Teaching Spatial Augmented Reality: a Practical Assignment for Large Audiences

Brett Ridel, Patrick Reuter, Nadine Couture

Teaching Spatial Augmented Reality:
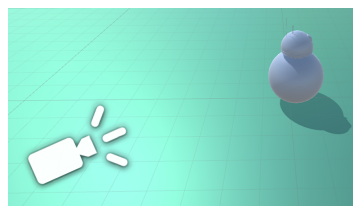# a practical assignment for large audiences

Brett Ridel[1], Patrick Reuter[1,2,3,4] and Nadine Couture[2,4]

[1]Inria Bordeaux Sud-Ouest, France   [2]LaBRI, Bordeaux, France   [3]University of Bordeaux, France
[4]ESTIA Research, Bidart, France

(a) Real-world setup

(b) Scene representing the real-world setup

(c) Virtual scene

Figure 1: *(a) Photo of the real-world setup that we simulate in the assignment, consisting of a video projector and a white diffuse styrofoam object. (b) Overall view of the scene representing the real-world setup, consisting of a simulated projector and the non-textured 3D model of the real-world styrofoam object. (c) Overall view of the virtual scene consisting of a virtual camera and the textured 3D model of the real styrofoam object.*

Abstract
We present a new methodology to teach spatial augmented reality in a practical assignment to large audiences. Our approach does not require specific equipment such as video projectors while teaching the principal topics and difficulties involved in spatial augmented reality applications, and especially calibration and tracking. The key idea is to set up a scene graph consisting of a 3D scene with a simulated projector that "projects" content onto a virtual representation of the real-world object. For illustrating the calibration, we simplify the intrinsic parameters to using the field of view, both for the camera and the projector. For illustrating the tracking, instead of relying on specific hardware or software, we exploit the relative transformations in the scene graph. We implemented our teaching methodology in Unity3D and tested it within a three-hour assignment to 24 and 20 master-level students in two consecutive years. We show the positive feedback that we received and discuss our plans for further improvement.

CCS Concepts
•*Social and professional topics* → *Computer science education;* •*Computing methodologies* → *Mixed / augmented reality; Tracking; Camera calibration;*

1. Motivation

Spatial augmented reality, also known as projection mapping, makes it possible to augment the real world with virtual information by means of a light emitter such as a video projector. First introduced by Raskar [RWLB01], one of the major advantages is that the users look at the real world and not on a screen as in see-through augmented reality, thus making the presence of a computer seamless. Today, the technology is mature and used in applications for industry, art [BCK*04], cultural heritage [RRL*14], and others [MZD*06].

Hence, it is timely to establish effective methods to teach spatial augmented reality to students of the different domains. Indeed, most teaching efforts in augmented reality address see-through augmented reality where, in terms of equipment for a practical assignment, a computer, tablet, or phone in conjunction with a camera is

sufficient. See for example the tutorials and examples on Udemy[1], Instructables[2], and others[3]. They all use Unity3D for representing the scene, Vuforia[4] for the tracking, and some script programming in C sharp. These examples involve only freely available software tools combined with hardware that most people have available anyway, resulting in a widespread use in teaching.

On the other hand, teaching spatial augmented reality necessarily requires the adjustment of a video projector, which is an equipment that is not available for every student in larger classes, thus requiring new teaching methodologies. Indeed, the motivation for writing this present paper came from the preparation and realization of a three-hour lecture for spatial augmented reality within a three-day course about augmented reality in general, targeted to groups of more than 20 students.

In this paper, we present a new methodology that we conceived at this occasion for teaching spatial augmented reality in a practical assignment to large audiences. The major benefit of our approach is that it does not require a specific equipment such as a video projector for every student while teaching the principal concepts and issues involved in spatial augmented reality applications, and especially calibration and tracking. Moreover, like the successful examples for teaching see-through augmented reality, our approach only relies on freely available software tools with no additional hardware, while still being a practical assignment.

The key idea is to set up a 3D scene with a projector that simulates the real-world projector used in the spatial augmented reality application, and a virtual representation of the real-world object to project virtual information on. So in fact, this 3D scene represents a view of the real-world setup. The image to project is calculated with a virtual camera that captures a textured version of the virtual object. We implemented our teaching methodology in Unity3D as an example.

This paper is organized as follows. In Section 2, we recall the principles of spatial augmented reality and the involved difficulties. In Section 3, we first show the overall idea and the targeted example of the assignment, before discussing the involved exercises to simulate a complete spatial augmented reality application step-by-step. In Section 4, we provide results of a preliminary evaluation of our practical assignment based on the feedback of the students, before we conclude and discuss potential improvements in Section 5.

