

## Research Article

# ipProjector: Designs and Techniques for Geometry-Based Interactive Applications Using a Portable Projector

**Thitirat Siriborvornratanakul and Masanori Sugimoto**

*Interaction Technology Laboratory, Department of Electrical Engineering and Information Systems, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan*

Correspondence should be addressed to Thitirat Siriborvornratanakul, thitirat@itl.t.u-tokyo.ac.jp

Received 1 September 2009; Revised 23 November 2009; Accepted 14 December 2009

Academic Editor: Jungong Han

Copyright © 2010 T. Siriborvornratanakul and M. Sugimoto. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We propose an interactive projection system for a virtual studio setup using a single self-contained and portable projection device. The system is named ipProjector, which stands for Interactive Portable Projector. Projection allows special effects of a virtual studio to be seen by live audiences in real time. The portable device supports 360-degree shooting and projecting angles and is easy to be integrated with an existing studio setup. We focus on two fundamental requirements of the system and their implementations. First, nonintrusive projection is performed to ensure that the special effect projections and the environment analysis (for locating the target actors or objects) can be performed simultaneously in real time. Our approach uses Digital Light Processing technology, color wheel analysis, and nearest-neighbor search algorithm. Second, a paired projector-camera system is geometrically calibrated with two alternative setups. The first uses a motion sensor for real-time geometric calibration, and the second uses a beam splitter for scene-independent geometric calibration. Based on a small-scale laboratory setting, experiments were conducted to evaluate the geometric accuracy of the proposed approaches, and an application was built to demonstrate the proposed ipProjector concept. Techniques of special effect rendering are not concerned in this paper.

## 1. Introduction

Recently, virtual studio setups have become popular for modern studio productions. Techniques such as studio camera tracking and 3D graphic rendering have been integrated into a conventional studio setup, so professional special effects can be created at lower cost. The main problem with current virtual studio setups is that these special effects are invisible during film recording. Therefore, the special effects cannot be shown to live audiences in live broadcasting. In addition to actors and moderators, it is difficult to respond correctly to invisible content.

One solution based on existing technologies uses a projector. By combining a projector and a camera, any visible or invisible special effect can be directly superimposed onto a studio surface. The visible special effects support live broadcasting, provide realism during virtual studio recording, and allow direct interaction with actors. The invisible special effect, which can only be seen by a specific

camera, works very well as a supporting system for virtual studio production. Moderators can embed hidden scripts for verification. In an advanced virtual studio system like that proposed in [1], the hidden information is used to accurately track studio cameras and to render real-time 3D graphics in an arbitrary studio environment.

Using a projector and a camera for live broadcasting with special effects is one type of real-time interactive projection application. By using a projector-camera paired system (a *pro-cam* system), we propose a system that includes hardware designs and software implementations. The system is self-contained, portable, and performs real-time projection and real-time environment analysis simultaneously.

The self-contained portable device allows easy setups in any arbitrary studio environment. Furthermore, the shooting and projecting angles of the camera and projector can be determined freely based on the director's consideration. This is important in actual film shooting where the highest priority is to get the best images from the desirable

angle shot. The simultaneous real-time projection and real-time environment analysis allows advanced virtual studio techniques to be performed while special effects are being continuously projected. Environment analysis in this context refers to the analysis of three targets: the studio scenery and objects, and inside actors. For example, captured images are analyzed to locate an inside actor and his postures so that special effects can be created and projected in response to his actions. Note that the concept proposed in this paper could also be applied to other portable interactive projection systems.

This paper focuses on two problems that are important in implementing real-time interactive projection applications. The first problem is projecting real-time special effects so that they do not interfere with normal camera capturing. The second problem is calibrating a pro-cam system precisely to ensure correct transformations between a projector and a camera in a portable setup, which implies use in an unknown environment.

The first problem is also called the nonintrusive projection problem. In normal projection, projected images can be captured by a camera; discriminating between projected contents and real contents in captured images is still difficult. Nonintrusive projection guarantees that projected special effects will be visible to humans and normal studio cameras but invisible to the calibrated camera. This is necessary to avoid projective interference (as seen by the camera) that may lead to an incorrect environment analysis. The technique proposed in this paper involves Digital Light Processing (DLP) projector, RGB color space sampling and nearest-neighbor search algorithm.

With respect to the second problem, there are two related calibrations that are often concerned by researchers using a pro-cam system: geometric and radiometric calibrations. Both calibrations require completely different calibration approaches. In this paper, we focus on the first calibration to achieve a geometrically calibrated pro-cam system. With the calibrated system, we are able to project special effects onto desired locations. A motion-sensor-based calibration technique and a beam-splitter-based calibration technique are investigated in the following sections.

In this paper, we show two complete designs of the self-contained portable projection device. Based on the developed designs, solutions of the nonintrusive projection problem and the geometric calibration problem are proposed. Experiments were conducted to evaluate the accuracy of the proposed approaches in a small-scale laboratory setting; a white board and color magnets were used to represent the studio projection surface and the studio target objects, respectively. Techniques of special effect rendering are not concerned in this paper; therefore, the projected special effects are the simple “+” pattern. The proposed devices and approaches can be applied in a full-scale virtual studio environment by implementing additional target detection algorithms for the desired studio target objects or actors.

The rest of this paper is laid out as follows. Section 2 explains recent advances in interactive projector systems and then discusses related research in the two problem areas as mentioned earlier. Section 3 shows the techniques using an

off-the-shelf DLP projector for nonintrusive projection; the initial setup for pro-cam synchronization, necessary camera settings and detailed analysis results regarding our DLP projector model are written in the section. Section 4 presents two alternative system configurations (in combination with the initial setup shown in Section 3) for real-time pro-cam geometric calibration based on perspective transformation model. On one hand, Section 4.1 applies an additional motion sensor for calibration on a planar or slanted surface. On the other hand, Section 4.2 uses a beam splitter and introduces a complete portable design consisting of a projector, camera and beam splitter; the precise geometric calibration is achieved on both planar and nonplanar surface and suitability of the design regarding the nonintrusive projection is observed in the section. In Section 5, performances of the proposed nonintrusive projection and geometric calibration approaches are experimentally confirmed and an application is built to demonstrate the overall concept of the ipProjector. Finally, Section 6 concludes this article along with a plan for future works.

## 2. Related Work

*2.1. Interactive Projector Systems.* Cotting and Gross [2] introduced an environment-aware display system that automatically avoids projections onto nonsurface objects. Their system performs real-time interactions, but is limited to fixed projectors and fixed cameras mounted on a ceiling. In addition, surfaces are restricted to flat table surfaces whose distance to the ceiling is unchanged. In Cao et al. [3, 4], interactive mobile projector systems were developed. While their projection device is self-contained, it requires a camera mounted separately in a workspace for 3D positioning purpose. In CoGAME [5], images projected from a handheld projector control the movement of a robot. A camera with an IR filter sees only three IR LEDs attached to the robot, and so other visual information about the environment is disregarded completely. The latest SixthSense prototype [6] is a mobile pro-cam device that offers meaningful interactions with different objects found in the real world. However, the projector and camera are not calibrated in their system and environment analysis is performed even though there is projective interference. Consequently, geometric accuracy is limited and color markers are used to help locate target objects like fingertips and desired projection areas. Unlike these systems, our proposed interactive system is truly self-contained and geometrically calibrated. It can perform real-time projection and real-time environment analysis simultaneously without any projective interference.

