133, 2003. 1, 2
- [20] N. Takao, J. Shi, and S.Baker. Tele-graffiti. Technical Report CMU-RI-TR-02-10, Robotics Institute, Carnegie Mellon University, 2002. 2
- [21] P. Wellner. Interactivng with paper on the DigitalDesk. *Communications of the ACM*, 36(7):86–97, 1993. 1
- [22] R. Yang, D. Gotz, J. Hensley, H. Towles, and M. Brown. PixelFlex: A Reconfigurable Multi-Projector Display System. In *Proceeding of IEEE Visualization 2001*, pages 167–174, San Diego, CA, 2001. 1
- [23] H. Zhou, Z. Zhang, and T. Huang. Visual Echo Cancellation in a Projector-Camera-Whiteboard System. In *Proceedings of ICIP*, pages 2885–2888, 2004. 1, 2, 4, 6