



Interactive AR Models in Participation Processes

Jonas Hansert, Mathias Trefzger^(✉), and Thomas Schlegel

Institute of Ubiquitous Mobility Systems, Karlsruhe University of Applied Sciences, Moltkestr. 30, 76133 Karlsruhe, Germany
{iums, mathias.trefzger}@hs-karlsruhe.de

Abstract. It is getting more and more important to enable stakeholders from different backgrounds to collaborate efficiently on joint projects. Physical models provide a better understanding of spatial relationships while using video mapping of suitable visualizations enables a meaningful enrichment of information. We therefore developed a demonstrator using a physical architectural model as base and projected additional data via video mapping onto it. In this paper, we describe the initial situation and the requirements for the development of our demonstrator, its construction, the software developed for this purpose, including the calibration process as well as the implementation of tangible interaction as a means to control data and visualizations. In addition, we describe the whole user interface and lessons learned. Ultimately, we present a platform that encourages discussions and can enrich participation processes.

Keywords: Video mapping · Spatially augmented reality · Architectural model

1 Introduction

1.1 Initial Project and Background Information

The campus of Karlsruhe University of Applied Sciences is undergoing an architectural rearrangement. To support the planning phase for these changes, we developed a demonstrator that is able to show changes of the buildings as well as changes in the mobility to, from and on the campus. The ambitious overarching goal of the project was the establishment of a CO₂ neutral campus until 2030. An additional goal was enhancing the quality of stay and several other requirements and constraints were also in place. Due to the multitude of goals and responsibilities, many different participants and stakeholders were involved in the planning process. Those were civil engineers, architects, traffic engineers, computer scientists, students, employees, representatives of the city of Karlsruhe, the local transport and transportation sharing companies and many more.

Therefore, the field of participants was very heterogeneous. Because of that, we searched for a suitable possibility to bring these different disciplines together and to provide a basis for a joint discussion. Our approach was the construction of an interactive demonstrator using interactive spatially augmented reality. For the development of the demonstrator, we pursued a prototyping approach.

The demonstrator was used for meetings with stakeholders, presentations, explorative surveys and participation workshops. During the planning process, new ideas were constantly being added, which influenced the planning. In the course of the project the concept was adapted and further developed in several steps. Thus, an important requirement for the demonstrator was to be able to visualize new ideas and measures quickly and easily. In addition, it should also be ensured that as much information as possible can be understood and memorized by those involved. The demonstrator should be able to display spatial data, architectural data, data concerning the mobility on campus and visualizations of key figures.

Looking for a suitable medium for this purpose, we decided to develop a construction with an architectural model as centerpiece onto which information is projected with video mapping. Already the old Egyptians used physical models to design and communicate [1]. Physical models make it easier to understand and evaluate forms and at the same time present spatial relationships and proportions. According to Stanford Hohausser, architectural models are the most easily understood presentational technique [2].

Especially for people that are not familiar with a project it can be difficult to perceive a planning clearly. Architectural models can directly communicate ideas to the stakeholders and public and therefore facilitates understanding. At the same time, the campus reconstruction project does not only focus on architectural changes. Due to the goal to achieve a CO₂ neutral campus, mobility and energy information is also relevant. These data are related to buildings and other architectural aspects of the campus, but cannot easily be displayed on an architectural model on its own. The video mapping can augment the architectural model and integrate additional data with architectural information.

1.2 Requirements

In the course of half a year, several events, meetings, explorative surveys and participation workshops on campus were held with different participants. The demonstrator should promote discussion and encourage participation. After each meeting there were new ideas that were to be incorporated into the planning. For this reason, the demonstrator had the requirement that it should be possible to incorporate changes of the visualizations at short notice.

