

# PROJECTOR-BASED AUGMENTED REALITY FOR QUALITY INSPECTION OF SCANNED OBJECTS

J. Kern, M. Weinmann, S. Wurstthorn

Karlsruhe Institute of Technology (KIT), Institute of Photogrammetry and Remote Sensing (IPF), Karlsruhe, Germany  
(jens.kern, martin.weinmann, sven.wurstthorn@kit.edu)

Commission II, WG II/4

**KEY WORDS:** Thematic mapping and monitoring, 3D, Augmented Reality, sensor and system calibration, projector-camera

## ABSTRACT:

After scanning or reconstructing the geometry of objects, we need to inspect the result of our work. Are there any parts missing? Is every detail covered in the desired quality? We typically do this by looking at the resulting point clouds or meshes of our objects on-screen. What, if we could see the information directly visualized on the object itself? Augmented reality is the generic term for bringing virtual information into our real environment. In our paper, we show how we can project any 3D information like thematic visualizations or specific monitoring information with reference to our object onto the object's surface itself, thus augmenting it with additional information. For small objects that could for instance be scanned in a laboratory, we propose a low-cost method involving a projector-camera system to solve this task. The user only needs a calibration board with coded fiducial markers to calibrate the system and to estimate the projector's pose later on for projecting textures with information onto the object's surface. Changes within the projected 3D information or of the projector's pose will be applied in real-time. Our results clearly reveal that such a simple setup will deliver a good quality of the augmented information.

## 1. INTRODUCTION

With augmented reality (AR), our vision of real objects can be extended with additional computer generated information. If the virtual information is spatially registered to the real object in real-time, the user will be able to interact with spatial data in his natural environment and scale (Wurstthorn et al., 2004). The augmentation can be achieved with the help of a head-mounted display (HMD, also called smart glasses or generally NED for near eye display (Hainich and Bimber, 2011) ) like Microsoft's HoloLens (Microsoft, 2017) or a video-based solution by adding information to the video stream of a tablet's back camera. With a video projector, a user does not need to wear a HMD or have to hold a tablet to view additional data. A group of users can jointly see the same projected information on an object with a sufficiently diffuse surface reflectance behavior. In (Ridel et al., 2014), it is proposed to highlight inscriptions on antique artifacts with projector-based augmented reality. The support of surgeries by projecting information on patients is addressed in (Nicolau et al., 2011; Lahanas et al., 2014; Kilgus et al., 2014).

In this paper, we show how we use augmented reality techniques to project information on the surface of objects that fit approximately into a cube of 1 m edge length to let one or more users naturally inspect this spatially registered information on the object itself. In this way, users can focus their work on the object instead of permanently changing views between the object and a display. For instance, the 3D information of color-coded differences between two scanning results of a Buddha statue shown on the left side of Figure 1 can be projected directly onto the real object on the right side. The 3D information used for augmentation does not necessarily need to be static like in the above example.

One could project a point cloud resulting from a scanning task in real-time in order to give the user a direct feedback on regions on the object that still need to be scanned. The focus in this paper

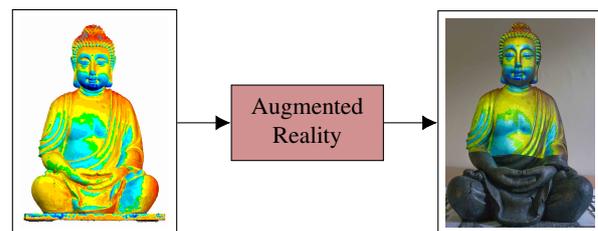


Figure 1. Schematic representation of a projector-based augmented reality application. Some color coded 3D information related to an object (left) is directly projected onto the object itself (right) to simplify an inspection task for the user.

is not on a distinct application but on the configuration, calibration and registration of a low-cost projector system enabling augmentation with any 3D information, static or dynamic, of small objects.

The registration of object and projection is done with a combination of a camera and an off-the-shelf projector and the use of a board with a pattern of coded targets. The working object is simply put on the pattern board and registered once at the beginning of a session. The projector system itself is mounted on a moveable tripod to enable a quick setup of different views. The pose of the projector system relative to the pattern board is computed in real-time. The main contributions of this paper are:

- A simple calibration method of the projector-camera system with a board of coded fiducial markers.
- The use of the calibration board for simple and real-time registration of object and projected information.
- A depth buffer based masking to avoid unwanted visual effects of the projection due to low quality of calibration or pose estimation.

After briefly summarizing related work in the field of augmented reality in Section 2, we describe the proposed methodology for projector-based augmented reality in Section 3. Subsequently, in Section 4, we present the involved setup and we demonstrate its performance for the example of an indoor scene. The insights gained via our experiments are summarized in Section 5. Finally, in Section 6, we provide concluding remarks as well as suggestions for future work.

## 2. RELATED WORK

Research on augmented reality (AR) started with early works on display technology for virtual reality (Cruz-Neira et al., 1993) or head-mounted displays (Sutherland, 1968) before the terms *augmented reality* (Azuma, 1997) or *mixed reality* (Milgram and Kishino, 1994) have been defined.

The user's vision can be augmented in different ways dependent on the technology, the application and the environment like prepared or unprepared indoor spaces or even outdoor spaces. With video-based AR, the environment is captured by a camera. Its image is shown on a screen fixed to the camera that can be augmented with computer-generated 3D models that are simply transformed into every camera frame. The camera's inner and exterior orientation (its pose) have to be known at the moment the transformation takes place to fulfill the real-time constraint defined for AR (Azuma et al., 2001).

