A Spatial AR System for Wide-area Axis-aligned Metric Augmentation of Planar Scenes

Michael Horňáček*, Hans Küffner-McCauley, Majesa Trimmel, Patrick Rupprecht, Sebastian Schlund

Human Centered Cyber Physical Production and Assembly Systems, Institute for Management Sciences, TU Wien, Vienna, Austria

Abstract

Augmented reality (AR) promises to enable use cases in industrial settings that include the embedding of assembly instructions directly into the scene, potentially reducing or altogether obviating the need for workers to refer to instructions in paper form or on a screen. *Spatial* AR, in turn, is a form of AR whereby the augmentation of the scene is carried out using a projector, with the advantage of rendering the augmentation visible to all onlookers simultaneously without calling for each to wear some form of head-mounted display. Care must be taken in carrying out spatial AR, however, to warp the images to be projected in a manner that they appear free of distortions to the viewer, since the geometry of the scene as it relates to the geometry of the projector plays a role in how the pixels of the projector's image plane map to points in the scene. For planar scene geometry (such as a floor, wall, or table), this can be done in a cumbersome manual process called keystone correction, often using software bundled with the projector.

We propose a spatial AR system for wide-area metric augmentation of planar scene surfaces that produces the effect of keystone correction analytically, by using a projector equipped with a steerable mirror and a downwards-facing camera. Our system renders the placement of augmentations in the scene more

*Corresponding author
Email address: michael.hornacek@tuwien.ac.at (Michael Horňáček)

Preprint submitted to CIRP Journal of Manufacturing Science and Technology July 24, 2021

intuitive than manual keystone correction in two ways. First, (i) placement of the desired augmentations is carried out in accordance with the axes of an image of the scene acquired by a camera placed to face downwards towards the scene plane; second, (ii) the desired dimensions of the projected augmentations is specified in metric terms. Our system produces such augmentations accurately, with compelling time savings relative to the manual approach.

Keywords: Spatial augmented reality (SAR), computer vision, projector calibration, Industry 4.0, Pilotfabrik

---

1. Introduction

Augmented reality (AR) [1, 2] promises to enable use cases in industrial settings that include the embedding of assembly instructions directly into the scene [3, 4, 5, 6, 7], potentially reducing or altogether doing away with the need for workers to refer to instructions in paper form or on a screen. Typically, AR works by embedding the augmentation in an image of the scene acquired from the viewpoint of a single individual, with the resulting augmented image in turn displayed using some form of head-mounted display. Reliance on head-mounted displays, however, has two adverse consequences: (i) a head-mounted display must be worn by each individual wishing to partake in the augmentation, and (ii) such a head-mounted display—in some cases taking the form of a helmet in order to house multiple sensors in support of accurately tracking the viewpoint of the viewer relative to the scene—can be obtrusive. In turn, *spatial* AR[8] is a form of augmented reality carried out not by embedding the augmentation in an image of the scene as with a head-mounted display, but by projection to the scene itself, thus eliminating both aforementioned problems. Yet considering a planar surface to be augmented (e.g., a floor, wall, or table), unless the projector faces the surface frontally, the bounds of a projected rectangular image will not appear rectangular, but will instead appear distorted. Such distortions can be eliminated by carrying out a cumbersome manual process called keystone correction to appropriately warp the image to be projected, often using software bundled with the projector.

Our contribution is to propose a wide-area spatial AR system for planar scenes that produces the effect of keystone correction analytically, and—using the X- and Y-axes of an image of the scene acquired by a downwards-facing camera as a proxy—in a manner aligning the axes of the augmentation with those of the proxy image. Moreover, our system enables specifying the dimensions of augmentations in metric terms. We achieve this by warping the image to be projected using a plane-induced homography computed to produce the effect of projecting the image not from the actual projector viewpoint, but in accordance with the viewpoint of a *virtual* projector (i) facing directly downwards to the scene plane and (ii) rotated to place the axes of the image plane of the virtual projector in line with those of the camera. This facilitates placement of augmentations by intuitively placing them in accordance with the axes of an image of the scene, and eliminates the need for manual keystone correction. Finally, our system is able to handle a projector equipped with a steerable mirror (without need for explicitly modeling the action of the steerable mirror on the projector), thereby enabling wide-area factory floor applications exceeding the immediate field of view of the projector without needing to rely on multiple projectors.