In cooperation with the partners a number of requirements were identified. Some of them were already defined at the beginning of the project, but some were only determined during the course of the project. The requirements are briefly summarized below:

- Different groups of people should be able to understand and memorize the planned measures as easily as possible
- The demonstrator should be suitable to be independently operated by people for self-information. (Later, the requirement was added, that it should additionally be usable for presentations.)
- Interaction with the model should be self-explanatory
- Discussions should be encouraged to strengthen the participation process
- It should be possible to change visualizations on the model quickly

- The demonstrator should attract attention and arouse interest
- The demonstrator should be transportable and able to be installed in rooms with a height of at least three meters

2 Related Work

Generally, Augmented Reality on interactive architecture models is generated by projections from above [3]. Depth, thermal, or infrared cameras provide real-time interaction possibilities with physical models [4].

Video mapping is an established technology that is used for entertaining, cultural [5] and somewhat less frequently for educational and planning purposes [6, 7]. A series of projects with Tangible-User Interfaces (TUIs) and Video mapping were developed for collaboration at the beginning of the millennium [3, 8] and have since been refined and implemented for a multitude of purposes. Mainly, video mapping is used to map pictures and video files onto real architecture.

Interactive projection can be used in augmented workplaces in production. For example, Korn et al. [9] shows, that assembly workers can understand manuals faster if they are projected directly on the worktop. They use hand detection for interaction with their system. A time consuming calibration is necessary.

Huber et al. automated the calibration process in the project LightBeam. It can project images on any planar object. It can be controlled with tangibles detected by a camera. The calibration works fast and without markers. Unfortunately, the accuracy is not fine enough for our use case [10].

Narazani et al. use 3D printed architectural prototypes with a conductive wireframe and combined them with touch gestures [11]. They use augmented reality on a mobile phone to show additional buildings and floors on existing buildings and use some common gestures for interaction.

3 Construction of the Demonstrator

3.1 Hardware

Basic components of the demonstrator are an Optoma 4K550st Short Throw beamer, an Intel RealSense D435 depth camera, the architectural model plus the remaining tabletop around the model. The beamer provides 4500 lumens and the camera provides 30 frames per second (FPS) at a resolution of 3840×2160 Pixels.

In addition, a scaffold can be attached around the table to which the beamer and the depth image camera can be attached (see Fig. 1).

3.2 Dimensions

The total table size of the demonstrator is 2.49 m length and 1.40 m width. Center piece of the demonstrator is an architectural model with 1,40 m length an 1,00 m width. The model was built in scale 1:500. The relationship between the height and

width as well as the orientation of the table ultimately depends on the architectural model. The horizontal alignment of the model, the beamer and limited room height influenced the entire design of the demonstrator. The dimensions of the table allow for up to 30 people to have a good view of the model.

The used beamer is able to cover the whole table with a distance of 200 cm from beamer lens to the table surface. The height of the table is 60 cm. This meets the requirement to fit into a room with a height of at least three meters. The height is chosen because both public buildings and apartments in old buildings usually have this minimum height.

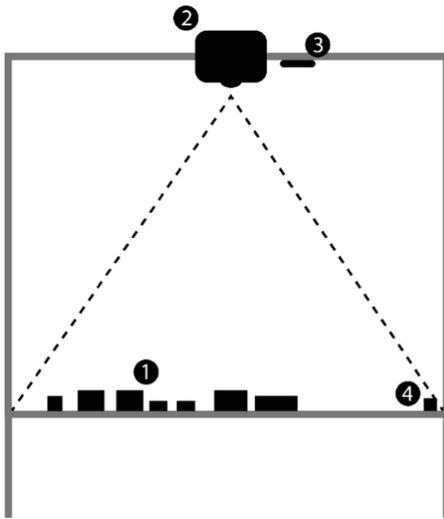


Fig. 1. Construction of the demonstrator. Left: The beamer (2) projects onto the architectural model (1) and rest of the table. The depth camera (3) recognizes the four markers in every edge of the table to calibrate the picture onto the table. It also recognizes the marker on the tangible (4). Right: The Demonstrator in construction.