In AR, much of the research is focused on tracking the device's pose in real-time with the help of one or more cameras. The simplest camera-based tracking method is marker-based with the typical planar square targets from projects like ARToolKit (Kato and Billinghurst, 1999; Abawi et al., 2004) or ArUco (Garrido-Jurado et al., 2014, 2016). If placing markers is not an option, natural features like keypoints could be used (Comport et al., 2006; Klein and Murray, 2009). Typical smartphone or tablet hardware can do the processing of such system setups today. More sophisticated methods use 3D building models (Urban et al., 2013, 2016) or even textured city models for camera-based tracking within these models (Reitmayr and Drummond, 2006). Video-based augmented reality enables occlusions by virtual objects since the real environment is only visible through the display.

Modern HMDs like Microsoft's HoloLens (Furlan, 2016) use four cameras in combination with a depth sensor to track the device within an unprepared environment and simultaneously create a sparse 3D model in a SLAM like manner. Environments are persistently stored between sessions and continuously updated. This enables the HoloLens to accurately determine its pose and robustly place virtual objects, which Microsoft calls "holograms", within the real environment. In order to enable occlusions with this technology, the HMD hardware has to be prepared for it (Gao et al., 2013). In a collaborative AR project, every participating person has to wear such a device. First, given the high price for the HoloLens or other HMDs, this will increase the costs for the project. Second, HMDs are still not comfortable enough for typical working times (Schwerdtfeger et al., 2011).

Using tablets or smartphones instead of HMDs may reduce cost. But the drawback of these devices is that they have to be held by one hand while the other hand is used for interactions on the screen. This way, the user has to lay down the tablet for interactions with his real environment or the tablet has to be fixed somewhere requiring users to always come back to see any augmented information on the tablet's screen.

Projector-based AR, also called spatial AR (Bimber and Raskar, 2005), enables a group of users to see the same augmented information. Bandyopadhyay et al. (2001) define the term "shader lamp" for projector-based textures. The AR setup can be prepared first without any action required by the actual users later on. The best usage scenario is an indoor space that could be specially prepared like surgery rooms (Nicolau et al., 2011; Lahanas et al., 2014; Kilgus et al., 2014), manufacturing workplaces (Otto et al., 2014; Sand et al., 2016) or exhibition rooms (Ridel et al., 2014). The geometry and texture of projection surfaces can be measured in advance or, with the help of depth sensors, in real-time (Hoang et al., 2017). Occlusions are not possible with projector-based augmented reality.

Table 1 gives an overview of the three different types for augmentation with a comparison of some characteristics like quality of perception, calibration or usability.

Typical projectors are made for flat uniformly colored surfaces with good reflectance capabilities. The geometry of more complex surfaces that differ from that flat model has to be known in order to correct the thereby resulting distortions. Pantuwong et al. (2016) use a projector in combination with a Microsoft Kinect depth camera to interactively augment the surface of sand in a sandbox with a color-coded height visualization and flowing water simulation. The surface of sand is ideal in this case because it has smooth edges. Every region of the sandbox can be measured by the depth camera with a nadir view and the sand's texture is more or less homogeneous.

Many applications require a setup of more than one projector (Bimber and Raskar, 2005; Hainich and Bimber, 2011; Otto et al., 2014; Jones et al., 2014; Hoang et al., 2017) to increase the projected area or to cover an object with projections from several sides. Multi-projector setups are more complex for geometric calibration and need additional photometric correction at least in the overlapping areas (Brown et al., 2005; Hainich and Bimber, 2011).

Typical office projectors are not bright enough for outdoor usage at normal daylight conditions. Many of them are based on DLP (Digital Light Processing) or LCD technology. They have a narrow depth of field because of the large apertures that are used to maximize the overall brightness. This causes problems for complex surfaces especially on surface parts that are outside the depth of field. Oyamada and Saito (2008) present a method for reducing the effect of defocus blur with a single projector-camera system. Laser projectors do not have these disadvantages because they operate with one laser beam that covers an area at a high frequency. They are able to project precise monochrome labels and signature on surfaces. EXTEND3D (2017) provide such laser-based projector-camera systems for commercial industry applications.

Calibration of projector-camera systems solves the relative pose of camera and projector and the intrinsic parameters of both devices. Audet and Okutomi (2009) use flat patterns of fiducial markers with binary codes for automatic identification of single markers to calibrate their projector-camera system. One half of the markers is printed and the other half is projected onto the white areas between the printed markers. Yang et al. (2016) propose a pattern composed of a printed black dot pattern and projected white dots. Both dot patterns are randomly distributed. These patterns can work on large distances without adopting the marker sizes, and they are said to give more stable results for the intrinsic parameters of camera and projector because circular targets allow more precise image measurements. Resch et al. (2016)

	Video-based	See-through	Projector-based
<b>Device</b>	Smartphone, Tablet	HMD, Smart Glasses	DLP, LCD, LASER
<b>Quality of perception</b>	display covers real environment; enables occlusions	transparency of virtual objects, like sunglasses	dependent on surface structure and color
<b>Calibration</b>	once	once per session	once
<b>Number of users per device</b>	few	single person	few to many
<b>Usability</b>	one hand to hold, both hands for interactions	hands free, may be uncomfortable on longer sessions	needs prepared environment
<b>Outdoor capability</b>	dependent on display brightness		only at night or with laser projector

Table 1. Some characteristics for the three different types of augmentation: video-based, see-through and projector-based.

use a self-calibration approach with the help of the known geometry of an object or surface that is later used for projecting the information onto. Jones et al. (2014) use a setup of multiple projectors, each in combination with a Microsoft Kinect. They use an automatic calibration process with the help of a projected code sequence observed by all units that yields dense correspondences in both 2D and 3D with the help of the depth images. After solving the extrinsic parameters of each unit, a unified 3D model of the room is formed by combining the depth maps.

