AIR: Anywhere Immersive Reality with User-Perspective Projection

JungHyun Byun, SeungHo Chae, YoonSik Yang and TackDon Han

Media System Lab., Department of Computer Science, Yonsei University, Korea, Republic of

Abstract
*Projection-based augmented reality (AR) has much potential, but is limited in that it requires burdensome installations and prone to geometric distortions on display surface. To overcome these limitations, we propose **AIR**. It can be carried and placed anywhere to project AR using pan/tilting motors, while providing the user with distortion-free projection of a correct 3D view.*

Categories and Subject Descriptors (according to ACM CCS): I.3.8 [Computer Graphics]: Computer Graphics—Applications
H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems—Artificial, augmented, and virtual realities

1. Introduction

Many augmented reality (AR) applications are built on head-worn or hand-held see-through displays, accompanied by issues of the limited field of view (FOV) and the user fatigue from bulky equipments [BR05]. Projection-based AR in contrast utilizes a projector to solve these issues. Located off-site, a projector directly augments graphics of virtually unlimited sizes and FOV onto physical objects [Ceb13] immersing users' surroundings, without diverting their attention from the real world [BWZ14]. However, these features come at a price [BEK05]. First, direct augmentation with projection is prone to geometric distortion. Second, a detached projector renders graphics that are inconsistent with the user's viewpoint.

The most crucial problem with projection AR is that the overall quality of AR relies heavily on the characteristics of the surface [RWF98]. A diffuse object with smooth geometry is required for an ideal projection, but that is rarely the case in projection AR [RWC*98]. Thus, the projection area is constrained to the size and shape of the objects' surfaces [BR05]. Projection on unideal surfaces with bumps and gaps would otherwise yield distortions.

Another problem in projection AR as well as most AR systems is inconsistency between the real scene and rendered graphics. Graphics are rendered on captured images from the perspective of the device, not the user. The former is called device-perspective rendering (DPR), and the latter user-perspective rendering (UPR) [TIS13]. DPR diminishes the sense of presence of virtual contents and this sense diminishes further as the camera-eye discrepancy, in terms of FOV, distortion, and viewpoint, increases. Because in projection AR, a projector is located independently of the user [BR05], the risk of diminished realism is much more severe.

To address these issues and realize immersive projection AR, we propose *Anywhere Immersive Reality (AIR)*, a novel projection system for any environment with user perspective rendering. The proposed system is a steerable platform with two cameras and a pro-

jector. It is designed to be carried and placed anywhere, rotated in any direction, and project contents even on uneven surfaces, while providing a correct 3D perspective view without distortions.

2. Related Work

2.1. Immersive Projection Display

In spatially immersive display (SID), users are surrounded by planar or curved screens, onto which multiple projectors display. Studies like [CNSD93], or [BJW12] implemented stationary versions of SID which require users to be at predesignated point for immersion. [BWZ14] and [JSM*14] extended this approach to a room-scale, non-planar geometry. Users can move around the room and be provided with correct 3D virtual views regardless of the surface geometry. However, since these SID approaches require extensive workloads of physical installations, off-line calibrations, and many high-cost hardware, they are not appropriate for the everyday user.

In this study, we also aim to realize an immersive projection system, but on an accessible budget. Our system is a pan-tilting platform, on which a projector-camera system for AR is built, with an additional camera unit for user tracking. Rotating the platform makes it possible to project virtually anywhere without the necessity of multiple units of costly hardware.

2.2. Portable Projection System

Most works on projection AR have predominantly focused on infrastructure-based spatial-aware systems, which infer the location and orientation of the device in 3D space based on the notion of one or more environmental sensors [RVBB*06]. Recently, the miniaturization of pico-projectors made it possible for systems to be portable, serving as an enabling technology for ubiquitous computing [Hub14]. For a system to be truly portable, it should be not only

spatially aware, but also geometrically aware, i.e., the system constructs the 3D structure of the world, and self-contained [MIK*12].

We also aim to design such a portable projection system. However, most of these systems are handheld, which can cause fatigue and limitation in interactions as users have to hold on to the system at all times. To solve these problems, the proposed system can be carried and placed, or held, and provide a seamless display.

2.3. Projection Distortion Correction

Many approaches have been proposed to overcome the effect of surface geometry on projection. In [RB01] and [XHH13], systems pre-warp images using the homography between a projector and its display surface, so that the projected image appears as it is desired. These systems, however, are limited as it can only correct oblique images on planar surfaces from a certain fixed device viewpoint.