*2.2. Nonintrusive Projection.* This topic is a subset of the embedded imperceptible-pattern projection problem. Prototypes of an infrared projector were proposed in [7, 8] to project infrared and visible light simultaneously. An infrared pattern is fixed by using an internal mask inside a projector in [7], but is variable in [8]. Unfortunately, the work of Lee et al. [8] requires many internal changes inside a DLP projector that can be accomplished only by

a commercial manufacturer. While an infrared projector is under investigation, there are existing solutions proposed for this problem. For the office of the future [9], structured light can be embedded into a DLP projector by making significant changes to the projection hardware. However, this implementation is impossible unless it is incorporated into the design of the projector or full access to the projection hardware is available. In [1, 10, 11], a code image is projected at high speed with its neutralized image, which integrates the coded patterns invisibly due to limitations of the human visual system. According to these papers, projecting and capturing at 120 Hz can guarantee a hidden code. Commonly available projectors usually perform projections at a maximum rate of 87 Hz.

For this paper, we apply an approach based on the DLP characteristics. Using the camera classification approach proposed in [12] and the nearest-neighbor search algorithm, we are able to perform nonintrusive projection using an off-the-shelf DLP projector.

*2.3. Real-Time Pro-Cam Geometric Calibration.* When a projector and a camera are rigidly fixed to each other, some have assumed that the geometric registration between them is roughly constant [13]. However, as the angle of the projector moves from the perpendicular or as a surface becomes nonplanar, this approach will no longer guarantee good geometric registration. Projecting a known pattern onto a surface is a classical approach to solve this problem that gives precise calibrations for both planar surfaces [14–16] and irregular surfaces [17–19]. However, the computational cost is high for complex surfaces, and patterns must be re-projected when a component of the system (e.g., a projector, camera or surface) moves. Similar approach is applied in the catadioptric projectors [20] whose projected light is refracted/reflected by refractors/reflectors; geometric registration between the two devices is obtained by projecting a series of known patterns and allowing the camera to sense them. A real-time approach that does not interrupt normal projection was proposed in [21] by attaching four laser-pens to a pro-cam system. Although detecting bright laser points sounds easier than detecting points projected by a projector, locating small laser points in a messy camera image is still difficult. In [22], Johnson and Fuchs proposed a real-time approach that does not interrupt normal projection, requires no fixed marker and can be applied to a complex surface. By matching feature points found in the projected image and the predicted captured image, the pose of the projector is tracked and the calibration is achieved in real time. However, the camera is stationary and separated from the projector in their system.

This paper involves two calibration approaches that can be implemented as a single self-contained device. The first approach uses one additional motion sensor, and geometric calibration is achieved in real time on a planar or slanted surface. The second approach uses a beam splitter to coaxialize the projector and camera. Geometric calibration is independent of the scene, so both planar and irregular surfaces can be used as projection surfaces.

### 3. Nonintrusive Projection

As mentioned in the introduction, it is important that a real-time environment analysis has no interference from any projected contents. If it does, the system might consider the projected content as a real target and generate special effects in response to that false detection.

Recently, internal characteristics of a DLP projector have received lot of attentions from research communities. On one hand, in [23, 24], the dithered illumination pattern corresponding to the DMD chip (which operates at 10 000 Hz) is observed and utilized using a very high speed camera (whose maximum speed is 3000 fps). On the other hand, characteristics of the color wheels (which rotate at 120 Hz) can be investigated and used by a camera with slower capturing speed. Our nonintrusive projection is based on the latter one; characteristics of the color wheels are utilized here for nonintrusive projection purpose. The proposed approach has three main advantages: it requires no internal change to the projector or the camera, it can be applied to any off-the-shelf DLP projector, and it supports embedded variable light patterns in the future without further hardware modifications.

In the following sections, we explain in detail how to analyze the characteristics of the color wheels inside the DLP projector and how to use these characteristics for nonintrusive projection. Note that a beam splitter (as described in Section 4.2) has not yet been applied in these sections.

*3.1. DLP Projector Analysis.* Because each DLP projector model owns unique characteristics of the inside color wheels, DLP projector analysis has to be performed before using an unknown DLP projector model for the proposed nonintrusive projection approach. To understand the overall characteristics of the color wheels without full access to a DLP chip and its controller, we applied the camera-based classification method proposed in [12]. In this section, we briefly explain the classification steps and show the classification results of our DLP projector.

First, this classification method requires a camera with an external trigger feature to synchronize it with a DLP projector. Synchronization between the projector and the camera is performed by tapping the vertical sync signal (5 V, 60 Hz) from the computer to the projector. By using the tapped signal as a trigger, our camera remains synchronized to the projector. In addition, the shutter of the camera must be set to open for a very short period in order to sense the fast characteristics of the color wheels. In our setup, the camera is set to expose for only 0.55 ms.

The following devices were used for our pro-cam synchronization: a HP MP2225 DLP projector (XGA 1024 × 768 projection resolution) with a D-sub 15 pin connector, a Dragonfly Express camera (VGA 640 × 480 captured resolution) connected through a FireWire 800 (IEEE1394B) port, and an ELECOM VSP-A2 VGA splitter. The camera is equipped with a Tamron 13VM308AS lens. The synchronization setup is illustrated in Figure 1.

Second, we analyzed the overall sequences of the color wheels inside our DLP projector by projecting single-color

images (corresponding to the colors of each available color wheel of the projector) at maximum intensity through all possible starting exposure times. Figure 2 was created by allowing the synchronized camera to sense these projected colors with different starting exposure times.

Third, we analyzed detailed mirror flip sequences for all 256 values in the selected color channel (i.e., red, green or blue) within a narrow starting exposure period. From Figure 2, the selected color channel is the red channel, which is the first channel appearing in the sequences. Mirror flip sequences were then obtained by projecting uniform red images with intensity values ranging from 0 to 255, and with the starting exposure times ranging from 0 to 2 ms. Figure 3 was created by allowing the camera to sense these red projections. A starting exposure time of 1.4 ms was finally chosen because it provides the best distributed red ramps, as shown in Figure 3.

Following the explained steps, we are able to synchronize the camera with the projector at the appropriate starting exposure time. For our selected starting exposure time, the camera can only see the red light of the DLP projector (correspondences between projected red intensities and red intensities seen by the camera are shown in Figure 3). If red intensities of the projected special effects are similar to those of the projected background color (as seen by the camera), the system cannot differentiate the projected special effects from the background in captured images. Hence, further environment analysis is not interfered by the real-time special effect projection.

**3.2. Environment Illumination.** To perform the DLP analysis and use it for nonintrusive projection (as mentioned in Section 3.1), the shutter of the camera is set to expose for only 0.55 ms, which is too short for the camera to sense the environment properly (as shown in Figure 4(b)) unless there is light emitting from the projector (as shown in Figure 4(c)). In [12], the 256 red intensities in the selected timeslot were classified into three sets: *white*, *black* and *grey*. *White* refers to colors whose projection fully turns mirrors inside the DLP projector and transmits lights toward a surface. *Black* refers to colors whose projection does not flip the mirrors and transmits no light toward a surface. *Grey* refers to unreliable states between *white* and *black*.

For environment analysis purpose, we need to illuminate the environment while projecting nonintrusive special effects. Thus, only colors whose red value contained in the *white* set should be projected. Figure 4(c) depicts an image seen by the camera when Figure 4(d), whose red intensity is maximal for all pixels, was projected. Note that the intensity of Figure 4(b) was enhanced here to allow the environment to be seen.

**3.3. Color Space Sampling and Color Conversion.** The method of projecting only colors whose red intensity contained in the *white* set (as mentioned in Section 3.2) works only for a DLP projector model that does not have interdependent color channels in mirror flipping. Instead of depending on the red channel and risking effects from the other color channels,

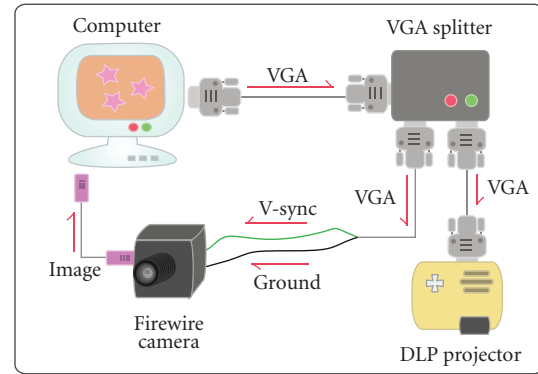


FIGURE 1: Setup for projector-camera synchronization.

