# Auto-Calibration for Screen Correction and Point Cloud Generation

***Jason Deglint*, Andrew Cameron*, Christian Scharfenberger*,***
***Mark Lamm**, Alexander Wong*, David Clausi****

***Department of Systems Design Engineering**
**University of Waterloo**
**Waterloo, Canada**

****Christie Digital Systems**
**Canada Inc.**
**Kitchener, Canada**

## Abstract

***Showing a checkerboard at different poses for camera projector calibration is impractical for large scale applications such as projection mapping onto buildings. We use an automatic calibration technique that projects Gray code structured light patterns, which, extracted by the camera, build a dense correspondence for calibration. Two applications benefit from automatic calibration: 3D model generation and screen correction.***

## Author Keywords

auto-calibration; self-calibration; learned geometry; active vision; camera; projector; structured light; screen correction; image warping; 3D model; point cloud; stereoscopic device; stereo vision; projection mapping

## 1. Introduction

Projection mapping is when an image is projected onto a known 3D surface to change the appearance of that surface. Some examples of projection mapping are illustrated in Figure 1. By knowing the geometry of a scene, as well the position of a projector relative to the scene, the projector can map images onto that scene. For example, in Figure 1a, a blank car can be made to have red, orange, or other paint appearance, with racing-stripes and other features easily to be added. More large scale examples can be seen in Figures 1b to Figures 1e, where stadiums and buildings are used as a target projection surfaces.

In order to project content onto an object's surface, the projector parameters such as the intrinsic and extrinsic calibration parameters must be known. The intrinsic parameters include the focal length, the principal point, and the lens distortion. The extrinsic parameters involve knowing the relative translation and rotation of the projector with respect to the 3D surface. Another requirement to project content that wraps perfectly the object's surface is the known geometry of a 3D surface. The latter is often accomplished using 3D laser scanning, which is both extremely time consuming and expensive. In order to facilitate the calibration procedure and

(a)

(b) (c) (d) (e)

**Fig. 1**: Examples of projection mapping.

to avoid costly laser scanning of 3D surfaces, camera projector systems are increasingly gaining importance in the multimedia market given their potential to extract the surface geometry using structured light [1]. While commercially available camera projector systems are made for short range projections and are calibrated by the manufacturer, the intrinsic and extrinsic parameters of the camera(s) and projector(s) in custom-designed systems are usually unknown and need to be calibrated on site.

Existing approaches to camera projector calibration aim to manually extract features from a checkerboard board shown at different poses to determine the camera parameters [2, 3] as a first step. Once the intrinsic calibration parameters are determined, one can then calibrate the entire camera projector system [4]. The standard calibration techniques are sufficient if the setup does not change such as the car scenario (Fig-

