

MARKER-BASED LOCALIZATION OF THE MICROSOFT HOLOLENS IN BUILDING MODELS

P. Hübner, M. Weinmann, S. Wursthorn

Institute of Photogrammetry and Remote Sensing, Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
(patrick.huebner, martin.weinmann, sven.wursthorn)@kit.edu

Commission I, WG I/7

KEY WORDS: Augmented Reality, HoloLens, Building Information Modeling, Localization, Visualization

ABSTRACT:

Mobile augmented reality devices for indoor environments like the Microsoft HoloLens hold potential for the in-situ visualization of building model data. While the HoloLens has sufficient real-time inside-out tracking capacity to provide a spatially correct and stable visualization of virtual content relative to its surroundings, the placement of virtual objects normally has to be done actively by the user. Beyond reliable tracking capacity, augmenting an indoor environment with corresponding building model data requires a one-time localization of the AR platform inside the local coordinate frame of the building model to be visualized. To this aim, we present a simple marker-based localization method for the HoloLens that is sufficient to overlay indoor environments with virtual room-scale model data with a spatial accuracy of few centimeters. Furthermore, an evaluation method suitable for the proposed scenario is presented, that does not rely on the HoloLens "Live Capture" camera which turned out to produce deviant placements of virtual content in relation to the perception of the user wearing the HoloLens device.

1. INTRODUCTION

Augmented reality (AR) allows for the spatially correct visualization of virtual content with respect to the surrounding environment. This opens the possibility of visualizing spatially related data directly in-situ on the location that the data refers to. One possible field of application in this context is the visualization of BIM (Building Information Modeling) data. Ren et al. (2016) exemplify various use-cases of mobile augmented reality in the context of BIM throughout the entire life cycle of a building project. Chu et al. (2018) examine the potential of AR-based in-situ usage of BIM data in terms of efficiency of on-site work and conclude that AR-based access to BIM data can improve the time efficiency of on-site work tasks by up to 50 %.

With the HoloLens (Microsoft, 2018a), Microsoft has released a mobile head-mounted AR platform that is capable of high-quality inside-out tracking and spatially stable visualization of virtual content (so-called "holograms") in indoor environments. (Liu et al., 2018) concluded in their technical evaluation of this AR device that it is generally capable of tracking with a position accuracy of few centimeters and sub-centimeter accuracy in the case of slow movements.

In the context of most existing HoloLens applications, individual virtual objects are placed in the environment of the user of the HoloLens device. This normally happens either by active manual placement of suchlike virtual objects by the users via gesture input or alternatively by automatic placement. Automatic placement in most cases is limited to placing virtual objects in open spaces on the floor, on tables or on walls in a way that aims to integrate them naturally in the surroundings of the user.

To use this augmented reality device for the in-situ visualization of BIM data, components of a building model like e.g. walls, doors or cables inside the walls can be represented as suchlike "holograms". In this case, the placement of those virtual room components has

to be done in a way, that ensures that the virtual components of the building accurately overlay their material pendants in the surrounding of the user.

Furthermore, in this context, the requirement for the accurate placement of virtual objects is much stricter than in use-cases where virtual objects are visualized that have no pendant in the material environment of the device, because aberrations from the intended position of a virtual object are clearly noticeable.

This can partly be overcome by the way the virtual content is visually represented. Instead of visualizing distinct wireframe structures of virtual objects overlaying their physical pendants, objects that are already physically present and visible in the surroundings of the user could be merely highlighted in a diffuse way. In this case, deviations between the pose of the physical object and its virtual representation would not be that noticeable. Physical objects that are not visible for the user like e.g. pipelines inside walls however should be rendered in a more distinct way that makes them clearly recognizable.

While the placement of a virtual building model to be visualized on-site can of course be done manually by the user via adjusting the position and orientation of the virtual content with e.g. gesture input, an automatic placement method that correlates the virtual building model with the real building surrounding the user is more desirable. To this aim, the pose of the HoloLens inside the respective building model corresponding to the surrounding indoor environment has to be determined.

Once the pose of the HoloLens with respect to this building model is known, holograms representing the BIM data can be placed correctly. As the HoloLens is capable of keeping the apparent position of virtual content spatially stable independent of the movement of its user (Vassallo et al., 2017), this localization of the HoloLens device inside the building model is a one-time process that does not have to be performed continuously in real-time.

In this paper, we show that such a one-time localization with respect to model data is sufficient for the spatially correct augmentation of a large room with virtual objects representing visible components (e.g. wall geometry, furniture and plug sockets) and non-visible components (e.g. pipelines and cables inside the walls) of the respective room.

After briefly summarizing related work, we present a simple marker-based localization method to this aim. Subsequently, we present a method for evaluating the spatial accuracy of the placement of virtual content achieved by means of this localization. Furthermore, we discuss the inapplicability of the HoloLens "Live Capture" camera for evaluation purposes concerning the placement accuracy of virtual content. After a detailed discussion of our acquired results, we provide concluding remarks and suggestions for further research.

2. RELATED WORK

3D models can be used for model-based estimation of camera poses. Reitmayr and Drummond (2006) used a textured 3D city model to provide a robust camera-based tracking solution. Algorithms like this need an initial pose of the camera within the model. In this outdoor case, location from global navigation satellite systems (GNSS) and orientation from an inertial measurement unit (IMU) is used. Côté et al. (2013) use just the video stream of a panoramic camera system to estimate its pose within a given model together with a point cloud from a terrestrial laser scanner to provide enough data for robust tracking. An initial pose has to be set up manually in order to start real-time tracking. This system requires powerful hardware that restricts its mobile use.