1.1. Related Work

Keystone correction for planar scenes can be carried out using a 2D homography [9]—an invertible transformation that preserves colinearity, hence alternatively termed a 'colineation'—and is ultimately the approach we take as well. The intuition for why it is that a transformation that maps lines to lines can serve to model the appropriate image warp can be drawn from considering a planar chessboard pattern: looking at an image of a chessboard acquired from an oblique angle, one observes that lines parallel in the chessboard appear to meet in respective vanishing points; looking at an image of the same chessboard acquired frontally with respect to the plane of the chessboard, lines parallel in the chessboard appear parallel in the image (i.e., they are said to meet 'at infinity').

To warp the former image (where lines parallel in the scene meet in respective vanishing points) such that the lines of the chessboard appear as in the latter image (where lines parallel in the scene meet at infinity), a transformation that maps lines to lines would be sufficient.

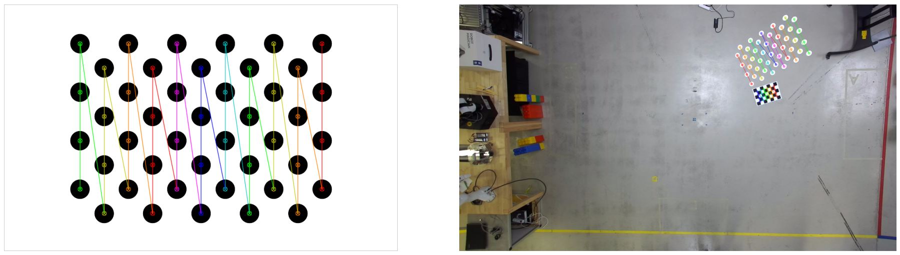
One way to compute a homography is by identifying at least four correspondences between pixel positions in two images of a planar surface. The keystone correction approach of Sukthankar *et al.* [10] reduces to using a segmentation approach to identify the four corners of projection screen in an image acquired by a camera, and computing a homography that maps the resulting four corners of the projection screen to the four corners of the projector's image plane. Tilt sensors can be used to recover the projector's gravity vector, and be integrated in the computation of a homography as in Raskar and Beardsley [11].

An early spatial AR system explicitly using a projector with a steerable mirror is the IBM Everywhere Displays prototype of Pinhanez [12]; the system, however, has keystone correction carried out manually. Rather than rely on a steerable mirror to support spatial AR to multiple locations using a single projector, some works mount the projector and a camera on a rigid rig and subject the rig to motion [13, 14, 15]. Steering only a mirror, however, places humbler requirements on the system from a hardware standpoint than if an entire camera-projector rig is to be subject to rigid motion.

Mention circles pattern: [16]

2. Approach

Correcting for projective distortions of the sort outlined in Section 1 can be achieved by modeling the manner in which the respective rays through the pixels of the projector's image plane fan out into the scene (i.e., by 'calibrating' the projector) and the geometry of the scene itself (i.e., by recovering the scene plane relative to the projector) within at least the projector's field of view. This is because the scene point 'illuminated' by a pixel in the projector's image plane can be recovered by intersecting its corresponding ray with the geometry of the



(a) Circles pattern image, in image plane of projector (detections overlain).

(b) Projector calibration image (one for each target location), in image plane of camera (detections overlain).

Figure 1: Recovering 2D positions in support of projector calibration. (a) 2D positions of the 2D-3D correspondences to be used for calibrating the projector are obtained by detecting—in the image plane of the projector—the circle centers in the circles pattern image, projected by the projector to each of the target locations in the scene plane. (b) For each such target location, an image is acquired from the viewpoint of the camera and the circle centers of the projected circles pattern are detected, in the image plane of the camera. A chessboard pattern to be used for recovering the local scene plane is placed near the projected pattern, whose corners are likewise detected. Detected 2D projected circles pattern center points and chessboard corners overlain for illustration.

scene surface. To model this interaction, we (i) carry out a one-time projector calibration, which in our approach calls for additionally calibrating a camera facing downward to the scene plane and includes recovery of the scene plane as a convenient side effect. Next, we use the relative camera-projector-scene plane geometry to (ii) compute a plane-induced homography that warps the image to be projected in a manner that it appear undistorted to the viewer, and placed in alignment with the axes of a proxy image of the scene. These two points are treated in Sections 2.1 and 2.2, respectively.

