

Browsing Spatial Photography for Dissemination of Cultural Heritage Research Results using Augmented Models

Florian Niebling¹ and Jonas Bruschke¹ and Marc Erich Latoschik¹

¹Human-Computer Interaction, University of Würzburg, Germany

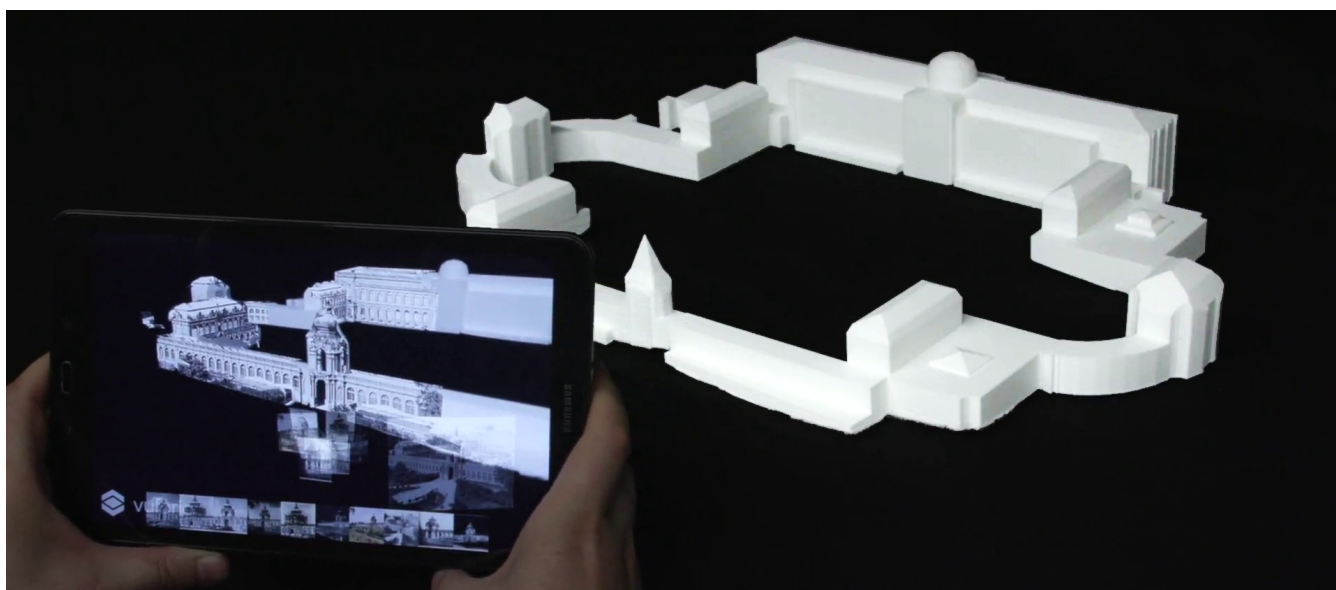


Figure 1: Hand-held Augmented Reality for spatial browsing of photographs registered to physical model.

Abstract

Both digital and physical 3D models of buildings as well as historical photographs of architecture are used for a wide range of needs, from research in humanities and information technologies, museum contexts and library studies, to touristic applications. Spatially oriented photographs play an important role in visualizing and browsing contemporary as well as historical architecture, starting with the ground-breaking Photo Tourism project [SSS06].

We present a technique to combine physical, 3D-printed models of buildings with spatially registered historical photographic documents in a hand-held Augmented Reality (AR) environment. Users are enabled to spatially explore historical views of architecture by selecting photos from a collection of images, which are then utilized as textures for the physical model rendered on their respective mobile device. We compare different methods to spatially select photos registered to a physical model in hand-held AR.

CCS Concepts

•Human-centered computing → Mixed / augmented reality; Geographic visualization;

1. Introduction

Photographs are an essential source for historical research in the Digital Humanities (DH). Numerous digital image archives, con-

taining vast numbers of photographs, have been set up within digitization projects. Our research focuses on creating methods and

techniques to enable DH researchers to address spatial investigations and to make their findings available to the general public.