Kopsida and Brilakis (2016a) provide a survey of different markerless solutions for pose estimation with respect to a known 3D model on mobile devices with 2D images, with a monocular algorithm for simultaneous location and mapping (SLAM) like ORB-SLAM (Mur-Artal et al., 2015) or with a combination of RGB images and their corresponding depth images (RGBD). They claim the RGBD solution should be the most robust solution. Kopsida and Brilakis (2016b) use the dense point cloud from the RGBD camera to refine its pose estimation by aligning the point cloud to the BIM model with an iterative closest point (ICP) algorithm (Besl and McKay, 1992; Segal et al., 2009). Still, the initial registration of the point cloud and the BIM model has to be done manually since ICP will only be suitable if a good a priori alignment is available.

Breunig et al. (2017) introduce a collaborative platform for 3D multi-scale modeling and planning for subway track infrastructure projects where a helmet-mounted tracking system consisting of three fish-eye cameras is used for a SLAM-based self-localization solution within a building. The untextured 3D building model is used both for initialization of the SLAM algorithm with an absolute pose and for correcting the SLAM drift by comparing visible edges in the camera images with rendered edges from the 3D model (Urban et al., 2016). Synthetic camera views at many positions and orientations of the 3D model are generated off-line for successfully estimating the absolute pose within the model (Urban et al., 2013).

Fonnet et al. (2017) discuss the use of a HoloLens device for heritage BIM (hBIM) inspection. They propose a method to register a simplified BIM model with the mesh generated by the device's internal spatial mapping procedure with an ICP algorithm.

The mesh will be filtered for this purpose so that it will only contain basic 3D structures of the building like walls, floors and ceilings. Again, the user will need to manually set up an initial pose first.

Since the early days of video-based AR, markers have been used as a practicable solution for the registration of virtual content (Kato and Billinghurst, 1999). Reitmayr and Schmalstieg (2003) have shown how a marker-based solution can be used for mobile indoor applications in AR. They covered the building environment of interest with a sufficient number of coded markers in order to use them for indoor navigation. Wang et al. (2014) and Kwon et al. (2014) show the use of coded markers in the context of BIM. Patti et al. (2017) use QR codes as markers. These codes can be used both for identification and registration. Garrido-Jurado et al. (2014) give an overview over different kinds of coded markers in use. They define a highly reliable coded marker which is called "ArUco". Hübner et al. (2018) evaluate such ArUco markers for tracking of relative poses. Kern et al. (2017) use a pattern of ArUco markers for precise pose estimation of a projector-camera-system for AR.

Koch et al. (2012) use natural markers that are already available in the building like device IDs, exit signs or position marks of fire extinguishers.

3. METHODOLOGY

To achieve a spatially correct augmentation of an indoor scene with virtual building model content on the HoloLens, the spatial relationship between the environment in which the device is operating and its corresponding model has to be known. This comes down to determining the pose \mathbf{T}_{Model}^{App} of the building model in the coordinate frame of the HoloLens application.

In this paper, \mathbf{T}_A^B generally denotes the pose of an object A in a coordinate frame B and consists of a 4×4 matrix in the form

$$\mathbf{T}_A^B = \begin{pmatrix} \mathbf{R}_A^B & \mathbf{t}_A^B \\ \mathbf{0}^\top & 1 \end{pmatrix} \in \mathbb{R}^{4 \times 4} \quad (1)$$

Here, \mathbf{R}_A^B is an orthogonal 3×3 matrix, that describes the orientation of the respective object A in relation to the coordinate frame B while the vector \mathbf{t}_A^B describes its position.

In the following Section 3.1, we present a simple, marker-based method to the aim of determining this building model pose \mathbf{T}_{Model}^{App} . Subsequently, the practical realization of this localization method with room-scale model data of a laboratory room is described in Section 3.2. Finally, in Section 3.3, an evaluation method suitable for quantifying the placement accuracy of virtual building model content that can be achieved with the proposed localization method is presented.

3.1 Localization method

To the aim of overlaying indoor environments with virtual building model data as holograms, the pose of the HoloLens device with respect to the building model has to be determined. This equates to determining the pose \mathbf{T}_{Model}^{App} of the building model in the HoloLens application coordinate frame that leads to a correct alignment between the virtual and physical building geometry.

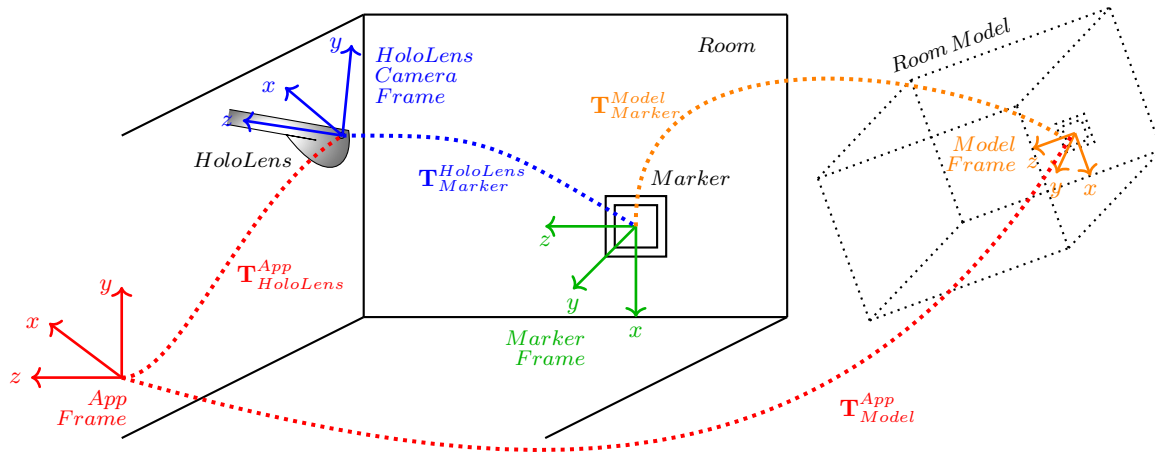


Figure 1. A room can be augmented with corresponding model data using a Microsoft HoloLens by determining the pose \mathbf{T}_{Model}^{App} of the virtual room model inside the HoloLens *App* coordinate frame. This pose can be determined via marker-based localization of the HoloLens camera. The pose $\mathbf{T}_{HoloLens}^{App}$ of the HoloLens device itself with respect to the *App* frame can be queried via the HoloLens SDK.

