A robust photometric calibration framework for projector-camera display system

Wenhai Zou (邹文海)* and Haisong Xu (徐海松)

State Key Laboratory of Modern Optical Instrumentation, Zhejiang University, Hangzhou 310027, China

**E-mail: zwh_gd2001@yahoo.com.cn*

Received August 4, 2008

A novel photometric calibration framework is presented for a projector-camera (ProCam) display system, which is currently under booming development. Firstly, a piecewise bilinear model and five 5-ary color coding images are used to construct the homography between the image planes of a projector and a camera. Secondly, a photometric model is proposed to describe the data flow of the ProCam display system for displaying color images on colored surface in a general way. An efficient self-calibration algorithm is correspondingly put forward to recover the model parameters. Aiming to adapt this algorithm to different types of ProCam display system robustly, a 3×7 masking coupling matrix and a patches image with 1024 color samples are adopted to fit the complex channel interference function of the display system. Finally, the experimental results demonstrate the validity and superiority of this calibration algorithm for the ProCam display system.

OCIS codes: 100.5010, 100.2000, 120.2040, 330.0330.

doi: 10.3788/COL20090706.0479.

In the last decade, projectors have broken out of the traditional role as awkward, simplex output devices and have been transformed to be portable[1], environment-sensing, interactively-communicating, and multifunctional display systems. A good example of these display systems is the projector-camera (ProCam) display system, of which the basic unit is a projector attached with a camera. Depending on the modern vision art, the ProCam display system has been commonly used in many developed applications, such as keystone correction[2], smart presentations[3], and tiled display[4]. Moreover, current efforts in this area have made it clear that the application outlook of the ProCam display system will be highly exploited in the near future by new ideas like controlling the appearance of three-dimensional (3D) object[5], iLamp and radion frequency identity and geometry (RFIG) lamp[6], 3D multi-spectral scanner[7], etc. Without exception, the ProCam display systems involved in these applications are all needed to be geometrically and photometrically pre-calibrated to achieve the inner characteristics of the display system, and to ensure the expected accurate results. Moreover, the screen for the ProCam display system is not constrained to a high-quality white screen but has been extended to an ordinary curved and colored surface[5,7].

Concerned about the photometric calibration problem of the ProCam display system, there are much existed work on the projector's and the camera's calibration, respectively. In Refs. [8, 9], several general methods were introduced to estimate the input transfer function of a camera and the precision optical instruments, for example, photometers or colorimeters were naturally used to obtain the transfer function of the projectors. However, due to the limited range of the measurement at a time using the precision optical instruments, these methods are inefficient and cannot recover the spatial variation of projectors accurately. Therefore, the calibrated cameras are recently introduced to estimate the projector's transfer function and several techniques have been developed to calibrate the ProCam display system[10,11]. In these methods, the projector is calibrated beforehand using the precision optical instrument and meanwhile the camera is pre-calibrated under ideal environment, and then the calibrated camera is used to estimate the projector's transmission function on a perfectly-white screen. These techniques can obtain relatively accurate calibration results but require fussy processes. More recently, online self-calibration techniques were developed, which did not depend on other calibration aids but extracted the calibration information from the user-projected imagery and their camera-outputs at different settings[5,12]. These techniques show high efficiency but suffer from the simple assumptions of color mixing between the projector and the camera or being only applied to the monochrome display surfaces and grayscale images. In this letter, a general model and a novel algorithm are proposed to extend the photometric calibration method to be adaptive to different types of ProCam display system, such as its projector operating in different modes for displaying color images on colored surface for different application purposes.

The ProCam display system is composed of a projector (VGA NEC LT 30+) with a native resolution of 1024 × 768 pixels and a camera (HITACHI HV-D30) with a resolution of 768 × 576 pixels. The images are projected onto the screen via a display card (RADEON R9200SE), and those images from the camera are captured by an 8-bit Matrox Meteor II/Multi-channel frame-grabber. We assumed that the color channel numbers of the projector and camera both were three (e.g., red (R), green (G), blue (B)).

