Virtual Archaeology Review, **10**(21): 1-13, 2019
https://doi.org/10.4995/var.2019.11867

Received: May 22, 2019
Accepted: July 4, 2019

A 4D INFORMATION SYSTEM FOR THE EXPLORATION OF MULTITEMPORAL IMAGES AND MAPS USING PHOTOGRAMMETRY, WEB TECHNOLOGIES AND VR/AR

UN SISTEMA DE INFORMACIÓN 4D PARA LA EXPLORACIÓN DE IMÁGENES Y MAPAS MULTITEMPORALES UTILIZANDO FOTOGRAMETRÍA, TECNOLOGÍAS WEB Y VR/AR

Ferdinand Maiwald[a,*], Jonas Bruschke[b], Christoph Lehmann[c], Florian Niebling[b]

[a] Institute of Photogrammetry and Remote Sensing, TU Dresden, Helmholtzstraße 10, 01069 Dresden, Germany. ferdinand.maiwald@tu-dresden.de

[b] Human-Computer Interaction, Universität Würzburg, Am Hubland, 97074 Würzburg, Germany. jonas.bruschke@uni-wuerzburg.de; florian.niebling@uni-wuerzburg.de

[c] Centre for Information Services and High Performance Computing, TU Dresden, Chemnitzer Straße 46b, 01187 Dresden, Germany. christoph.lehmann@tu-dresden.de

Highlights:

- Strategies for a completely automated workflow from image repositories to four-dimensional (4D) access approaches.
- The orientation of historical images using adapted and evaluated feature matching methods.
- 4D access methods for historical images and 3D models using web technologies and Virtual Reality (VR)/Augmented Reality (AR).

Abstract:

This contribution shows the comparison, investigation, and implementation of different access strategies on multimodal data. The first part of the research is structured as a theoretical part opposing and explaining the terms of conventional access, virtual archival access, and virtual museums while additionally referencing related work. Especially, issues that still persist in repositories like the ambiguity or missing of metadata is pointed out. The second part explains the practical implementation of a workflow from a large image repository to various four-dimensional (4D) applications. Mainly, the filtering of images and in the following, the orientation of images is explained. Selection of the relevant images is partly carried out manually but also with the use of deep convolutional neural networks for image classification. In the following, photogrammetric methods are used for finding the relative orientation between image pairs in a projective frame. For this purpose, an adapted Structure from Motion (SfM) workflow is presented, in which the step of feature detection and matching is replaced by the Radiant-Invariant Feature Transform (RIFT) and Matching On Demand with View Synthesis (MODS). Both methods have been evaluated on a benchmark dataset and performed superior to other approaches. Subsequently, the oriented images are placed interactively and in the future automatically in a 4D browser application showing images, maps, and building models. Further usage scenarios are presented in several Virtual Reality (VR) and Augmented Reality (AR) applications. The new representation of the archival data enables spatial and temporal browsing of repositories allowing the research of innovative perspectives and the uncovering of historical details.

Keywords: historical photographs and maps; photogrammetry; repository; documentation; Geographic Information System (GIS); orientation and matching

Resumen:

Esta contribución muestra la comparación, investigación e implementación de diferentes estrategias de acceso a datos multimodales. La primera parte de la investigación se estructura en una parte teórica en la que se oponen y explican los términos de acceso convencional, acceso a los archivos virtuales, y museos virtuales, a la vez que se hace referencia a trabajos relacionados. En especial, se señalan los problemas que aún persisten en los repositorios, como la ambigüedad o la falta de metadatos. La segunda parte explica la implementación práctica de un flujo de trabajo desde un gran repositorio de imágenes a varias aplicaciones en cuatro dimensiones (4D). Principalmente, se explica el filtrado de imágenes y, a continuación, la orientación de las mismas. La selección de las imágenes relevantes se hace en parte manualmente, pero también con el uso de redes neuronales convolucionales profundas para la clasificación de las imágenes. A continuación, se utilizan métodos fotogramétricos para encontrar la orientación relativa entre pares de imágenes en un marco proyectivo. Para ello, se presenta un flujo de trabajo adaptado a partir de *Structure from Motion*, (SfM), en el que el paso de la detección y la correspondencia de entidades es sustituido por la Transformación de entidades invariante a la radiancia (*Radiant-Invariant Feature Transform*, RIFT) y la Correspondencia a demanda con vistas sintéticas (*Matching on Demand with View Synthesis*, MODS). Ambos métodos han sido evaluados sobre la base

*Corresponding author: Ferdinand Maiwald, ferdinand.maiwald@tu-dresden.de

de un conjunto de datos de referencia y funcionaron mejor que otros procedimientos. Posteriormente, las imágenes orientadas se colocan interactivamente y en el futuro automáticamente en una aplicación de navegador 4D que muestra imágenes, mapas y modelos de edificios. Otros escenarios de uso se presentan en varias aplicaciones de Realidad Virtual (RV) y Realidad Aumentada (RA). La nueva representación de los datos archivados permite la navegación espacial y temporal de los repositorios, lo que permite la investigación en perspectivas innovadoras y el descubrimiento de detalles históricos.

Palabras clave: fotografías y mapas históricos; fotogrametría; repositorio; documentación; Sistema de Información Geográfica (SIG); orientación y correspondencia

1. Introduction

The access to extensive repositories of historical photographs and maps using classical database systems presupposes a comprehensive knowledge about structure and semantics of the respective database objects and is aimed above all at experienced users. Without knowledge of the depicted buildings as well as the concrete structuring of data, like keywords, ontologies and metadata, search queries often do not lead to the desired results. Consequently, new representation strategies and the usage of informative models and systems become more and more accepted by the users.

Especially, Geographic Information Systems (GIS) tools combined with 3D models of buildings or heritage sites enable access to a much broader community. Technologies such as Web3D, Augmented Reality (AR) or Virtual Reality (VR) can then be used to enhance the understanding of the virtual 3D environment.

This research aims to investigate and prototypically develop access to such repositories via a spatiotemporal localization of historical photographs and maps within a 3D model of the city center of Dresden/Germany. The project is conducted by nine researchers with various disciplinary backgrounds such as information science, education technology, art and architectural history, photogrammetry, media informatics, and computer science. The central component is a web-based 4D browser, which includes temporal data in addition to the 3D spatial information of the media and building objects. The spatial visualization of location-based media and objects via a 3D city model is intended to enable access to urban history information as simply and intuitively as possible to an extended circle of users. That means, that browsing the data is not limited to a search bar and its output, but gets expanded into four dimensions using 3D navigation methods and a responsive time slider. A tackled scenario is to provide the interface for large-scale collections of urban images. Backed by spatial information, diverse visualization methods enable investigation of, e.g., the statistical distribution of images in respect to their orientation or the depicted objects.

If all objects are located in a higher-level coordinate reference system, the 3D model should serve as the basis for a location-dependent AR representation on mobile devices. Our research group aims to provide tools for Cultural Heritage education in different settings using AR technology. The focus is on the communication of research results concerning historical photography of architecture to the general public through exhibitions, as well as guided and unguided tours. Bringing geo-located photography and historic models into the current cityscape through hand-held AR, tourists are enabled to engage with the historical situation of the photographer. In addition to location-based AR, tools for the exploration of large amounts of photographic documents in exhibitions are developed, combining 3D printed models of architecture with geo-located photographs in hand-held as well as see-through AR settings. Using the computed spatial orientation of photographs with respect to the respective 3D models of buildings, 3D printed models can be augmented with historical images to provide historical textures.