2.1. Recovering Geometry

Calibration of a projector (or camera) in the sense we employ the term here[1] renders one able to project a scene point $\mathbf{X} \in \mathbb{R}^3$ to its corresponding

[1]We are referring to a geometric calibration; not, e.g., to a color calibration.

pixel $\mathbf{x} \in \mathbb{R}^2$ in the projector's (or camera's) image plane, or to compute the 'back-projection' of $\mathbf{x}$, i.e., the ray from the projector's (or camera's) center of projection through $\mathbf{x}$ along which $\mathbf{X}$ must lie. Such a calibration can be expressed in terms of (i) a 3×3 calibration matrix $\mathtt{K}$ derived from the projector's (or camera's) focal length f and principal point $\mathbf{p}_0 = (x_0, y_0)^\top \in \mathbb{R}^2$ [9], and (ii) the coefficients of a lens distortion model used to correct for radial or tangential distortions caused by the lens system [17]; the calibration matrix $\mathtt{K}$ is then

$$\mathtt{K} = \begin{bmatrix} f & 0 & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{1}$$

where f and $(p_x, p_y)^\top$ are both expressed in units of pixels.

In practice, calibrating a camera relies on (i) establishing 2D-3D correspondences between pixels in the camera's image plane and corresponding points in the scene, and on (ii) using those correspondences as input to an optimization procedure that relies on bundle adjustment [18] to output the calibration matrix $\mathtt{K}$, the associated lens distortion model coefficients, and, for each calibration image, the pose (i.e., position and orientation) of the camera relative to the 3D points [9, 19]. To calibrate a camera, images of a calibration surface such as a chessboard pattern are used to identify the 2D-3D correspondences $\{\mathbf{x}_{i,j} \leftrightarrow \mathbf{X}_{i,j}\}$, $i \in \{1, \ldots, n_{\mathrm{pt}}\}, j \in \{1, \ldots, n_{\mathrm{im}}\}$, where n_{pt} gives the number of correspondences obtained from one calibration image of the pattern, n_{im} the number of such images, and $\mathbf{X}_{i,j} = \mathbf{X}_{i,j'}$, for $j, j' \in \{1, \ldots, n_{\mathrm{im}}\}$. Bundle adjustment is used to obtain the lens distortion model coefficients, calibration matrix $\mathtt{K}$, and rigid body transformations $(\mathtt{R}_j, \mathbf{t}_j) \in SE(3)$ that minimize the reprojection error

$$\sum_{i,j} d\big(\mathbf{x}_{i,j}, \mathtt{K}(\mathtt{R}_j \mathbf{X}_{i,j} + \mathbf{t}_j)\big), \tag{2}$$

where $\mathbf{X}'_{i,j} = \mathtt{R}_j \mathbf{X}_{i,j} + \mathbf{t}_j \in \mathbb{R}^3$ expresses $\mathbf{X}_{i,j}$ in the camera coordinate frame of the camera corresponding to the j^{th} calibration image, $(\mathbf{x}'_{i,j}, 1)^\top \sim \mathtt{K}\mathbf{X}'_{i,j} \in \mathbb{P}^2$ gives the projection $\mathbf{x}'_{i,j} \in \mathbb{R}^2$ of $\mathbf{X}'_{i,j}$ to that image, and d is a distance function that (i) corrects for lens distortion using the recovered lens distortion model coefficients [17] and (ii) returns a distance with respect to the two undistorted pixel positions. The inverse rigid body transformation $(\mathtt{R}_j^{-1}, -\mathtt{R}_j^{-1}\mathbf{t}_j)$ gives the pose of the j^{th} camera relative to the 3D points of the calibration pattern.