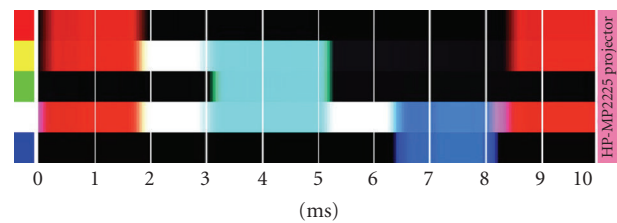


FIGURE 2: Color-wheel sequences of the HP MP2225 as seen by the synchronized camera with starting exposure times ranging from 0 to 10 ms.

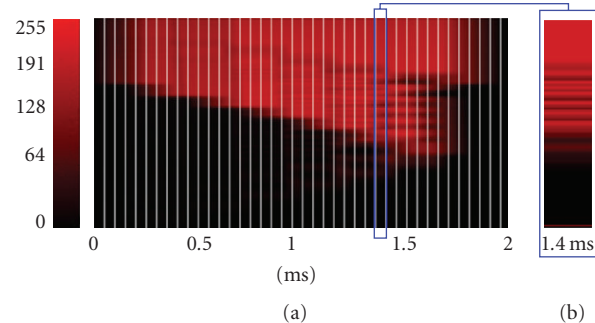


FIGURE 3: (a) Detailed mirror flip sequences for all 256 values in the red channel with the starting exposure times ranging from 0 to 2 ms. (b) Mirror flip sequences at the selected starting exposure time.

we propose a DLP-model-independent classification that involves all color channels (i.e., red, green and blue) of projected images. Each color in the RGB color space was projected onto the white surface and sensed by the camera. Using the camera classification method, these projected colors were classified into the three sets previously discussed. Based on this proposed classification, we are also able to determine whether our DLP projector has interdependent color channels in mirror flipping.

From the entire RGB color space containing  $256 \times 256 \times 256$  colors, we selected only  $11 \times 11 \times 11$  colors and obtained 115 colors belonging to the *white* set. Because the amount of light passing through the camera is decreased by

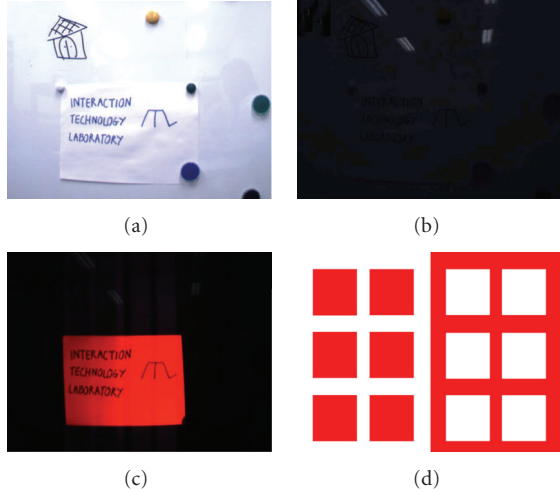


FIGURE 4: (a) is an environment seen by a camera with normal shutter settings. (b) is the environment (a) as seen by the synchronized camera. (c) is the environment (a) seen by the synchronized camera when image (d) is being projected from the DLP projector.

the beam splitter setup (as detailed in Section 4.2), we set a high threshold value for the *white* set. Thus, there are few colors categorized into this set. In detail, we projected single-color images (corresponding to the colors in the selected color space) and allowed the camera to sense the projected colors. For one projected color, values of all pixels inside the projection area (as seen by the camera) are averaged before that color is classified into the appropriate set. Because of this, we can ensure that the different resolutions between the projector and camera will not affect our processes of color space sampling and color classification.

For converting an arbitrary color into the most similar color in the *white* set, we used an approximate nearest-neighbor search algorithm called Best Bin First (BBF) [25]. Using this, we can convert an arbitrary color into the *white* set, and the working environment is then always illuminated. In Figure 5, we randomly generated 400 colors and converted them into the *white* set using the BBF algorithm. The converted image was then projected and sensed by the synchronized camera. Conversion of the same image to the *black* set is also shown for comparison. From images captured by the camera shown in Figure 5, it is clear that our classification and conversion approach can perform efficiently, and the surface remains well illuminated when projecting colors in the *white* set.

#### 4. Real-Time Pro-Cam Geometric Calibration

Geometric calibration is the first problem encountered by most pro-cam systems. Geometric mapping between camera and projector coordinates is necessary to find corresponding positions between the two coordinate systems, and to project images back to desired locations on an actual surface. The three objects affecting this calibration are the projector, the camera and the surface. Any relative movement between

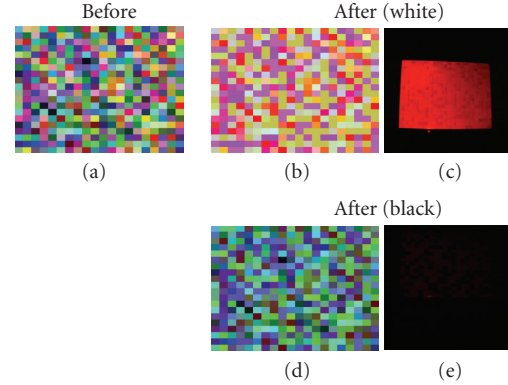


FIGURE 5: Color conversion using BBF algorithm. (a) is the randomly generated 400 colors before conversion. (b) and (d) are the (a) image after converted to the *white* and *black* set, respectively. (c) and (e) are the blank surface seen by the synchronized camera when image (b) and (d) are being projected, respectively.

any pair of the three objects causes changes in the pro-cam geometric mapping. By attaching the camera rigidly to the projector, there is no relative movement between the projector and the camera but this cannot prevent relative movements with the surface. On a planar surface, when the angle of a projector moves away from perpendicular, the geometric mapping changes. Nonoverlapping fields of view of the projector and the camera may make geometric mapping impossible in some positions. On an irregular surface, there is an additional serious problem as 3D shapes of the surface create parallax effects between projector and camera coordinates. This problem is difficult to recover from unless the 3D geometries of the surface are known.

In this paper, geometric mapping between camera and projector coordinates is computed by perspective transformation [26] whose computation effort is lighter than Euclidean calculation does. Based on the fact that all points seen by the camera lay on some unknown plane, the perspective transformation between the two coordinates can be established by a  $3 \times 3$  homography matrix. Suppose that  $(X, Y)$  is a pixel in projector coordinates whose corresponding pixel in camera coordinates is  $(x, y)$ , the perspective transformation from  $(x, y)$  to  $(X, Y)$  can be expressed with eight degrees of freedom in

$$(X, Y) = \begin{pmatrix} h_1x + h_2y + h_3 & h_4x + h_5y + h_6 \\ h_7x + h_8y + h_9 & h_7x + h_8y + h_9 \end{pmatrix}, \quad (1)$$

where  $\vec{h} = (h_1 \cdot \cdot \cdot h_9)^T$  is constrained by condition  $|\vec{h}| = 1$ . The same transformation as written in (1) can be expressed in homogeneous coordinates as

$$\begin{pmatrix} Xw \\ Yw \\ w \end{pmatrix} = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}. \quad (2)$$