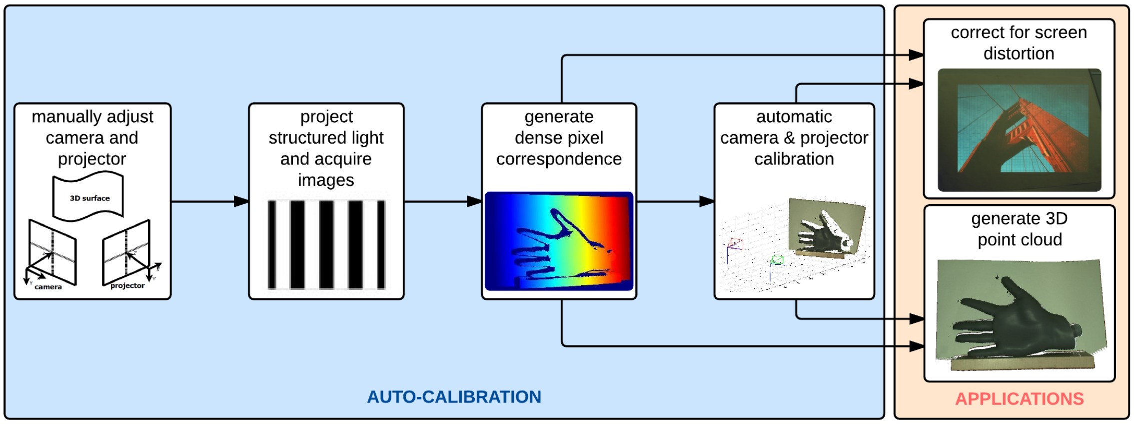



**Fig. 2**: The main steps in the auto-calibration process as well as two main applications.

ure 1a) or if the projection target is in a close proximity of the system. However, in scenarios where the camera projector setup or the 3D scene changes frequently or where the target object is farther away from the camera projector system, manual camera projector calibration may become inapplicable, time consuming as well as repetitious and tedious. In addition, manually extracted features for camera calibration may be error prone and, hence, inappropriate for camera projector calibration because of their potential to introduce and propagate projection errors. Therefore, a more flexible and faster method that can extract features and determine the calibration parameters in an automated fashion would be of great interest.

In this paper we present a new calibration method [5, 6], known as self-calibration or auto-calibration. The main advantage of this method is its simultaneous camera projector calibration, without the need for a manual checkerboard extraction procedure. In addition, this auto-calibration method can be used to learn the 3D surface of a scene by encoding the scene using structured light, and finding the relationship between the scene and the camera projector setup. The latter is of great importance for applications such as screen correction and projection mapping based on 3D models since it does not require any costly 3D laser scans at all. In this paper, we also want to present two potential applications, point cloud generation and screen correction, which benefit from this auto-calibration technique.

## 2. Methodology

The overall architecture of the automatic camera projector calibration [5, 6] can be broken down into four main stages as shown in Figure 2: i) manual adjustment of camera and projector, ii) simultaneous projection and acquisition of structured light sequences, iii) generation of dense pixel correspondence between projector image and camera image, and iv) automatic calibration of the camera and projector pair.

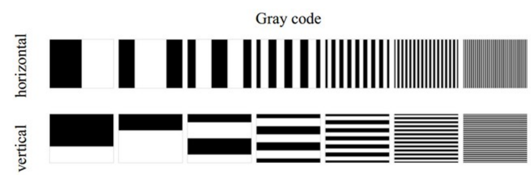



**Fig. 3**: Horizontal and vertical gray code binary pattern [5].

The first step involves properly aligning the camera and projector pair. This is achieved by manually changing the zoom, focus, and any other parameters for both the camera and projector. In addition, it is necessary that both the projector and camera cover the same region of interest on the scene.

Once the system is properly set up, the scene can be encoded by projecting a series of structured light images and capturing an image for each projection. Various structured light schemes may be used, each with its own advantages. Single-shot acquisitions have been demonstrated by exploiting color and non-repeating sequences in the structured light patterns, such as de Bruijn sequences [7]. However, these typically draw vertical bands, sacrificing horizontal scan resolution. Furthermore, ambient lighting and the color of the surface being scanned can both interfere with reliable extraction of color from the camera's image of the structured light pattern.

In our approach, a Gray code binary pattern was used for a time multiplexed code, as seen in Figure 3. Gray code patterns are more robust, requiring only the delineation of intensity. Using this coding scheme every pixel in the projector has a unique code based on binary sequence. The Gray code can then be extracted in the camera image and allows for the generation of dense pixel correspondences between the projector and camera pair [1]. In order to adequately encode the sur-


(a) (b) (c)

**Fig. 4**: (a) Projecting a gray code structured light sequence and capturing images with a camera in order to build a 3D model of a hand, (b) the 3D model of a hand, and (c) a 3D model of a car.

face, both horizontal and vertical series of Gray code binary images are needed.

Using the dense pixel correspondences established between the projector and camera images, a mathematical model is then used to create a cost function which is optimized to estimate the intrinsic and extrinsic camera parameters. This model is based on the concepts of epipolar geometry [8].