The use of depth sensors is common in AR projects because they provide 3D models of the near environment in real-time. This is the environment the user interacts with like placing virtual objects on real tables or walls. The depth information enables occlusion of virtual objects by real objects which helps users to experience precise spatial interaction of both virtual and real objects. Furthermore, depth sensors help to make egomotion tracking of the devices themselves more robust if combined with other sensors. For instance, the HoloLens HMD has a built-in depth sensor for tracking and mapping support and hand gesture recognition. Google’s Project Tango uses smartphones or tablets with a built-in depth sensor (Google Inc., 2017). There is even a small attachable depth sensor called “Structure Sensor” for tablets available (Kakadiaris et al., 2016).

### 3. METHODOLOGY

Our augmented reality system consists of a rigid combination of an off-the-shelf color projector and a camera system. The camera is used for a simple marker-based pose estimation of the projector relative to the 3D object and for calibration of the inner orientation (IO) of the projector. The projector restricts the system for indoor use. The system itself is mounted on a studio tripod with wheels that allow movements on the floor within the range of length of the cables used for data transfer and power supply to camera and projector. The inclination can be adjusted but has to be fixed with a locking screw. Most of the pose adjustments during a session will be the movements around the object and turn around the vertical axis in order to project information on different parts of the object.

Since most off-the-shelf projectors have a lens that allows manual adjustment of focus or even zoom, the system needs a robust and automatic calibration procedure. Accordingly, calibration will be described first (Section 3.1). Afterwards, we will describe registration and the steps needed for optimizing the rendering pipeline that allows us to project virtual 3D models onto the object’s surface (Section 3.2). A schematic overview of our system is provided in Figure 2. Photometric corrections of the projections due

to the object’s varying texture or reflectance are not covered in this paper.

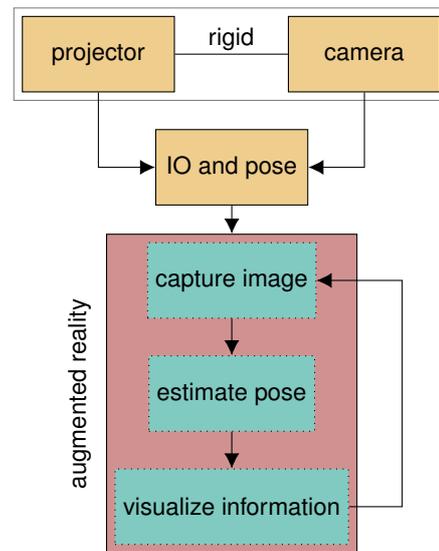


Figure 2. Flowchart of required steps to combine real and virtual world with a projector-based augmented reality application.

#### 3.1 Calibration

To calibrate the projector-camera system, a marker-based method as presented in (Audet and Okutomi, 2009) is used. In general, the calibration of the interior orientation of the projector and the relative pose between projector and camera can be estimated using standard camera calibration algorithms. One of the patterns is, as usual with many camera calibration techniques, printed on paper. The other is projected using the projector. The printed pattern is a chessboard-like arrangement of fiducial markers in which each square marker can be uniquely identified by different binary codes with the squared areas. We use the ArUco marker system specially optimized for pose estimation (Garrido-Jurado et al., 2014, 2016). This allows us to perform the calibration without requiring all markers to be visible at once. The projected pattern complements the empty spaces of the printed one with ArUco markers if they are overlaying properly. To facilitate the detection, the color of the projected pattern is inverted and has significantly smaller markers than the printed one. See the picture of the calibration board on top of the flow chart in Figure 4. The projected white markers cover only a fraction of the board due to the limitation of the projector’s lens settings.

At the beginning, the desired patterns are uniquely generated in

a common coordinate system. The scale of the printed pattern has to be known in order to perform a metric evaluation. Both patterns, the printed one and the projected one, need to have the same resolution or the scale between these patterns has to be well-known. Starting the calibration routine, the projected image is displayed as full-screen and has to lie mainly in the same area as the printed pattern. Both patterns are now superimposing in parts, so that a few transformation steps are required to solve this problem (see Figure 2).

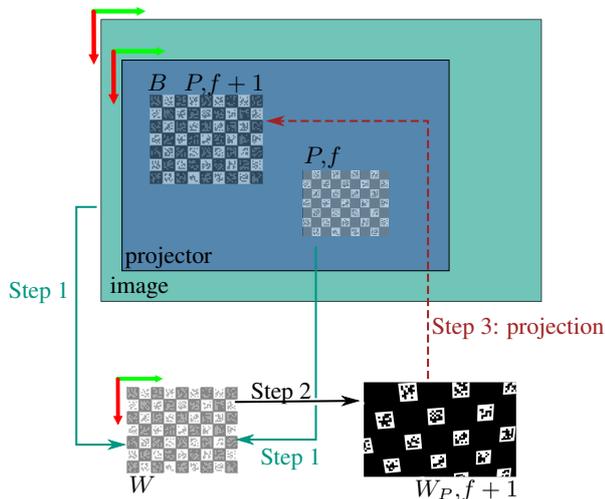


Figure 3. Use of a homography to warp the projected pattern. After applying the transformation, the projected pattern complements the printed one.

First, the image to be projected is generated in a resolution four times larger than the resolution of the projector to decrease aliasing effects while transforming the image.