2. Spatial augmented reality and the involved difficulties

The core of spatial augmented reality is the superimposition of a digital image onto a physical, real-world object, by using projected light. To simplify the writing, we refer in the remainder of this paper to a video projector as the light emitter as most commonly used. The digital image to superimpose is computed by deriving a 2D image from the 3D scene of the real-world object to project on. Using

the correct viewing parameters to derive this 2D image is essential for a perfect superimposition, which is quite easy to understand: intuitively, the 2D image must be identical to what an observer would see at the position and orientation of the projector.

There are two major steps for setting up a spatial augmented reality application, namely the calibration and tracking, that we briefly explain in the following. For more technical details, we refer the interested reader to Bimber and Raskar [BR05].

Calibration In order to compute the correct image, the video projector has to be calibrated. This means that its intrinsic parameters, such as the optical center and the focal distance, have to be determined. For the reader less familiar with these optical terms, this roughly corresponds to the viewing cone of the projector. Although intrinsic to the projector and independent from the spatial augmented reality application, these parameters are rarely provided by the manufacturers as they change according to the chosen display resolution and thus the involved aspect ratio, as well as the zoom factor if available. Most often, this calibration step involves the use of a camera in addition to the video projector [AO09, MT12, YNM16].

Object tracking The object tracking is specific to the augmented reality application. Indeed, this step consists in determining the position and orientation of the real-world object, relative to the position and orientation of the video projector, on which the digital image is projected. Object tracking is an active research field with hard constraints, especially in a dynamic setting where the object and/or the projector are moving [SCT*15]. Practical approaches are either vision-based with one or several cameras (ARToolkit, Vuforia, OptiTrack, etc.), or sensor-based (Valve Lighthouse, Razer STEM, etc.).

3. The assignment

To circumvent the use of additional material, our assignment is based on a 3D scene with a simulated projector that "projects" content onto a virtual representation of a real-world object, and the elements are organized hierarchically in a scene graph. Our assignment is targeted to Unity3D as it is freely available, and only basic knowledge of this 3D authoring tool is sufficient.

For the sake of pedagogical effectiveness, we conceived a simple yet representative example where both the real-world object and the projector are at fixed positions: the projection of a texture onto an object made of styrofoam with diffuse reflection. This example can then be extended as desired with dynamic elements, for example. We suggest that the lecturer of the assignment brings one single projector and the real-world object in order to render the explanations more tangible (see Figure 1(a)). We chose the BB8 character since an approximate replica can be easily built with two styrofoam balls that are cheap and easy to obtain, and a textured 3D model is freely available. Recall that in spatial augmented reality applications, it is essential to have both the real-world object together with its virtual representation: either the real-world object is acquired to obtain the 3D model, or, starting from the 3D model, a real-world physical object is fabricated, for example by 3D printing.
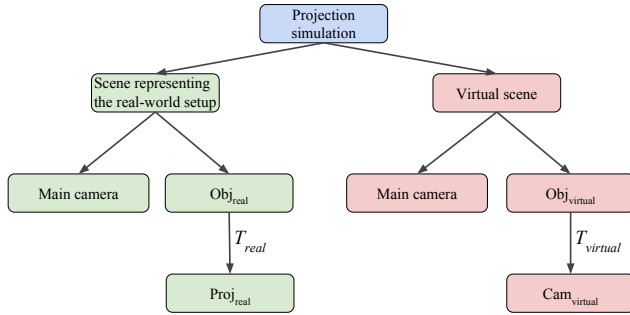
The guiding principle is to make the students understand that for

---

Figure 2: *Illustration of the scene graph used in the practical assignment. The tracking consists in making correspond the transformations $T_{virtual}$ and T_{real}.*

Figure 3: *Scene graph used in the first exercise of the practical assignment for rendering on a virtual representation of a real planar screen. The aspect ratio of the screen must correspond to the aspect ratio of the virtual camera.*

a correct superimposition of the virtual information over the real-world object, the projected image in the real world has to be computed properly according to the prior calibration and tracking. To this end, we let the students create a scene graph that consists of two related 3D scenes as described below and illustrated in Figure 2.