Hence, one main aspect for the different presented access strategies is the precise selection of relevant images and their spatialization in a 3D space. The historical photographs origin from the photo library of the Saxon State and University Library Dresden (SLUB), which contains about 1.8 million images of over 80 institutions at this point in time. Access to the images is granted via the website (http://deutschefotothek.de), which provides keyword search and additional content filters like photographer, depicted people and topic.

Our research group wants to extend these conventional search possibilities using machine learning methods regarding the visual content of the images. In cooperation with the Competence Center for Scalable Data Services and Solutions (ScaDS) different deep learning approaches were tested for the filtering of historical images.

Once filtered, the images have to be oriented in 3D space in relation to the depicted 3D city models. For this purpose, different photogrammetric methods were applied on the images and models. An interactive Direct Linear Transformation (DLT) and manual placing of images in 3D space have already been tested in the prototype application. Since a completely automated process is preferable, the photogrammetric work focuses on the relative orientation and thus an optimized Structure from Motion (SfM) workflow for historical images. Difficulties arise due to different camera types, acquisition techniques, image media, digitization artefacts, and many more. Especially for feature detection and matching, the common used methods have to be adapted.

This article aims to provide an overview of parts of current research and development work in the ongoing interdisciplinary research project. It shows how multimodal access especially on image repositories has been realized in the past and how it will be implemented in the future. The workflow from conventional repositories as databases over filtering and spatializing the images to presentation and interactive browsing in a web application as well as in a VR/AR environment are presented. The main focus is on the photogrammetric works, since these are the key part for the orientation and positioning of images in the different applications.

Some approaches and ideas have already been described in (Maiwald, Henze, Bruschke, & Niebling, 2019) but are now updated and extended by more recent work results.

2. Multimodal access on repositories

There are a lot of different ways to get access to multimodal data resources such as plans, maps, images, etc., from different points in time. Conventional interactive access to repositories with the requirement of being on site stands in contrast to the virtual access using online collections. Listed under the collective term *virtual museums*, 3D web applications and technologies such as AR, VR, and Mixed Reality (MR) are considered and compared.

2.1. Conventional access

When working with specific historical data, it is often necessary to crawl archives and find source documents piece by piece, especially, since neither the digitization nor the annotation of huge data archives are fully completed. However, this requires an excellent knowledge of the topic and some information may not be found at all. Especially, the following four issues still persist in many libraries and archives (Evens & Hauttekeete, 2011):

- Complex process of digitization.
- Identification of correct metadata.
- Business models/financing.
- Intellectual property rights management.

This leads to the two main problems of finding and afterward publishing the relevant data. Even in recent research, it is often necessary to collect analog data and process it. For the purpose of reconstruction, the archive documentation research is often considered as the first step (Bitelli, Dellapasqua, Girelli, Sbaraglia, & Tinia, 2017). One example is the modeling of the Larz Anderson Estate using mainly photographs, journals, and drawings scattered in several American archives (Ackerman & Glekas, 2017). It was also possible to reconstruct and texturize the collapsed dome of the San Pietro Church in Tuscania using an automatic SfM approach with 23 images (Beltrami et al., 2019), while the Great Mosque of Aleppo could be reconstructed with image data on 16 CD-ROMs (Grussenmeyer & Al Khalil, 2017). For the automatic 3D modeling of the Zwinger in Dresden, a dataset of over 800 archival images had to be reduced to around 50 images manually (Maiwald et al., 2017) for the subsequent SfM approach, and recently, parts of the Fortezza Vecchia could be recreated using archive images and plans from different analog sources (Bevilacqua, Caroti, Piemonte, & Ulivieri, 2019).

Consequently, the traditional process of archive work from exploring and finding data to digitizing and publishing the sources is still present nowadays.

2.2. Virtual access using online collections

As already stated, digital information is and will be essential for automated workflows as well as the dissemination of information and knowledge in society. Therefore, (Ross & Hedstrom, 2005) proposed six reasons why digital preservation plays an important role in cultural heritage institutions. These are in brief:

- protection and conservation of cultural memory is a societal good;
- international scientific collaborations benefit from the availability of data repositories;
- accountability of those institutions;
- re-use of digital information as an economic benefit;
- effective and affordable strategies as a move from an industrial to a knowledge economy;
- sustainable digital libraries depend upon the availability of preservation tool and services.

When looking at today's archives, data is usually presented in one or two dimensions. Hence, objects, books, photographies, and other information is displayed as text or can be depicted as images. Examples for comprehensive online collections are Europeana (https://www.europeana.eu), Artstor (https://library.artstor.org) and Arachne (https://arachne.dainst.org), but there exist many more. Usually, the websites providing access to these repositories show a search bar in which users can type keywords. Those lead to a resulting page (Fig. 1) where numerous filters (e.g. media classification, contributor, license, etc.) can be applied.

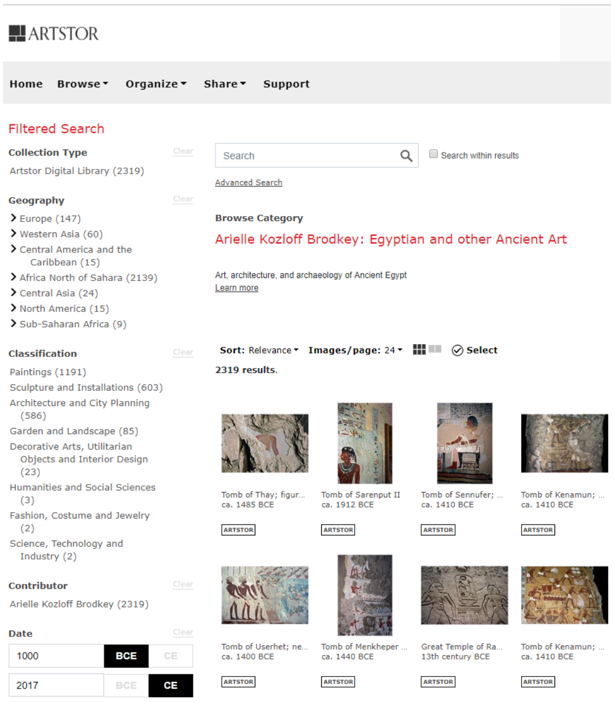

Figure 1: Example search result for the category "Arielle Kozloff Brodkey: Egyptian and other Ancient Art" on https://library.artstor.org.

Nonetheless, challenges such as low data quality, difficult access, usability problems, and poor availability of information in certain areas of interest still persist today (Münster, Kamposiori, Friedrichs, & Kröber, 2018). Considering the search bar, metadata plays an important role in finding the relevant object or data file but also limits the search possibilities. Problems arise when the metadata is missing, incorrect, or insufficient. Then, it will not be possible for a user to find the specific object. Filters can help to narrow the search query but are often not enough to simplify the results.