The HoloLens application coordinate frame *App* is the coordinate frame in which the pose of the HoloLens device as determined by its inside-out tracking system is described. The definition of the *App* coordinate frame is derived from the pose of the device in the moment of starting the respective HoloLens application.

The current pose $\mathbf{T}_{HoloLens}^{App}$ of the HoloLens device in this *App* frame can be queried by HoloLens applications via the HoloLens SDK. Furthermore, an arbitrary virtual object – a so-called "hologram" – can be positioned by setting its pose $\mathbf{T}_{Hologram}^{App}$ in this *App* frame via the SDK. In this manner, building model data can be positioned in the *App* frame as holograms if the correct pose \mathbf{T}_{Model}^{App} of the building model in the *App* frame can be determined.

An easy way to achieve this is by placing a marker in the building environment that is to be augmented with model data. The pose $\mathbf{T}_{Marker}^{Model}$ of this marker in the coordinate frame of the corresponding building model has to be known. If the pose of this marker can also be determined in the HoloLens *App* frame ($\mathbf{T}_{Marker}^{App}$), the model pose \mathbf{T}_{Model}^{App} results in

$$\mathbf{T}_{Model}^{App} = \mathbf{T}_{Marker}^{App} \mathbf{T}_{Marker}^{Model} = \mathbf{T}_{Marker}^{App} \mathbf{T}_{Marker}^{Model^{-1}} \quad (2)$$

The marker pose $\mathbf{T}_{Marker}^{App}$ in the *App* frame can be determined by observing the marker with the HoloLens "Live Capture" camera and thus determining its pose $\mathbf{T}_{Marker}^{HoloLens}$ in the local coordinate frame of this camera. The camera pose in the *App* frame in the moment of capturing the image of the marker can in turn be queried via the HoloLens SDK. The marker pose $\mathbf{T}_{Marker}^{App}$ in the *App* frame then results to:

$$\mathbf{T}_{Marker}^{App} = \mathbf{T}_{HoloLens}^{App} \mathbf{T}_{Marker}^{HoloLens} \quad (3)$$

A schematic overview of this marker-based method for localizing the HoloLens device inside building models corresponding to its surrounding indoor environment is depicted in Figure 1.

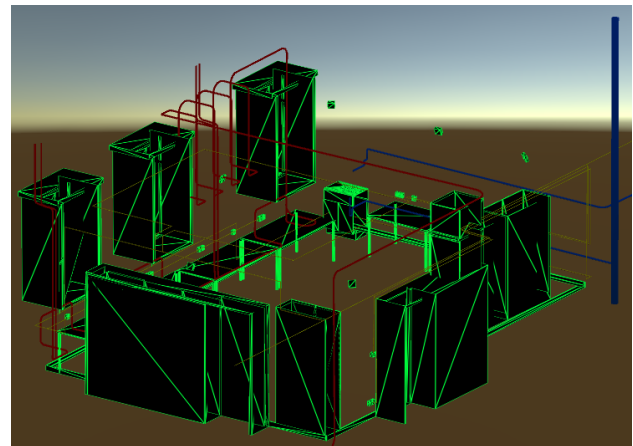


Figure 2. Room model: components like tables, cabinets, plug sockets and wall-mounted cameras are depicted in black and green wire frame, while infrastructure pipelines with fictive course inside the walls are depicted in red for heating pipes, blue for water pipelines and yellow for power supply lines.

3.2 Implementation

To demonstrate that the Microsoft HoloLens is capable of visualizing large room-scale virtual data with good spatial accuracy and stability, a simple marker-based localization method as described in Section 3.1 has been implemented.

For this purpose, a building model of a laboratory room of the dimensions of about 8 m × 5 m × 3 m was created. This room model includes the room geometry like walls, windows and doors as well as furniture and various infrastructure pipelines with fictive courses inside the walls. An overview of the room model is depicted in Figure 2. For reasons of clarity, the wall geometry has been omitted in this depiction.

The room which this model represents is equipped with various ArUco markers (Garrido-Jurado et al., 2014) of different sizes, as can be seen in Figures 3, 4 and 5. The local coordinate frame used for the creation of the room model equates to the local coordinate

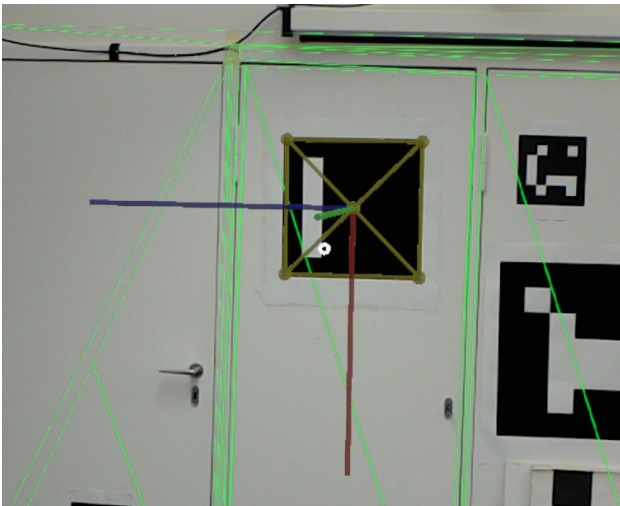


Figure 3. The ArUco marker used for localization augmented with a virtual representation of the marker and the coordinate frame of the room model.