Hereinafter, the common world target coordinates of the projected and printed patterns ( $W$ ) are  $\in \mathbb{R}^3$  (see Figure 3). Coordinates  $p \in \mathbb{R}^2$  which refer to printed components on the calibration board are denoted with index  $B$  and projected components with index  $P$ . In the captured image at frame  $f$  of the camera, the projected points  $p_{P,f}$  and the points  $p_{B,f}$  of the printed pattern on the board are detected.

The homography  $H_B^W$  is a perspective transformation model for planes which transforms the image points of the board to the image plane of the world coordinates. The  $Z$ -component of the world coordinates in the board can be omitted. The resulting points are called  $p_W \in \mathbb{R}^2$ . The homography  $H_P^W$  transforms the image points of the projected pattern into the image plane of the world coordinates. Both homographies are calculated using a RANSAC-based method (see Step 1 in Figure 3). First, the camera image points of the printed markers are transformed to the target position in world coordinates:

$$p_{W_B} = H_B^W \cdot p_B \quad (1)$$

Then, the camera image points of the projected markers are transformed to the world coordinate system:

$$p_{W_P} = H_P^W \cdot p_P \quad (2)$$

In a further step, the previously projected pattern is warped so that the projected points are displayed at the desired position  $p_{B,f+1}$  on the printed pattern of the next frame ( $f + 1$ , see Step 2 in

Figure 3):

$$p_{W_{P,f+1}} = H_{P,f}^W \cdot H_W^B \cdot p_{W_{P,f}} \quad (3)$$

Since the image for projection is still in higher resolution, it has to be smoothed and scaled to the size of the projector resolution to reduce aliasing effects before it can be displayed.

Now the printed and projected markers are no longer superimposed (see Step 3 in Figure 3) which allows a more accurate detection of the projected markers and images can be captured for calibration. To compute the calibration, all transformations made on the projected pattern have to be stored. The previously presented transformations are calculated and applied during the capturing of the calibration images. The calibration images and the corresponding stored transformations are required for the following determination of the interior orientation of the projector and the relative pose between projector and camera. Figure 4 schematically shows the used calibration method of the projector-camera system. The interim results are highlighted in green (dotted) and implemented functions to estimate the calibration in red (dashed).

The calibration of a projector can be implemented by using camera calibration methods. For this, object points and corresponding image points of the projector are needed. The calibration of the relative pose between projector and camera is computed using a stereo camera calibration method. For this, object points and corresponding image points of the camera and the projector and their interior orientations are needed. The matrix  $V = H_{P,f}^W \cdot H_W^B$  represents the stored homography applied to the projected image during capturing the calibration images. First, the printed markers are detected in the image and  $H$  is estimated. The matrix  $H$  is the renewed calculation of the homography  $H_W^B$ , using optimized settings for the marker detection with subpixel accuracy of the edges. The visible projected markers are detected with subpixel accuracy, too. Because of the narrow depth of field of the projector, the defocused markers are bigger than the focused markers. To take this effect into account and minimize the influence of inaccurate edge detection, the marker centers are used for calibration (Audet and Okutomi, 2009). The marker centers which are still distorted by the camera are used as image points of the camera. In a further step, the distortion of the camera is removed to get the object points (world coordinates) after applying the transformation  $H$ . Now the marker centers in the image are only distorted by the projector. The image points of the projector are calculated from the generated markers which are used for projection. For this purpose, the visible projected markers are detected and the corresponding marker centers are determined. The subsequently applied transformation  $V$  delivers the image points of the projector. To perform the calibration, the described steps have to be applied to every captured calibration image. The calibration itself is carried out using the OpenCV function `calibrateCamera()` based on methods from (Zhang, 2000) and (Bouguet, 2004). It resolves the interior orientation of the projector.

Now the interior orientations of the projector and the camera are known and the relative pose can be calibrated subsequently. For that, the OpenCV function `stereoCalibrate()` can be called with the beforehand determined object points, image points of the camera and the projector. The interior orientations of the camera and the projector have to be handed over and labeled as fix while calling the function.

With the relative pose between camera and projector and the known

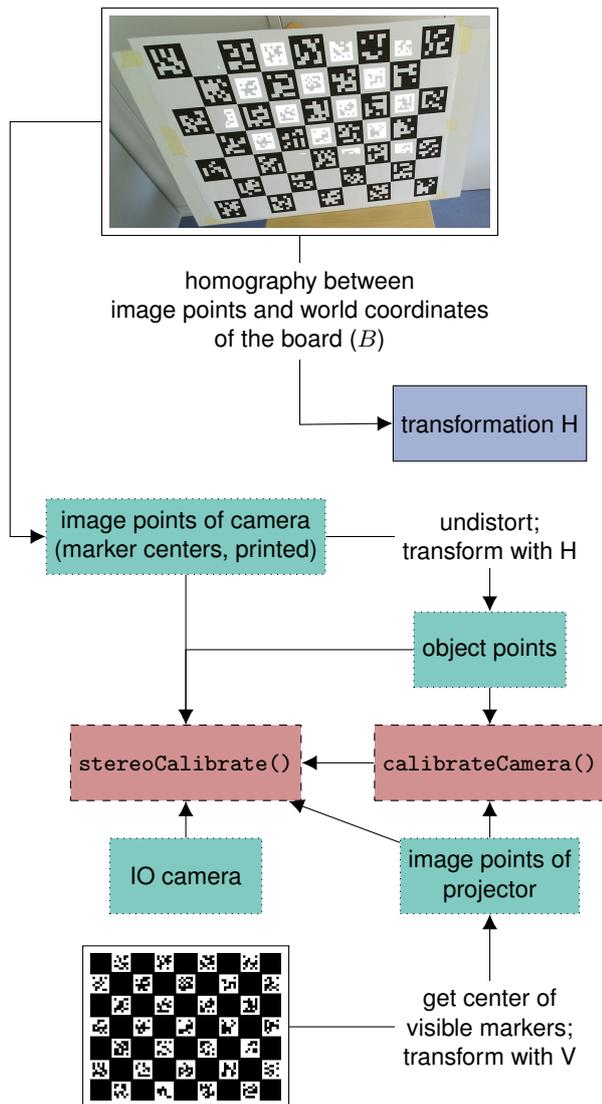


Figure 4. Flowchart with the steps that are required to calibrate the relative pose between the camera and the projector.