Additionally, different types of users are following diverse search strategies and for most of the users, Google is still the main entrance for search queries (Beaudoin & Brady, 2011). It has been investigated by different research groups that a beginner uses far more simple terms than an expert for metadata-based search queries and that online data retrieval and access is still a challenge for many people (Beaudoin & Brady, 2011; Fleming-May & Green, 2016; Münster et al., 2018; Yoon & Chung, 2011). In the presented research, access to historical images is granted via the photo library of the SLUB, where similar issues are present. These issues are meant to be investigated and improved during the project. In addition, the mentioned repository holds maps, drawings, and scaled plans, which are also valuable for diverse multimodal representation strategies in informative models and systems.

2.3. Virtual museums

Since the junior research group develops various applications resulting out of multimodal online repositories, this chapter presents related work in the field of cultural heritage summarized under the term *virtual museum*.

While there is no standard definition for this term, the Virtual Multimodal Museum (https://www.vi-mm.eu) states that “a virtual museum (VM) is a digital entity that draws on the characteristics of a museum, in order to complement, enhance, or augment the museum through personalization, interactivity, user experience and richness of content”. (Styliani, Fotis, Kostas, & Petros, 2009) identified four exhibition types as:

- Web3D exhibitions;
- VR exhibitions;
- AR exhibitions;
- MR exhibitions;

which shall be defined and presented with recent examples in the field of cultural heritage in the following. Considering the fact that also a two-dimensional (2D) representation of an exhibit can be regarded as a virtual museum the research in this chapter focuses on examples in at least 3D.

A Web3D exhibition could be defined as a virtual museum that is made accessible through the World Wide Web platform and browser independent. Advantages of such a representation, especially concerning the field of cultural heritage, are (Slater & Sanchez-Vives, 2016):

- Worldwide access and exploration of cultural sites.
- Preservation by digitization.
- Restoration and experience of cultural sites.
- Modeling of cultural sites under different future conditions.

Originally defined for VR applications, these advantages also apply to Web3D exhibitions of cultural sites and objects.

Web3D emerged especially due to the rise and distribution of fast Internet connections and additionally the development and maintenance of standards. Starting with the Virtual Reality Modeling Language (VRML) in 1997 by the Web3D Consortium (http://www.web3d.org), this standard has been replaced by the Extensible 3D (X3D) graphics and Humanoid Animation (H-Anim) standards, which are open, free, extensible, and interoperable

This allows the creation of different applications such as 3D WebGIS for analysing the ancient Maya kingdom of Copan, Honduras (Girardi, von Schwerin, Richards-Rissetto, Remondino, & Agugiaro, 2013) or the visualization of high resolution 3D models from Andalusian universities’ cultural heritage in the project Atalaya3D (Melero, Revelles, & Bellido, 2018). In contrast to the depicted technology, the 4D cities project used Java and a Java applet to grant web-based access on time-varying city models (Schindler & Dellaert, 2012) related to the presented research.

In addition, the digitization and visualization of cultural heritage play an important role in the field of Web3D exhibitions. Examples are various projects of CultLab3D (https://www.cultlab3d.de), models of the 3D-ICONS project (Koutsoudis, Arnaoutoglou, Tsaouselis, Ioannakis, & Chamzas, 2015), and of the INCEPTION project (Maietti, Di Giulio, Piaia, Medici, & Ferrari, 2018).

While it is easier to distinguish Web3D and extended reality applications, the terms VR, AR, and MR are more difficult to be separated. A lot of researchers see VR as the possibility to navigate and interact (selecting and moving objects) in a 3D environment (Gutierrez, Vexo, & Thalmann, 2008; Guttentag, 2010), with the addition that the user is immersed in a completely virtual environment (Milgram, Takemura, Utsumi, & Kishino, 1995). AR is often seen as a type of VR (Burdea & Coiffet, 2003; Guttentag, 2010), but should be in fact classified as an independent term. The user sees the real environment through, e.g. a see-through head-mounted display (HMD) and additional virtual information are added to the real world (Milgram et al., 1995). MR in the scheme of the Reality-Virtuality Continuum (Fig. 2) can be seen as everything between the two extrema, completely real environment, and completely virtual environment (Milgram et al., 1995; Styliani et al., 2009).

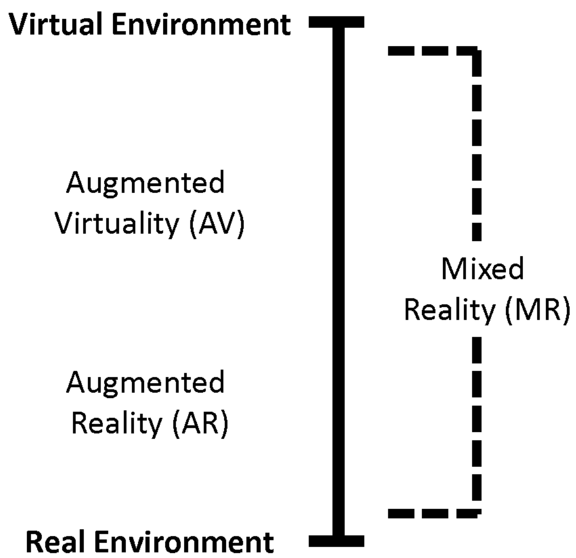

Figure 2: Representation of the Reality-Virtuality Continuum according to Milgram et al. (1995).

Some recent examples are a VR application of an Ottoman fortress (Tschirschwitz et al., 2019), research concerning the effectiveness of the AR application in Augusta Raurica (Armingeon, Komani, Zanwar, Korkut, & Dornberger, 2019), or the MR application of the reconstruction of one building of the Bergen-Belsen concentration camp (Oliva et al., 2015).

3. Workflow and access strategies

3.1. Overview

This section presents strategies for a completely automatic workflow from a repository of historical images up to different 4D applications using web technologies and VR/AR. Some of the methods are already implemented, while others still have to be evaluated for achieving superior results (Fig. 3).

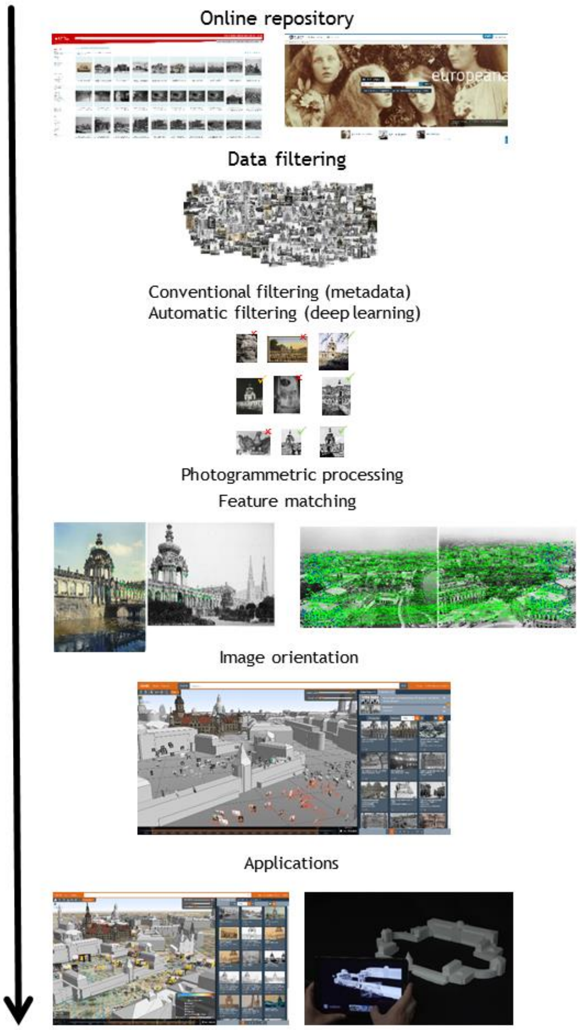

Figure 3: Overview of the presented workflow from an online repository to applications presenting informative models.