$\vec{h}$  can be computed from four corresponding pixels between the two coordinates (four correspondences ensure that no

three points is collinear). When there are more than four corresponding pixels found between the two coordinates ( $\delta > 4$  in (3)), the RANSAC method is applied for estimating the values of  $\bar{h}$  in the following equation:

$$\begin{pmatrix} X_1 w & X_2 w & \cdots & X_\delta w \\ Y_1 w & Y_2 w & \cdots & Y_\delta w \\ w & w & \cdots & w \end{pmatrix} = \begin{pmatrix} h_1 & h_2 & h_3 \\ h_4 & h_5 & h_6 \\ h_7 & h_8 & h_9 \end{pmatrix} \cdot \begin{pmatrix} x_1 & x_2 & \cdots & x_\delta \\ y_1 & y_2 & \cdots & y_\delta \\ 1 & 1 & \cdots & 1 \end{pmatrix}. \quad (3)$$

The purpose of the following sections (i.e., Sections 4.1 and 4.2) is to find at least four corresponding pixels between camera and projector coordinates and use them to compute the updated  $\bar{h}$  values in real time. After that, geometric mapping from any  $(x, y)$  to  $(X, Y)$  and vice versa is achieved using  $\bar{h}$  and  $[\bar{h}]^{-1}$ , respectively. Sections 4.1 and 4.2 explain two alternative approaches for finding a set of corresponding pixels between the two coordinates based on two different setups. Both approaches are implemented as a single self-contained device and support portable use. The first approach uses an additional motion sensor and can perform real-time geometric calibration on a planar or slanted surface. The second approach uses a beam splitter so that the geometric mapping is independent of the surface.

By the way, considering the frame buffer architecture of the projector, there is one frame delay time before an image sent to the projector will be projected out. Any movement regarding the projector, camera or surface during this delay time causes geometric errors between the calculated geometries and the actual geometries (of projected images) appearing on the surface. In our system, frequency of the projection cycle is 60 Hz (as mentioned in Section 3.1) which equals to the maximum 16.67 ms projection delay time. For the application proposed in Section 5, this delay time is relatively short compared to other processing times. Therefore, we decided to neglect the effect of this delay from our calculation. However, for an interactive system that requires the millisecond precision in projection, additional techniques such as motion estimation should be applied.

**4.1. Motion-Sensor-Based Approach.** This section describes the approach using a motion sensor to find a set of corresponding pixels between camera and projector coordinates on a planar or slanted surface. Two tilt sensors fixed to a projector were first proposed in [16]. Acquiring the tilt angles from both sensors in real time allows the correct estimation of the world's horizontal and vertical directions without using markers. Dao et al. [27] extended the sensor-based idea to an accelerometer combined with a digital compass. Their system measures the inclined angle of a projector directly in both vertical and horizontal axes, and then creates an interactive application by using real-time keystone correction.

A sensor eliminates the need for fiducial markers but still allows a single self-contained device. Our configuration for this calibration approach is shown in Figure 6; a projector, camera and motion sensor are fixed together on a wooden

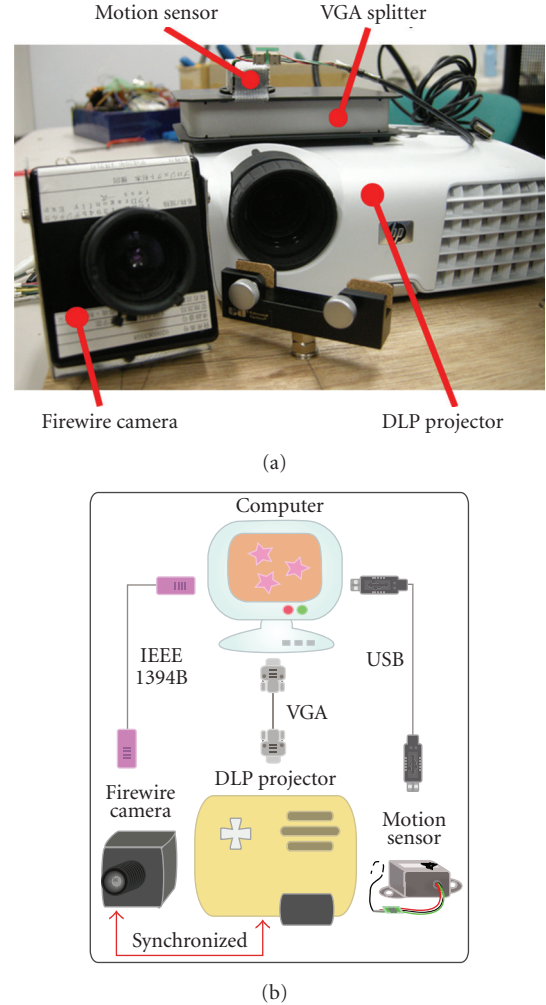


FIGURE 6: Projector-camera-sensor configuration.

base so that their relative positions and orientations cannot be changed. The purpose is to obtain pairs of corresponding  $(X, Y)$  and  $(x, y)$  in real time. Unlike previous researches that use the sensors to compute the rotation matrix of a projector or to correct the keystone distortion, we directly correlate the sensor values with camera coordinates in order to find updated geometric mapping between camera and projector coordinates in real time. Keystone correction is not concerned in our system.

In this paper, a NEC/TOKIN MDP-A3U9S 3D motion sensor is used with a data update rate of 125 Hz. The relative pitch and roll angles are calculated from three acceleration values,  $AccX$ ,  $AccY$  and  $AccZ$  (acceleration values along  $X$ ,  $Y$ , and  $Z$  axis, resp.), read from the accelerometer embedded in our motion sensor. Setting the reference angles is simple—the projector is moved until the images appear rectangular on a surface and then a key or button is pressed. The reference can be reset whenever a user prefers or feels significant geometric errors (between the calculated geometries and the actual geometries appearing on the surface) in the calibration. Five consecutive samples of the 3D acceleration

values acquired from the sensor are averaged in real time before being used to compute the relative pitch and roll angles. Averaging adds a delay but is recommended for a smoother calibration. The calculation of the relative pitch and roll angles can be summarized as follows:

$$\begin{aligned} \text{pitch}_{\text{rel}} &= \arccos\left(\frac{\text{Acc}X_{\text{avg}}}{\sqrt{\text{Acc}X_{\text{avg}}^2 + \text{Acc}Y_{\text{avg}}^2 + \text{Acc}Z_{\text{avg}}^2}}\right) \\ &\quad - \text{pitch}_{\text{ref}}, \\ \text{roll}_{\text{rel}} &= \arctan\left(\frac{\text{Acc}Y_{\text{avg}}}{\text{Acc}Z_{\text{avg}}}\right) - \text{roll}_{\text{ref}}, \end{aligned} \quad (4)$$

where  $\text{pitch}_{\text{ref}}$  and  $\text{roll}_{\text{ref}}$  refer to the reference pitch and roll angles.  $\text{Acc}X_{\text{avg}}$ ,  $\text{Acc}Y_{\text{avg}}$  and  $\text{Acc}Z_{\text{avg}}$  are the average of five consecutive  $\text{Acc}X$ ,  $\text{Acc}Y$  and  $\text{Acc}Z$  values read from the sensor, respectively.