Scene representing the real-world setup The first scene, illustrated in Figure 1(b), is a virtual representation of the real-world setup of Figure 1(a), consisting of the virtual counterparts of the real-world video projector $Proj_{real}$ and the real-world object Obj_{real} to augment. The expected result of the assignment, the augmentation, will become visible in this scene as explained further below, and therefore, our practical assignment does not require specific hardware such as a video projector. We use a green color to identify this scene representing the real-world setup throughout this paper.

Virtual scene The second scene, the *virtual scene*, is created in order to compute the image to project over the real-world object for the augmentation (see Figure 1(c)). It consists of a 3D model of the real-world object $Obj_{virtual}$, textured as desired, and a virtual camera $Cam_{virtual}$. We use a red color to identify this *virtual scene* throughout this paper.

For a correct superimposition, first, concerning the **calibration**, the intrinsic parameters of the virtual camera must correspond to the intrinsic parameters of the projector. In virtual environments and video games, it is common practice to use the field of view (FOV) as the only intrinsic parameter of the virtual camera. As a projector can be considered as an inverse camera [MT12], in our assignment, we simplify the intrinsic parameters of the virtual camera $Cam_{virtual}$ and the projector $Proj_{real}$ to the FOV. This allows the students to manipulate only this one unique, representative and intuitive parameter for the result of the calibration. Note again that the calibration is usually far more complex, and we simplify it for a better understanding.

Second, concerning the **tracking**, the transformation (position and orientation) of the virtual camera relative to the 3D model must correspond to the one between the real-world projector and the real-world object. We exploit the hierarchical organization of our two scenes in the scene graph. Indeed, the relative transformations between the scene graph nodes make it possible to simulate the tracking by copy-pasting the transformation T_{real} between the
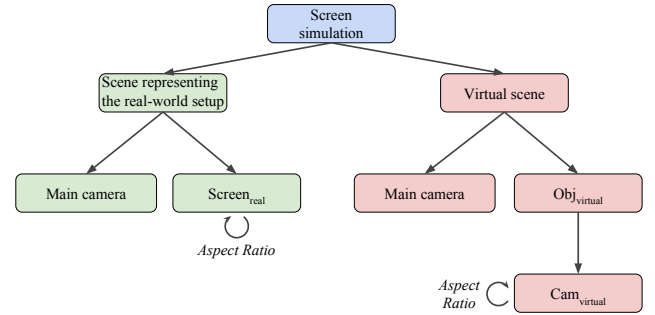
projector $Proj_{real}$ and the representation of the real-world object Obj_{real} to the transformation $T_{virtual}$ between the virtual camera $Cam_{virtual}$ relative to the 3D model $Obj_{virtual}$ (see again Figure 2).

Our assignment consists of two exercises that construct the aforementioned example step-by-step. In addition to distributing a detailed exercise sheet (provided in the additional files for review), we suggest the lecturer to advise the students directly to keep them at a similar pace for a collective experience. The first exercise consists in rendering a view of the 3D object on planar screens of different sizes and aspect ratios. The second exercise consists in creating the actual example, the projection of an augmentation onto a 3D object representing the real-world object.

3.1. Exercice 1: Rendering on planar virtual screens

For getting the students on track and familiar with the working environment in Unity, and also for overriding some default parameters that could hinder further comprehension, we designed a first exercise where a view of the 3D object is rendered on a virtual representation of a planar screen. To this end, the students create the *virtual scene* that first consists of the virtual 3D object $Obj_{virtual}$, the virtual camera $Cam_{virtual}$, and a main camera to visualize the scene. Then, the students create the *scene representing the real-world setup*, and they add a virtual representation of a real squared screen $Screen_{real}$ as a simple *quad*. This quad is then textured with the acquired image of the virtual camera $Cam_{virtual}$. Note that for now, the rendering is done on a virtual representation of a planar screen, and there is not yet a virtual representation of the projector and the real-world object. The resulting scene graph of this first exercise is illustrated in Figure 3.