Before photometric calibration, the geometric mapping, which is defined as $m(x, y) = (u, v)$ from the pixel (x, y) in the image plane of the projector to the corresponding pixel (u, v) of the camera, should be constructed. If m is divided into proper numbers of subsec-

tions, each section can be simply fitted by using a bilinear interpolation function, which is defined in matrix form as

$$[u \quad v]^{\mathrm{T}} = \mathbf{A} \cdot [xy \quad x \quad y \quad 1]^{\mathrm{T}}, \tag{1}$$

where $\mathbf{A}$ is a 2 ×4 matrix including the unknown coefficients of the mapping. By projecting 625 square patches uniformly spaced in the display domain and scanning the patches efficiently by a 5-ary color coding, 625 corresponding points in the two image planes are obtained using just five projected images. Every four groups of nearest-neighboring points are adopted to calculate the coefficient matrix using the least square method and the final geometric mappings are stored as a look-up table. Compared with the binary coding algorithm[5], the mean and root mean square (RMS) errors of this algorithm are 0.93 and 0.24 pixels, respectively, which are accurate enough for the photometric calibration.

Supposing that the projector, the camera, and the screen surface are all time-invariant, the data flow of a single pair ProCam display system can be abstracted as illustrated in Fig. 1. Let r, g, and b define the digital values for one pixel of the image, an input image $\mathbf{I}_{\mathrm{in}}(x, y)$, where $\mathbf{I}_{\mathrm{in}} = [r_{\mathrm{in}},\ g_{\mathrm{in}},\ b_{\mathrm{in}}]^{\mathrm{T}}$, is firstly transformed by the spatially invariant channel response functions, f_r^{p}, f_g^{p} and f_b^{p}, of the projector and mapped to the transformed image $\mathbf{I}_{\mathrm{tr}}(x, y)$, where $\mathbf{I}_{\mathrm{tr}} = [r_{\mathrm{tr}},\ g_{\mathrm{tr}},\ b_{\mathrm{tr}}]^{\mathrm{T}} = [f_r^{\mathrm{p}}(r_{\mathrm{in}}),\ f_g^{\mathrm{p}}(g_{\mathrm{in}}),\ f_b^{\mathrm{p}}(b_{\mathrm{in}})]^{\mathrm{T}}$. Then, the transformed image $\mathbf{I}_{\mathrm{tr}}(x, y)$ is modulated by the spatially uniform channel interference function, f_{m}, between the projector and the camera to create the projected image $\mathbf{I}_{\mathrm{p}}(u, v) = f_{\mathrm{m}}(\mathbf{I}_{\mathrm{tr}}(x, y))$, where $\mathbf{I}_{\mathrm{p}} = [r_{\mathrm{p}},\ g_{\mathrm{p}},\ b_{\mathrm{p}}]^{\mathrm{T}}$. Based on the assumption that the linear reflectance model can accurately characterize a wide range of physical surfaces[13], the transformed image is further modulated by a spatially dependent but input-independent intensity variation, $p(u, v)$, of the projector and the spatially variant surface's reflectance function, $\mathbf{S}(u, v)$, to form the displayed image $\mathbf{I}_{\mathrm{d}}(u, v)$=$\mathbf{S}(u, v) \times p(u, v) \times \mathbf{I}_{\mathrm{p}}(u, v)$, where $\mathbf{I}_{\mathrm{d}} = [r_{\mathrm{d}},\ g_{\mathrm{d}},\ b_{\mathrm{d}}]^{\mathrm{T}}$ and $\mathbf{S}$=$[s_{\mathrm{r}},\ s_{\mathrm{g}},\ s_{\mathrm{b}}]^{\mathrm{T}}$. This image is reflected by the screen and captured by the camera to produce the image $\mathbf{I}_{\mathrm{c}}(u, v)$=$\mathbf{I}_{\mathrm{d}}(u, v) \times c(u, v)$, where c represents the spatially intensity variance of the camera, and $\mathbf{I}_{\mathrm{c}} = [r_{\mathrm{c}},\ g_{\mathrm{c}},\ b_{\mathrm{c}}]^{\mathrm{T}}$. Finally, the captured image is further processed by the spatially invariant channel response functions, f_r^{c}, f_g^{c} and f_b^{c}, of the camera to obtain the output image $\mathbf{I}_{\mathrm{out}} = [r_{\mathrm{out}},\ g_{\mathrm{out}},\ b_{\mathrm{out}}]^{\mathrm{T}} = [f_r^{\mathrm{c}}(r_{\mathrm{c}}),\ f_g^{\mathrm{c}}(g_{\mathrm{c}}),\ f_b^{\mathrm{c}}(b_{\mathrm{c}})]^{\mathrm{T}}$. Thus, the output image is a function of the input image $\mathbf{I}_{\mathrm{in}}$ and the spatial coordinates (x, y), and the equation can be finally described as