This approach requires offline calibration. However, the calibration data are compatible with the system if there is no change in the relative positions or orientations of the three devices. Suppose that the offline calibration is achieved using  $N$  sample images captured from the camera, and all sample images share a set of  $n$  points to be calibrated. A set of calibration data provided by one sample image can be written as  $(p, r, x_1, y_1, x_2, y_2, \dots, x_n, y_n)$ . Let  $p$  and  $r$  refer to the relative pitch and roll angles, and  $(x_i, y_i)$  represents the 2D camera coordinate of an  $i$ th observed point in the sample image. For  $N$  sample images captured from different angles and orientations, we have

$$\begin{aligned} A &= \begin{bmatrix} p^{(1)} & r^{(1)} & 1 \\ p^{(2)} & r^{(2)} & 1 \\ \vdots & \vdots & \vdots \\ p^{(N)} & r^{(N)} & 1 \end{bmatrix}, \\ B &= \begin{bmatrix} x_{1(1)} & y_{1(1)} & x_{2(1)} & y_{2(1)} & \cdots & x_{n(1)} & y_{n(1)} \\ x_{1(2)} & y_{1(2)} & x_{2(2)} & y_{2(2)} & \cdots & x_{n(2)} & y_{n(2)} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{1(N)} & y_{1(N)} & x_{2(N)} & y_{2(N)} & \cdots & x_{n(N)} & y_{n(N)} \end{bmatrix}. \end{aligned} \quad (5)$$

The offline calibration is finished when an adjustment matrix ( $\beta$ ) is obtained by using linear least squares to solve

$$(A^T A)\beta = A^T B. \quad (6)$$

At any time  $\tau$  during the online calibration,  $(x_i, y_i)$  is updated from the relative pitch and roll angles (calculated as explained earlier) by

$$[x_{1(\tau)}, y_{1(\tau)}, x_{2(\tau)}, y_{2(\tau)}, \dots, x_{n(\tau)}, y_{n(\tau)}] = [p_{(\tau)}, r_{(\tau)}, 1]\beta. \quad (7)$$

Following the above explanation, even though the position and orientation of the pro-cam system (relative to the surface) are not known, the system is able to obtain camera

coordinates  $(x_{i(\tau)}, y_{i(\tau)})$  of the  $n$  pre-defined points (whose projector coordinates  $(X_{i(\tau)}, Y_{i(\tau)})$  are known). Using the corresponding  $(x_{i(\tau)}, y_{i(\tau)})$  and  $(X_{i(\tau)}, Y_{i(\tau)})$  to compute  $\vec{h}$  as shown in (3) (when  $\delta = n$ ), real-time geometric mapping between camera and projector coordinates is achieved.

**4.2. Beam-Splitter-Based Approach.** A beam splitter is an optical device that reflects half of the incoming light and transmits the other half. There are few researches concerning pro-cam systems using a beam splitter. In [7], two cube beam splitters are used to construct an IR projector and a multi-band camera. However, the beam splitters are used for internal hardware architecture purposes not calibration purposes. Fujii et al. [28] briefly described the idea of scene-independent geometric calibration using a plate beam splitter attached to an off-the-shelf projector. The calibration technique proposed in this section was inspired by this work. Both their research and ours operate in the visible light spectrum; however, the camera settings are completely different. In their research, a camera uses a normal shutter speed and works independently to a projector. In our research, as described in Section 3, the camera is accurately synchronized with the DLP projector and its shutter is opened for only 0.55 ms. In this section, we investigate a concrete portable design and suitability of the beam splitter regarding our nonintrusive projection approach. Many related factors are introduced and observed. The beam splitter used in our configuration is a TechSpec plate beam splitter 48904-J, whose dimensions are  $75 \times 75$  mm.

Using a beam splitter to coaxialize the two devices ensures that any surface visible to a camera can also be projected upon. The shapes of the surface do not affect the geometric mapping or cause parallax between the two coordinates. This means that geometric mapping ( $\vec{h}$ ) needs to be computed only once using (3); recomputation is not necessary if there is no change in the relative positions or orientations among the projector, camera and beam splitter. The coaxial concept is illustrated in Figure 7(a) and our beam splitter configuration is shown in Figure 7(d). The distance from the front edge of the wooden base to the projector lens is 13 cm. In addition to the design proposed in [28], we added a curtain made from a light absorbing black-out material to achieve the more practical portable design. This curtain not only eliminates reflections of the environment at the left side of the projector (as shown in Figure 7(c)) but also prevents the projector's reflected light from interfering with the environment (as shown in Figure 7(b)). Note that the camera was not in the high shutter speed mode when capturing Figure 7(c).

As explained in Section 3, the nonintrusive projection shortens the exposure time of the camera significantly. Furthermore, the beam splitter setup allows only half of the projected light to be transmitted to the surface. As a result, the amount of light passing through the camera lens in this configuration is quite limited and may result in inaccurate environment analyses. Therefore, we conducted three experiments using the camera setting explained in Section 3 and investigated factors related to the practicality of

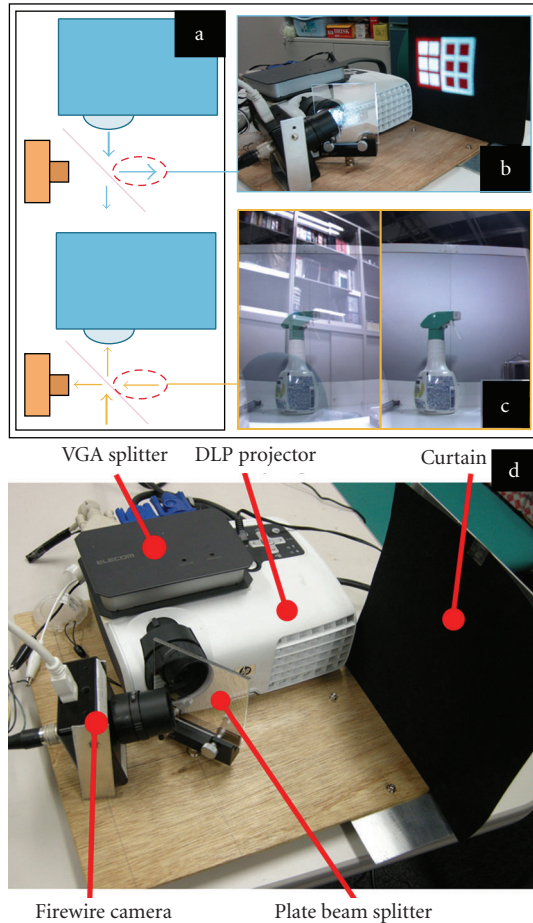


FIGURE 7: (a) The concept of pro-cam coaxialization using a beam splitter. (b) The use of a curtain to block projector's reflected light. (c) Captured image without and with the curtain (left and right). (d) Our beam splitter configuration.

this beam splitter configuration. Only red value is concerned in our experiments because the projector projects the red light at the selected timeslot.

In the first experiment, we investigated how much the beam splitter reduces the amount of light seen by the camera. The three experimental setups are (1) no beam splitter, and camera directly sees a surface, (2) a beam splitter in front of the projector lens, but the camera still sees a surface directly, and (3) our configuration as shown in Figure 7(d). With a distance of 50 cm from the wooden base to the surface, we projected uniform red images with intensity values ranging from 0 to 255 onto a whiteboard, and allowed the camera to sense these red projections. Figure 8 shows the red intensities seen by the camera in the three setups. Comparing the first and second setup, the red intensities seen by the camera are reduced by 53.78% when the beam splitter is placed in front of the projector lens. Comparing the second and third setup, the red intensities are reduced by 38.39% when the camera sees the surface through the beam splitter. In total, by comparing the first and third setup, our configuration reduces the red intensities seen

by the camera to about 71.52% of its original intensities compared with the conventional pro-cam setup. This means that amount of light seen by the camera will be quite limited with this configuration. Therefore, a projector model whose brightness is not strong enough might be difficult to be used in our proposed system.