The starting point of the applications is an online repository containing mainly historical images but also drawings, maps, and plans. The workflow proceeds with the filtering of the relevant data. Images that are not desirable are, e.g., stereo images, close range and indoor scenes, and images, where people are the main motive.

Filtering can be carried out manually by keyword search and a further selection of several images, but due to the already stated problems when using metadata (see 2.2), a completely automated image retrieval is preferable. Thus, machine learning approaches are also used for filtering.

The resulting images are oriented using a broad range of photogrammetric methods. The main goal is the relative orientation of a large image mosaic which is already in the same (global) coordinate system as the application or has to be transformed using automatically retrieved parameters for a Helmert transformation. Therefore, a historical image data set has been created and is used for the evaluation of different feature detection and matching methods.

Oriented images, 3D city models, and maps are spatialized and made accessible in the 4D web browser application (http://4dbrowser.urbanhistory4d.org) of the research group where the user can jump into the perspective of the photographer. Additionally, different points in time with changing building models and maps can be visited. Thus, filtering of images can be carried out spatially and temporally leading to representation in 4 dimensions.

Different AR and VR applications are developed in the project where the oriented images and the 3D city models are used. HMD-based fully immersive VR with gamification aspects as well as handheld AR is employed to allow users free exploration of historical photography in a spatial setting.

3.2. Filtering

When working in large image repositories, a first step is to filter the images which are suitable for the use of photogrammetric methods. Especially close-up, indoor or portrait photographs should be neglected.

Filtering online collections is still considered a difficult task, because training a deep learning network often requires a large amount (500-1000 images per category) of training data (Deng et al., 2009). Whereas, for general object categories such as *buildings* a lot of training images exist, for a more concrete category like a specific building in a city, there is mostly no annotated data. A second problem is the changing nature and style of illustrations and pictures over different epochs, e.g. as drawings, paintings, photos, etc. Note that from the perspective of a neural net there is already a difference between pictures from a mobile phone and a normal camera.

Recent experiments with pre-trained networks lead to promising results, as these have to be slightly modified only in the upper layers. The advantage is three-fold: the training time is reduced drastically, the neural net has a structure that is already proven for the task and training is possible with smaller amounts of data. Thereby, the crucial requirement is to find an appropriate pre-trained network. For the application at hand, the pre-trained networks VGG16 and VGG19, which are deep convolutional neural nets for image classification, were used (Simonyan & Zisserman, 2014).

Only the dense layers on top of the network are modified and retrained, while the convolutional layers are used with their pre-trained weights. More precisely, on top of the pre-trained convolutional part of the network two dense layers were added: one consisting of 256 units (with a ReLU activation function) and the top layer consisting of one unit (with a sigmoid activation function). The next step could be a hyperparameter optimization, e.g. targeting the number of units in the penultimate layer.

The overall accuracy rate was approximately 70-80%. Nevertheless, one issue still is a high false-negative rate (approximately 40%), i.e. pictures of interest were classified as irrelevant. The current work in progress tackles especially this issue by dividing the whole pipeline into single modules that are optimized separately. For example, one module classifies a picture as "building/architecture" or "other". Within these pre-filtered pictures, another module classifies as "building of interest" or "irrelevant building".

One further important aspect is the assessment and quantification of uncertainty of the estimated accuracy- and error-rates, especially as the input datasets are small. This is realized with statistical methods such as resampling methods, e.g. bootstrapping (Efron & Tibshirani, 1994).

Through the whole optimization, it is important to make the decision process of the convolutional neural networks visible. This is realized by means of so-called heat maps, that identify areas of most relevance for the final classification decision. More precisely, classification activation maps are used (Selvaraju et al., 2017).

At the moment, the output of the neural network helps to classify the images, but a manual selection has to be performed afterward to remove the outliers, which are not useful for the following methods.

3.3. Photogrammetry

The filtered images need to be oriented and georeferenced in order to be placed in the applications. While it was possible to run a complete SfM procedure using Agisoft PhotoScan v. 1.3.1 for a small subset (11 of 499) of images, the major number of images from the repository could not be integrated into such a workflow (Maiwald et al., 2017). Most of the times the first step of orienting the images failed due to a lack of homolog points found. Other existing applications such as VisualSFM (Wu, 2013), Meshroom (Moulon, Monasse, & Marlet, 2012), and Agisoft Metashape produced similar results for various historical image sets (Fig. 4) as shown in Table 1.

(a) (b) (c)

Figure 4: Historical imagery of urban sites in Dresden, (Germany): a) Dinglinger House, b) Georges's Gate, and c) Crown Gate of the Zwinger.

Even though Agisoft Metashape and Meshroom were able to find a solution for the orientation of around half the images of the George's Gate a visual control of the cameras revealed that the calculated orientations were most of the times incorrect. A significantly smaller number of correct orientations (around 5) has to be assumed for this building, similar to the others. Sometimes tie points for multiple models have been created but could not be merged into a single building model.

Several reasons negatively effecting the keypoint detection and matching could be determined amongst others as image differences due to (Maiwald, 2019):

- original image medium;
- digitization;
- different camera models;
- acquisition time;
- acquisition technique,

resulting in image artefacts, scene occlusion, and radiometric variations.

Table 1: Comparison of three different SfM softwares used on three historical image data sets. The table shows respectively the used modes, average feature points found, tie points used for the creation of a sparse cloud and finally the number of oriented images in comparison to total images.

	George's Gate	*Zwinger*	*Dinglinger House*
Software	Agisoft Metashape (proprietary)		
Mode	Highest Accuracy, Generic Preselection		
Average feature points found	39127	214919	39354
Tie points (Sparse cloud)	3109	3027	1968
Oriented images	17/32	3/34	5/14
Software	Meshroom (open source)		
Mode	Ultra Accuracy, SIFT descriptor		
Average feature points found	4251	26642	3877
Tie points (Sparse cloud)	3651	4282	232
Oriented images	19/32	14/34	3/14
Software	VisualSFM (open source)		
Mode	Default Accuracy, SIFT descriptor		
Average feature points found	8966	8893	9187
Tie points (Sparse cloud)	272	128	49
Oriented images	2/32	4/34	2/32

Additionally, image pairs may have, on the one hand, very similar perspectives inconvenient for a photogrammetric reconstruction, or, on the other hand, large viewpoint changes hampering the common feature matching approach based on SIFT (Lowe, 2004) and SURF (Bay, Tuytelaars, & Van Gool, 2006). One approach for buildings, which *still exist*, could be the generation of a strong image network by using contemporary images. This increases the reliability and the number of matching points between the historical images and the recent images (Barazzetti, Erba, Previtali, Rosina, & Scaioni, 2013; Maiwald et al., 2017).