$$\begin{aligned} i_{\mathrm{out}}\ (\mathbf{I}_{\mathrm{in}}, m(x, y)) &= f_i^{\mathrm{c}}(k_i(m(x, y)) \\ &\times (f_{\mathrm{m}}(f_r^{\mathrm{p}}(r_{\mathrm{in}}(x, y)), f_g^{\mathrm{p}}(g_{\mathrm{in}}(x, y)), f_b^{\mathrm{p}}(b_{\mathrm{in}}(x, y))))_i), \\ i &= r, g, b \end{aligned} \tag{2}$$

where $k_i(m(x, y)) = c(m(x, y)) \cdot s_i(m(x, y)) \cdot p(m(x, y))$ represents the spatial variation for each color channel. Hence, the task of the calibration for the projector-camera based display display system is to recover the function of f^{c}, f^{p}, and f_{m} and the coefficient $\mathbf{K}$=$[k_r,\ k_g,\ k_b]^{\mathrm{T}}$.

Generally, f^{c}, f^{p}, and f_{m} are all nonlinear functions. Firstly, the function f^{c} can be determined offline only once, and a calibration chart or auto-calibration techniques[12] on the screen may be used and the range of the captured color $\mathbf{I}_{\mathrm{c}}$ is constrained to be [0,1]. Once this is done, the look-up table ($\mathbf{LUT}^{\mathrm{c}}$) mapping the output color $\mathbf{I}_{\mathrm{out}}$ to the captured color $\mathbf{I}_{\mathrm{c}}$ can be constructed. Then, because the coefficient $\mathbf{K}$ represents the spatial variation of the input image, it can be efficiently scaled by projecting a flat-white (r=g=b=255) image and recovered by normalizing the captured colors to the objective white point. Since the camera is color corrected in prior and adopted as the proxy for the viewer in the ProCam display system, the objective white point is simply assumed to be [1, 1, 1] here. Finally, a series of monochromic uniform patches are further projected and a series of uniform patches for their projection output are recovered by using $\mathbf{LUT}^{\mathrm{c}}$ and $\mathbf{K}$ to obtain the channel response function f^{p}. Meanwhile, inspired by the existed studies for the characterization of liquid crystal display (LCD) or digital light prosessing (DLP) projectors[14,15], the calibration of the function f_{m} is to find a $3 \times L (L = 7)$ matrix $\mathbf{M}$ to map the color $\mathbf{I}_{\mathrm{tr}}$ to the corresponding color $\mathbf{I}_{\mathrm{p}}$:

$$\mathbf{I}_{\mathrm{p}} = \mathbf{M} \cdot \mathbf{g}, \tag{3}$$

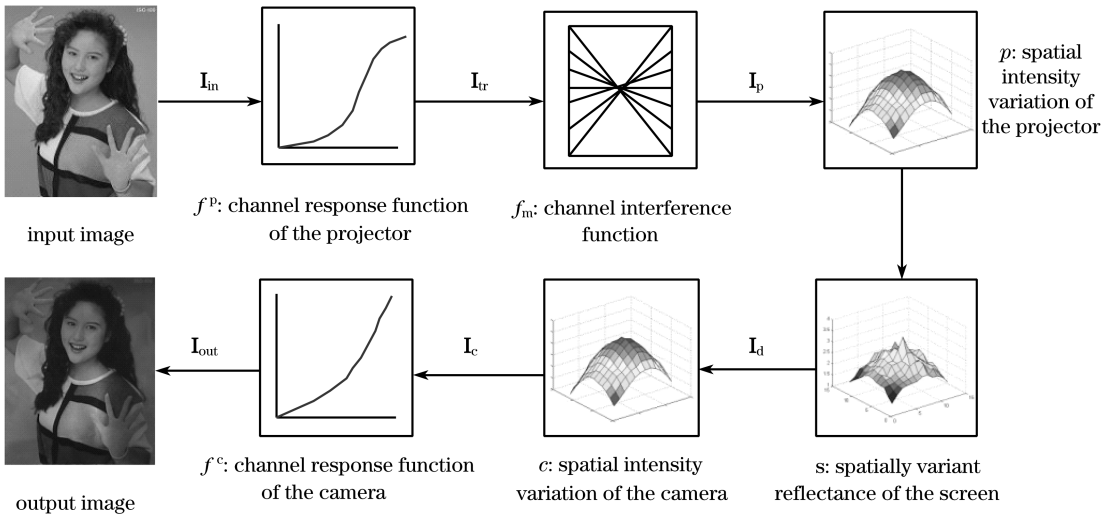

Fig. 1. Data flow chart of a single pair projector-camera display system.

where $\mathbf{g}=[r_{\rm e}, g_{\rm e}, b_{\rm e}, c_{\rm e}, m_{\rm e}, y_{\rm e}, k_{\rm e}]^{\rm T}$ is a color vector extended from $\mathbf{I}_{\rm tr}$ using the masking model approach[16]. Given that n color samples are captured and let $\mathbf{I}_{\rm p}^{i}$ and $\mathbf{g}^{i}$ $(i = 1,\ 2,\ 3, \cdots,\ n)$ be the ith captured color $\mathbf{I}_{\rm p}$ and the extended color vector of the ith patch, respectively, then the matrix $\mathbf{M}$ can be solved using Moore-Penrose pseudo-inverse as

$$\mathbf{M} = (\mathbf{Q}\mathbf{G}^{\rm T})(\mathbf{G}\mathbf{G}^{\rm T}), \tag{4}$$

where $\mathbf{Q}=[\mathbf{I}_{\rm p}^{1}, \mathbf{I}_{\rm p}^{2},\ \mathbf{I}_{\rm p}^{3}, ...,\ \mathbf{I}_{\rm p}^{n}]$ and $\mathbf{G} =[\mathbf{g}^{1},\ \mathbf{g}^{2},\ \mathbf{g}^{3}, ...,\ \mathbf{g}^{n}]$. In our experiment, a calibration pattern including 1024 color patches, in which 512 colors were evenly (8×8×8) sampled from the device color space (RGB), was produced, as shown in Fig. 2(a), for the recovery of the matrix $\mathbf{M}$. Figure 2(b) shows the test screen, of which the reflectance is spatially variant. The image in Fig. 2(c) is the variant pattern when the calibration pattern is projected onto the test screen and captured by the camera. Figure 2(d) shows the transformed pattern inversely computed from the variant pattern using the obtained function $\mathbf{LUT}^{\rm c}$ and the coefficient $\mathbf{K}$ as

$$i_{\rm tr}(x, y) = {\rm LUT}_{i}^{\rm c}(i_{\rm out}(x, y))/k_{i}, \qquad i = r, g, b. \tag{5}$$

Note that the transformed pattern was hardly dependent of the screen's reflectance and spatial intensity variance of the projector and the camera when compared with the variant pattern, which successfully demonstrated the assumption of linear reflectance for usual physical surfaces.

In order to verify the accuracy of the estimated parameters, some standard digital images selected from ISO/JIS-SCID are used as the testimages, and one of which is given in Fig. 3(a). Based on the geometric mapping and the photometric calibration results when the projector is operated in the "sRGB" display mode, the predicted output image of the camera for the standard image can be rendered as Fig. 3(b). Figure 3(c) shows the expected output image for the standard image, which is obtained by directly projecting the standard image onto the test screen and then capturing it by the camera of ProCam display system. By comparing Fig. 3(b) with Fig. 3(c) pixel by pixel, the mean ($\mathbf{E}_{\rm mean}$) and RMS ($\mathbf{E}_{\rm RMS}$) errors of the absolute difference for the channel R, G and B among them are defined as