4 Software and Interaction Technique

We used python for our implementation. The pyrealsense2 library was used to read the images from the RealSense depth camera. All computer vision tasks are solved by using the open source computer vision library OpenCV. For the detection of markers, we used the ArUco library.

For an exact projection of content on our table and detection of the position of tangibles, a correlation between the coordinate systems of the camera, projector and table is necessary. One of the issues of computing this affine transformation, is finding some similar points in the camera image, projector image and the real world. This can be done manual with a large amount of time. Because a new calibration is necessary after moving the construction as well as after the effects of any other vibration, manual calibration is not useful in our use case.

We solved this by ArUco markers [12] painted on predefined positions on our table and ArUco markers projected on predefined positions in the projector image on the table and by detecting their position in the camera image.

In order to use only a minimum amount of space on the table for the markers, we decided to print them as small as possible even if the detection of them is more difficult.

For calibration of the system and transformation of the content we used OpenCV, a common open source computer vision framework. In the following part, we want to give an overview of the algorithm we implemented for calibration.

The registration procedure starts with reading the position of the markers on the table and the markers in the projected image from a JSON based configuration file. The ArUco dictionary, the size of the markers and the marker IDs together with the position of their center are defined in this file too. The unit of the positions can be selected in the configuration file but has to remain the same throughout the project.

In the next step, one color frame is read. We convert it to a grayscale image and call the detection algorithm of the ArUco library. This algorithm returns a list of all detected markers. We filter this list for the markers defined in the configuration and repeat this steps until the four markers are detected.

To avoid inaccuracy if one corner is not detected exactly, we compute the mean value of all x and y positions for every marker. These are the centers of the markers.

Now we have four positions on the table and their associated position in the camera image. With that we can compute an affine transformation between the table and the camera and the inverse direction. This is represented as a 4×4 Matrix as a combination of a Matrix R and a translation vector T.

$$A \begin{pmatrix} x \\ 1 \end{pmatrix} = \begin{pmatrix} x' \\ 1 \end{pmatrix} \text{ with } A = \begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}$$

Now we have to get the transformation between the projector and the camera. To compute it, we project four markers to positions defined in the configuration file. These positions are predefined so that they are on the table and do not cover any parts of the architectural model. Then we use the algorithm described before to compute the affine transformations between the projector and the camera.

Only the affine transformation between the table and the projector and the inverse direction is missing at this point. This can be computed by a multiplication of the transformation matrices between the camera and the projector and between the table and the camera:

$$\begin{aligned} A_{\text{projector} \rightarrow \text{table}} &= A_{\text{camera} \rightarrow \text{projector}} \cdot A_{\text{table} \rightarrow \text{camera}} \\ A_{\text{table} \rightarrow \text{projector}} &= A_{\text{camera} \rightarrow \text{table}} \cdot A_{\text{projector} \rightarrow \text{camera}} \end{aligned}$$

With the transformation between the table and projector $A_{\text{table} \rightarrow \text{projector}}$ we transform all media in the startup process of our solution. Depending on the number of images, length of the videos and their resolution this may take several minutes. The results of the calibration process are saved and the calibration process can be skipped as long as the construction has not been moved.

```

{
  "path": "/home/katze/",
  "start_image": "start.png",
  "table": {
    "width": 24900,
    "height": 14000,
    "marker": {
      "size": 500,
      "marker": [0,1,2,3],
      "position": [[200,200], [200,200], [200,200], [200,200]]
    }
  },
  "projector": {
    "width": 3840,
    "height": 2160,
    "marker": {
      "size": 120,
      "marker": [5,6,7,8],
      "position": [[200,250], [300,300], [300,200], [400,200]]
    }
  },
  "areas": [
    {
      "name": "area1",
      "position_x": 1000,
      "position_y": 1000,
      "length": 20000,
      "width": 15000
    }
  ],
  "content": [
    {
      "name": "image1",
      "area": "area1",
      "url": ["image1.png", "image2.png"],
      "type": "slideshow",
      "duration": [5, 10],
      "conditions": [
        {
          "type": "marker",
          "id": 4,
          "position_x": 0,
          "position_y": 5000,
          "width": 1500,
          "height": 4000
        }
      ]
    }
  ]
}