Note also that when rendering to a texture, Unity's default camera has a FOV of 60 degrees with a 1:1 aspect ratio. For this reason, the students start the exercise with a squared quad representing the screen. Then, when they succeeded to render the image to the squared screen (see Figure 4(a)), we ask them to modify the screen to a rectangular one by augmenting only its width without changing its height. This results in a deformation of the displayed image (see Figure 4(b)) which makes the students understand that the aspect ratio of the virtual camera $Cam_{virtual}$ used to generate the im-
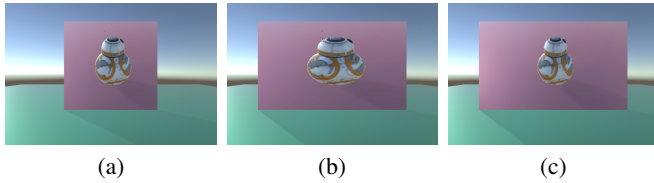
Figure 4: *Exercice 1: Rendering of a virtual image on a screen represented by a quad. (a,c) The rendering is correct when taking into account the intrinsic parameters. (b) Otherwise, the rendering is deformed.*
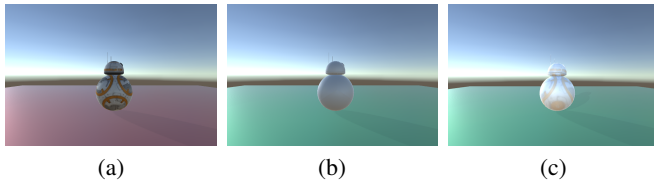
Figure 5: *Exercice 2: Expected results for the same viewing or projection parameters. (a) The textured object of the virtual scene. (b) The diffuse object of the scene representing the real-world setup before augmentation, and (c) after augmentation.*

age of the screen must correspond to the aspect ratio of the screen $Screen_{real}$. Consequently, we let the students adjust the aspect ratio of the virtual camera $Cam_{virtual}$ (see Figure 4(c)). Recall that the aspect ratio is one part of the so important intrinsic parameters to take into account in spatial augmented reality.

At the end of this first exercise, in order to prepare the second exercise, we ask the students to discard the quad representing the screen, but to maintain the virtual camera $Cam_{virtual}$. This has the side effect that this virtual camera keeps a different aspect ratio than the default aspect ratios of virtual cameras and projectors in Unity3D which are the same - otherwise the calibration step of the second exercise, represented by making coincide the aspect ratios, would be trivial and the students could neglect its importance.

3.2. Exercice 2: Projection to augment a 3D object

The second exercise consists in extending the preceding observations to the projection on a real-world object. To this end, starting from the scene graph established in the first exercise, the students have to add a virtual representation of the object Obj_{real} and the projector $Proj_{real}$ to the *scene representing the real-world setup* (Figure 5). The simulated projector $Proj_{real}$ projects the image acquired by the virtual camera $Cam_{virtual}$, which is identical to the one used in the previous exercise. Note that it is essential that both the real-world object Obj_{real} and the virtual object $Obj_{virtual}$ have the same geometry and the same scale. Concerning the material, we suggest to use a textured virtual object $Obj_{virtual}$ in the *virtual scene*, and an untextured diffuse object Obj_{real} in the *scene representing the real-world setup*. In this way, in addition to the technical aspects of making understand the principal concepts of spatial augmented reality, this exercise provides a nice and concrete application, namely color restoration. Indeed, in ancient artifacts, color

pigmentations may have disappeared due to weathering, and with spatial augmented reality, the original colors can be projected on the real artifact [RRL*14].

As a first step towards a superimposition, the students are asked to manually position and orient the projector $Proj_{real}$ in Unity so that its projection covers the entire targeted surface of the object Obj_{real}. As in real spatial augmented reality settings, the students see that the manual positioning is difficult and almost always results in an incorrect superimposition. Here, the tracking comes into play: the virtual camera $Cam_{virtual}$ of the *virtual scene* has to be placed exactly like the projector $Proj_{real}$ of the *scene representing the real-world setup*, relative to the respective object. Recall that in a scene graph, the transformations of the nodes are specified relative to their parent nodes. Consequently, as the virtual camera $Cam_{virtual}$ is a child node of the virtual object $Obj_{virtual}$, and as the projector $Proj_{real}$ is a child node of the virtual representation of the real-world object Obj_{real}, the tracking is done by copying the transformation T_{real} to the transformation $T_{virtual}$.

After this step, the superimposition is still not correct. Indeed, recall that after the first exercise, the virtual camera $Cam_{virtual}$ has a different aspect ratio than the projector, and thus different intrinsic parameters. As the intrinsic parameters have to coincide, the last step takes into account the calibration: it consists in modifying the aspect ratio of the virtual camera $Cam_{virtual}$ so that it corresponds to the one of the projector $Proj_{real}$. After this, the students see that the superimposition is correct.

Even though the visual result would be identical, it is essential to note that the virtual camera $Cam_{virtual}$ has to be adjusted according to the projector $Proj_{real}$, both for the calibration and the tracking. Indeed, in a physical spatial augmented reality setup, it is easy to manipulate the parameters of the virtual camera, whereas the projector has fixed intrinsic parameters. Similarly, the positioning of the real-world projector often has strong physical constraints, and so we rather modify the position and orientation of the virtual camera, and not the ones of the projector.