For destroyed or lost buildings, experiments using exclusively geometrical features show that orientation is possible for single image pairs if the feature detection and matching step is individually observed (Maiwald, Schneider, Henze, Münster, & Niebling, 2018).

Thus, the feature detection and feature matching step of the end-to-end SfM tools has to be broken up, since most of the times SIFT or a variation of it is used. Testing different feature matching methods requires a dataset with ground truth orientation. Since no dataset using exclusively historical images was available, an open source image dataset (https://dx.doi.org/10.25532/OPARA-24) oriented using the properties of the Trifocal Tensor was created (Maiwald, 2019) (Fig. 5).

Figure 5: Examples of image pairs of the proposed oriented dataset. The image pair in the top row (Hk_1/2_3) was very difficult to be matched, while the bottom pair (Mb_1/1_3) had 100% correct matches by two different methods.

Evaluating some already successful feature matching methods such as MSER (Matas, Chum, Urban, & Pajdla, 2004) and RIFT (Li, Hu, & Ai, 2018) showed that feature matching between historical image pairs is difficult but sometimes possible. It has been identified that most of the times the outlier removal method seems to be the crucial point. Especially, the outlier removal method RANSAC (Fischler & Bolles, 1981) can be improved if there is only a small number of *correct matches out of total matches* (also referred to as *matching score*). Alternatives are, e.g., PROSAC (Chum & Matas, 2005) or FSC (Wu, Ma, Gong, Su, & Jiao, 2015). The number of correct matches can be determined using the properties of three-view geometry and the already given Trifocal Tensor.

Recent results using MODS (Mishkin, Matas, & Perdoch, 2015) and RIFT with advanced outlier removal could improve the matching score significantly and thus, could be integrated into the image orientation of a SfM workflow (Fig. 6). Especially, the combination of both methods can be useful to find matches between a lot of different historical image pairs (Table 2).

It is planned to add single historical images consecutively, compare them to the already existing oriented images and try to find matches using the shown methods. The accuracy of the image orientation is improved by running a subsequent bundle adjustment. Since SfM is a scale-invariant method the retrieved image mosaics have to be scaled, generally performed with scale bars or field measurements. Usually, this is not possible when using historical images. Thus, in the web application, the relatively oriented image mosaic has to be adjusted to fit into the scene.

Table 2: Matching score (%) for image pairs of the proposed image data set processed using the different feature matching methods SIFT, RIFT and MODS.

Image pair	*SIFT*	*RIFT*	*MODS*	*RIFT+MODS*
So_1/1_3	41.89	0.00	23.36	23.36
Hk_1/2_3	0.00	25.00	0.00	25.00
So_2/2_3	0.00	25.00	10.00	25.00
So_1/2_3	50.01	0.00	31.76	31.76
So_2/1_2	0.88	0.00	36.36	36.36
Hk_1/1_3	9.62	37.50	0.00	37.50
Zw_1/2_3	53.72	50.00	39.71	50.00
Zw_2/2_3	0.08	62.50	18.18	62.50
Zw_2/1_2	0.00	69.26	46.39	69.23
Mb_2/2_3	0.00	72.70	17.31	72.70
Mb_1/ 2_3	5.10	80.43	0.00	80.43
Hk_2/1_2	4.30	0.00	83.64	83.64
Hk_2/2_3	3.80	84.21	53.85	84.21
Mb_2/1_2	1.40	0.00	88.88	88.88
Mb_2/1_3	1.40	0.00	92.31	92.31
Zw_1/1_3	59.35	83.33	96.97	96.97
So_2/1_3	41.30	70.59	98.68	98.68
Hk_1/1_2	8.77	100.00	42.31	100.00
Zw_2/1_3	11.85	100.00	52.17	100.00
Mb_1/1_2	2.60	100.00	70.00	100.00
So_1/1_2	75.06	100.00	81.73	100.00
Zw_1/1_2	84.87	100.00	93.22	100.00
Hk_2/1_3	1.40	84.62	100.00	100.00
Mb_1/1_3	3.90	100.00	100.00	100.00

One possible implementation is an interactive orientation using homolog points between a detailed 3D building model (possibly with an underlying historical map) and a single image. The orientation of the image can be estimated using a minimum of six points and the DLT (Hartley & Zisserman, 2003). In the following, the relative orientations can be used for placing the whole image mosaic.

Another approach could be the use of gamification elements implemented in the VR application (see Section 3.5). Users are engaged to locate the position of the photographer of a historical image in relation to a 3D model competitively against other users. The resulting image orientations could be averaged (including outlier removal) and used as a prior orientation for completely automatic workflow.

Approaches for automatic registration of 3D models and images are already investigated in many different studies (Callieri, Cignoni, Corsini, & Scopigno, 2008; Franken et al., 2005; Pani Paudel, Habed, Demonceaux, & Vasseur, 2015) and will be tested for reliability in the future.

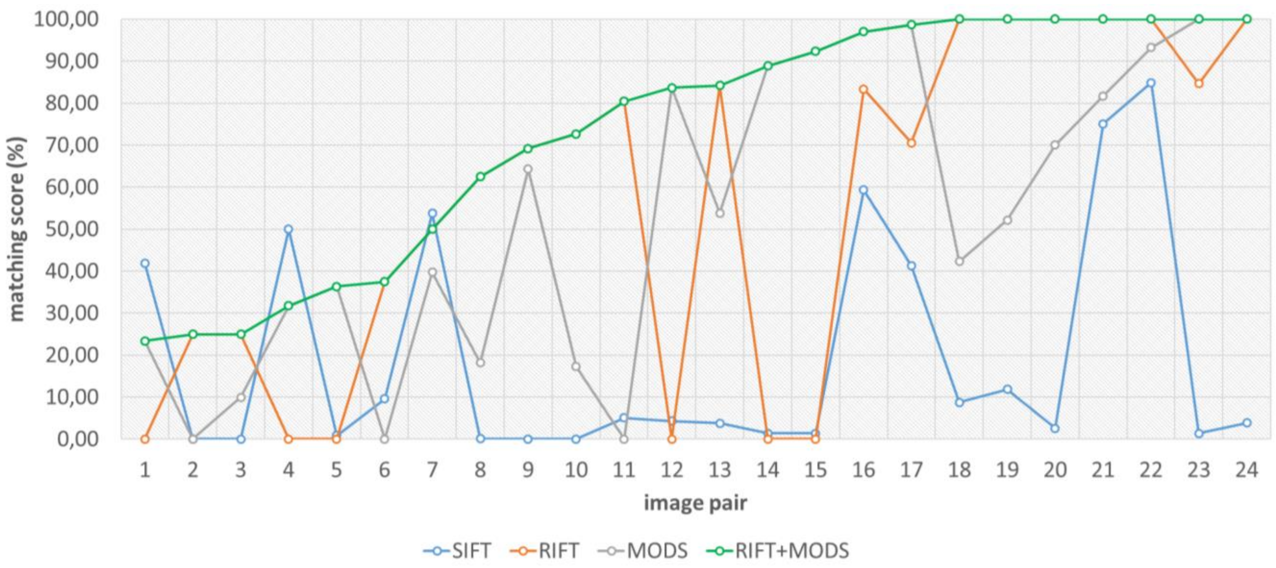

Figure 6: Ascending matching score for three different feature matching methods used on 24 image pairs of a benchmark dataset. The combination of RIFT and MODS showed the best results on the historical images.