Digital repositories fulfill diverse purposes in humanities research as well as in a frame of museum or touristic transfer. A possible technological basis for creating access to large scale image repositories is the spatial and temporal aggregation of data, combining historical photographs and digital 3D models [NMBL18].

3D visualizations correspond with human viewing habits and allow intuitive presentation of further information, such as the location, appearance and surroundings of an object, and showing photos at the respective camera position where they were taken. Spatial user interfaces employing AR technology are used in museum contexts in outdoor, room-scale, as well as model-sized installations for that purpose.

In this paper, we present a visualization and interaction technique for browsing collections of spatially registered photography of architecture, combining historical photographs with physical 3D-printed models of buildings. We also provide an initial comparison of different methods for spatial selection of photos registered to a physical model in hand-held AR.

2. State of the Art

Photogrammetric techniques make it possible to get metric object information by means of indirect measurements within photographs or photogrammetric images. Thus, photogrammetric recordings have been used for documentation and automated measurement for more than 150 years. In contrast to hand measurements or tachometry, photographs or special photogrammetric recordings offer a pictorially and non-interpreted object documentation along with the possibility of obtaining geometric information.

Structure-from-Motion (SfM) workflows [HZ03] have been employed to extract further 3D information from 2D photos. In addition to merely location-based browsing, Snavely et. al. use camera pose (location, orientation, and field of view) and sparse 3D scene information from SfM to create interfaces for browsing collections of photographs [SSS06].

Conventional browsing methods — such as keyword-based or faceted search — have their limitations when metadata is incomplete or imprecise or the user has little knowledge about the object they are looking for. To access historical photographs by location and orientation, new interfaces are required. Schindler et. al. generate time-varying digital 3D models to index large collections of images serving as a tool of historical discovery, revealing information about locations, dates, and contents of historical images [SD12]. 3D pointclouds, digital models and textures have since been brought into desktop applications for easily accessible navigation and browsing of spatial and temporal photo repositories [BNM*17].

Projector-based Spatial AR (SAR) combines real-world objects with virtual illuminations in physical space [RWF98]. Textured physical models using SAR have been used to disseminate findings from cultural heritage (CH) research in museums. Ridet et. al. introduce spatial interaction methods for projected AR to highlight details on CH artifacts [RRL*14].

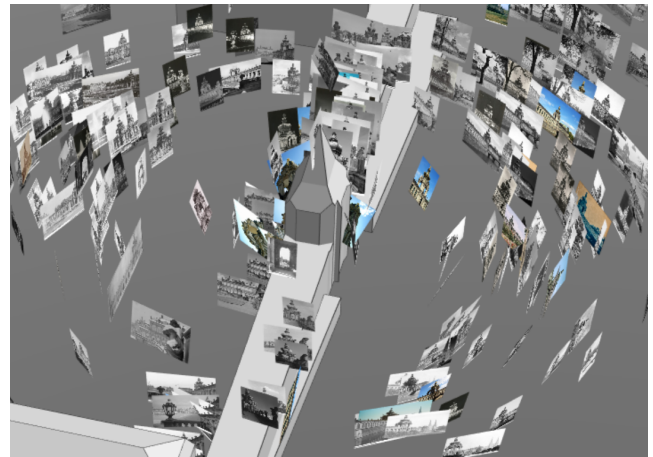


Figure 2: Photographs registered to digital 3D model.

3. Methods

We enable spatial access to repositories of historical photography of architecture. Position and orientation of historical images is put into the context of (historical) 3D models of the depicted buildings. We combine physical 3D-printed models of architecture with spatially oriented historic images in a hand-held AR environment.

3.1. Spatial AR Browsing

A small subset of historical photos in the utilized media repository *Deutsche Fotothek* are spatially oriented in the coordinate system of the underlying digital 3D model created by a previous SfM workflow [MVS*17], providing location, orientation, and field of view of the camera for each photo. We limit the images to a region of Dresdner Zwinger, with focus on the Kronentor (Crown Gate) in particular (cf. Figure 2). An Augmented Reality (AR) view is created consisting of a physical 3D-printed model, and virtual texturing with historical depictions of the buildings. A blank 3D-printed model of Dresdner Zwinger is used to provide a basis for texturing with historic images. The AR view is generated using hand-held devices (cf. Figure 1), allowing for free movement of the users with respect to the physical model, interaction using multi-touch, and also potentially for collaboration between users employing multiple tablets.