inner orientations of both devices, the projector-camera system is ready for performing visual augmentations.

### 3.2 Augmentation

Figure 5 shows the projection of the rendered model onto the statue. The ArUco chessboard below the statue defines the world coordinate system in which the pose of the statue is known. It is the same pattern used for calibration. These markers are now used to estimate the pose of the projector-camera system in the world coordinate system. The virtual objects are then projected. The difference between the texture of the Buddha statue and the projected texture is clearly visible. Figure 2 schematically shows the procedure of a projector-based augmented reality application. At the start of the application, the software loads the required calibration parameters and settings. Now the actual augmented reality application can be executed. The first step of the AR loop is the capturing of a camera image and the detection of visible markers. The pose of the camera can be estimated using the known world coordinates of the markers and the image coordinates of the detected markers. The relative pose  $M_C^P$  between camera ( $C$ ) and projector ( $P$ ) is known from calibration. Using the actual pose  $M_C$  of the camera, the pose  $M_P$  of the projector



Figure 5. Live view with representation of the detected markers. Useful to control the field of view and the camera settings.

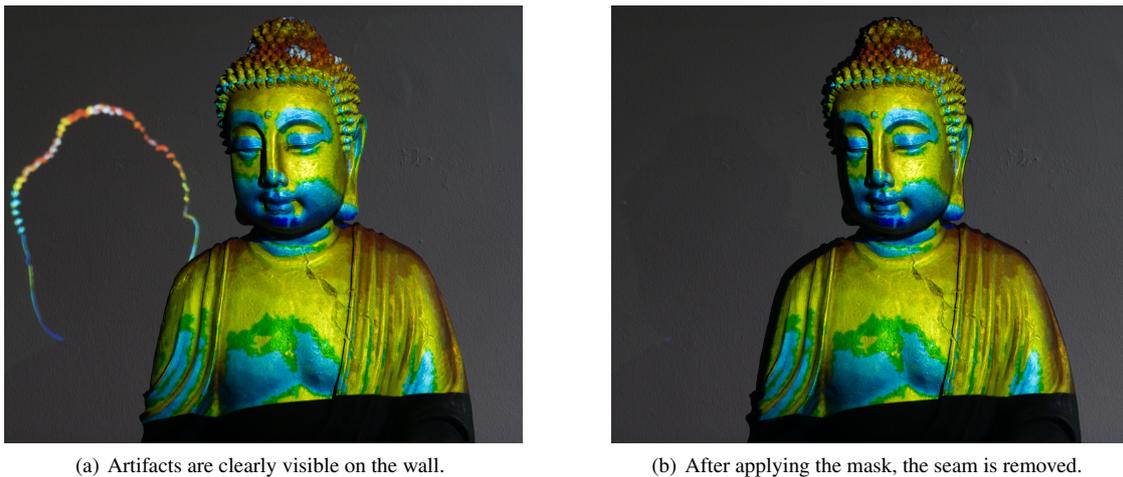
can be calculated as follows:

$$M_P = M_C \cdot M_C^P \quad (4)$$

with

$$M = \begin{bmatrix} R & t \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 4} \quad , \quad (5)$$

where  $R \in \mathbb{R}^{3 \times 3}$  represents a rotation matrix and  $t \in \mathbb{R}^3$  represents a translation vector. To complement the real world with the virtual information, an image of the virtual world has to be captured and projected by the projector. The process to generate a raster image of a virtual scene containing points, lines, surfaces and their normals and textures, etc. is called rendering. First, the virtual world containing all the desired information has to be generated in the same scale as the real world. Then the camera in the virtual world has to be positioned using the estimated projector pose in the real world. The inner orientation of the projector is represented by a projection matrix which contains also the near and far plane settings. This matrix warps the virtual objects perspective so that the raster image can be generated. After the rendering process, the generated image is projected into the real world and, in the ideal case, it fits seamlessly into the environment. While running the augmented reality application, the steps in the red box in Figure 2 are repeated continuously to project the virtual information in real-time onto the correct position in the real world. In our application, the open source 3D graphics toolkit OpenSceneGraph (OSG) (OSG Community, 2016) is used to visualize the virtual data. It can load 3D data in many different formats. Such 3D data is used to render the projected texture for the augmentation of the real object. The more accurate the geometry of the 3D data matches the real object geometry, the more accurate the projected texture will be. The real object's position within the marker system has to be measured and set up within the software. This process is quite simple if the calibration board is used for this task.



(a) Artifacts are clearly visible on the wall.

(b) After applying the mask, the seam is removed.

Figure 6. Due to the remaining errors of calibration and pose estimation of the projector-camera system, colored artifacts can be observed (left image). These artifacts can be removed through applying a depth buffer-based mask (right image). The lower part is black due to the restricted area that the projector can cover.