Figure 4. The real room augmented by the room model from Figure 2 after localizing the HoloLens via the ArUco marker shown in Figure 3.

frame of one of those ArUco markers with corresponding points of origin.

Thus, the localization of the HoloLens device in the local coordinate frame of the room model can be achieved by capturing an image of the respective ArUco marker in the room with the HoloLens "Live Capture" camera and determining the pose of this camera relative to the marker. The current pose of the HoloLens camera in the coordinate frame in which the HoloLens tracks itself can be queried via the HoloLens SDK. In this way, the pose of the ArUco marker corresponding to the pose of the room model can be determined in the HoloLens frame which in turn allows for the correct placement of virtual objects representing room components.

Figure 3 for example shows the ArUco marker used for localization augmented with a virtual representation of this marker and its coordinate frame which corresponds to the coordinate frame of the room model. Figure 4 shows the room augmented with the model depicted in Figure 2 from the perspective of a camera placed directly in front of this localization marker looking into the room.

3.3 Evaluation scheme

For evaluating the spatial accuracy of the overlay between the real room and its virtual representation as visualized via the HoloLens

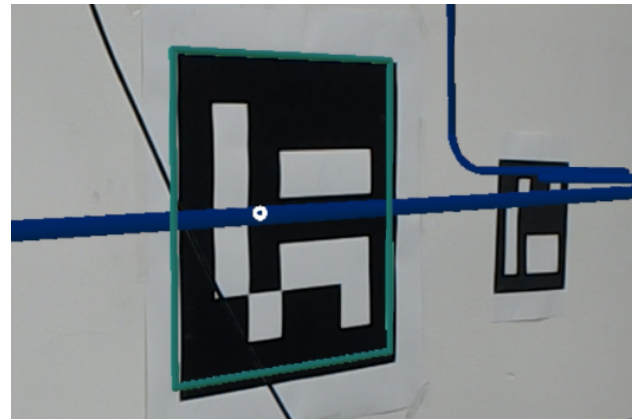


Figure 5. The ArUco marker used for evaluating the spatial accuracy of its overlay with the room model (blue rectangle).

device, another ArUco marker placed diametrically opposed to the one used for localization was augmented by a blue rectangle as part of the room model. This evaluation marker and its augmenting blue frame can be seen in Figure 5. The spatial accuracy of the placement of the virtual room model can be quantified by determining the spatial deviation between the corners of the virtual rectangle and the respective corners of the physical evaluation marker.

However, for evaluating the apparent accuracy of the overlay of this evaluation marker, it is not advisable to use the "Live Capture" camera of the HoloLens. We will show conclusively in Section 4.1 that the position of virtual content with respect to the real objects in the background as depicted via the camera can differ considerably from its apparent position as perceived by the user wearing the HoloLens.

Consequently, an evaluation procedure has to be applied, that directly evaluates the impression as perceived by the user of the HoloLens device. Vassallo et al. (2017) who evaluate the spatial stability of virtual content over time observed through the HoloLens achieve this by measuring the perceived position of the corners of a rectangular virtual object by contacting them with a stylus tracked with a motion capture system. This of course limits the possible evaluation distance of the virtual content to the range in which manual contacting can be performed by the user. In our case, the spatial accuracy of the placement of room-scale virtual content had to be evaluated. Thus, a contact-free measuring method was applied.

To this end, the corners of the evaluation marker and its virtual pendant were targeted with the laser pointer of a tachymeter and their cartesian coordinates in the local coordinate frame of the tachymeter were sequentially determined. A value for the spatial accuracy of the overlay between the physical and the virtual evaluation marker can be derived by averaging the Euclidean distances between the coordinates of their four corresponding corners.

In this way, deviations between the virtual and physical marker rectangle in the direction perpendicular to the wall which the physical marker is affixed to cannot be measured directly because the distance meter of the tachymeter always measures the distances to the wall surface regardless of the virtual marker being placed before or inside the wall. Thus, the projections of the corners of the virtual rectangle on the wall surface are actually measured and compared with the physical marker corners.

Large deviations of the placement of the virtual marker rectangle perpendicular to the wall surface contribute to the determined value of the placement accuracy, in so far as they cause deviations of the position of the virtual corners projected on the wall surface. Actually, this apparent position deviation on the wall surface causes the impression of the virtual rectangle being positioned in front or behind the wall surface in the perception of the user wearing the HoloLens device. Thus, measuring the position of the marker corners projected on the wall surface can be considered as justified for quantifying the placement accuracy of virtual objects as perceived by the user.

4. RESULTS AND DISCUSSION

When evaluating the apparent spatial placement accuracy of virtual objects that can be achieved with a head-worn AR device like the Microsoft HoloLens, it is of importance to take into account that the three-dimensional impression of a virtual object can essentially only be experienced by an operator using the respective device and its near-eye display.

Head-worn AR devices may provide means of capturing images which themselves include renderings of the virtual objects augmenting the scene as is the case with the HoloLens and its "Live Capture" camera. However, it is by no means certain that the apparent position of virtual objects depicted in suchlike images is in accordance with the augmented scene the user is experiencing while looking through the near-eye display of the respective head-mounted AR device.