Both [Pin01] and [WBIH12] utilize pan-tilting platforms for a rotatable viewpoint, but only the latter can adapt to dynamic environments and users, and correct complex distortions accordingly. However, pre-construction of the room geometry is required for the system, and localization of the user via sound demands even further installations and calibrations. Beyond these limitations, our system uses only real-time, raw depth data, and still project immersive, distortion-free AR content without prior knowledge of the space.

The main contributions of this paper are summarized as follows:

- We propose the mathematical modeling of such a system, from calibration to implementation.
- We propose the concept of a portable projection system, and its possible applications.
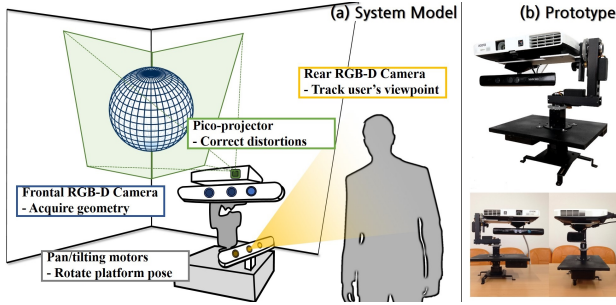
3. Proposed System

Figure 1: *Proposed (a) system model and (b) its prototype*

3.1. System Overview

Figure 1 describes the system model of *AIR*. The proposed system is based on a projector-camera system. It is held up by a platform, to which servo motors are attached vertically and horizontally, so that it can pan or tilt freely. We added a rear-facing RGB-D camera that serves as a dedicated user tracker for user-perspective rendering.

We built a prototype on a personal computer with an Intel i7-3770 CPU, 8GB RAM, and an NVIDIA GTX 960 GPU. The system consists of an Epson projector, a Microsoft Kinect v2, a Microsoft Kinect 360, servo motors, and an Arduino board for control.
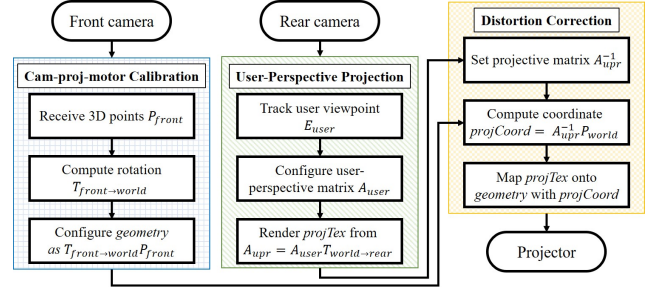
3.2. System Flow

Figure 2: *System workflow*

In this section, we present the mathematical model of *AIR*, from calibrations to implementations. We extend the algorithm for a projector-camera system in [RWC*98] to incorporate rotatable motors and an additional camera, and handle dynamic environments and users. The three core steps to realize user-perspective projection are i) camera-projector-motor calibration, ii) user-perspective projection, and iii) distortion correction.

3.2.1. Camera-projector-motor Calibration

AIR consists of two cameras, two motors, and one projector, each of which have their own coordinate systems. The first step is to configure them in one common *world* coordinate system. We setup the origin of the *world* system as the pose of front-facing camera facing front, with no rotation. For simplicity, we assume that internal parameters A_{front} and A_{rear} of the front and rear camera are known.

First, we model the movements of the pan/tilting motors. Because their rotational axes are not exactly perpendicular and parallel to the ground, we calculate the rotation matrix $R(\theta)$ of angle θ with an arbitrary, unit axis vector (x, y, z) as

$$\begin{bmatrix} \cos\theta + x^2(1-\cos\theta) & xy(1-\cos\theta) - z\sin\theta & xz(1-\cos\theta) + y\sin\theta \\ yx(1-\cos\theta) + z\sin\theta & \cos\theta + y^2(1-\cos\theta) & yz(1-\cos\theta) - x\sin\theta \\ zx(1-\cos\theta) - y\sin\theta & zy(1-\cos\theta) + x\sin\theta & \cos\theta + z^2(1-\cos\theta) \end{bmatrix}.$$

A 3D point P_{world} in *world* coordinates and its corresponding P_{front} perceived by the front camera can be modeled as

$$P_{world} = R(\beta)R(\alpha)P_{front}, \tag{1}$$

with panning angle α and tilting β. Matrix product $R(\beta)R(\alpha)$ can be thought of as external parameters of the front camera, denoted as $T_{front \to world}$. We use a checkerboard of a known size at a fixed pose and solve for x, y, z with least mean square method for two rotational axes while rotating the system about each of those axes.