In addition to these two exercices, we suggest that the lecturer shows a demonstration of the real spatial augmented reality setting with the brought along physical projector and real-world BB8 character, including the calibration and tracking steps. In this way, the students can make the link between these practical steps in the real setting and the underlying theoretical concepts that they experienced in the assignment. We also suggest that the lecturer distributes this present paper that can foster further understandings and provide more insight, as we report below.

4. Preliminary evaluation

We designed the described practical assignment for a three-hour lecture to teach spatial augmented reality. It was part of a three-day elective course about augmented reality in general, called "Augmented Reality Engineer", that we are teaching for the second consecutive year at the ESTIA Institute of Technology. In order to validate the proposed methodology, we conducted a preliminary evaluation for the 24 students of the first time of our intervention. We asked the students to fill out an anonymous online questionnaire at the end of the three-day course. The participants (3F, 21M, age

range from 22 to 24 years) were at master-degree level. They are knowledgeable in Computer Science and have all deliberately chosen to participate in the assignment.

In total, the students replied to five questions, consisting of three closed and two open ones. The three closed questions were asked via a Multiple Choice Questionnaire (MCQ) to qualify the truthfulness of the following three statements:

1. I understood why this practical assignment was part of the three-day course *(options: not at all/just about/absolutely)*.
2. I enjoyed this practical assignment *(options: not at all/a little bit/quite/absolutely)*.
3. I have retained some content from this practical assignment *(options: yes/no)*.

For the first statement, among the three options, the students answered 0% "not at all", 12.5% "just about" and 87.5% "absolutely". For the second statement, among the four options, the students answered 0% "not at all", 4,2% "a little bit", 12,5% "quite", and 79,2% "absolutely". For the third statement, the students answered 100% yes.

The two open questions, where the students could provide free answers in paragraph, were:

4. From my point of view, the main interest of this practical assignment was to ...
5. I propose to improve this practical assignment by ...

We used a discourse analysis [GKL95] to analyse the corpus of the answers. It is a specialized method for discourses related to the Multiple Correspondence Analysis (MCA), followed by a hierarchical clustering analysis. For the first statement, we can distinguish seven semantic universes. The number of occurrences relative to each universe are: 8 for "discovering SAR", 6 for "practical assignment", 5 for "AR in a research project", 4 for "projection and mapping", 4 for "exploring cultural domain", 3 for "exploring Unity3D", 2 for "discovering an aspect of AR", and 1 for "Too cool!". Note that "projection and mapping" are always associated with "discovering SAR", and that the two main answers are exactly what the lecturers wanted to transmit: it is a "practical assignment" for "discovering SAR". The result of the discourse analysis is thus satisfying regarding our objectives. For the second statement, we can distinguish four semantic universes. The number of occurrences relative to each universe are: 21 for "nothing to change", 1 for "go faster!", 1 for "make it longer", and 1 for "provide a more useful example for augmented reality".

Summing up, the results obtained with this summative assessment are considered as very good when regarding the closed questions. This is confirmed by the semantic universes given by the students to describe the main interest of the lecture.

When we taught the practical assignment for the second time, we also provided to our students a preliminary version of this present paper, and we gathered their remarks. It came out that the reading of the paper was very welcomed by the students, and some of them stated that it clarified even more some points from the assignment. Some students also claimed that they appreciated to see their practical assignment from a pedagogical point of view.

5. Conclusion and future work

In this paper, we presented a practical assignment to teach spatial augmented reality to large audiences. As we simulate the projection to an object in 3D, the assignment does not require any specific equipment such as a video projector for the students, while teaching the major concepts involved in spatial augmented reality. This is possible thanks to the hierarchical organization of the objects, the camera, and the projector in a scene graph. To our knowledge, our work is the first approach in this direction, and we hope that it will raise a discussion about teaching methods for spatial augmented reality in larger classrooms. Indeed, teaching spatial augmented reality becomes important nowadays: referring to the market and industrial needs, and according to the Emerging Technology Hype Cycle from Gartner[5], it is today timely to train engineers for augmented reality. This "new job" was first introduced in 2012 at the ISMAR Symposium on Mixed and Augmented Reality, and his/her role is to set up, monitor, and sometimes control the augmentations, just as light or sound technicians set-up, monitor and control lighting and sound [CCN*12].