```

Fig. 2. Listing 1: Example of a configuration file

On the top of our tangible we printed an ArUco marker. We use a separate thread for detection and all related computations for performance reasons. After marker detection we filter the list of detected markers for the used IDs. Then we use the transformation between the camera and the table $A_{camera \rightarrow table}$ for the computation of the position on the Table. The interaction that is triggered when the tangible is in a certain position is defined in the configuration file. As a next step, the interaction is therefore determined from this file.

We decided to use a configuration file in JSON format [Listing 1]. The path of all media is defined at first. It is still possible to use subfolders in this path. Then an image shown after startup and the position of the calibration markers are defined. We define multiple areas of the table in a sequence. They can be used to change the content on just a part of the table. In a sequence of content, every entry contains a name, an area, a type and a sequence of conditions. Depending on the type, additional properties can be necessary (Fig. 2).

The condition contains the type, in our example always marker, but it is possible to expand it and to add hand detection or object detection. Additionally, it contains an ID and a rectangular area on the table, where the object with the given ID has to be detected. If multiple conditions are defined, we combine them with an AND operation. An OR operation is possible by adding the same content with the second condition.

4.1 Integration of Visualizations

Visualizations that should be projected on the model and worktop around were created with common graphic programs. The image must have the same proportions as the real world model. A resolution of 10 pixel per mm can be chosen, for example. We created a template to simplify and speed up the creation process. This was done to meet the requirement that visualizations should be quickly realizable. The template contained a 2D plan of the architectural model and the dimensions of the table. In the template, the information could be visualized and finally be exported.

The images can then be used in the following steps to create animations. In the next step, the images and videos only have to be stored in the file folder of the program and the file names have to be written in the JSON file.

After each insertion of new images, the transformations of the images and videos must be recomputed. For this purpose, the variable “-calc-transformation” must be passed to the program. After the calibration run, the projection starts.

5 Information Presentation and Visualization

One major challenge concerning the demonstrator was the balancing act of not showing too much information at once to the spectators while giving them a good insight into relevant planning data. Also the visualizations should be designed in a way that the most important information draws the attention first. As shown in Fig. 3 the proposed and planned measures on the campus were visualized using icons on the architectural model itself. Because of the limited space directly on the model, which would have had a huge impact on readability, we included the frame around the model to display

information. The space around the model was used to give additional information about measures located in the model. Furthermore, we halved the space below the model. On the left half, we showed pictures of the modifications of campus buildings. The right half was mainly used to show graphs with the impact of the planned measures on CO₂ developments, the modal split and the energy mix. The timeline below this space provided a clear structure of the planned measures in each phase. The following explains the structure and design of the user interface.