Suppose we have a pixel coordinate $\boldsymbol{x_i}$ in the camera image, and a corresponding pixel coordinate in $\boldsymbol{u_i}$ in the projected image. These two points are related by the Fundamental matrix, $\boldsymbol{F}$, which states that $\boldsymbol{u_i}^T \boldsymbol{F} \boldsymbol{x_i} = 0$ [8]. Using this relationship we can build the cost function, as seen in Equation 1, which is the Sampson approximation of the reprojection error [6, 8]. Here, $(\bullet)_k$ is the $k$-th component of a vector, and all $n$ camera and projector coordinates are in homogeneous coordinates.

$$\sum_{i=1}^{n} \frac{\left(\boldsymbol{u_i}^T \boldsymbol{F} \boldsymbol{x_i}\right)^2}{\left(\boldsymbol{F}\boldsymbol{x_i}\right)_1{}^2 + \left(\boldsymbol{F}\boldsymbol{x_i}\right)_2{}^2 + \left(\boldsymbol{F}^T\boldsymbol{u_i}\right)_1{}^2 + \left(\boldsymbol{F}^T\boldsymbol{u_i}\right)_2{}^2} \quad (1)$$

To incorporate the constraints we add penalty terms for the focal lengths and the principal points. After building this objective function the Levenberg-Marquardt Algorithm can be used to find the intrinsic and extrinsic parameters of the camera and projector [5, 6]. These include the focal lengths and principal points of the camera and projector, and the relative position and orientation of the projector with respect to the camera. Knowing these intrinsic and extrinsic parameters results in a calibrated camera and projector pair that can be used to extract the 3D geometry of objects and to wrap projected content accordingly.

## 3. Applications

There are two main applications that use the dense pixel correspondence and the calibrated camera projector pair. They are 3D model generation and screen correction.

### 3.1. Point Cloud Generation

A calibrated camera projector system allows the reconstruction of 3D data from two or more views of a scene. In order to accomplish this structured light patterns are displayed by the projector and corresponding images are captured by the camera.

The model produced by a structured lighting scan consists of a collection of 3D points in space, or a point cloud. In the case of large or complex objects, sometimes multiple scans may be required in order to sample the entire surface of interest. For instance, after scanning, the scanned object or the scanning system may be repositioned to scan more of the object. The resultant point clouds can then be fused together to provide a larger or denser point cloud than can be acquired from a single scan.

### 3.2. Screen Correction

Given a 3D surface model such as a point cloud acquired as described above, any desired two-dimensional imagery can be distorted or "pre-warped" such that by projecting it on the scanned surface, the shape of the surface distorts it back into the expected image. This is subject to the limitations arising from the surface geometry.

For our approach, we use the extracted 3D geometry of the surface, as described in Section 3.1, to find a 2D region, rectangular from the viewpoint, which lies entirely inside the surface region illuminated by the projector. This region will contain the corrected image content once it has been prewarped by software and then distorted by the surface. The pre-warping of the image occurs by first mapping the original, undistorted image coordinates to the coordinate range of the desired viewport region. Then, these coordinates are cast through the viewport onto the 3D surface model, where they are projected into the projector's viewport. Once the image coordinates are in the projector coordinate space, the image is re-sampled at the desired output resolution and displayed.

## 4. Results

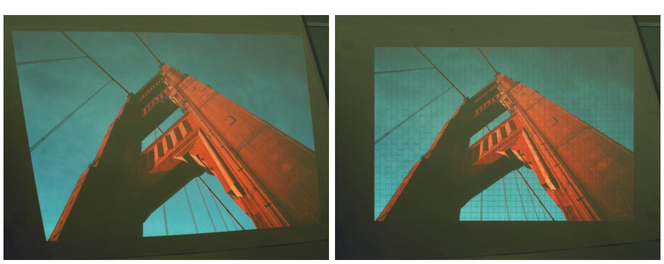
The proposed method has been tested using a Christie Matrix StIM projector and a Flea3 Point Grey camera. The projector has a resolution of $1920 \times 1200$ pixels and the camera a resolution of $1624 \times 1224$ pixels. In the experiments, the intrinsic and extrinsic calibration parameters were determined and used for two different experiments. In the first experiment, we build a 3D point cloud, while in the second experiment, we want to explore the performance of the system to correct for screen distortion. Both of these experiments will be discussed in the following subsections.