3.2. Texturing with Registered Photographs

Photographs represent a projective transformation of real-world 3D scenery onto a 2D plane using a camera with a given optical system.

We create historic textures for registered models of architecture by UV mapping vertices of the respective digital 3D model to the historic photography. UV coordinates are calculated by casting rays from the position P of the photographer to each vertex V in the model, through the projective plane of the currently processed image I . Collision detection of rays with models in the scene is performed to detect visibility of vertices from position P . Parts of a building that are not contained in I , because they are either outside

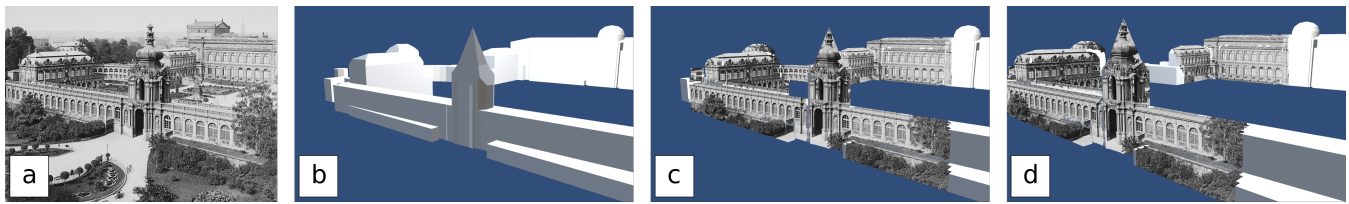


Figure 3: (a) Historic depiction of Dresdner Zwinger. (b) LOD2 CAD model provided by the city municipality. (c) Projection of image to model and usage as texture. (d) Changing viewpoint with applied texture, structures hidden in the original image are left blank.

of the view frustum of the camera (cf. right hand side of the model in Figure 3c) or hidden by other structures (cf. center-back part of Figure 3d), cannot be textured with information contained within I and are therefore left blank. Calculation of UV coordinates for each photo and model can either be pre-computed or performed in real-time, depending mainly on the complexity of the 3D models in the scene.

Employing photos as textures allows for the usage of very coarse 3D models of buildings, as details are provided by comparably high resolution photos. Small variances that often occur during lifetime of a historic building can then be approximated by the same coarse 3D model. In this work, we use an LOD2 model — a coarse building model with standard roof structures — of Dresdner Zwinger containing just 1889 triangles (cf. Figure 3a-c).

3.3. Hand-held AR with Physical 3D Models

Digital 3D models procured from the city municipality are 3D-printed into physical models to be used in a hand-held video see-through AR setup containing a 10.1 inch Android tablet (Samsung Galaxy Tab A 10.1, cf. Figure 1). In the presented case, the physical 3D model is measuring roughly 0.55m x 0.20m x 0.09m. Tracking of the physical model is provided by Vuforia *Model Targets with Advanced Recognition* in a Unity application, combining the digital model of the building with a media repository of spatially registered historical photography.

We render the textured digital model (cf. Section 3.2) on top of the video of the tracked physical model, discarding parts that are not contained in the respective photo selected by the user. Spatial movement of the tablet can be performed to view buildings from a perspective diverging from the perspective of the photo, still allowing the user to perceive the historical appearance of the building depicted in the historical image (cf. Figure 3d and Figure 4).

3.4. Interaction Methods

We performed an initial evaluation of different interaction methods as part of our user-centered design approach to allow selection of images to be used as textures. Users were handed the task to find certain given depictions of the Crown Gate by moving the tablet with respect to the physical model, employing the interaction methods described in the following.

- *Direct selection of images* by direct touching of photos in the 3D AR view. Even small changes in orientation during interaction can induce large changes in position of the displayed photos.