In the case of the Microsoft HoloLens, we demonstrate in Section 4.1 that the apparent position of virtual objects in augmented images captured with the "Live Capture" camera can differ noticeably from their apparent position as experienced by the user wearing the device. Therefore it is imperative for an adequate and objective evaluation of the positioning accuracy of virtual objects to always measure the apparent position of a hologram directly as the user is experiencing it and not by means of an additional capturing mechanism like the HoloLens "Live Capture" camera.

Our evaluation procedure as presented in Section 3.3 takes this into account. The evaluation results derived by this evaluation scheme are presented in Section 4.2. Subsequently, we provide a detailed discussion of our findings in Section 4.3.

4.1 Observations on the HoloLens camera

Besides its tracking cameras, the Microsoft HoloLens is equipped with an additional camera, that allows for capturing images and videos that include visualizations of the holograms (i.e. virtual objects) the user wearing the device is seeing. The position of those virtual objects relative to the physical surroundings in the images of this camera however is not in every case identical to their position as it is perceived by the user looking through the see-through display of the device. The position of holograms overlaying physical objects can thus appear deviated in images captured with the HoloLens camera while the user wearing the device perceives them as correctly overlaying his surroundings.

This effect appears when the object to be overlaid with virtual content is not in the center of the image, i.e. when the cursor symbolizing the gaze direction of the user is not placed roughly on the object of interest. In those cases, when the user does not look directly at a hologram, but it is still within the field-of-view that can be augmented with virtual content by the HoloLens display,

the apparent position of the respective virtual object in the image deviates noticeably from its apparent position as observed by the user.

This offset between the content of the images captured with the HoloLens camera and the perception of the user wearing the device is demonstrated in Figure 6. In this experiment, a blue rectangle was placed as a hologram manually by the user in such a way that it exactly overlays the ArUco marker depicted in the images. Then the user was standing in front of the marker looking directly at it. While the user gradually turned his head to the left and respectively to the right, images of the marker and the virtual rectangle augmenting it were captured with the camera. Some of these images labeled with the respective angle of horizontal rotation of the device are depicted in the figure.

While it is clearly visible in Figure 6, that the virtual rectangle as depicted in the images increasingly shifts from the physical marker while the user is turning his head sideways, no such offset was perceived by the user himself through the see-through display of the HoloLens device. The user rather perceived an unvaryingly correct augmentation of the marker over the whole rotation range allowed by the field-of-view of the HoloLens display.

This aberration between the images captured with the HoloLens camera and the perception of the user has implications on the adequacy of this camera for various possible usage scenarios. Firstly, the spatial accuracy of the placement of virtual objects can appear worse in screenshots and videos captured with the HoloLens camera than it actually is. This can cause problems in the context of demonstration and documentation of HoloLens applications where the precise placement of virtual content is of importance. Furthermore, the spatial accuracy of the placement of virtual objects cannot be evaluated automatically via those images under these circumstances (e.g. by placing virtual markers next to physical ones and comparing their offset as determined from the images with the actual reference offset).

This shift of holograms in the images of the HoloLens camera normally is not that conspicuous because in most use-cases a hologram of interest is focused by the user (i.e. placed roughly in the center of the image) so that the offset between what the user sees through the HoloLens display and what the camera image shows is minor. Furthermore, holograms normally do not represent and overlay physically existing objects. Positional offsets of few centimeters are thus not that noticeable in prevalent use-cases.

In our use-case however, we intend to overlay physical room geometry with a corresponding virtual room model. In this scenario of overlaying physical objects with virtual content, even small deviations in the apparent position of virtual objects stand out noticeably. Furthermore, we deal with large-scale virtual objects that, even when directly focused, extend over the bounds of the region around the cursor where no offsets of the position of virtual content are visible in the images.

The inadequacy of spatially extended virtual objects for correct representation in the images captured with the HoloLens camera is demonstrated in Figure 7. Here, a blue virtual grid was placed manually over a spatially extended checkerboard pattern to achieve a precise overlay in the perception of the user wearing the HoloLens device. While offsets between the checkerboard pattern and its virtual pendant are clearly visible in this image, the user perceived no such deviations.

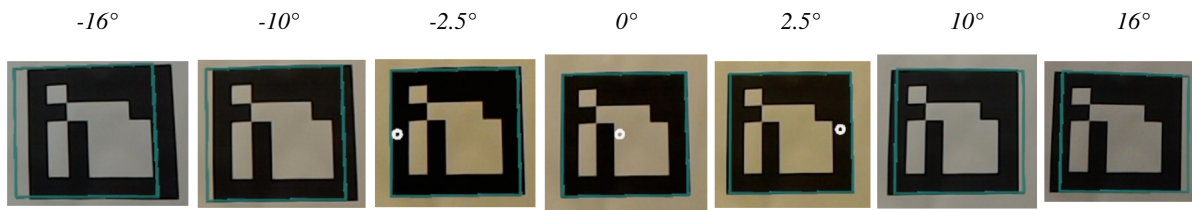


Figure 6. An ArUco marker is augmented with a manually positioned, blue virtual border. The images have been recorded with the HoloLens camera positioned in front of the marker while the camera is rotated subsequently to the left and to the right respectively. With increasing angle, an offset between the virtual rectangle and the physical marker is apparent that is not perceived by the user wearing the HoloLens device.