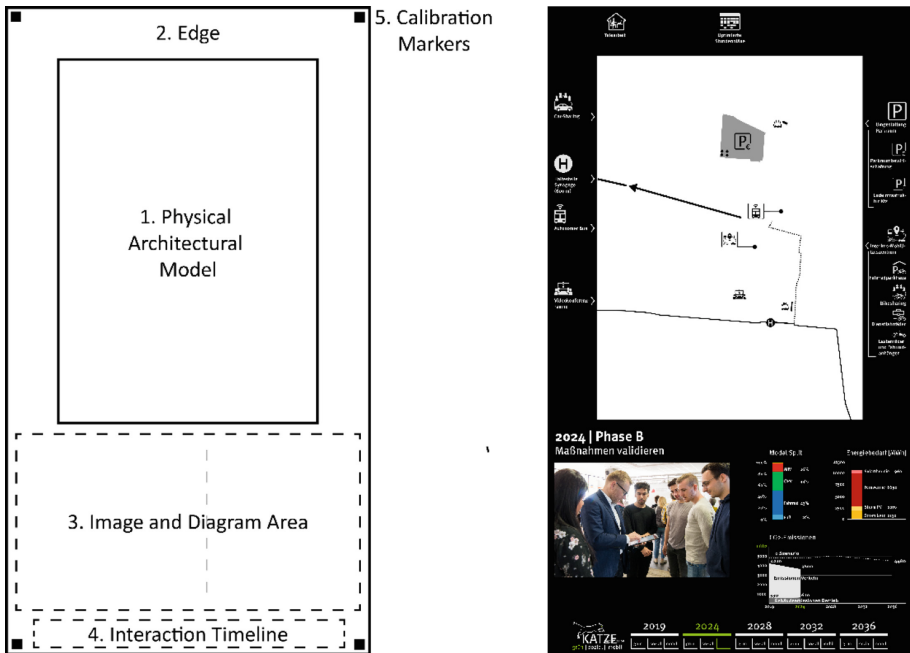


Fig. 3. Layout of the user interface. Left: Schematic representations of the divided areas. Right: Example of a displayed slide.

5.1 Architectural Model

The physical model is the centerpiece of the demonstrator. Some buildings can be exchanged to show the planned structural changes. A striking feature of the model is the characteristic forest around the campus, which is permanently implemented on the model. On the one hand, this ensured a high recognizability of the model. On the other hand, this limited the possibilities of visualizations. Icons were used to locate and represent the results of planned measures.

Areas (e.g. parking lots, solar plants) and paths (routes to the campus, bus route) were also used for visualization. Icons with a size of 4 by 4 cm turned out to be easily recognizable. Due to the relatively close proximity of the buildings to each other, it was not possible to use a font with a clearly legible size directly on the model. For this reason, we used the frame around the model for any written information.

5.2 Frame

The frame was used to show additional information to the displayed information in the model. It serves as a summary of the measures for the viewers due to its bullet word based structure. The respective measures are displayed iteratively, triggered by interaction and highlighted in green to draw attention to the specific issues to be discussed in this step. In addition, the icons from the architecture model are also displayed on the margin in order to support the reception of information for the participants. An arrow indicates the connection between the measures displayed in the model and the labels on the border. A connection with a dash was omitted due to the large proportion of forest area around the campus, which is not a good projection surface. In addition, many connecting lines quickly create a restless and muddled presentation.

5.3 Image and Diagram Area

The left half area of the image and diagram area is used for the projection of images related to the respective measures. Pictures are used as eye catchers and support the message of the presented measures and diagrams. The right half is mainly used for showing diagrams of the modal split, energy mix and CO₂ emissions. In the starting phase of our project, the image and diagram area were strictly separated. During the first presentations we saw that the information provided overstrained many participants in the planning process. We therefore decided to show diagrams only on as a summary.

5.4 Interaction Timeline

The timeline enables interaction and allows a user to navigate through different years of campus development. Shown five phases, starting 2019 until 2036. Each phase is divided into three categories: green, social and mobility measures (see Fig. 4). By moving the tangible onto the respective measures, the respective images, figures and diagrams are shown. The selected field is always highlighted in order to give feedback to the spectators. Operation via a timeline was chosen because it is self-explanatory: people who have not yet been involved in the project can easily also discover the phases. Furthermore, the interaction using a tangible is also suitable for presentations in front of an audience. It can be used to move through the timeline and each category successively. The placement of the timeline determines the operating side and viewing direction.



Fig. 4. Interactive Timeline that can be controlled via a tangible.

5.5 Coloring

Initially the project partners set the requirement to only use grey tones, black and white for the complete design. This turned out not to be appropriate, as the visualizations and fonts were hardly visible on the physical model. So it was necessary to use bright colors in the model. We then tested two color schemes: one with dark font on a white background and an inverted color scheme. The inverted version turned out to be a much more popular variant and could offer further advantages. Of 18 people interviewed 16 preferred the inverted variant and two were undecided. The inverted version also has other advantages: The black border around the model highlights the architectural model. In addition, the black border ensures an invisible transition beyond the edge of the table.