We investigated the distance in the second experiment. Because the exposure time of the camera is very short, the farther away the surface is, the less light the camera will see. If an environment is insufficiently illuminated, it will be difficult to use any image processing techniques. In this experiment, we constantly projected a red image (whose red intensity value was 255) onto a whiteboard located at different distances. Figure 9 shows the experimental results. At a distance of 20 cm, the environment was illuminated with a very bright red light. At a distance of 130 cm, the environment was too dark for the camera to see properly. Note that the brightness of our DLP projector is 1400 ANSI lumens and the distance written in this context refers to a distance from the surface to the front edge of the wooden base.

Finally, we performed an experiment to measure the maximum distance over which our configuration can perform environment analyses correctly. We projected a pure red image (with red intensity values equal to 255) at different distances from 20 to 130 cm. The surface was a whiteboard containing five color magnets inside the projection area. We applied 2D Gabor filters [29] to each captured image to evaluate the accuracy of the object detection at each distance. The Gabor filters failed to detect all objects at a distance of 130 cm. Because the projected color used here is the brightest color possible to be seen by the camera, we conclude that the maximum distance at which this beam splitter configuration can be used in the nonintrusive projection mode is 120 cm.

## 5. Experimental Results and Evaluations

In this section, we discuss the experiments conducted to evaluate the proposed approaches. All experiments were performed using a Dell Inspiron 1150 Mobile Intel Pentium 4 laptop with a processor running at 2.80 GHz. The projector's focus was adjusted manually in all experiments.

For the nonintrusive projection, we evaluated whether our sampled colors (from the *white* set) can illuminate an environment enough for an environment analysis. At a specific distance, we projected each color from the *white* set onto the whiteboard holding five color magnets, and allowed the camera to sense the whiteboard. After applying 2D Gabor filters to the captured images, the projected colors causing incorrect detection were counted. Table 1 shows the experimental results. Most of the misdetection was caused by noises added in the dark environment. Overall, five magnets were detected satisfactorily until the distance reached 110 cm.

To determine the accuracy of the pro-cam geometric calibration on planar surfaces, we measured the geometric error of the three approaches compared with a ground truth. Apart from the two proposed approaches (see Section 4), we added the single calibrated approach for comparison purposes; this approach lets the camera sense a surface



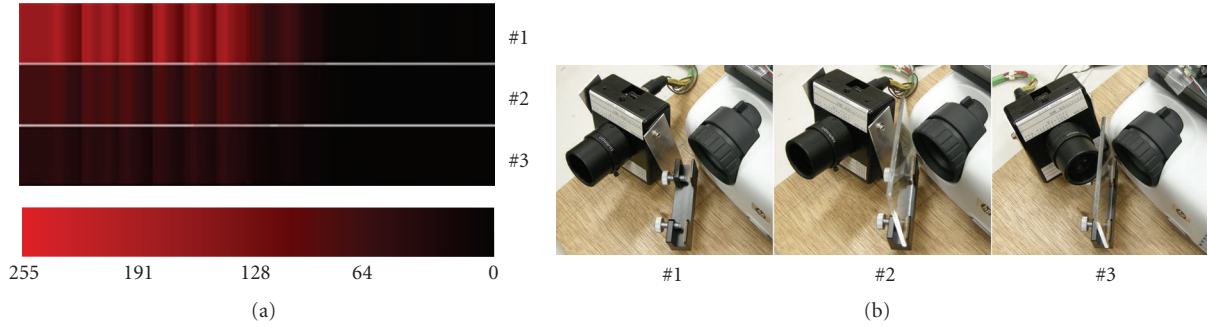


FIGURE 8: Red intensities seen by the camera in three different setups.

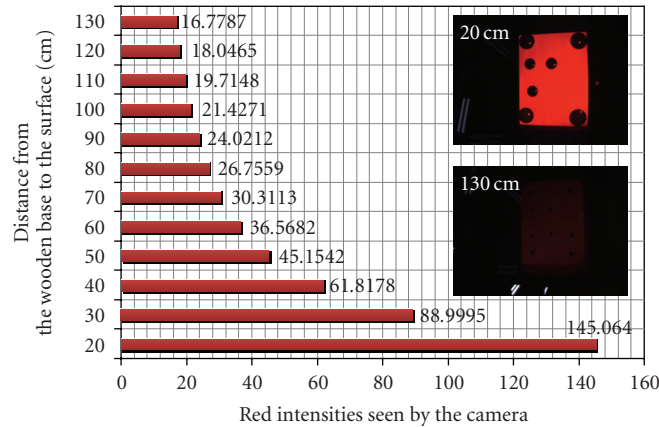


FIGURE 9: Red intensities seen by the camera at different distances to the surface using the beam splitter setup.

TABLE 1: Accuracy evaluation of the nonintrusive projection regarding the distance to the surface.

Distance (cm)	Number of misdetected colors	Distance (cm)	Number of misdetected colors
20	0	80	19
30	0	90	29
40	1	100	39
50	1	110	48
60	4	120	80
70	19	130	111

Note: There are 115 colors in the white set.

directly and assumes a static geometric mapping between projector and camera coordinates. The motion sensor offline calibration was performed using 10 sample images containing 16 calibrated points ( $N = 10$  and  $n = 16$  according to Section 4.1). In the experiments, camera coordinates generated by the three approaches were compared with actual camera coordinates. Five experiments were conducted with different orientations of the projector for each approach and each experiment was performed using 16 tested points (the number of tested point written here is not the  $n$  value used in the motion sensor offline calibration). Except for the beam splitter approach, whose device setup is different, the

experiments were conducted simultaneously. Note that the resolution of the camera coordinates was  $640 \times 480$  pixels.

According to the experimental results shown in Figure 10, the beam-splitter-based approach is the most accurate and provides the narrowest range of geometric errors in both axes. For the other two approaches, error values derived from the same experiment cluster together. The clustering locations in the five experiments are similar in both approaches because they were generated from the same projector orientations. However, the distribution of error values in the single calibrated approach is wider than that for the motion-sensor-based approach. Besides, the accuracy of the two proposed calibration approaches does not fall over time. The accuracy of the motion-sensor-based approach may decrease when the current projector orientation differs significantly from the reference orientation.

In addition to the beam-splitter-based calibration approach, we conducted the same experiment with 16 tested points on five nonplanar surfaces that cannot be calibrated using the other two calibration approaches. Figure 11 shows the experimental results and the experimental surfaces captured by a separate camera. Using the beam splitter setup, geometric errors are small even for these difficult surfaces.

Comparing the motion sensor and the beam splitter setups, the latter provides more precise geometric calibration and allows calibration on nonplanar surfaces. This is good for a portable system where the geometry of a surface is

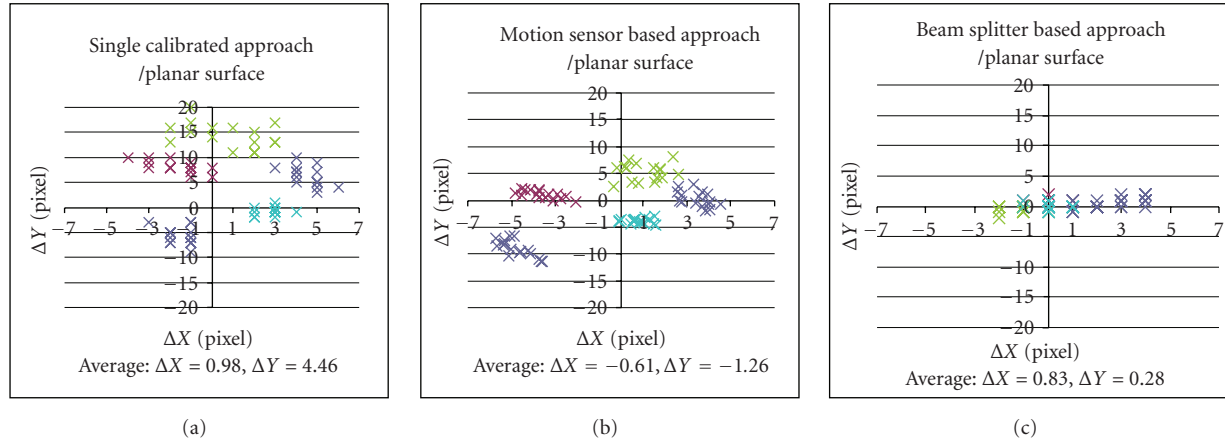


FIGURE 10: Geometric error (camera coordinates) of the single calibrated approach, the motion-sensor-based approach and the beam-splitter-based approach on planar surfaces.