### 3.3 Depth buffer-based mask

Due to errors in the calibration of the interior orientation of the projector and camera, their relative pose and the pose estimation of the projector-camera system results in differences between desired and actual state of the projected information. The inaccuracy is especially visible at the edges of the real object. Unwanted artifacts are pictured on surfaces behind the augmented object, which disturbs the subjective impression (see Figure 6a). These artifacts can be removed by applying a depth buffer-based mask. Figure 7 describes the developed method. First, the virtual objects are rendered and the resulting picture frame is buffered. While rendering the image, the depth buffer can be prompted as a grayscale image by using an OpenGL query. After binarization, the depth buffer delivers a mask which separates object information from the background. Now the mask is decreased with a morphological operator, so that it fades out the edge areas. After applying the mask, the projected image has no interfering seams (see Figure 6b). It can be observed that this method is simple and fast, however, it has a great influence on the subjective impression of the augmentation.

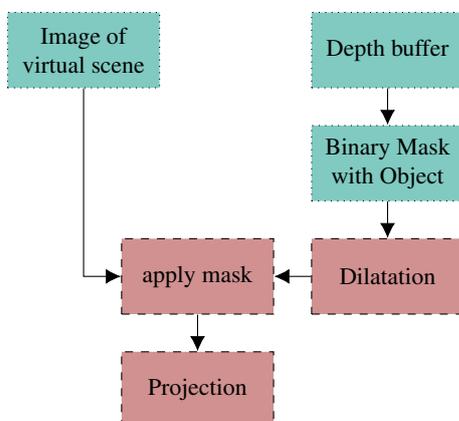


Figure 7. Process of applying a depth buffer-based mask on the projected image to remove unwanted seams.

## 4. EXPERIMENTAL RESULTS

We built a system which integrates OpenCV, ArUco and OSG to calibrate a projector-camera system and to test it in a specific application. For the setup we used the full HD sensor (1920 × 1080) from the low-cost RGBD camera Intel RealSense R200 and a NEC V311W (1280 × 800 color DLP) projector, with a footprint size of 0.82 mm @ 2 m of a focused projected pixel. To test the system, a texture is projected onto a Buddha statue with a height of 70 cm. The texture is a result of the rendering pipeline described in Section 3. With our system, we are able to project any 3D model onto real objects in this way.

In Figure 1, the right side of the diagram shows how the projector can cover the statue at a distance of approximately 2 m. The lower part of the statue is black because it is outside the projection area. In Figure 8, two close-ups of the resulting overlay of our projector-based augmented reality system are shown. The overlaid texture is used for demonstration only. It visualizes color-coded differences between two meshes resulting from different measurement techniques. The differences range from blue over green to red for -10 mm over 0 mm to 10 mm difference between those two registered meshes. The reference model has been reconstructed from 86 scans created with a commercial scanning device (Breuckmann stereoSCAN<sup>3D</sup>-HE) based on fringe projection resulting in a dense and precise mesh. The other model has been reconstructed from 170 images with classic photogrammetric bundle adjustment and meshing.

The accuracy of the registration of the overlaid texture and the object can be seen in the detailed view in Figure 8(a). Overall, the configured system achieved a reprojection accuracy of known world coordinates better than 4 mm in a hemisphere of 4 m diameter. For the calibration, 15 different projector-camera views of the calibration board have been used.

While using the system in an indoor laboratory, the projected texture is clearly visible and sufficiently bright. Despite the dark surface of our test object, it was no problem to distinguish the different colors. Also the narrow depth of field of the projector is no disadvantage here. The projected texture has large-area structures, whereby the blurring of the projection is not significant.



(a) Detailed front view of the head. (b) Detailed side view of the head.

Figure 8. Detailed view of the projected color-coded representation of a 3D comparison between the model of the statue and a point cloud. The detailed structures of the head allow to evaluate the quality of the projection.

## 5. DISCUSSION

We used two different low-cost RGBD cameras, the Intel RealSense R200 and the Microsoft Kinect V2 for testing purposes. For now, we did not use the depth sensing capabilities of these cameras in the methodology described in this paper. Even if the Kinect delivers images in higher quality, the RealSense is more suitable for the implemented application (Kern, 2016). In addition, the RealSense is more compact and there are more possibilities to fix the camera to other objects. Another decisive advantage of using the RealSense is the ability to configure parameters like the exposure time, which is necessary for good calibration results.

The used off-the-shelf DLP office projector provides little luminous intensity, so it can only be used in darkened rooms. In addition, the footprint of unfocused pixels is bigger than the footprint of focused ones, so it is not suitable for high-precision applications. Also the projection of text or numbers is not possible because of the narrow depth of field. However, the projector is ideally suited for the application shown in this paper. The projector generates a color image from sequentially projected monochrome frames using a color wheel. In conjunction with rolling shutter cameras, there are some unpleasant effects, making it necessary to set the optimal camera and projector settings for best calibration results.

The moveable tripod enables the user to quickly change the projected view due to the registration with the marker board used for calibration. No other interference is needed by the user as long as there are enough markers within the camera's sight. A successful pose estimation is clearly visible by the resulting projection onto the object. Because of the real-time tracking loop, the desired result can be simply adjusted by the user by moving the tripod back and forth. If the freedom of movements with the projector-camera system on its tripod is too restricted by the length of cables, the laptop computer can be put on a rack directly fixed to the tripod.

Having all markers on just one plane is not the most robust solution for pose estimation. This was motivated by the fact that we wanted to keep the considered scenario as simple as possible for our first approach towards projector-based AR. If more robustness and accuracy is required, the space covering the object

could be easily prepared with more markers on more planes, e.g. on the walls in a corner of a room.