Calibration of a projector can be carried out in precisely the same manner as calibrating a camera insofar as step (ii) is concerned; the major difference in projector calibration relative to the calibrating a camera concerns the manner in which 2D-3D correspondences are identified, i.e., between pixels in the image plane of the projector and the corresponding points in the scene. What remains of this section is concerned primarily with the recovery of 2D-3D correspondences in support of calibrating cameras and projectors. A consequence of the approach we take to identifying the 3D points of the 2D-3D correspondences we use for projector calibration is, for each target location, recovery of the pose of the scene plane relative to the coordinate frame of the camera.

Camera calibration. We recover 2D-3D correspondences in support of calibrating the downwards-facing camera by relying on a planar calibration surface to automatically identify correspondences between the 3D points on the calibration surface and their 2D correspondences in the image plane. The classical calibration surface is a chessboard pattern. The 3D corner points of the chessboard are obtained *a priori* in a coordinate system defined in the plane of the chessboard[2], requiring knowledge only of the dimensions of the chessboard pattern and of the length of a side of a chessboard square. The corresponding 2D points are obtained, in the same order, using a specialized algorithm [20]. A set of calibration images is acquired, each with the calibration pattern visible in a different part of the image plane, and such that the center and all corners and edges of the image plane are covered, the camera's autofocus setting be off, and the camera's zoom factor remain fixed. 2D-3D correspondences are then recov-

[2] E.g., $(0,0,0),(1.5,0,0),(3,0,0),\ldots,(9,7.5,0)$ for a chessboard with 7×6 corners (8×7 squares), with each square of length and width of 1.5 unit, respectively. Note that the units of the chessboard's 3D points give the units of the camera calibration, and—owing to how our projector calibration relies on the camera calibration—of the projector calibration as well.

ered for each calibration image, and the resulting list is passed on as input to an optimization procedure that relies on bundle adjustment to yield the camera calibration matrix $\mathtt{K}_{\mathrm{cam}}$.

Projector calibration. As with camera calibration, projector calibration relies on 2D-3D correspondences, yet we obtain them in this case by *projecting* a calibration pattern. The pattern we project is one of circles (cf. Figure 1(a)), and we rely on an algorithm to detect the circle pattern center points in the circles pattern image in the image plane of the projector [20], giving the 2D positions of our 2D-3D correspondences for calibrating the projector. We project the circles pattern image to each of the target locations, and use the the calibrated downwards-facing camera to acquire a projector calibration image for each. Given a projector calibration image acquired using camera, we detect the circle centers of the *projected* circles pattern (cf. Figure 1(b)); given the scene plane (the recovery of which we shall return to in the paragraph that follows) and such a 2D circle center $\mathbf{x}$, we obtain its 3D correspondence by intersecting the back-projection $\mathtt{K}_{\mathrm{cam}}^{-1}(\mathbf{x}^\top, 1)^\top \in \mathbb{P}^2$ of $\mathbf{x}$ with the recovered scene plane (cf. Figure 2). Since the algorithm that yields 2D circle centers does so in a consistent ordering, we thus obtain the 2D-3D correspondences between the projector's image plane and the scene required for projector calibration, yielding the projector calibration matrix $\mathtt{K}_{\mathrm{proj}}$.

We recover the scene plane via spatial resection by applying a PnP algorithm [21] to the 2D-3D correspondences obtained using a chessboard pattern. Note that this step is separate from camera calibration, yet could well be carried out using the same calibration pattern used in the camera calibration step.[3] While a single image of such a chessboard pattern placed on the floor could be sufficient if the floor is even, we place a chessboard pattern in close proximity to the projected circles pattern in each projector calibration image in order to

[3] The critical point is that the pattern should ideally be coplanar with the local scene plane, meaning its height above the scene plane should not exceed a few millimeters.