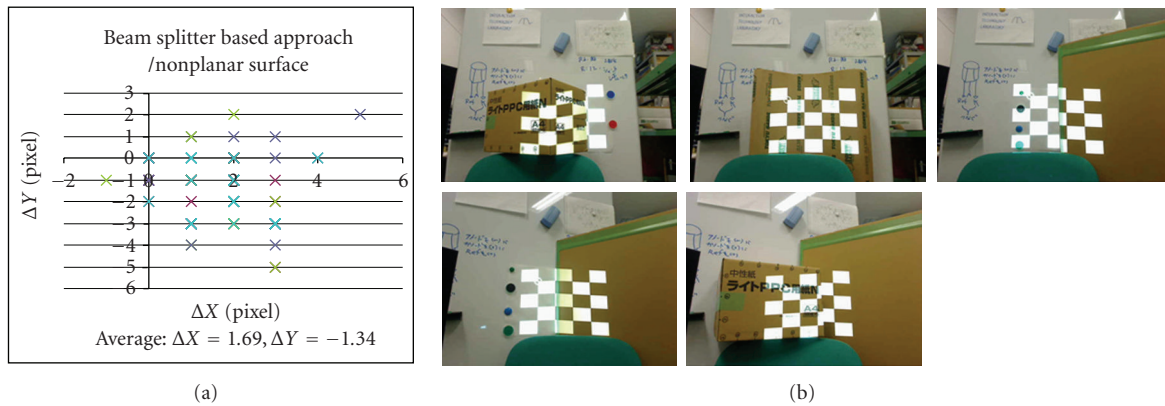


FIGURE 11: Geometric error (camera coordinates) of the beam-splitter-based approach on nonplanar surfaces.

not known. However, considering the quality of captured images, direct capturing can sense the environment more effectively than indirect capturing. Therefore, by using the motion sensor setup, image processing can be performed more efficiently.

We measured the geometric error of the perspective transformation by reprojecting known points onto an actual surface. Using a planar surface, three experiments were conducted using the geometric mapping created from 4, 9, and 16 correspondences, respectively. During each experiment, a chessboard pattern with 25 inner corners was projected. We located each inner corner in camera coordinates, applied the perspective transformation to map those camera coordinates to projector coordinates using (3), and drew the transformed projector coordinates back on the projected images. In this way, we can measure the geometric error between the actual projector coordinates and the transformed projector coordinates of each tested point on the actual surface. Figure 12 shows the experimental results. While experimenting, we moved neither the surface

nor the devices, and the projection area appearing on the surface had the dimensions  $294 \times 222$  mm.

Finally, we built a small-scale application using the beam splitter setup calibrated with 16 correspondences. The purpose was to demonstrate the entire concept of our interactive portable projector. Any objects are detected using 2D Gabor filters; nonintrusive projection is used to draw text or special effects upon the detected objects or other areas of the surface. Large-scale programs using the same procedure can be built for live broadcasting of virtual studio productions. Instead of the Gabor detector, human detector, gesture recognition or motion analysis algorithms might be used to detect an actor and interpret his actions accurately in real time; robustness of the system regarding moving targets will depend on the selected detection algorithm. The rendering technique might be used to virtually paint the studio, and animation techniques could be utilized to draw attractive characters interacting with the detected actor.

Figure 13 shows the application in action, including snapshots of the surface (taken with a separate camera)

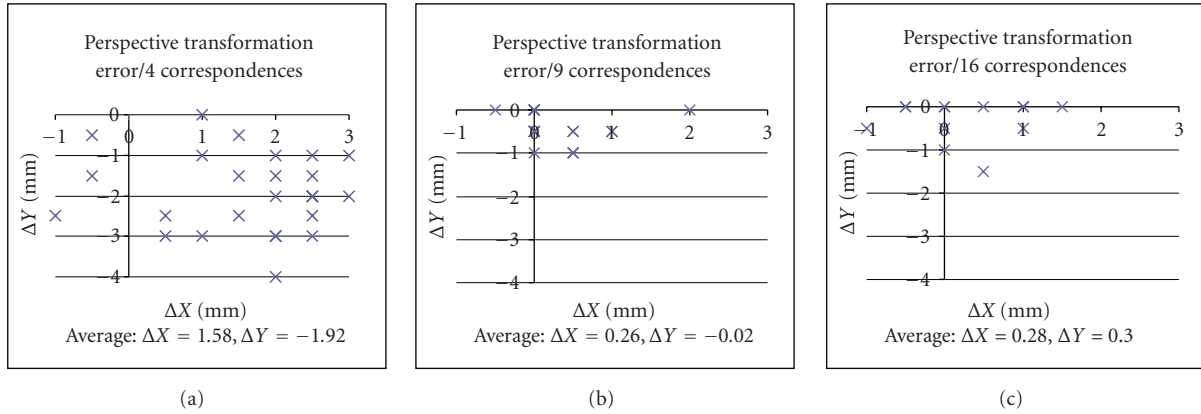


FIGURE 12: Geometric error of perspective transformation on an actual surface (world coordinates) using the number of calibrated correspondences of 4, 9, and 16.

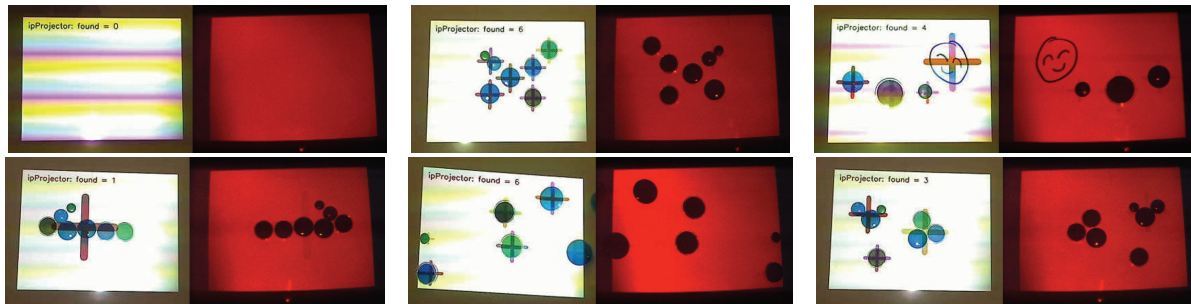


FIGURE 13: Demonstration of the ipProjector concept. Left images are snapshots of the surface taken by a separated camera. Right images are images captured by the system camera.

and images captured by the system camera. The number of detected objects is indicated in red while the other detected objects are marked with a “+” sign at the detected centroid. The “+” signs are drawn in different colors but our separate camera cannot sense them properly. The reader is encouraged to look closely at the images on the right. It can be seen that the system camera sensed traces of the projected contents in some images. This is due to the range of the threshold set during the color classification (Section 3.3). However, these traces are not clear enough to intrude on any environment analysis. Note that the captured images were enhanced here for better visualization.

## 6. Conclusion and Future Work

In this paper, we investigated hardware setups and software techniques for creating a real-time interactive projection application, including live broadcasting from an advanced virtual studio. The projection device developed in this paper is self-contained and portable, and can be installed or used easily in an existing studio environment. The nonintrusive projection problem is solved by using an off-the-shelf DLP projector together with DLP color wheel analysis, color space sampling and approximate nearest-neighbor search.