5.6 Font Sizes

During the development process, different font sizes were tested and defined. In the model itself, the use of lettering did not make sense, because the existing areas on the model are too small, so that no sufficiently large lettering is possible. Only the margin is therefore used to provide textual information. The required font size was tested with ten people. The people stood at the bottom edge of the table and gave feedback which font size was legible for them. The tested writing was thus between 2.09 and 2.29 meters away from the test persons. We asked whether the persons could read a word with the font size. And afterwards we asked which font size they preferred. Eight people stated that they could read the font with 46 pt, but preferred 54 pt. Two people preferred a larger font size. Due to the limited edge and the possibility to walk around the table, a font size of 54 pt. was finally chosen.

6 Evaluation

We conducted a study with eight participants (3 female, 5 male) to evaluate the interface and to investigate, how much information could be memorized. For this purpose, we showed the participants the different visualizations of the mobility measures in phase 2024. This phase was chosen because it contained the most information of all phases. While the visualizations on the demonstrator were running, a previously recorded audio file was played, to simulate an oral presentation. After the five-minute presentation, the test persons had to fill out a questionnaire. We used six questions to check, how many facts were remembered by the participants. Those were:

- Which means of transport will be the most used to the campus?
- Which means of transport will be the least used to the campus?
- In which way will the university cover most of its energy needs?
- What is the main source of the university's CO₂ emissions?
- By what percentage is the CO₂ requirement of the university to be reduced from 2019 to 2024?
- Please tick which measures have been planned in this phase to reduce the CO₂ requirements of the university.

Of those six, the first four answers were not contained in the audio information and could only be perceived by looking at the diagrams (see Fig. 3).

The result was that the test persons could answer these questions very poorly or not at all. To evaluate if the participants were able to capture the information shown on the three diagrams. They were asked about the most and least used means of transport (3 out of eight correct answers and no correct answer). The question of how the university can cover the largest part of its energy requirements could not be answered either. Only the information that most of the university's emissions are produced by transport could be answered correctly by seven of the respondents.

Better results could be achieved when asked about the planned measures in this phase. Here the test persons were able to classify 67% of a list of 18 proposed measures as correct or incorrectly correct. At this point it was noticeable that six of the test persons had always left three of the measures presented in the presentation unchecked. These measures were sub-items of the measure "Mobility Centre". This leads to the conclusion that an enumeration of such sub-items was not observed. We are convinced that this information would have been memorized better without the use of sub-items.

Furthermore, participants were asked to rate the clarity of the interface. On a scale between clear (100%) and unclear (0%) the interface was rated as rather clear (59%). The test persons generally liked the demonstrator (85% - good (100%), bad (0%)). Four people would prefer a presentation with the demonstrator to a PowerPoint presentation, two were undecided and two would prefer a PowerPoint presentation.

The results show that there was a tendency to display too much information. The test persons were able to remember a large part of the information on the first visualizations, where no diagrams were displayed. In contrast, the last visualization that used three different diagrams contained too much information. For this reason, a maximum of one diagram should be displayed at a time.

The evaluated variant is therefore not ideally suited in this form for presentations where each visualization is only displayed for a certain period of time. In order to be able to capture all the information presented, it is necessary that users can go through the information at their own pace. This is also possible through interaction with the timeline and was the primary use case at the beginning of the project.

7 Conclusion and Discussion

In this paper, we presented a demonstrator utilizing interactive spatially augmented reality that can be used in planning processes. It was designed so that the visualizations are easily and quickly adaptable and to support an iterative planning process over several workshops. The developed user interface is appropriate for the use case that participants can navigate through the different project phases in their own pace. For presentations with a tight time frame, the interface has a novelty effect, but visualizations should be chosen carefully, if viewers are supposed to memorize information. The demonstrator as a medium of participation was well received by stakeholders and other participants. Future phases and scenarios could be prepared and presented in an understandable way with simple visualizations: The relationships of location and time information were clearly worked out by many spectators and project participants. The involvement of the demonstrator strengthened the interdisciplinary participation process and helped to develop new ideas. For this reason, the use of similar setups in comparable projects can be recommended.