## 6. CONCLUSIONS AND OUTLOOK

Even with our simple setup, small objects like our Buddha statue can be augmented with an off-the-shelf projector and a camera. Any kind of 3D information can be projected on known objects this way a long as this information is registered in 3D to the object. For the registration estimation of the projector-camera system, the flat calibration board with its coded fiducial markers is robust enough for a good quality. To compensate possible insufficient calibration of intrinsic parameters of the projector's lens, our depth buffer-based masking can remove unwanted seams.

Users are able to change the projected view by moving the tripod on its wheels around the object. The rendering process is continuously running in a loop adapting to all changes, be it the virtual 3D information or the projector's pose.

Though we did not take photometric corrections due to varying properties of the object's surface into account, the different colors of color-coded texture in our example projection are still distinguishable. We started our project with less information in mind like the binary information to highlight "good" or "bad" regions, e.g. in terms of regions that have already been scanned or still need to be scanned or in terms of regions with a sufficient point density and regions with a point density below a given threshold.

In future work, using more than one projector would result in a more flexible system. This would for instance allow to completely augment the statue with projected information. The user can then inspect the object from all directions without moving the system. Furthermore, different parts of the object can be augmented simultaneously because less parts are occluded, especially for more complex objects like our test statue. Starting with two projectors, it would be necessary to take photometric corrections into account, at least for the overlapping regions.

In a laboratory, the projector could also be tracked by an installed (optical) tracking system. This could lead to more accurate results but requires a large financial investment and different calibration methods. To project readable labels, a laser projector is indispensable. The actual system could be expanded by a laser projector to add this feature. The better way is to configure a system only using a color laser projector. This would solve problems like low luminous intensity and shallow depth of field.

## References

- Abawi, D. F., Bienwald, J. and Dörner, R., 2004. Accuracy in Optical Tracking with Fiducial Markers: An Accuracy Function for AR-ToolKit. In: *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality, ISMAR '04*, IEEE Computer Society, Washington, DC, USA, pp. 260–261.
- Audet, S. and Okutomi, M., 2009. A user-friendly method to geometrically calibrate projector-camera systems. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, pp. 47–54.
- Azuma, R., 1997. A survey of augmented reality. *Presence: Teleoperators and Virtual Environments* 6(4), pp. 355–385.
- Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S. and MacIntyre, B., 2001. Recent advances in augmented reality. *IEEE Computer Graphics and Applications* 21(6), pp. 34–47.