$$\mathbf{E}_{\rm mean} = \frac{1}{N}\sum_{j=1}^{N}\Delta\mathbf{E}_{j},$$

$$\mathbf{E}_{\rm RMS} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\Delta\mathbf{E}_{i} - \mathbf{E}_{\rm mean})^{2}},$$

$$\Delta\mathbf{E} = |\mathbf{I}_{\rm O} - \mathbf{I}_{\rm R}|, \tag{6}$$

where N is the number of the compared pixels, $\mathbf{I}_{\rm R}$, $\mathbf{I}_{\rm O}$ are the referenced and objective color, respectively. The results are listed in the sub-column of "Pro." under "sRGB" in Table 1. Hereby, the rendered result is very close to the expected output image, which verifies the feasibility of this framework. Moreover, as seen from the pseudo-color in Fig. 3(d), taking the channel B for example and actually the same for other channels, the rendered error of Fig. 3(b) mainly occurred on the edges of the details, which is mainly due to the limited resolution of the camera used in this experiment. Furthermore, in order to demonstrate the superiority, this algorithm is also run in the "presentation" mode and compared with the 3×3 color mixing matrix presented in Ref. [4]. As shown in Table 1, when the projector operated in the "sRGB" mode, the simple 3×3 matrix is enough for encoding the color mixing between the projector and the camera. However, because the DLP projector is enhanced with the white channel in the "presentation" mode, the proposed 3×7 masking coupling matrix outperforms the simple method and produces robust calibration results.

In conclusion, a detailed photometric model for the ProCam display system is presented, completely considering the vignette effect and the channel interfer-

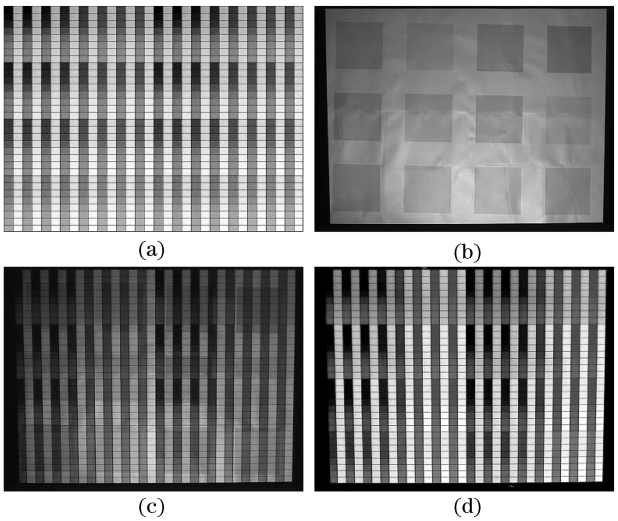

Fig. 2. (a) A calibration pattern used to recover the matrix $\mathbf{M}$ and (b) a screen with chromatic patches used to test our calibration method. The calibration pattern (a) is projected onto the test screen and captured by the camera, resulting in (c). From this, a transformed pattern (d) is computed by using Eq. (5).

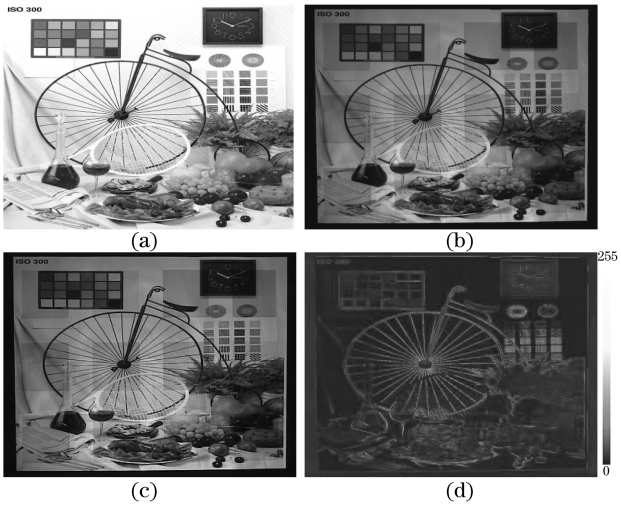

Fig. 3. (a) The standard image used to test our algorithm. (b) The predicted result based on the calibration results. (c) The standard image projected and captured by our display system. (d) The pseudo color image to show the absolute difference for the channel B between (b) and (c).