Acknowledgements. This work was conducted within the scope of two research projects. The first is “KATZE” as part of the idea competition “Mobility concepts for the emission-free campus” and was funded by the German Federal Ministry of Education and Research. The second is “View-BW – Visualization of the energy transformation Baden-Württemberg” that was funded by the German Federal Ministry of Environment (Funding ID: BWED19004). We like to thank Prof. Robert Pawlowski, Prof. Jan Riel, Prof. Jochen Eckart, Prof. Susanne Dürr, Jonas Fehrenbach, Isabelle Ginter and Lena Christ for their contribution to this work. We would also like to thank our stakeholders, the students and all other participants for their good cooperation over the course of the project.

References

1. Smith, A.: *Architectural Model as Machine – A New View of Models From Antiquity to the Present Day*. Elsevier, Oxford (2004)
2. Hohauser, S.: *Architectural and Interior Models*, p. 6. Van Nostrand Reinhold, New York (1970)
3. Piper, B.; Ratti, C.; Ishii, H.: *Illuminating clay: a 3-D tangible interface for landscape analysis*. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2002)
4. Schubert, G.: *Interaktionsformen für das digitale Entwerfen: Konzeption und Umsetzung einer rechnergestützten Entwurfsplattform für die städtebaulichen Phasen in der Architektur*. Dissertation. Technical University of Munich (2014)

5. Catanese, R.: 3D architectural videomapping. In: International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XL-5/W2 (2013)
6. Alonso, L., et al.: CityScope: a data-driven interactive simulation tool for urban design. use case volpe. In: Morales, A.J., Gershenson, C., Braha, D., Minai, A.A., Bar-Yam, Y. (eds.) ICCS 2018. SPC, pp. 253–261. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-96661-8_27
7. Grignard, A., et al.: Simulate the impact of the new mobility modes in a city using ABM. In: ICCS 2018 (2018)
8. Ishii, H., et al.: Augmented urban planning workbench: overlaying drawings, physical models and digital simulation. In: Proceedings of the 1st International Symposium on Mixed and Augmented Reality, p. 203. IEEE Computer Society (2002)
9. Korn, O., Schmidt, A., Hörz, T.: The potentials of in-situ-projection for augmented workplaces in production: a study with impaired persons. In CHI 2013 Extended Abstracts on Human Factors in Computing Systems (CHI EA 2013), pp. 979–984. Association for Computing Machinery, New York (2013). <https://doi.org/10.1145/2468356.2468531>
10. Huber, J., Steimle, J., Liao, C., Liu, Q., Mühlhäuser, M.: LightBeam: interacting with augmented real-world objects in pico projections. In: Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia, MUM 2012, pp. 16:1–16:10 (2012). <https://doi.org/10.1145/2406367.2406388>
11. Narazani, M., Egtebas, C., Jenney, S.L., Mühlhaus, M.: Tangible urban models: two-way interaction through 3D printed conductive tangibles and AR for urban planning. In: Adjunct Proceedings of the 2019 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2019 ACM International Symposium on Wearable Computers (UbiComp/ISWC 2019 Adjunct), pp. 320–323. Association for Computing Machinery, New York (2019). <https://doi.org/10.1145/3341162.3343810>
12. Garrido-Jurado, S., Muñoz-Salinas, R., Madrid-Cuevas, F.J., Marín-Jiménez, M.J.: Automatic generation and detection of highly reliable fiducial markers under occlusion. Pattern Recogn. **47**(6), 2280–2292 (2014). <https://doi.org/10.1016/j.patcog.2014.01.005>