### 4.1. Point Cloud Generation

In the first experiment a 3D point cloud of the surface is to be obtained. This is accomplished by projecting a Gray code structured light sequence onto a 3D model of a hand, as seen in Figure 4a. Using the calibrated system and the dense pixel correspondence between the stereo images a 3D point cloud of the hand was generated, as seen in Figure 4b. In addition, a 3D point cloud of a car was created by using the same technique. A rendered image of this 3D point cloud can be seen in Figure 4c. Therefore, by using a calibrated stereo pair and by projecting structured light onto the scene, the 3D information of that scene can be determined.

### 4.2. Screen Correction

In the second experiment, given a projector relationship to a 3D surface, we want to project an image that appears corrected to a viewer. Here the viewer sees a distorted projected image on a planar surface, as seen in Figure 5. Since the viewer is at some oblique angle to the normal of the surface, the image will appear distorted. By using the calibrated camera projector pair and the dense pixel correspondences a 3D model can be built. The screen correction algorithm then finds a 2D region, rectangular from the viewpoint, and generates a pre-warped image that can be projected. This pre-warped image is then projected onto the planar surface and will appear corrected to the viewer. An example of this can be seen in Figure 5.



(a) (b)

**Fig. 5**: The distorted image in (a) is corrected in (b).

## 5. Conclusions

Projection mapping is used to project an eye-point corrected image onto a known 3D surface. In order to accomplish this task there are two requirements: to have a calibrated camera projector pair, and to know the 3D geometry of the scene. The present method of calibrating a camera projector pair involves using a checkerboard calibration chart, which can be very time consuming and tedious if repeated often. In addition, the standard method of determining a 3D surface is to use 3D laser scanning, which is also time consuming and expensive. By using an auto-calibration technique both these issues are resolved. The auto-calibration can simultaneously calibrate the system and determine the 3D geometry of a scene. This is accomplished by using Gray code structured light which creates a dense pixel correspondence between the projector image and the camera image. By knowing these relationships, a screen correction algorithm can also pre-warp an image so it appears corrected to a viewer.

## 6. Acknowledgments

The study was funded by the Canadian Mitacs Accelerate Cluster program, the Natural Sciences and Engineering Research Council (NSERC) of Canada and the Ontario Ministry of Economic Development and Innovation.

## 7. References

[1] J. Salvi, J. Pags, J. Batlle, Pattern codification strategies in structured light systems, PATTERN RECOGNITION 37 (2004) 827–849.

[2] Z. Zhang, A flexible new technique for camera calibration, IEEE Trans. Pattern Anal. Mach. Intell. 22 (11) (2000) 1330–1334.

[3] J. Y. Bouguet, Camera calibration toolbox for MATLAB.

[4] D. Moreno, G. Taubin, Simple, accurate, and robust projector-camera calibration, in: 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT), 2012 Second International Conference on, 2012, pp. 464–471.

[5] S. Yamazaki, M. Mochimaru, T. Kanade, Simultaneous self-calibration of a projector and a camera using structured light, in: Proc. Projector Camera Systems, 2011, pp. 67–74.

[6] R. Hartley, C. Silpa-Anan, Reconstruction from two views using approximate calibration, in: Proc. 5th Asian Conf. Comput. Vision, Vol. 1, 2002, pp. 338–343.

[7] L. Zhang, B. Curless, S. Seitz, Rapid shape acquisition using color structured light and multi-pass dynamic programming, in: 3D Data Processing Visualization and Transmission, 2002. Proceedings. First International Symposium on, 2002, pp. 24–36.

[8] R. I. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, 2nd Edition, Cambridge University Press, ISBN: 0521540518, 2004.