The next step is the calibration of the rear-facing camera. The projector-camera system is rotated to a certain angle so that there is enough overlap between the front camera and rear camera for calibration. Using the coordinates of the checkerboard in equation 1 as the baseline, the relations among the *world*, front camera, and rear camera coordinate systems can be expressed as

$$P_{world} = T_{rear \to world}P_{rear} = R(\beta)R(\alpha)P_{front}. \tag{2}$$

Finally, the calibration between the front-facing RGB-D camera and the projector is performed. We used the calibration method
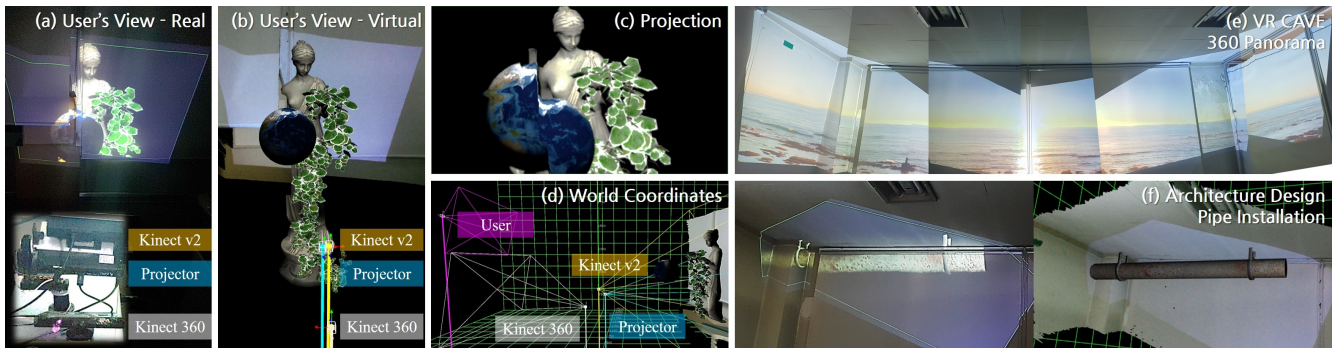
Figure 3: *(a) AIR provides the user with correct projection of (b) a 3D perspective view. (c) The projection is pre-warped to compensate for distortions. (d) Cameras and a projector are registered in world. Applications for AIR include (e) VR Cave and (f) architecture design.*

proposed in [KMK07]. A projector can be thought of as a reverse-camera model with the same parameters. The projector projects chessboards, x_{proj}, of various, pre-known sizes to different planar surfaces. Applying standard projective geometry, we obtain

$$sx_{proj} = A_{proj}T_{front \to proj}P_{front}. \quad (3)$$

Since depth data P_{front} are available, we can solve for the internal parameters of projector A_{proj} and external parameters $T_{front \to proj}$ of the rotations and translations with regard to the front camera.

3.2.2. User-Perspective Projection

Providing the user with a correct 3D perspective view and motion parallax is the key factor in realizing immersive AR. We setup the equation of user-perspective rendering [TIS13]. Instead of an actual tablet, here the screen is located in the virtual world. The system tracks poses of the user's head from the rear camera. The user's viewpoint is denoted as $E_{user} = [e_x, e_y, e_z]^T$, and its perspective projection matrix A_{user}. The corresponding equation is as follows:

$$A_{upr} = A_{user}T_{world \to rear}, \text{where } A_{user} = \begin{bmatrix} -e_z & 0 & e_x & 0 \\ 0 & -e_z & e_y & 0 \\ 0 & 0 & 1 & -e_z \end{bmatrix}. \quad (4)$$

3.2.3. Distortion Correction

Unlike mobile AR that uses a tablet screen, we use an off-site projector. This causes discrepancies in viewpoints, projection matrices and so on, which result in distorted projection. To correct these distortions, we implemented a two-pass rendering pipeline for projective texture mapping [Eve01], where we simulate the modeling of a calibrated projector-camera system. First, the system receives the geometry of the real world, and virtual objects are rendered from the user's perspective. Then, the rendered result is used as the texture to be mapped projectively onto the geometry. The final rendering is correct in terms of both user-perspective and geometric distortions. Figure 2 shows the overall process.