We presented our assignment as being tightly related to the use of Unity3D because of its widespread use, its graphical user interface combined with a high-level scripting language, and since it is freely available. However, our assignment can be tailored to other environments and languages, according to the desired student profiles, the level of difficulty, and available time. Indeed, we only rely on standard computer graphics bricks such as importing meshes and textures, multiplying transformation matrices, positioning cameras, rendering to textures, and shader programming. Hence, our assignment can also be implemented, for example, in C++ with OpenGL, according to the students' program.

In the future, according to our received feedbacks, we plan to extend our practical assignment to dynamic settings with moving projectors and/or objects. Dynamic settings introduce real-time constraints for the tracking: in our assignment, introducing an artificial latency for the tracking could simulate these constraints and provide maximum bounds for deriving the appropriate tracking solutions. We also plan to integrate interaction techniques so that the students can influence the augmentation. Moreover, we want to propose different surface materials for the real-world object in order to show the importance of reflection. We are convinced that the organization in a scene graph makes it comfortable to extend the assignment for our future plans.

Acknowledgments

References

[AO09] AUDET S., OKUTOMI M.: A user-friendly method to geometrically calibrate projector-camera systems. In *IEEE CVPR Workshops (Procams 2009)* (June 2009), pp. 47–54. doi:10.1109/CVPRW.2009.5204319. 2

---

5 Technology Research Gartner Inc., www.gartner.com

[BCK*04] Bimber O., Coriand F., Kleppe A., Bruns E., Zollmann S., Langlotz T.: Superimposing pictorial artwork with projected imagery. In *ACM SIGGRAPH 2004 Sketches* (2004), SIGGRAPH '04, ACM. URL: http://doi.acm.org/10.1145/1186223. 1186321, doi:10.1145/1186223.1186321. 1

[BR05] Bimber O., Raskar R.: *Spatial Augmented Reality: Merging Real and Virtual Worlds*. A. K. Peters, Ltd., Natick, MA, USA, 2005. 2

[CCN*12] Clay A., Couture N., Nigay L., De La Rivière J.-B., Martin J.-C., Courgeon M., Desainte-Catherine M., Orvain E., Girondel V., Domenger G.: Interactions and systems for augmenting a live dance performance. In *2012 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (Nov. 2012), pp. 29–38. doi:10.1109/ISMAR-AMH.2012.6483972. 5

[GKL95] Ghiglione R., Kekenbosch C., Landré A.: *L'analyse cognitivo-discursive*. Presses universitaires de Grenoble, 1995. 5

[MT12] Moreno D., Taubin G.: Simple, accurate, and robust projector-camera calibration. In *Proceedings of 3DIMPVT 2012* (2012). URL: http://dx.doi.org/10.1109/3DIMPVT. 2012.77, doi:10.1109/3DIMPVT.2012.77. 2, 3

[MZD*06] Miyake R., Zeman H., Duarte F., Kikuchi R. R. E., G. L., C. V.: Vein imaging: a new method of near infrared imaging, where a processed image is projected onto the skin for the enhancement of vein treatment. *Dermatol Surg. 32*, 8 (2006), 1031–1038. 1

[RRL*14] Ridel B., Reuter P., Laviole J., Mellado N., Couture N., Granier X.: The revealing flashlight: Interactive spatial augmented reality for detail exploration of cultural heritage artifacts. *J. Comput. Cult. Herit. 7*, 2 (June 2014), 6:1–6:18. URL: http://doi. acm.org/10.1145/2611376, doi:10.1145/2611376. 1, 4

[RWLB01] Raskar R., Welch G., Low K.-L., Bandyopadhyay D.: Shader lamps: Animating real objects with image-based illumination. In *Proceedings of the 12th Eurographics Workshop on Rendering Techniques* (London, UK, 2001), pp. 89–102. URL: http: //dl.acm.org/citation.cfm?id=647653.732300. 1

[SCT*15] Siegl C., Colaianni M., Thies L., Thies J., Zollhöfer M., Izadi S., Stamminger M., Bauer F.: Real-time pixel luminance optimization for dynamic multi-projection mapping. *ACM Trans. Graph. 34*, 6 (Oct. 2015), 237:1–237:11. URL: http://doi.acm. org/10.1145/2816795.2818111, doi:10.1145/2816795. 2818111. 2

[YNM16] Yang L., Normand J. M., Moreau G.: Practical and precise projector-camera calibration. In *2016 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)* (Sept 2016), pp. 63–70. doi:10.1109/ISMAR.2016.22. 2