The proposed solution ensures that the environment is always illuminated, while projected content does not intrude on any environment analysis. For the real-time geometric calibration problem, two approaches with different setups were proposed. On a planar surface, the motion-sensor-based approach can update the geometric mapping in real time. For a more accurate calibration approach that can be applied to planar and nonplanar surfaces, we proposed the beam-splitter-based setup. Related factors of this setup regarding the nonintrusive projection and further image processing were also investigated.

The device and techniques proposed in this paper will help the creation of special effects that appear in response to actors or other studio objects in real time. By integrating the proposed concepts with appropriate target detection algorithms and special effect rendering techniques, an existing virtual studio setup will then support live broadcasting in an actual studio production. For example, in the weather forecast virtual studio, positions of the actor and his hands should be detected so that the weather map can be rendered and projected in response to those detected positions. In the future, we plan to create a robust interactive projection application that benefits from the proposed device and techniques. Furthermore, we are interested in adding human

detection or hand gesture recognition to the system to incorporate the system with a real human.

## References

- [1] A. Grundhöfer, M. Seeger, F. Häntsch, and O. Bimber, "Dynamic adaptation of projected imperceptible codes," in *Proceedings of the 6th IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR '07)*, 2007.
- [2] D. Cotting and M. Gross, "Interactive environment-aware display bubbles," in *Proceedings of the Annual ACM Symposium on User Interface Software and Technology (UIST '06)*, pp. 245–254, 2006.
- [3] X. Cao and R. Balakrishnan, "Interacting with dynamically defined information spaces using a handheld projector and a pen," in *Proceedings of the Annual ACM Symposium on User Interface Software and Technology (UIST '06)*, pp. 225–234, 2006.
- [4] X. Cao, C. Forlines, and R. Balakrishnan, "Multi-user interaction using handheld projectors," in *Proceedings of the Annual ACM Symposium on User Interface Software and Technology (UIST '07)*, pp. 43–52, 2007.
- [5] K. Hosoi, V. N. Dao, and M. Sugimoto, "CoGAME: manipulation by projection," in *Proceedings of the International Conference on Computer Graphics and Interactive Techniques, emerging technologies (SIGGRAPH '07)*, 2007.
- [6] P. Mistry, P. Maes, and L. Chang, "WUW—Wear ur world—a wearable gestural interface," in *Proceedings of the Conference on Human Factors in Computing Systems (CHI '09)*, pp. 4111–4116, 2009.
- [7] K. Akasaka, R. Sagawa, and Y. Yagi, "A sensor for simultaneously capturing texture and shape by projecting structured infrared light," in *Proceedings of the 6th International Conference on 3-D Digital Imaging and Modeling (3DIM '07)*, pp. 375–381, 2007.
- [8] J. C. Lee, S. Hudson, and P. Dietz, "Hybrid infrared and visible light projection for location tracking," in *Proceedings of the Annual ACM Symposium on User Interface Software and Technology (UIST '07)*, pp. 57–60, 2007.
- [9] R. Raskar, G. Welch, M. Cutts, A. Lake, L. Stesin, and H. Fuchs, "The office of the future: a unified approach to image-based modeling and spatially immersive displays," in *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '98)*, 1998.
- [10] H. Park, M. H. Lee, B. K. Seo, Y. Jin, and J. I. Park, "Content adaptive embedding of complementary patterns for non-intrusive direct-projected augmented reality," in *Proceedings of the 12th International Conference on Human-Computer Interaction (HCI '07)*, 2007.
- [11] M. Waschbüsch, S. Würmlin, D. Cotting, F. Sadlo, and M. Gross, "Scalable 3D video of dynamic scenes," *Visual Computer*, vol. 21, no. 8–10, pp. 629–638, 2005.
- [12] D. Cotting, M. Naef, M. Gross, and H. Fuchs, "Embedding imperceptible patterns into projected images for simultaneous acquisition and display," in *Proceedings of the 3rd IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR '04)*, pp. 100–109, 2004.
- [13] S. Borkowski, O. Riff, and J. L. Crowley, "Projecting rectified images in an augmented environment," in *Proceedings of the IEEE International Workshop on Projector-Camera Systems (PROCAMS '03)*, 2003.
- [14] M. Fiala, "Automatic projector calibration using self-identifying patterns," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, 2005.
- [15] R. Raskar, J. van Baar, and J. X. Chai, "A low-cost projector mosaic with fast registration," in *Proceedings of the IEEE Asian Conference on Computer Vision (ACCV '02)*, 2002.
- [16] R. Raskar and P. Beardsley, "A self-correcting projector," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '01)*, vol. 2, pp. 504–508, 2001.
- [17] J. van Baar, T. Willwacher, S. Rao, and R. Raskar, "Seamless multi-projector display on curved screens," in *Proceedings of the Workshop on Virtual Environments (EGVE '03)*, 2003.
- [18] R. Raskar, J. van Baar, P. Beardsley, T. Willwacher, S. Rao, and C. Forlines, "iLamps: geometrically aware and self-configuring projectors," in *Proceedings of the ACM Transactions on Graphics (SIGGRAPH '03)*, no. 3, pp. 809–818, July 2003.
- [19] W. Sun, X. Yang, S. Xiao, and W. Hu, "Robust checkerboard recognition for efficient nonplanar geometry registration in Projector-Camera systems," in *Proceedings of the ACM/IEEE 5th International Workshop on Projector Camera Systems (PROCAMS '08)*, 2008.
- [20] Y. Ding, J. Xiao, K.-H. Tan, and J. Yu, "Catadioptric projectors," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPR '09)*, pp. 2528–2535, 2009.
- [21] A. Kushal, J. van Baar, R. Raskar, and P. Beardsley, "A handheld projector supported by computer vision," in *Proceedings of the 7th Asian Conference on Computer Vision (ACCV '06)*, pp. 183–192, 2006.
- [22] T. Johnson and H. Fuchs, "Real-time projector tracking on complex geometry using ordinary imagery," in *Proceedings of the International Workshop on Projector-Camera Systems (PROCAMS '07)*, 2007.
- [23] S. J. Koppal and S. G. Narasimhan, "Illustrating motion through DLP photography," in *Proceedings of the IEEE International Workshop on Projector-Camera Systems (PROCAMS '03)*, pp. 9–16, 2009.
- [24] S. G. Narasimhan, S. J. Koppal, and S. Yamazaki, "Temporal dithering of illumination for fast active vision," in *Proceedings of the 10th European Conference on Computer Vision (ECCV '08)*, pp. 830–844, 2008.
- [25] J. S. Beis and D. G. Lowe, "Shape indexing using approximate nearest-neighbour search in high-dimensional spaces," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '97)*, pp. 1000–1006, 1997.
- [26] R. Sukthankar, R. G. Stockton, and M. D. Mullin, "Smarter presentations: exploiting homography in camera-projector systems," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV '01)*, vol. 1, pp. 247–253, 2001.
- [27] V. N. Dao, K. Hosoi, and M. Sugimoto, "A semi-automatic realtime calibration technique for a handheld projector," in *Proceedings of the ACM Symposium on Virtual Reality Software and Technology (VRST '07)*, 2007.
- [28] K. Fujii, M. D. Grossberg, and S. K. Nayar, "A projector-camera system with real-time photometric adaptation for dynamic environments," in *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '05)*, vol. 1, pp. 814–821, 2005.
- [29] S. E. Grigorescu, N. Petkov, and P. Kruizinga, "Comparison of texture features based on Gabor filters," *IEEE Transactions on Image Processing*, vol. 11, no. 10, pp. 1160–1167, 2002.