Figure 7: Graphical user interface of the 4D browser prototype showing image clusters, a historical map, 3D models, a visualization of orientations and the listed search results.

3.4. Browser access

The registered images can be viewed in a 4D browser-based prototype application (http://4dbrowser.urbanhistory4d.org). As an entry point, the user is lead to the starting page including information on the project, the data and navigation through the platform. Additionally, a *help* site is accessible explaining all the available tools and features. When starting to explore the data, a contemporary 3D city model and a historical map are displayed in a 3D viewport that forms the main part of the graphical user interface (GUI). The viewport also shows all the images hierarchically clustered with respect to their location using single-linkage clustering (Fig. 7).

Observing a single image allows jumping into the perspective of the photographer using the predefined orientations. This enhances amongst others the understanding of the position of the photographer and allows the uncovering of historical details (Schindler & Dellaert, 2012).

The opacity of the images can be regulated revealing the underlying 3D building model and eventually detect changes in the object geometry (Bruschke et al., 2017). Further information (author, date, owner, etc.) belonging to the respective image can be displayed in an additional window. Multiple images can be added to a collection allowing comparison and, in the future, the user-dependent recording of scientific workflows.

Searching for special images can be done using a conventional search bar and the respective metadata. In addition, the browsing for images is enhanced by directly showing the output of the search request in the 3D viewport. Thus, the user is not only limited to a single output list used in classical online repositories, but is able to directly verify the request using the 3D view.

Furthermore, one can narrow the image search using the timeline on the bottom of the application adding the 4D aspect. The timeline is not only applicable for historical images, but also for the 3D models and for currently four underlying historical maps. Another search possibility is the filtering of images by depicted buildings by selecting the corresponding 3D objects. It is planned to generate additional 3D building models for different points in time manually and by using photogrammetric methods.

In regard to the usage by experts such as art and architectural historians, advanced statistical visualization methods can be applied to the images to support answering specific research questions (Bruschke, Niebling, & Wacker, 2018). This includes conventional heat maps but also innovative visualization methods that consider the orientation of the images (Fig. 8).

Figure 8: "Fan" visualization of image orientations in the 4D browser.

The browser application utilizes *AngularJS* as a basic front end web framework to build single-page applications (SPA). Regarding 3D graphics, *three.js* is used as a higher-level API for WebGL. The 3D models are stored in the *OpenCTM* format that features a high compression rate and enables fast transmission. Owing to the SPA approach, all loaded 3D content is kept in the background when the user switches to another subpage.

When switching back to the main interactive 3D view, the 3D content has not to be loaded again and can be displayed instantly. The back end is a REST API on the basis of *Node.js* and *Express.js*. It handles not only all database requests, but also routines to automatically retrieve metadata from the original image repository. The data itself is stored compliant to the CIDOC Conceptual Reference Model (CRM). This results in strongly connected data that is predestined to be stored not in a relational database but in a graph database, in this case *Neo4j* (Bruschke & Wacker, 2014).

3.5. VR and AR access

The registered images are also used in different prototypical AR and VR applications. In a HMD-based fully immersive VR environment users are challenged to find the specific locations and orientations of multiple images. One scenario for the dissemination of CH research results includes the use of gamification elements, where several users are encouraged to replicate the position and orientation of photographers of historical images faster than their opponents. After finding the positions of some already spatialized images with respect to a full-scale 3D model, the users could be motivated to help to find and improve the location of unregistered as well as only roughly registered images (Fig. 9).

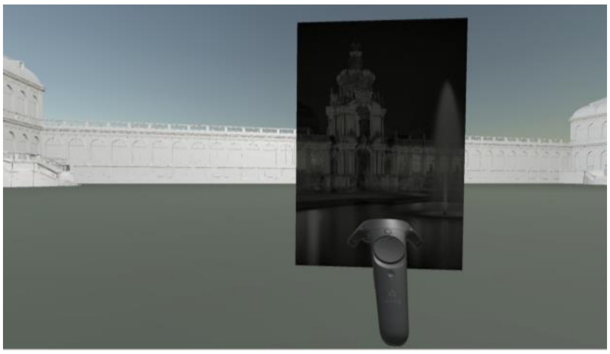

Figure 9: Fully immersive VR environment where users are challenged to find the location and orientation of historical images.

In addition, historic textures for registered models of architecture are created by UV mapping vertices of the respective 3D model to the registered historical photography in both VR and AR. The usage of photos as textures enables us to use coarse 3D models of buildings acquired from the city municipality, as details are provided by historical high-resolution photos.

The UV coordinates are calculated by casting rays from the known position *P* of the photographer to each vertex *V* in the model, through the projective plane of the currently processed image *I*. To verify visibility of vertices in the geometry of buildings from position *P*, we perform collision detection of rays with models in the scene. Parts of a building that are not contained in *I*, because they are either outside of the view frustum of the camera or hidden by other structures, cannot be textured with information contained within *I* and are left blank.

Employing photos as textures allow for the usage of very coarse 3D models of buildings, as details are provided by comparably high resolution photos.

In addition to the usage in VR environments, we use the introduced photo projection method to create AR installations combining 3D printed models of architecture with historical photography using 10.1" Android tablets (Niebling, Bruschke, & Latoschik, 2018). Tracking of the 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 (Fig. 10).

Figure 10: AR application blending registered images in front of a 3D printed model of the Zwinger.

In the AR installation, the textured digital model is rendered 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. The user is then still able to perceive the historical appearance of the visible parts of the building depicted in the historical image. In addition, the handheld AR devices can be used to spatially access and browse the catalog of images in AR allowing selection of photography according to the positions of the spatially registered documents. Pedagogical approaches to enhance user's experiences and knowledge transfer are currently investigated (Niebling, Maiwald, Barthel, & Latoschik, 2018).

4. Conclusions

The contribution shows in the first part that there are still a lot of different ways to access archival data. While digital methods are on the rise, the conventional analog access to repositories is nonetheless used for recent research. Especially, the issue that only a fraction of all museum objects is digitized at the moment enables the searching success only for experts of the respective topic.

But, access will expand in the future due to increasing digitization and the use of digital technologies. Hence, different access strategies from the digital online repository up to the virtual museum are presented and compared. An increased amount of Web3D and AR/VR exhibitions could be demonstrated. Recent examples show that advantages in these technologies are, e.g., preservation, protection, accessibility, sustainability, and reusability of the digitized data. 3D technologies and extended reality enhance the understanding of objects not only in cultural heritage.

The steps from a repository to these access methods are not only theoretically discussed. The contribution shows how a fully automatic workflow can look like and how it is prototypically developed in the project. Filtering a large amount of archival data is considered as a first step. While usually a manual selection of appropriate images is performed, this contribution also shows new technologies based on deep learning for the automatic choice of photographs. Historical images are very challenging for these technologies and thus, these methods have to be evaluated further but can be a suitable addition for completely automatic workflow.

Photogrammetric evaluation can be seen as the core of this research. The precise orientation of images is necessary for the different applications. For some image pairs, the orientation can be calculated automatically using end-to-end software packages such as Agisoft Metashape. Though, for many image combinations, the first step of the orientation fails due to the lack of homolog points in image pairs. Therefore, the use of advanced feature matching methods particularly MODS and RIFT is proposed. These methods performed superior in comparison to other approaches and serve as a basis for the perspective relative orientation of the images.