3.3. Applications

In this section, we describe some possible applications for *AIR*. Projection-based AR has a lot of potentials, as it can directly augment physical objects with graphics, promoting interactive cooper-

ation of multiple users. Figure 3 shows our early implementation of *AIR* including 360 panorama and pipe installation applications.

Previously to be immersed in 360 images, systems like VR-CAVE [CNSD93] were needed, which require large spaces, screens, and multiple projectors, or bulky head mounted displays (HMD), which restrict shared experiences. To ease these limitations, we developed a 360 panorama viewer (Figure 3(e)) on *AIR* with minimum hardware. It can project correct rotating views of a 360 image on surfaces and users can enjoy them together.

Another field in which *AIR* can be utilized is architecture design (Figure 3(f)). Previous HMD-based systems required off-line reconstruction of the room geometry and forced solo experience. In contrast, *AIR* system can be placed on site and project actual-scale virtual objects, such as a pipe. Clients and designers together can get a good sense of how it will turn out before installation.

4. Evaluation and Results

In [BJW12] to evaluate the effect of distortion correction, authors computed root mean square errors of absolute differences of corrected projections pixel-by-pixel. However, as stated in the paper, directly measuring pixel differences is susceptible to varying color conditions, while our only concern here is geometric distortion.

To measure errors without the influence of color conditions, we compute differences in structural compositions, point-by-point. First, we setup the base image by projecting a checkerboard on a planar surface. Then, we placed different obstacles before the surface and computed the correct projections. We matched corresponding corner points P_i of the checkerboards with those of the base image and computed the dislocations of the pixels using

$$\frac{1}{N} \sum_{i=1}^{N} |P_{base} - P_i|, \text{ where } N: \text{ the number of corners.} \quad (5)$$

Figure 4 shows results of the evaluation. Only projection regions of original 1920 by 1080 images are cropped. Numbers on the caption are average dislocation errors in pixel.

We tested our method on obstacles of various geometric attributes. Results show that our method corrects distortions in projection effectively even on (d), (e) complex geometry and (f) a
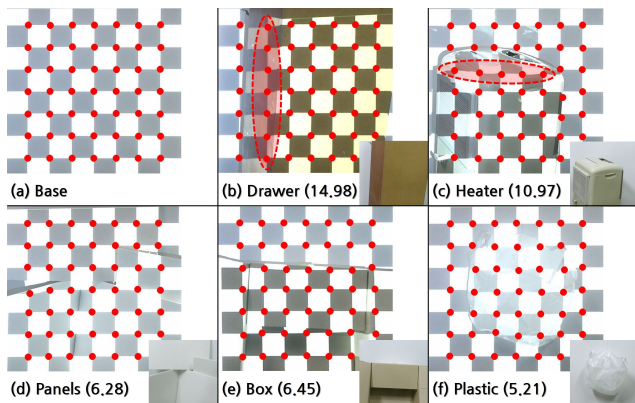
Figure 4: *Evaluation results of five obstacles with the base image. Corner points of the checkerboards are denoted in red dots.*

freely deformable object. Rigid (b) and curved geometry (c) caused the highest errors. This is an expected result, as objects with a flat angle to the camera lead to a complete lack of depth information. Specifically, for angles between $10°$ and $30°$ the Kinect delivers up to 100% invalid pixel [SLK15]. As one can see in (b) and (c), dislocations were greatest on slopes that are almost parallel to the depth camera (red dotted region), resulting in wrong depth value and wrong distortion correction accordingly. Since this is innate limitation of the Kinect, we leave this problem as our future work.

5. Conclusion and Future Work

We presented *AIR*, a steerable projection system for anywhere immersive reality. Our system can be carried, placed, and project anywhere without prior knowledge of the room geometry, and still deliver a correct 3D view of the virtual world to the user. We also implemented two applications in which the *AIR* system would be useful. An evaluation on projection distortion correction was conducted and the results confirm the usability of *AIR*.

We intend to further the research on *AIR* by integrating the system with natural user interactions so that users can be immersed in and interact with an AR space naturally and intuitively. We also intend to broaden and validate our search for application scenarios.