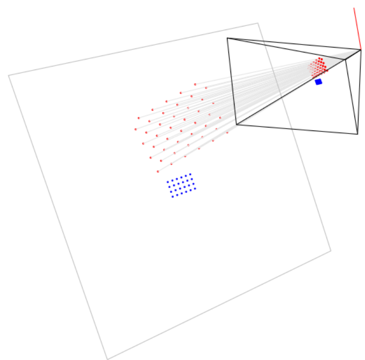
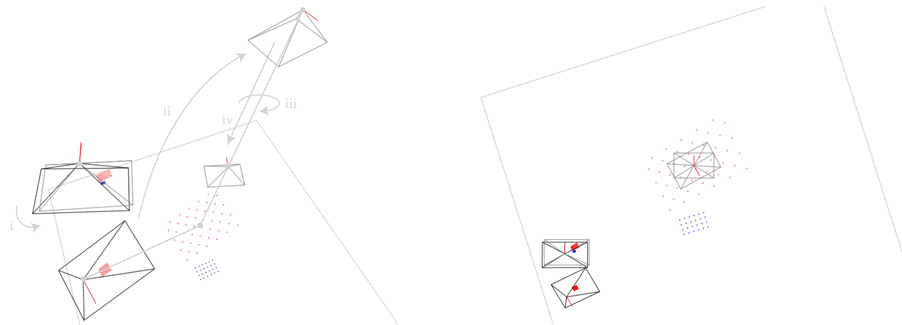


Figure 2: Scene plane (gray) recovered via spatial resection with respect to 2D-3D correspondences obtained using a chessboard pattern (blue); 3D circles pattern points—i.e., the 3D positions of the 2D-3D correspondences to be used for calibrating the projector—obtained by intersection with the scene plane of back-projections (likewise gray) of the 2D circles pattern center points detected in the image plane (red). Note that as in the figures that follow, the rendering in the figure corresponds to the projection calibration image in Figure 1(b), acquired by the downwards-facing camera (frustum of the camera in black, with up vector in red).

recover the scene plane locally to each target location, in order to account for the possibility of an uneven floor (cf. Figure 1(b)). Note that in principle, we could project a chessboard pattern instead of a circles pattern to obtain the 2D-3D correspondences needed for projector calibration; it is, however, because we rely on detecting the corners of a chessboard pattern to recover the scene plane that we opt instead for a alternative pattern.

The pose of the projector, for each projector calibration image, is provided relative to the 3D points of the pattern—and thus in the coordinate frame of the camera—alongside $\mathtt{K}_{\mathrm{proj}}$ by the aforementioned optimization procedure. Note that for a fixed projector with steerable mirror, given a projector calibration image, the recovered projector's pose is the pose the projector would have to have had to project to the given target location *in the absence of the mirror*. This is sufficient for our needs in Section 2.2, and it is in this sense that our system is able to handle a projector equipped with a steerable mirror, without need for modeling the steerable mirror explicitly.



(a) Oblique view.

(b) Downwards view.

Figure 3: The virtual camera is obtained by (i) rotating the camera (top left, black) about its center of projection such that its optical axis be made parallel with the normal vector of the scene plane. The virtual projector is obtained by (ii) rotating the projector (bottom left, black) about the point of intersection of its optical axis with the scene plane such that the optical axis be made parallel with the scene plane's normal vector, (iii) rotating the X- and Y-axes to align them with those of the virtual camera, and (iv) translating along the normal direction to achieve the desired metric projected image dimensions.

2.2. Correcting for Projective Distortion

If the projector is calibrated and its pose relative to the scene plane is known, a 'virtual' projector (with the same calibration $\mathtt{K}$ and lens distortion model coefficients) can be placed elsewhere relative to the scene plane. If we for a moment imagine that the projector—at its recovered pose—functions as a camera,[4] then (i) projecting an image to the scene plane *from the viewpoint of the virtual projector* and (ii) acquiring the resulting projected image from the viewpoint of the recovered projector gives the desired corrective warp. Projecting an image warped in this manner to the scene plane *from the viewpoint of the recovered projector* then has the same effect as projecting the original image to the scene plane from the viewpoint of the virtual projector. This warp can effected using