In the following, the global registration in relation to the 3D models can be carried out manually or using semi-automatic (DLT) and in the future, automatic methods.

The oriented images can be accessed in different applications. A 4D web browser is introduced as an alternative repository holding historical images, plans, buildings, and maps. It is possible to take the perspective of the photographer and view temporal changes of the city using 3D models and historical maps. The depicted visualization of orientations can be a useful tool for art and architectural historians.

Several examples on VR and AR applications provide different access methods for the images using a printed and a virtual 3D model enhancing the understanding of the scene. Ray casting can be used to texture the model using the historical image material and users are able to support the orientation of images in a prototypical gamification VR application.

Concluding, a completely automatic workflow from an image repository over filtering and photogrammetry to various access methods is outlined. In the future, the different parts should be connected in a single working environment or even attached to an existing repository to provide the different strategies to a broader range of users. While most of the examples were executed with archival data of the city of Dresden, the proposed methods could be integrated into other urban areas or even in rural sites combined with GIS.

Acknowledgments

The research upon which this paper is based is part of the junior research group UrbanHistory4D's activities which has received funding from the German Federal Ministry of Education and Research under grant agreement No 01UG1630.

This work was supported by the German Federal Ministry of Education and Research (BMBF, 01IS18026BA-F) by funding the competence center for Big Data "ScaDS Dresden/Leipzig".

References

Ackerman, A., & Glekas, E. (2017). Digital capture and fabrication tools for interpretation of historic sites. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, IV-2/W2*, 107–114. https://doi.org/10.5194/isprs-annals-IV-2-W2-107-2017

Armingeon, M., Komani, P., Zanwar, T., Korkut, S., & Dornberger, R. (2019). A case study: Assessing effectiveness of the augmented reality application in Augusta Raurica. In *Augmented Reality and Virtual Reality* (pp. 99–111). Cham: Springer. https://doi.org/10.1007/978-3-030-06246-0_8

Barazzetti, L., Erba, S., Previtali, M., Rosina, E., & Scaioni, M. (2013). Mosaicking thermal images of buildings. In *Videometrics, Range Imaging, and Applications XII; and Automated Visual Inspection*. Munich, Germany. https://doi.org/10.1117/12.2019985

Bay, H., Tuytelaars, T., & Van Gool, L. (2006). SURF: Speeded Up Robust Features. In A. Leonardis, H. Bischof, & A. Pinz (Eds.), *Computer Vision - ECCV 2006* (pp. 404–417). Berlin: Springer Heidelberg. https://doi.org/10.1007/11744023_32

Beaudoin, J. E., & Brady, J. E. (2011). Finding visual information: a study of image resources used by archaeologists, architects, art historians, and artists. *Art Documentation: Journal of the Art Libraries Society of North America, 30*(2), 24–36. https://doi.org/10.1086/adx.30.2.41244062

Beltrami, C., Cavezzali, D., Chiabrando, F., Iaccarino Idelson, A., Patrucco, G., & Rinaudo, F. (2019). 3D digital and physical reconstruction of a collapsed dome using SFM techniques from historical images. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W11*, 217–224. https://doi.org/10.5194/isprs-archives-XLII-2-W11-217-2019

Bevilacqua, M. G., Caroti, G., Piemonte, A., & Ulivieri, D. (2019). Reconstruction of lost architectural volumes by integration of photogrammetry from archive imagery with 3-D models of the status quo. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W9*, 119–125. https://doi.org/10.5194/isprs-archives-XLII-2-W9-119-2019

Bitelli, G., Dellapasqua, M., Girelli, V. A., Sbaraglia, S., & Tinia, M. A. (2017). Historical photogrammetry and terrestrial laser scanning for the 3D virtual reconstruction of destroyed structures: A case study in Italy. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-5/W1*, 113–119. https://doi.org/10.5194/isprs-archives-XLII-5-W1-113-2017

Bruschke, J., Niebling, F., Maiwald, F., Friedrichs, K., Wacker, M., & Latoschik, M. E. (2017). Towards browsing repositories of spatially oriented historic photographic images in 3D web environments. In *Proceedings of the 22nd International Conference on 3D Web Technology*. Brisbane, Australia. https://doi.org/10.1145/3055624.3075947

Bruschke, J., Niebling, F., & Wacker, M. (2018). Visualization of orientations of spatial historical photographs. In *Eurographics Workshop on Graphics and Cultural Heritage*. Vienna, Austria https://doi.org/10.2312/gch.20181359

Bruschke, J., & Wacker, M. (2014). Application of a Graph Database and Graphical User Interface for the CIDOC CRM. In *Access and Understanding–Networking in the Digital Era – CIDOC*. Dresden, Germany.

Burdea, G. C., & Coiffet, P. (2003). *Virtual reality technology*. John Wiley & Sons

Callieri, M., Cignoni, P., Corsini, M., & Scopigno, R. (2008). Masked photo blending: Mapping dense photographic data set on high-resolution sampled 3D models. *Computers & Graphics, 32*(4), 464–473. https://doi.org/10.1016/j.cag.2008.05.004

Chum, O., & Matas, J. (2005). *Matching with PROSAC – progressive sample consensus.* In *2005 IEEEC on Computer Vision and Pattern Recognition (CVPR)*. https://doi.org/10.1109/CVPR.2005.221

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*. https://doi.org/10.1109/CVPR.2009.5206848

Efron, B., & Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. Boca Raton: CRC press.

Evens, T., & Hauttekeete, L. (2011). Challenges of digital preservation for cultural heritage institutions. *Journal of Librarianship and Information Science, 43*(3), 157–165. https://doi.org/10.1177/0961000611410585

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM, 24*(6), 381–395. https://doi.org/10.1145/358669.358692

Fleming-May, R. A., & Green, H. (2016). Digital innovations in poetry: Practices of creative writing faculty in online literary publishing. *Journal of the Association for Information Science and Technology, 67*(4), 859–873. https://doi.org/10.1002/asi.23428

Franken, T., Dellepiane, M., Ganovelli, F., Cignoni, P., Montani, C., & Scopigno, R. (2005). Minimizing user intervention in registering 2D images to 3D models. *The Visual Computer, 21*(8-10), 619–628. https://doi.org/10.1007/s00371-005-0309-z

Girardi, G., von Schwerin, J., Richards-Rissetto, H., Remondino, F., & Agugiaro, G. (2013). The MayaArch3D project: A 3D WebGIS for analyzing ancient architecture and landscapes. *Literary and Linguistic Computing, 28*(4), 736–753. https://doi.org/10.1093/llc/fqt059

Grussenmeyer, P., & Al Khalil, O. (2017). From metric image archives to point cloud reconstruction: Case study of the Great Mosque of Aleppo in Syria. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W5*, 295–301. https://doi.org/10.5194/isprs-archives-XLII-2-W5-295-2017

Gutierrez, M., Vexo, F., & Thalmann, D. (2008). *Stepping into Virtual Reality*. London: Springer Science & Business Media.