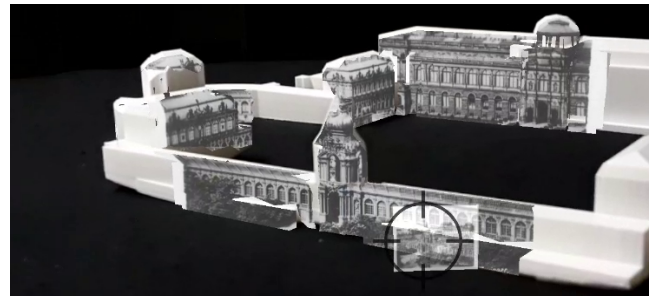


Figure 4: AR view of 3D-printed physical model (white) with historical texture. Notice the blank part in center-back of the model that were hidden by the Crown Gate in the front of the depiction and thus cannot be textured using information contained in the photo.

Since small changes happen all the time involuntary, induced mainly by instabilities while holding the tablet and having to move fingers to interact or by tracking instabilities, direct selection proved to be too imprecise. An additional problem is the high amount of overdraw in our dataset. Using ca. 500 historical images of the Crown Gate, many of these are partly or completely occluded by other images for most viewpoints, making them hard or even impossible to select.

- *Spatial distance selection* by automatically switching to the photo closest to the position of the tablet. Selecting the photo closest to the position of the tablet in 3D space suffers from poor conformity with user expectations, as it is often not intuitively clear which of the photos actually is closest to the user in all spatial dimensions. There often are photos *behind* but close to the tablet's position. Texturing then switches to photos that the user does not see since they cannot be displayed on the tablet's screen. Limiting selection to photos in the viewing direction of the user still performs many unexpected switches, as small involuntary movements are constantly performed with the tablet as noted before.
- *Reticle-based expand selection*. To avoid having to move fingers across the screen during interaction, we decided on selection by targeting photos with a reticle in the tablet's center, making the selection process more stable (cf. Figure 5a). To be able to select partially or completely hidden images, we use an adapted *expand* selection method [CWLJ12] with refinement to pre-select a number of images close to a ray cast through the reticle. In contrast to regular *expand* where pre-selected objects are more

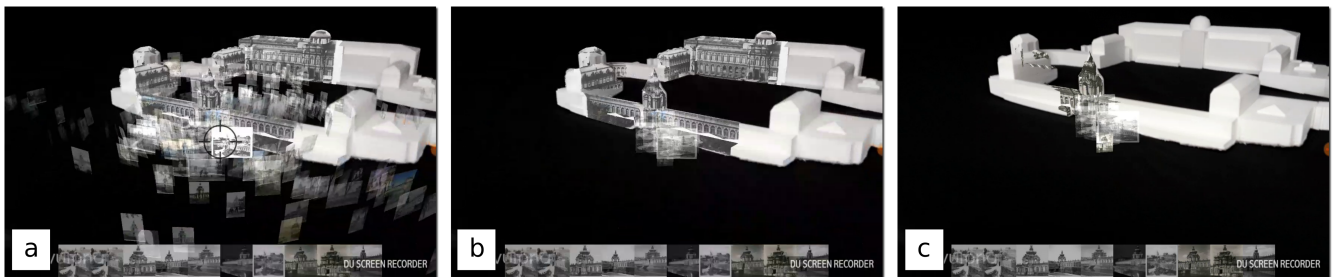


Figure 5: Hand-held device view. 3D AR rendering on top, selectable 2D image gallery on bottom: (a) Tap-and-hold to show all registered images and fill image gallery on bottom row with images close to center reticle. (b) Release tap to limit images to selected subset in 3D view. (c) Visualize different textures by touching images in gallery.

or less randomly placed on a regular grid obstructing the 3D view, we sort images by angular distance to the ray into a 2D image gallery on the bottom of the display (cf. Figure 5), conserving a certain amount of spatial information in the expanded images. We show available spatially oriented photos registered with the physical model using a tap and hold gesture anywhere on the tablet's display. An image can then be either selected and assigned as a texture by spatial movement of the tablet to position the image into the center of the display and releasing the tap, or by tapping one of the near-by images in the image gallery (cf. Figure 5c). When the tap is released, visualization of photos in the 3D AR view is limited to pre-selected images to reduce clutter (cf. Figure 5b).