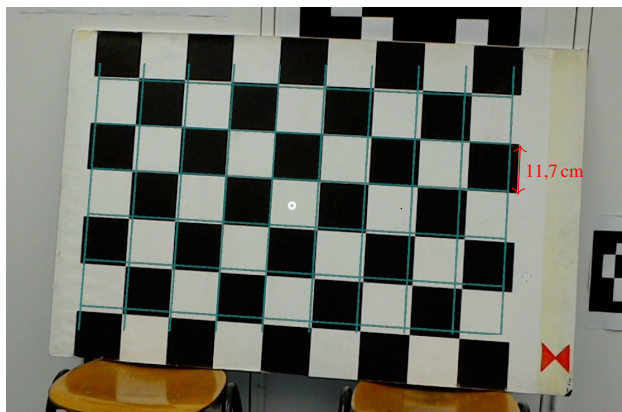


Figure 7. A checkerboard pattern augmented by its virtual pendant in blue. The increasing offset between the virtual and real pattern with distance from the position of the HoloLens cursor only appears in the images captured with the HoloLens "Live Capture" camera. The user wearing the device perceives an overall correct overlay.

4.2 Results

A quantitative evaluation of the augmentation of a room with building model data as described in Section 3.2 achieved by localizing the HoloLens device in the respective model frame with the marker-based method proposed in Section 3.1 was performed according to the evaluation procedure presented in Section 3.3.

In the course of doing so, the mean overlay error of the evaluation marker was measured from seven different points of view as depicted in Figure 8 after localizing the HoloLens device via the opposing localization marker. The points of view differ in the angle under which the evaluation marker is observed and their distance to the evaluation marker. This procedure of localizing the HoloLens and the subsequent measurement of the overlay accuracy from the seven respective points of view was repeated 15 times.

The results of this evaluation (see Table 1) show that the mean overlay accuracy achievable by placing the virtual room model automatically via marker-based localization of the HoloLens device amounts to 2.3 cm. While the mean overlay errors per localization averaged over all points of view cover a wide range of values from 1.4 to 4.1 cm, the overlay errors per view point averaged over all 15 localizations all fall in the narrow range between 2.0 and 2.5 cm.

		Point of View							∅
		1	2	3	4	5	6	7	
Localization	1	2.8	2.0	2.9	2.3	2.2	2.1	2.1	2.3
	2	1.9	1.7	1.9	1.3	1.2	1.3	1.4	1.5
	3	1.7	0.8	1.7	1.4	2.1	2.0	1.9	1.7
	4	1.1	0.6	0.9	1.8	3.2	3.2	3.0	2.0
	5	3.5	3.3	3.7	2.6	1.2	0.8	1.0	2.3
	6	3.3	2.7	3.4	3.9	4.8	4.9	5.1	4.0
	7	2.5	2.3	2.7	1.7	0.8	0.5	1.1	1.7
	8	0.2	0.3	0.4	1.1	2.6	3.0	2.4	1.4
	9	0.4	1.4	0.6	1.1	2.0	2.3	1.8	1.4
	10	3.2	3.1	3.2	2.7	2.9	2.2	2.5	3.5
	11	0.8	1.0	0.5	1.7	3.2	3.3	3.4	2.0
	12	3.7	4.0	4.0	3.7	4.2	4.3	4.8	4.1
	13	2.5	2.3	3.0	2.3	3.5	3.2	3.4	2.9
	14	1.4	1.8	1.7	0.7	1.9	2.0	1.5	1.6
	15	2.6	2.8	2.9	2.1	2.6	2.3	2.6	2.5
∅	2.1	2.0	2.2	2.0	2.5	2.5	2.5	2.3	

Table 1. Overlay error [cm] between the real and the virtual evaluation marker depicted in Figure 5 for all localizations and points of view (see Figure 8 for positions of the points of view).

4.3 Discussion

This paper demonstrates the general applicability of the Microsoft HoloLens device for the on-site visualization of building model data. First of all, it could be shown that large room-scale model data can be visualized as holograms without experiencing adverse effects like jittering or noticeably reduced rendering frame rate. Building model data in an on-site indoor usage scenario is mainly perceived on a per-room basis, i.e. in most cases only the data in the direct surrounding of the user, e.g. just for the room he currently is in, has to be visualized. So, the usage of large building-scale BIM data in this context should also be feasible, as it does not need to be visualized completely but mainly in room-scale parts around the current position of the user inside the building.

Furthermore, it could be demonstrated that correct, automatic on-site placement of building model data in indoor environments with sufficient spatial accuracy is feasible even with a rather simple marker-based approach. Of course, marker-based approaches are in practice not desirable in many usage scenarios, because they require physical preparation of the building environment via placement of markers. Fonnert et al. (2017) for example discuss a marker-based localization approach in the context of on-site AR visualization of cultural heritage BIM data (hBIM), but discard this method because it would not be feasible to equip a large number of different historical buildings with artificial markers.