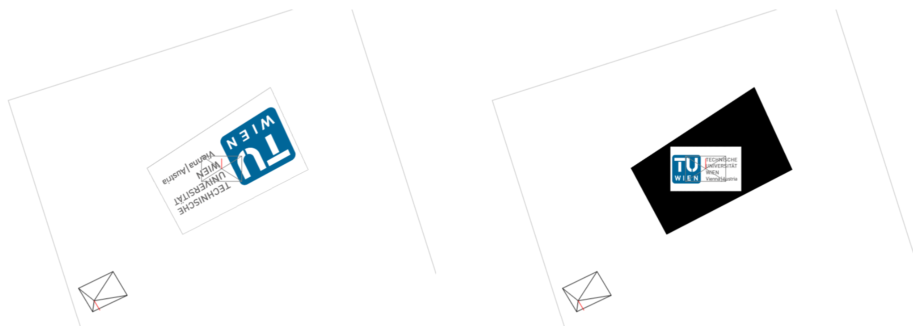
[4]Recall that the calibration matrix $\mathtt{K}$ enables computing both (i) the projection of a scene point to the image plane (the function of a camera), or (ii) the back-projection of a pixel in the image plane, giving a ray into the scene (along which a projector illuminates the scene with the given pixel).

a plane-induced homography, computed analytically as a function of the scene plane, the projector, and the virtual projector.

Virtual projector. The placement of the virtual projector determines from which pose the image to be projected is to *appear* to have been projected. This placement is produced according to a small handful of steps. First, we (i) rotate the camera about its center of projection to align its optical axis with the normal vector of the scene plane, giving a virtual camera likewise facing directly[5] downwards to the scene plane (cf. Figure 3). Next, we (ii) intersect the scene plane with the optical axis (i.e., the ray from the projector's center of projection through the center of the image plane) and rotate the projector's placement about that point of intersection, aligning the optical axis with the scene plane's normal vector and giving an initial virtual projector. Finally, we (iii) align the X- and Y-axes of the initial virtual projector with those of the virtual camera, which gives the virtual projector (cf. again Figure 3). The virtual projector is thus rendered fronto-parallel with the scene plane, enabling projection to the scene plane absent of projective distortions. We additionally (iv) adjust the height above the scene plane of the virtual projector, in order to satisfy desired projected image dimensions provided in metric units.

Owing to the manner in which we place the virtual projector, the virtual projector's axes and thus the augmentation are aligned with the axes of the downward-facing camera; the placement of the camera thereby intuitively determines the principal axes according to which augmentations are placed. Note further that a consequence of placing the virtual projector by rotating about the point of intersection of the projector's optical axis with the scene plane is that the center of the projector's image plane remains invariant to the placement of the virtual projector, i.e., a steerable mirror can be aimed with respect to a point projected from the center of the projector's image plane, further

[5]A physical camera placed to face downwards is almost certain to not face downwards precisely; in contrast, the virtual camera's optical axis is aligned exactly with the scene plane's normal vector, rendering it fronto-parallel with respect to the scene plane.



(a) Projecting original image. (b) Projecting our warped image.

Figure 4: Projection to the scene plane from the recovered projector viewpoint (bottom left, black; virtual projector in center, gray) of the original image and of the warped image. (a) Projecting the original image to the scene plane. (b) After warping the original image according to our plane-induced homography for the given target location, the image is projected in a manner that appears free of projective distortions, aligned with the axes of the virtual camera (via the virtual projector), and to have the desired dimensions in the scene plane, expressed in metric units. Note that background in the projected image is shown set to black.

facilitating placement of augmentations.

Plane-induced homography. Let $\mathtt{K}_{\mathrm{proj}}$ express the calibration matrix of the recovered projector and $(\mathtt{R}, \mathbf{t}) \in SE(3)$ the rigid body transformation that transforms points from the coordinate frame of the recovered projector to that of the virtual projector, for a given target location. Moreover, let $(\mathbf{n}^\top, -d)^\top$ give the scene plane, expressed in the coordinate frame of the recovered projector, where $\mathbf{n} \in \mathbb{R}^3$ is the scene plane's normal vector and $d = \mathbf{n}^\top \mathbf{X}$ for any point $\mathbf{X}$ in the plane, so that $(\mathbf{n}^\top, -d)(\mathbf{X}^\top, 1)^\top = 0$. The transformation that warps the image to be projected to the scene plane by the recovered projector such that it appear as if were projected to the scene plane by the virtual projector (cf. Figure 4(b)) is given the by the 3×3 matrix

$$\mathtt{H} = \mathtt{K}_{\mathrm{proj}} \left(\mathtt{R} - \frac{\mathbf{t}\mathbf{n}^\top}{d} \right) \mathtt{K}_{\mathrm{proj}}^{-1}, \tag{3}$$

a form of 'plane-induced' homography [9]. For convenience, we enable optional rotation of the image to be projected *before* applying $\mathtt{H}$, about the image center;

that rotation, parameterized in degrees, is thus in effect likewise carried out intuitively relative to the placement of the camera.