6. Acknowledgments

References

[BEK05] BIMBER O., EMMERLING A., KLEMMER T.: Embedded entertainment with smart projectors. *Computer 38*, 1 (2005), 48–55. 1

[BJW12] BENKO H., JOTA R., WILSON A.: Miragetable: freehand interaction on a projected augmented reality tabletop. In *Proceedings of the SIGCHI conference on human factors in computing systems* (2012), ACM, pp. 199–208. 1, 3

[BR05] BIMBER O., RASKAR R.: *Spatial augmented reality: merging real and virtual worlds*. CRC press, 2005. 1

[BWZ14] BENKO H., WILSON A. D., ZANNIER F.: Dyadic projected spatial augmented reality. In *Proceedings of the 27th annual ACM symposium on User interface software and technology* (2014), ACM, pp. 645–655. 1

[Ceb13] CEBULLA A.: Projection-based augmented reality. *ETH Zurich* (2013). 1

[CNSD93] CRUZ-NEIRA C., SANDIN D. J., DEFANTI T. A.: 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* (1993), ACM, pp. 135–142. 1, 3

[Eve01] EVERITT C.: Projective texture mapping. *White paper, NVidia Corporation 4* (2001). 3

[Hub14] HUBER J.: A research overview of mobile projected user interfaces. *Informatik-Spektrum 37*, 5 (2014), 464–473. 1

[JSM*14] JONES B., SODHI R., MURDOCK M., MEHRA R., BENKO H., WILSON A., OFEK E., MACINTYRE B., RAGHUVANSHI N., SHAPIRA L.: Roomalive: magical experiences enabled by scalable, adaptive projector-camera units. In *Proceedings of the 27th annual ACM symposium on User interface software and technology* (2014), ACM, pp. 637–644. 1

[KMK07] KIMURA M., MOCHIMARU M., KANADE T.: Projector calibration using arbitrary planes and calibrated camera. In *2007 IEEE Conference on Computer Vision and Pattern Recognition* (2007), IEEE, pp. 1–2. 3

[MIK*12] MOLYNEAUX D., IZADI S., KIM D., HILLIGES O., HODGES S., CAO X., BUTLER A., GELLERSEN H.: Interactive environment-aware handheld projectors for pervasive computing spaces. In *International Conference on Pervasive Computing* (2012), Springer, pp. 197–215. 1

[Pin01] PINHANEZ C.: The everywhere displays projector: A device to create ubiquitous graphical interfaces. In *International Conference on Ubiquitous Computing* (2001), Springer, pp. 315–331. 2

[RB01] RASKAR R., BEARDSLEY P.: A self-correcting projector. In *Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on* (2001), vol. 2, IEEE, pp. II–504. 2

[RVBB*06] RASKAR R., VAN BAAR J., BEARDSLEY P., WILLWACHER T., RAO S., FORLINES C.: ilamps: geometrically aware and self-configuring projectors. In *ACM SIGGRAPH 2006 Courses* (2006), ACM, p. 7. 1

[RWC*98] RASKAR R., WELCH G., CUTTS M., LAKE A., STESIN L., FUCHS H.: 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* (1998), ACM, pp. 179–188. 1, 2

[RWF98] RASKAR R., WELCH G., FUCHS H.: Spatially augmented reality. In *First IEEE Workshop on Augmented Reality (IWAR 98)* (1998), Citeseer, pp. 11–20. 1

[SLK15] SARBOLANDI H., LEFLOCH D., KOLB A.: Kinect range sensing: Structured-light versus time-of-flight kinect. *Computer Vision and Image Understanding 139* (2015), 1–20. 4

[TIS13] TOMIOKA M., IKEDA S., SATO K.: Approximated user-perspective rendering in tablet-based augmented reality. In *Mixed and Augmented Reality (ISMAR), 2013 IEEE International Symposium on* (2013), IEEE, pp. 21–28. 1, 3

[WBIH12] WILSON A., BENKO H., IZADI S., HILLIGES O.: Steerable augmented reality with the beamatron. In *Proceedings of the 25th annual ACM symposium on User interface software and technology* (2012), ACM, pp. 413–422. 2

[XHH13] XIAO R., HARRISON C., HUDSON S. E.: Worldkit: rapid and easy creation of ad-hoc interactive applications on everyday surfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2013), ACM, pp. 879–888. 2