Table 1. Mean (E_{mean}) and RMS (E_{RMS}) Errors of the Rendered Output Compared to the Expected Output for the Test Image in Fig. 3(a) on the Test Screen in Fig. 2(b) in the "sRGB" and "Presentation" Modes. The Results for the Simple 3 × 3 Mixing Matrix (Abbreviated as "Sim.") and the Proposed 3×7 Masking Coupling Matrix (Abbreviated as "Pro.") Are Listed for Comparison

Channel		sRGB		Presentation	
		Sim.	Pro.	Sim.	Pro.
R	E_{mean}	6.1	6.2	19.5	7.0
	E_{RMS}	10.7	10.8	24.4	11.4
G	E_{mean}	6.2	6.2	21.1	6.8
	E_{RMS}	11.0	11.5	25.4	11.4
B	E_{mean}	9.7	8.5	23.4	8.7
	E_{RMS}	15.0	14.7	29.6	15.2

ence of the display system for color images displayed on colored surface. An efficient calibration algorithm, adaptive to different types of ProCam display system, is developed to recover the parameters of the photometric model. This method does not depend on other calibration aids or physical measurements, but instead extracts the calibration information by projecting a few user-specified imageries on the color surface. The experimental results successfully verify the feasibility of this framework and especially its priority for the multi-channel projector including in the ProCam display system. However, due to the limited resolution of the camera in the experimental display system, the calibration result shows somewhat obvious error for the rendering of detail content in the image. Given that the high resolution camera is used in the ProCam display system, the calibration accuracy will be expected to be increased. On the other hand, how to further correct the color artifacts of the projection display, as shown in Fig. 3(c), based on this calibration framework will be the next issue of our research.

References

1. X. Zhao, Z. Fang, J. Cui, X. Zhang, and G. Mu, Acta Opt. Sin. (in Chinese) **27,** 913 (2007).
2. R. Raskar and P. Beardsley, in *Proceedings of IEEE Conference on CVPR* 626 (2001).
3. R. Sukthankar, R. G. Stockton, and M. D. Mullin, in *Proceedings of ICCV* 247 (2001).
4. G. Wallace, H. Chen, and K. Li, in *Proceedings of the Workshop on Virtual Environments* 293 (2003).
5. M. D. Grossberg, H. Peri, S. K. Nayar, and P. N. Belhumeur, in *Proceedings of IEEE Conference on CVPR* 4529 (2004).
6. R. Raskar, P. Beardsley, J. Baar, Y. Wang, P. Dietz, J. Lee, D. Leigh, and T. Willwacher, ACM Transactions on Graphics **23,** 406 (2004).
7. A. Mansouri, A. Lathuilière, F. S. Marzani, Y. Voisin, and P. Gouton, IEEE Multimedia **14,** 40 (2007).
8. T. Mitsunaga and S. K. Nayar, in *Proceedings of IEEE Conference on CVPR* 1374 (1999).
9. A. Raij, G. Gill, A. Majumder, H. Towles, and H. Fuchs, in *Proceesings of the IEEE International Workshop on Projector-Camera Display Systems* 100 (2003).
10. A. Majumder and R. Stevens, in *Proceedings of the ACM Symposium on Virtual Reality Software and Technology* 147 (2002).
11. W. Zou, H. Xu, B. Han, and D. Park, Chin. Opt. Lett. **6,** 499 (2008).
12. R. Juang and A. Majumder, in *Proceesings of the IEEE International Workshop on Projector-Camera Display Systems* 17 (2007).
13. M. Shi and G. Healey, J. Opt. Soc. Am. A **19,** 645 (2002).
14. Y. Kwak and L. MacDonald, Displays **21,** 179 (2000).
15. B. Bastani, B. Funt, and R. Ghaffari, in *Proceedings of the 10th AIC Color* 209 (2005).
16. N. Tamura, N. Tsumura, and Y. Miyake, Journal of the SID **11,** (2) 1 (2003).