4. Conclusion and Future Work

We have presented a method for browsing of and interaction with photographs in an environment featuring physical 3D models and hand-held AR. A method for texture-mapping of spatial photography to coarse 3D models was used to provide historical views of the depicted buildings, allowing for change in perspective by spatial movement of the user's mobile device. An initial comparison of interaction methods for spatial selection of images in hand-held AR was performed. We have shown an extension to the *expand* selection method with refinement, to keep some spatial information of the pre-selected images by ordering images in the 2D image gallery accordingly.

UV mapping of model vertices to image coordinates can be produced in real-time but has shortcomings. There are triangles where not all vertices are contained in the processed image in the general case, so interpolation will produce incorrect results at some point in the model. Discarding these incomplete triangles will create holes in the rendering of the model, depending on the level of subdivision of the CAD model (cf. Figure 4). Re-textured geometry modeling could be used to address these problems even in real-time [DH18].

The potential for combining multiple different photos into a single building texture to provide for visualization of hidden structures needs to be evaluated. An extension of the presented methods towards employing smart glasses is an interesting research objective approaching novel modes of interaction.

References

- [BNM*17] BRUSCHKE J., NIEBLING F., MAIWALD F., FRIEDRICHS K., WACKER M., LATOSCHIK M. E.: Towards browsing repositories of spatially oriented historic photographic images in 3d web environments. In *Proceedings of the 22Nd International Conference on 3D Web Technology* (New York, NY, USA, 2017), Web3D '17, ACM, pp. 18:1–18:6. URL: <http://doi.acm.org/10.1145/3055624.3075947>, doi:10.1145/3055624.3075947. 2
- [CWLJ12] CASHION J., WINGRAVE C., LAVIOLA JR J. J.: Dense and dynamic 3d selection for game-based virtual environments. *IEEE transactions on visualization and computer graphics* 18, 4 (2012), 634–642. 3
- [DH18] DONG S., HÖLLERER T.: Real-time re-textured geometry modeling using microsoft hololens. In *2018 IEEE Virtual Reality (VR)* (2018). 4
- [HZ03] HARTLEY R., ZISSERMAN A.: *Multiple View Geometry in Computer Vision*, 2 ed. Cambridge University Press, New York, NY, USA, 2003. doi:<https://doi.org/10.1017/CBO9780511811685>. 2
- [MVS*17] MAIWALD F., VIETZE T., SCHNEIDER D., MÜNSTER S., NIEBLING F.: Photogrammetric analysis of historical image repositories for virtual reconstruction in the field of digital humanities. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLII-2/W3* (Feb. 2017), 447–452. URL: <https://dx.doi.org/10.5194/isprs-archives-XLII-2-W3-447-2017>. 2
- [NMBL18] NIEBLING F., MAIWALD F., BARTHEL K., LATOSCHIK M. E.: 4d augmented city models, photogrammetric creation and dissemination. In *Digital Research and Education in Architectural Heritage* (Cham, 2018), Münster S., Friedrichs K., Niebling F., Seidel-Grzesińska A., (Eds.), Springer International Publishing, pp. 196–212. doi:10.1007/978-3-319-76992-9_12. 2
- [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. 2
- [RWF98] RASKAR R., WELCH G., FUCHS H.: Spatially augmented reality. In *First IEEE Workshop on Augmented Reality (IWAR'98)* (1998), Citeseer, pp. 11–20. 2
- [SD12] SCHINDLER G., DELLAERT F.: 4D Cities: Analyzing, Visualizing, and Interacting with Historical Urban Photo Collections. *Journal of Multimedia* 7, 2 (2012), 124–131. 2
- [SSS06] SNAVELY N., SEITZ S. M., SZELISKI R.: Photo tourism: Exploring photo collections in 3d. *ACM Trans. Graph.* 25, 3 (July 2006), 835–846. URL: <http://doi.acm.org/10.1145/1141911.1141964>, doi:10.1145/1141911.1141964. 1, 2