3. Evaluation

We evaluate our approach by augmenting 15 locations across the floorspace at the Pilotfabrik[6] of TU Wien, a collaborative space for research on Industry 4.0 topics situated in Vienna, Austria. We contrast our approach with a baseline approach involving manual keystone correction, by aiming with both approaches to place the same image for each location aligned with the principal axes of the floorspace, absent of projective distortions, and with the same metric dimensions (50 cm $\times$ 31.25 cm[7]). All experiments were carried out by the same technician, experienced in both approaches.

The hardware setup employed in the evaluation comprised a Panasonic PT-RZ660BE projector with a steerable mirror system—used in our experiments to point the projection to each of the 15 locations—manufactured by Dynamic Projection Institute [6, 7]. The steerable mirror system was bundled with the MDC-X software for steering the mirror, loading imagery, and optionally carrying out manual keystone correction, such that each position and (warped) image can be registered as a preset. In addition, we used a downward-facing Zed 2 stereo camera manufactured by Stereolabs, yet relied only on the left view. The floorspace used for our experiments measured dimensions of ca. 6 m $\times$ 4 m; the projector was mounted at approximately the center of this space, at a height of ca. 3.5 m.

Manual approach. ...

[6] `https://www.pilotfabrik.at/`

[7] The dimensions in pixels of the image we project are 960 $\times$ 600; we chose for our experiments to set the projected metric length of the horizontal axis of the image to 50 cm, which implies 31.25 cm for the vertical axis if aspect ratio is to be preserved.

Our approach. We began by carrying out a calibration of the camera, acquiring 10 camera calibration images (cf. Section 2.1) of a chessboard calibration pattern with 6×4 corners (7×5 squares) and feeding the images as input to our camera calibration module. Separately, for each of the 15 target locations, we produced a projector calibration image (cf. again Section 2.1) by projecting a 11×4 circles pattern image to the location in question using the steerable mirror, placing a checkerboard pattern beside the projected pattern, and acquiring the image using the downward-facing camera. We then fed these images alongside the output of the camera calibration module to our projector calibration module. For each of the target locations, the steerable mirror was made to point to that location, the circles pattern image was projected to the scene plane, an image was using the camera, and the location was registered in the MDC-X software as a preset. The output of the projector calibration module is a homography per input projector calibration image (cf. Section 2.2). Next, we warped the images to be projected to the respective locations using their corresponding homography, using a third dedicated custom module. These warped images were finally imported into the MDC-X software and associated with their respective location presets.

The total amount of time to carry out all the above steps amounted to ca. 20 min, with ca. 2 min going to acquisition of the camera calibration images, and ca. 5 min going to that of projector calibration images. The remainder of the time was spent running our modules or working with the MDC-X software. Note that once the camera is calibrated, that calibration can be reused if the camera's intrinsics remain fixed, in particular if no change is made to the zoom factor of the camera.

4. Conclusion

We presented a spatial AR system for planar scenes that produces the effect of keystone correction analytically, in a manner that enables intuitive placement the augmentations in accordance with the axes of an image of the scene acquired by a downwards-facing camera, and such that the desired dimensions of augmentations can be specified in metric terms. Moreover, we showed our system to be able to handle a projector equipped with a steerable mirror, enabling factory floor augmentation exceeding the bounds of the projector's own immediate field of view. Our evaluation demonstrated our approach to produce compelling results at less time than the more cumbersome traditional manual approach to keystone correction.

A natural extension of this work would be to address non-planar scenes. To handle non-planar scenes would call for a change in how scene geometry is recovered and how warping of the image to be projected is carried out; the methodology we proposed for projector calibration could, however, be left unchanged.