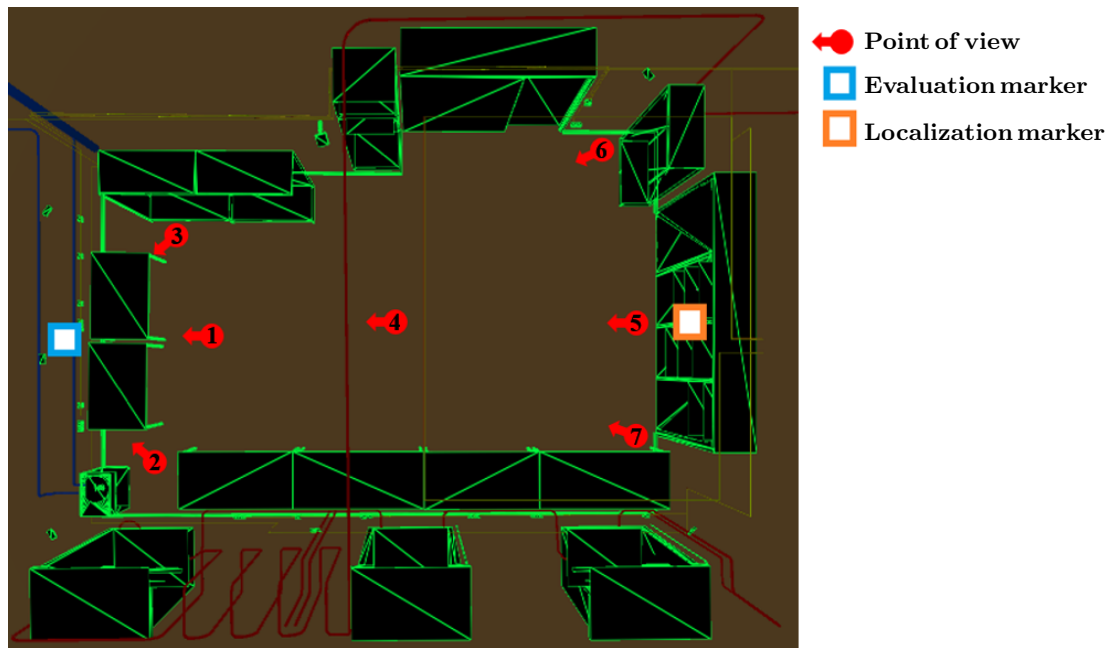


Figure 8. Placement of the seven different points of view that were used during the evaluation in relation to both ArUco markers.

Anyway, even if marker-based approaches for localization are not adequate for various scenarios, there certainly are use-cases, where markers can provide simple and efficient means for localization. In the context of historical buildings, a suchlike marker-based localization is suitable, if the number of buildings where the respective AR application should be used is not too large. For example, a single building can easily be equipped with an artificial marker in the entrance area for localizing the pose of the building model data with respect to the real world.

Other usage scenarios where a marker-based approach for augmented reality applications is feasible, concern environments, that are already equipped with marker-like planar signs that can be used for determining the location of a mobile AR client in a large-scale building model. In this context, e.g. the usage of door-plates as markers would be an option for applications in the field of facility management, with pre-known door-plate geometry providing means for pose estimation and room numbers on the door-plates providing for unambiguousness of location. Orlosky et al. (2014) for example use optical detection of door-plates in combination with magnetic tracking for localization in indoor emergency scenarios.

Irrespective of the way in which the localization is achieved – be it marker-based, via manual adjustment or via some sophisticated method of automatic localization like the approach proposed by Urban et al. (2013) – once the room model is localized as holograms in the environment of the user, it stays affixed over time independent of the position of the user inside the respective room (Vassallo et al., 2017). The evaluation results presented in Section 4.2 show little variation in the mean position error of the virtual room model content over the seven different points of view in the room on which the evaluation procedure was performed. The process of acquiring evaluation data from all seven points took about 30 minutes. During this time, no grave displacements of the apparent position of the holograms presenting the room model were experienced. It was also tested to leave the room through the door and enter it again. In doing so, the virtual room model still proofed to match the real room geometry after re-entering.

Ultimately, it can be concluded that the HoloLens as an AR device is principally usable for the on-site visualization of BIM data, if the correct pose of the model data corresponding to the physical building environment it represents can be determined. In contrast to tracking, this localization procedure does not have to be performed in real-time. It is rather sufficient to localize building model content just initially at the beginning or from time to time, e.g. when entering new rooms or if tracking is lost. Furthermore, it is possible to set so-called "Spatial Anchors" (Microsoft, 2018b) for the holograms representing the BIM data. This allows storing the poses of the respective holograms with respect to the spatial mapping data of the respective room they are placed in. Thus, when the HoloLens application is started a second time in a room, that contains "Spatial Anchors", the respective holograms can be loaded automatically and are visualized in the correct positions without having to perform a localization procedure with respect to the building model again.

5. CONCLUSION AND OUTLOOK

In the scope of this work, we demonstrated that the Microsoft HoloLens as a mobile augmented reality platform is suited for the spatially correct on-site visualization of building model data. Once the virtual building model content is placed to correctly overlay the physical structures of the indoor building environment it represents, the tracking capacity of the HoloLens is sufficient to keep its apparent position stable over time independent of the movements of the user wearing the device.

So, using the HoloLens as an AR client for the on-site inspection of BIM data comes down to a localization procedure for aligning the virtual building data with its physical pendants. In this paper, we presented a simple marker-based approach to this aim. Furthermore, an evaluation procedure was proposed, that measures the spatial precision of hologram placement directly as it is experienced by the user and does not rely on the HoloLens "Live Capture" camera that proofed to produce apparent positions of virtual content that differ from the impression the user is perceiving.

Although there exist use-cases suited for a marker-based localization approach in the context of AR-based in-situ presentation of BIM data, automatic localization methods that do not require the use of artificial markers or manual adjustment of the pose of a building model are a promising field of future research. In this context, the depth sensor and spatial mapping capability of the HoloLens could be made use of for detecting the part of a building model which the spatial mapping meshes best correspond to. Here, special emphasis has to be laid on how to deal with ambiguities inherent in certain building structures and how to remove those parts of the spatial mapping meshes that are not represented in corresponding building model data (e.g. furniture) and can disturb the correlation with building model structures.