Guttentag, D. A. (2010). Virtual reality: Applications and implications for tourism. *Tourism Management, 31*(5), 637–651. https://doi.org/10.1016/j.tourman.2009.07.003

Hartley, R., & Zisserman, A. (2003). *Multiple View Geometry in Computer Vision*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511811685

Koutsoudis, A., Arnaoutoglou, F., Tsaouselis, A., Ioannakis, G., & Chamzas, C. (2015). Creating 3D replicas of medium-to large-scale monuments for web-based dissemination within the framework of the 3D-Icons project. In *CAA2015* (pp. 971–978).

Li, J., Hu, Q., & Ai, M. (2018). RIFT: Multi-modal Image Matching Based on Radiation-invariant Feature Transform. *arXiv preprint arXiv:1804.09493*

Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision, 60*(2), 91–110. https://doi.org/10.1023/B:VISI.0000029664.99615.94

Maietti, F., Di Giulio, R., Piaia, E., Medici, M., & Ferrari, F. (2018). Enhancing heritage fruition through 3D semantic modelling and digital tools: the INCEPTION project. In *IOP Conference Series: Materials Science and Engineering, 364*. https://doi.org/10.1088/1757-899X/364/1/012089

Maiwald, F. (2019). Generation of a benchmark dataset using historical photographs for an automated evaluation of different feature matching methods. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2/W13*, 87–94. https://doi.org/10.5194/isprs-archives-XLII-2-W13-87-2019

Maiwald, F., Henze, F., Bruschke, J., & Niebling, F. (2019). Geo-information technologies for a multimodal access on historical photographs and maps for research and communication in urban history. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XLII-2/W11*, 763–769. https://doi.org/10.5194/isprs-archives-XLII-2-W11-763-2019

Maiwald, F., Schneider, D., Henze, F., Münster, S., & Niebling, F. (2018). Feature matching of historical images based on geometry of quadrilaterals. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLII-2*, 643–650. https://doi.org/10.5194/isprs-archives-XLII-2-643-2018

Maiwald, F., Vietze, T., Schneider, D., Henze, F., Münster, S., & Niebling, F. (2017). Photogrammetric analysis of historical image repositories for virtual reconstruction in the field of digital humanities. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 42*, 447–452. https://doi.org/10.5194/isprs-archives-XLII-2-W3-447-2017

Matas, J., Chum, O., Urban, M., & Pajdla, T. (2004). Robust wide-baseline stereo from maximally stable extremal regions. *Image and Vision Computing, 22*(10), 76–767. https://doi.org/10.1016/j.imavis.2004.02.006

Melero, F. J., Revelles, J., & Bellido, M. L. (2018). Atalaya3D: making universities' cultural heritage accessible through 3D technologies. In R. Sablatnig & M. Wimmer (Eds.), *EUROGRAPHICS Workshop on Graphics and Cultural Heritage*. https://doi.org/10.2312/gch.20181338

Milgram, P., Takemura, H., Utsumi, A., & Kishino, F. (1995). Augmented reality: A class of displays on the reality-virtuality continuum. In *Proceedings SPIE 2351, Telemanipulator and telepresence technologies*. https://doi.org/10.1117/12.197321

Mishkin, D., Matas, J., & Perdoch, M. (2015). MODS: Fast and robust method for two-view matching. *Computer Vision and Image Understanding, 141*, 81–93. https://doi.org/10.1016/j.cviu.2015.08.005

Moulon, P., Monasse, P., & Marlet, R. (2012). Adaptive structure from motion with a contrario model estimation. In *Asian Conference on Computer Vision* (pp. 257–270). https://doi.org/10.1007/978-3-642-37447-0_20

Münster, S., Kamposiori, C., Friedrichs, K., & Kröber, C. (2018). Image libraries and their scholarly use in the field of art and architectural history. *International journal on digital libraries, 19*(4), 367–383. https://doi.org/10.1007/s00799-018-0250-1

Niebling, F., Bruschke, J., & Latoschik, M. E. (2018). Browsing spatial photography for dissemination of cultural heritage research results using augmented models. In R . Sablatnig & M. Wimmer (Eds.), *Eurographics Workshop on Graphics and Cultural Heritage*. https://doi.org/10.2312/gch.20181358

Niebling, F., Maiwald, F., Barthel, K., & Latoschik, M. E. (2018). 4D augmented city models, photogrammetric creation and dissemination. In S. Münster, K. Friedrichs, F. Niebling, A. Seidel-Grzesińska (Eds.), *Digital Research and Education in Architectural Heritage* (vol 817, pp. 196–212). UHDL 2017, DECH 2017. Communications in Computer and Information Science. Cham: Springer. https://doi.org/10.1007/978-3-319-76992-9_12

Oliva, L. S., Mura, A., Betella, A., Pacheco, D., Martinez, E., & Verschure, P. (2015). Recovering the history of Bergen Belsen using an interactive 3D reconstruction in a mixed reality space the role of pre-knowledge on memory recollection. In *2015 Digital Heritage*. https://doi.org/10.1109/DigitalHeritage.2015.7413860

Pani Paudel, D., Habed, A., Demonceaux, C., & Vasseur, P. (2015). Robust and optimal sum-of-squares-based point-to-plane registration of image sets and structured scenes. In *2015 IEEE International Conference on Computer Vision*. https://doi.org/10.1109/ICCV.2015.237

Ross, S., & Hedstrom, M. (2005). Preservation research and sustainable digital libraries. *International Journal on Digital Libraries, 5*(4), 317–324. https://doi.org/10.1007/s00799-004-0099-3

Schindler, G., & Dellaert, F. (2012). 4D Cities: Analyzing, visualizing, and interacting with historical urban photo collections. *Journal of Multimedia, 7*(2), 124–131.

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In *2017 IEEE International Conference on Computer Vision*. https://doi.org/10.1109/ICCV.2017.74

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*

Slater, M., & Sanchez-Vives, M. V. (2016). Enhancing our lives with immersive virtual reality. *Frontiers in Robotics and AI, 3*, 74. https://doi.org/10.3389/frobt.2016.00074

Styliani, S., Fotis, L., Kostas, K., & Petros, P. (2009). Virtual museums, a survey and some issues for consideration. *Journal of Cultural Heritage, 10*(4), 520–528. https://doi.org/10.1016/j.culher.2009.03.003

Tschirschwitz, F., Büyüksalih, G., Kersten, T., Kan, T., Enc, G., & Baskaraca, P. (2019). Virtualising an Ottoman fortress - Laser scanning and 3D modelling for the development of an interactive, immersive virtual reality application. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 42*(2/W9), 723–729. https://doi.org/10.5194/isprs-archives-XLII-2-W9-723-2019

Wu, C. (2013). Towards linear-time incremental structure from motion. In *2013 International conference on 3D Vision*. Seattle, USA. https://doi.org/10.1109/3DV.2013.25

Wu, Y., Ma, W., Gong, M., Su, L., & Jiao, L. (2015). A novel point-matching algorithm based on fast sample consensus for image registration. *IEEE Geoscience Remote Sensing Letters, 12*(1), 43–47. https://doi.org/10.1109/LGRS.2014.2325970

Yoon, J., & Chung, E. (2011). Understanding image needs in daily life by analyzing questions in a social Q&A site. *Journal of the American Society for Information Science and Technology, 62*(11), 2201–2213. https://doi.org/10.1002/asi.21637