5. Acknowledgments

This work was supported by the Austrian Research Promotion Agency (FFG) through its endowed professorship in Human Centered Cyber Physical Production and Assembly Systems at TU Wien (FFG-852789).

References

[1] D. Van Krevelen, R. Poelman, A survey of augmented reality technologies, applications and limitations, International Journal of Virtual Reality 9 (2) (2010) 1–20.

[2] F. Zhou, H. B.-L. Duh, M. Billinghurst, Trends in augmented reality tracking, interaction and display: A review of ten years of ISMAR, in: 2008 7th IEEE/ACM International Symposium on Mixed and Augmented Reality, IEEE, 2008, pp. 193–202.

[3] S. Schlund, W. Mayrhofer, P. Rupprecht, Möglichkeiten der Gestaltung individualisierbarer Montagearbeitsplätze vor dem Hintergrund aktueller

technologischer Entwicklungen, Zeitschrift für Arbeitswissenschaft 72 (4) (2018) 276–286.

[4] T. Masood, J. Egger, Augmented reality in support of industry 4.0—implementation challenges and success factors, Robotics and Computer-Integrated Manufacturing 58 (2019) 181–195.

[5] M. Gattullo, G. W. Scurati, M. Fiorentino, A. E. Uva, F. Ferrise, M. Bordegoni, Towards augmented reality manuals for industry 4.0: A methodology, Robotics and Computer-Integrated Manufacturing 56 (2019) 276–286.

[6] P. Rupprecht, H. Kueffner-McCauley, S. Schlund, Information provision utilizing a dynamic projection system in industrial site assembly, Procedia CIRP 93 (2020) 1182–1187.

[7] P. Rupprecht, H. Kueffner-McCauley, M. Trimmel, S. Schlund, Adaptive spatial augmented reality for industrial site assembly, Procedia CIRP.

[8] O. Bimber, R. Raskar, Spatial Augmented Reality: Merging Real and Virtual Worlds, AK Peters/CRC Press, 2019.

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

[10] R. Sukthankar, R. G. Stockton, M. D. Mullin, Smarter presentations: Exploiting homography in camera-projector systems, in: Proceedings Eighth IEEE International Conference on Computer Vision (ICCV), Vol. 1, IEEE, 2001, pp. 247–253.

[11] R. Raskar, P. Beardsley, A self-correcting projector, in: Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 2, IEEE, 2001, pp. II–II.

[12] C. Pinhanez, The everywhere displays projector: A device to create ubiquitous graphical interfaces, in: International Conference on Ubiquitous Computing, Springer, 2001, pp. 315–331.

[13] J. Ehnes, K. Hirota, M. Hirose, Projected augmentation-augmented reality using rotatable video projectors, in: Third IEEE and ACM International Symposium on Mixed and Augmented Reality, IEEE, 2004, pp. 26–35.

[14] S. Borkowski, J. Letessier, J. L. Crowley, Spatial control of interactive surfaces in an augmented environment, in: IFIP International Conference on Engineering for Human-Computer Interaction, Springer, 2004, pp. 228–244.

[15] A. Butz, A. Kruger, Applying the peephole metaphor in a mixed-reality room, IEEE Computer Graphics and applications 26 (1) (2006) 56–63.

[16] P.-H. Chiu, S.-Y. Lin, L.-W. Chan, N.-H. Yu, Y.-P. Hung, Novel projector calibration approaches of multi-resolution display, in: 2011 IEEE International Conference on Multimedia and Expo, IEEE, 2011, pp. 1–6.

[17] D. C. Brown, Close-range camera calibration, Photogrammetric Engineering 37 (8) (1971) 855–866.

[18] B. Triggs, P. F. McLauchlan, R. I. Hartley, A. W. Fitzgibbon, Bundle adjustment—a modern synthesis, in: International Workshop on Vision Algorithms, Springer, 1999, pp. 298–372.

[19] Z. Zhang, A flexible new technique for camera calibration, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (11) (2000) 1330–1334.

[20] G. Bradski, The OpenCV library, Dr. Dobb's Journal of Software Tools 25 (2000) 120–125.

[21] T. Collins, A. Bartoli, Infinitesimal plane-based pose estimation, International Journal of Computer Vision 109 (3) (2014) 252–286.