References

- Besl, P. J. and McKay, N. D., 1992. A method for registration of 3-d shapes. *IEEE Transactions on Pattern Recognition and Machine Intelligence* 14(2), pp. 239–256.
- Breunig, M., Borrmann, A., Rank, E., Hinz, S., Kolbe, T., Schilcher, M., Mundani, R.-P., Jubierre, J. R., Flurl, M., Thomssen, A., Donaubaer, A., Ji, Y., Urban, S., Laun, S., Vilgertshofer, S., Willenborg, B., Menninghaus, M., Steuer, H., Wursthorn, S., Leitloff, J., Al-Doori, M. and Mazroobsemani, N., 2017. Collaborative multi-scale 3d city and infrastructure modeling and simulation. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLII-4/W4*, pp. 341–352.
- Chu, M., Matthews, J. and Love, P. E., 2018. Integrating Mobile Building Information Modelling and Augmented Reality Systems: An Experimental Study. *Automation in Construction* 85, pp. 305–316.
- Côté, S., Trudel, P., Desbiens, M.-A., Giguère, M. and Snyder, R., 2013. Live Mobile Panoramic High Accuracy Augmented Reality for Engineering and Construction. In: *Proceedings of the 13th International Conference on Construction Applications of Virtual Reality*, pp. 262–271.
- Fonnet, A., Alves, N., Sousa, N., Guevara, M. and Magalhães, L., 2017. Heritage BIM Integration With Mixed Reality for Building Preventive Maintenance. In: *Proceedings of the 24^o Encontro Português de Computação Gráfica e Interação (EPCGI)*, pp. 1–7.
- Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F. and Marín-Jiménez, M., 2014. Automatic Generation and Detection of Highly Reliable Fiducial Markers Under Occlusion. *Pattern Recognition* 47(6), pp. 2280–2292.
- Hübner, P., Weinmann, M., Hillemann, M., Jutzi, B. and Wursthorn, S., 2018. Combining independent visualization and tracking systems for augmented reality. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLII-2*, pp. 455–462.
- 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)*, San Francisco, USA, pp. 1–10.
- Kern, J., Weinmann, M. and Wursthorn, S., 2017. Projector-Based Augmented Reality for Quality Inspection of Scanned Objects. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 4(IV-2/W4), pp. 83–90.
- Koch, C., Neges, M., König, M. and Abramovici, M., 2012. BIM-based Augmented Reality for Facility Maintenance using Natural Markers. In: *Proceedings of the 2012 EG-ICE International Workshop on Intelligent Computing in Engineering*.
- Kopsida, M. and Brilakis, I., 2016a. BIM Registration Methods for Mobile Augmented Reality-Based Inspection. In: *Proceedings of the European Conference on Product and Process Modelling*, pp. 1–7.
- Kopsida, M. and Brilakis, I., 2016b. Markerless BIM Registration for Mobile Augmented Reality Based Inspection. In: *Proceedings of the International Conference on Smart Infrastructure and Construction*, pp. 1–6.
- Kwon, O.-S., Park, C.-S. and Lim, C.-R., 2014. A defect management system for reinforced concrete work utilizing BIM, image-matching and augmented reality. *Automation in Construction* 46, pp. 74–81.
- Liu, Y., Dong, H., Zhang, L. and Saddik, A. E., 2018. Technical Evaluation of HoloLens for Multimedia: A First Look. *IEEE Multimedia* 25(3), pp. 1–7.
- Microsoft, 2018a. Microsoft HoloLens. <https://www.microsoft.com/de-de/hololens>. last accessed: 07/2018.
- Microsoft, 2018b. Spatial Anchors. <https://docs.microsoft.com/en-us/windows/mixed-reality/spatial-anchors>. last accessed: 07/2018.
- Mur-Artal, R., Montiel, J. M. M. and Tardós, J. D., 2015. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. *IEEE Transactions on Robotics* 31(5), pp. 1147–1163.
- Orlosky, J., Toyama, T., Sonntag, D., Sarkany, A. and Lorincz, A., 2014. On-Body Multi-Input Indoor Localization for Dynamic Emergency Scenarios: Fusion of Magnetic Tracking and Optical Character Recognition with Mixed-Reality Display. In: *IEEE International Conference on Pervasive Computing and Communication Workshops (PERCOM WORKSHOPS)*, pp. 320–325.
- Patti, E., Mollame, A., Erba, D., Dalmasso, D., Osello, A., Macii, E. and Acquaviva, A., 2017. Combining Building Information Modelling and Ambient Data in Interactive Virtual and Augmented Reality Environments. *IT Professional*. in press.
- 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, pp. 1–10.
- Reitmayr, G. and Schmalstieg, D., 2003. Location based applications for mobile augmented reality. In: *Proceedings of the Fourth Australasian User Interface Conference on User Interfaces 2003 - Volume 18*, AUIC '03, Australian Computer Society, Inc., Darlinghurst, Australia, pp. 65–73.
- Ren, J., Liu, Y. and Ruan, Z., 2016. Architecture in an Age of Augmented Reality: Applications and Practices for Mobile Intelligence BIM-based AR in the Entire Lifecycle. In: *Proceedings of the International Conference on Electronic Information Technology and Intellectualization (ICEITI 2016)*, pp. 664–674.
- Segal, A., Haehnel, D. and Thrun, S., 2009. Generalized-icp. In: *Robotics: Science and Systems*, Vol. 2, pp. 1–8.
- 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. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 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.
- Vassallo, R., Rankin, A., Chen, E. C. S. and Peters, T. M., 2017. Hologram Stability Evaluation for Microsoft HoloLens. In: *Proceedings of SPIE 10136, Medical Imaging 2017: Image Perception, Observer Performance, and Technology Assessment*, pp. 1013614:1–6.
- Wang, X., Truijens, M., Hou, L., Wang, Y. and Zhou, Y., 2014. Integrating Augmented Reality with Building Information Modeling: Onsite Construction Process Controlling for Liquefied Natural Gas Industry. *Automation in Construction* 40, pp. 96–105.