- Bandyopadhyay, D., Raskar, R. and Fuchs, H., 2001. Dynamic shader lamps: painting on movable objects. In: *Proceedings of the IEEE and ACM International Symposium on Augmented Reality*, New York, NY, USA, pp. 207–216.
- Bimber, O. and Raskar, R., 2005. *Spatial Augmented Reality – Merging Real and Virtual Worlds*. A K Peters.
- Bouguet, J.-Y., 2004. Camera Calibration Toolbox for Matlab. [http://www.vision.caltech.edu/bouguetj/calib\\_doc/](http://www.vision.caltech.edu/bouguetj/calib_doc/), visited in April 2017.
- Brown, M., Majumder, A. and Yang, R., 2005. Camera-based calibration techniques for seamless multiprojector displays. *IEEE Transactions on Visualization and Computer Graphics* 11(2), pp. 193–206.
- Comport, A. I., Marchand, E., Pressigout, M. and Chaumette, F., 2006. Real-time markerless tracking for augmented reality: The virtual visual servoing framework. *IEEE Transactions on Visualization and Computer Graphics* 12(4), pp. 615–628.
- Cruz-Neira, C., Sandin, D. J. and DeFanti, T. A., 1993. Surround-screen projection-based virtual reality: The design and implementation of the cave. In: *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '93, ACM, New York, NY, USA, pp. 135–142.
- EXTEND3D, 2017. Laser and video projection for the industry by extend3d. <http://www.extend3d.de/en/index.php>, visited in April 2017.
- Furlan, R., 2016. The future of augmented reality: HoloLens - Microsoft's AR headset shines despite rough edges [Resources.Tools and Toys]. *IEEE Spectrum* 53(6), pp. 21–21.
- Gao, C., Lin, Y. and Hua, H., 2013. Optical see-through head-mounted display with occlusion capability. In: *Proc. SPIE*, Baltimore, MD, USA, Vol. 8735, p. 87350F.
- Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. J. and Marín-Jiménez, M. J., 2014. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition* 47(6), pp. 2280–2292.
- Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. J. and Medina-Carnicer, R., 2016. Generation of fiducial marker dictionaries using Mixed Integer Linear Programming. *Pattern Recognition* 51, pp. 481–491.
- Google Inc., 2017. Project Tango. <https://www.google.com/atap/projecttango/>, visited in April 2017.
- Hainich, R. R. and Bimber, O., 2011. *Displays: Fundamentals & Applications*. A K Peters LTD / CRC Press LLC.
- Hoang, T., Reinoso, M., Joukhadar, Z., Vetere, F. and Kelly, D., 2017. Augmented studio: Projection mapping on moving body for physiotherapy education. In: *Proceedings of CHI 2017 – The ACM CHI Conference on Human Factors in Computing Systems*, Denver, CO, USA, pp. 1–12.
- Jones, B., Sodhi, R., Murdock, M., Mehra, R., Benko, H., Wilson, A., Ofek, E., MacIntyre, B., Raghuvanshi, N. and Shapira, L., 2014. Roomalive: Magical experiences enabled by scalable, adaptive projector-camera units. In: *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology*, UIST '14, ACM, New York, NY, USA, pp. 637–644.
- Kakadiaris, I. A., Islam, M. M., Xie, T., Nikou, C. and Lumsden, A. B., 2016. iray: Mobile ar using structure sensor. In: *Proceedings of the 2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR-Adjunct)*, Merida, Mexico, pp. 127–128.
- Kato, H. and Billinghurst, M., 1999. Marker tracking and HMD calibration for a video-based augmented reality conferencing system. In: *Proceedings of the 2nd International Workshop on Augmented Reality (IWAR 99)*, San Francisco, CA, USA, pp. 85–94.
- Kern, J., 2016. Entwicklung eines Projektor-basierten Augmented-Reality-Systems. Master Thesis, Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology.
- Kilgus, T., Heim, E., Haase, S., Prüfer, S., Müller, M., Seitel, A., Fangerau, M., Wiebe, T., Iszatt, J., Schlemmer, H.-P., Hornegger, J., Yen, K. and Maier-Hein, L., 2014. Mobile markerless augmented reality and its application in forensic medicine. *International Journal of Computer Assisted Radiology and Surgery* 10(5), pp. 573–586.
- Klein, G. and Murray, D., 2009. Parallel tracking and mapping on a camera phone. In: *Proceedings of the IEEE International Symposium on Mixed and Augmented Reality - Arts, Media and Humanities. ISMAR-AMH*, Orlando, FL, USA, pp. 83–86.
- Lahanas, V., Loukas, C., Smailis, N. and Georgiou, E., 2014. A novel augmented reality simulator for skills assessment in minimal invasive surgery. *Surgical Endoscopy* 29(8), pp. 2224–2234.
- Microsoft, 2017. Microsoft HoloLens: a new way to see your world. <https://www.microsoft.com/microsoft-hololens/en-us/hardware>, visited in April 2017.
- Milgram, P. and Kishino, F., 1994. A taxonomy of mixed reality visual displays. *IEICE Transactions on Information Systems* E77-D(12), pp. 1321–1329.
- Nicolau, S., Soler, L., Mutter, D. and Marescaux, J., 2011. Augmented reality in laparoscopic surgical oncology. *Surgical Oncology* 20(3), pp. 189–201.
- OSG Community, 2016. OpenSceneGraph. <http://trac.openscenegraph.org/projects/osg/>, visited in April 2017.
- Otto, M., Prieur, M. and Rukzio, E., 2014. Using scalable, interactive floor projection for production planning scenario. In: *Proceedings of the Ninth ACM International Conference on Interactive Tabletops and Surfaces*, ITS '14, ACM, New York, NY, USA, pp. 363–368.
- Oyamada, Y. and Saito, H., 2008. *Defocus Blur Correcting Projector-Camera System*. Springer, Berlin, Heidelberg, pp. 453–464.
- Pantu Wong, N., Chutchomchuen, N. and Wacharawisoot, P., 2016. Interactive topography simulation sandbox for geography learning course. In: *Proceedings of the 2016 8th International Conference on Information Technology and Electrical Engineering (ICITEE)*, Yogyakarta, Indonesia, pp. 1–4.
- Reitmayr, G. and Drummond, T., 2006. Going out: Robust model-based tracking for outdoor augmented reality. In: *Proceedings of the International Symposium on Mixed and Augmented Reality (ISMAR)*, Santa Barbara, CA, USA.
- Resch, C., Keitler, P. and Klinker, G., 2016. Sticky projections – a model-based approach to interactive shader lamps tracking. *IEEE Transactions on Visualization and Computer Graphics* 22(3), pp. 1291–1301.
- Ridel, B., Reuter, P., Laviole, J., Mellado, N., Couture, N. and Granier, X., 2014. The Revealing Flashlight: Interactive Spatial Augmented Reality for Detail Exploration of Cultural Heritage Artifacts. *Journal on Computing and Cultural Heritage* 7(2), pp. 6:1–6:18.
- Sand, O., Büttner, S., Paelke, V. and Röcker, C., 2016. *smARt.Assembly – Projection-Based Augmented Reality for Supporting Assembly Workers*. Springer International Publishing, Cham, pp. 643–652.
- Schwerdtfeger, B., Reif, R., Günthner, W. A. and Klinker, G., 2011. Pick-by-vision: there is something to pick at the end of the augmented tunnel. *Virtual Reality* 15(2), pp. 213–223.
- Sutherland, I. E., 1968. A head-mounted three-dimensional display. In: *Proceedings of the AFIPS Fall Joint Computer Conference*, Vol. 33, pp. 757–764.
- Urban, S., Leitloff, J., Wursthorn, S. and Hinz, S., 2013. Self-localization of a multi-fisheye camera based augmented reality system in textureless 3d building models. In: *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Antalya, Turkey, Vol. II-3/W2, pp. 43–48.
- Urban, S., Wursthorn, S., Leitloff, J. and Hinz, S., 2016. MultiCol bundle adjustment: A generic method for pose estimation, simultaneous self-calibration and reconstruction for arbitrary multi-camera systems. *International Journal of Computer Vision* 121(2), pp. 234–252.
- Wursthorn, S., Coelho, A. H. and Staub, G., 2004. Applications for mixed reality. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Istanbul, Turkey, Vol. XXXV-B3, pp. 1049–1054.
- Yang, L., Normand, J. M. and Moreau, G., 2016. Practical and precise projector-camera calibration. In: *Proceedings of the 2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, Merida, Mexico, pp. 63–70.
- Zhang, Z., 2000. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22(11), pp. 1330–1334.