



# Image based rendering of large historical image collections

Evelyn Paiz-Reyes

## ► To cite this version:

Evelyn Paiz-Reyes. Image based rendering of large historical image collections. Computer Graphics Forum, May 2019, Genova, Italy. hal-02301561

**HAL Id: hal-02301561**

**<https://hal.science/hal-02301561>**

Submitted on 5 Nov 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Image Based Rendering of Large Historical Image Collections

Evelyn Paiz-Reyes 

Univ. Paris-Est, LASTIG GEOVIS, IGN, ENSG, F-94160 Saint-Mande, France

## Abstract

*This paper states an overview of my dissertation research centered on the continuous immersive visualization and navigation through time and space of large sets of historical photographs. The research aims for: (i) the treatment of scientific obstacles (e.g. data volume, heterogeneity, distortions, and uncertainties) that appear when old pictures are placed in today's environment; (ii) the visualization (saliently and spatially) of these photos. The main model of the study is image-based rendering IBR, because of its capacity to use imprecise or non-existent geometry (i.e. since a modern 3D scene may differ from a historical one, due to environmental changes over time). The findings of this work may contribute significantly by extending the current IBR models and providing a new innovative way to examine these massive and heterogeneous datasets.*

## CCS Concepts

• **Computing methodologies** → Image-based rendering; • **Information systems** → Geographic information systems; Digital libraries and archives; • **Human-centered computing** → User interface design;

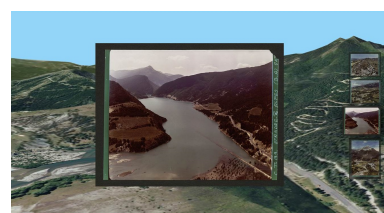
## 1. Introduction

I am fascinated by the way an image captures and preserves perpetually a piece of time. Since it is not possible to go back to a specific moment again, we may be able to create a virtual environment for it. For this reason, I chose to study Engineering in Computer Science followed by a Master in Color Science. At this moment, my research interests center on Computer Graphics, Image Processing, Photogrammetry, and Geographic Information Sciences. My dissertation research started on the 1<sup>st</sup> of October 2018 for three years long, with an expected graduation date in October 2021. From the academic point of view, the Ph.D. depends on the Doctoral School of Mathematics and Information Technology and Communication Sciences MSTIC from Paris-Est University. My research work is conducted at IGN, the French National Mapping Agency, within the GeoVIS team of the LaSTIG lab (Laboratory of Geographic Information Sciences). The focus of the team is the existing subjects related to the **geovisualization** and **interaction** with **spatial data** for **visual spatio-temporal analytics**.

The study is funded by the French Research Agency ANR, in the framework of the ALEGORIA ANR-17-CE38-0014-01 (Advanced Linking and Exploitation of digitized geOgraphic Iconographic heritAge) project (2018-2022). Its purpose is the **valorization** of **large iconographic collections** depicting the French territory at different times from the interwar period to the present day. The datasets are digitized partially, generally poorly documented and weakly georeferenced, but they are very rich in terms of context and represent a strong heritage. These images consist of hundreds of thousands of photographs and postcards from oblique aerial

imagery as well as terrestrial/street-level acquisitions, scattered within different institutions (showed in Table 1).




To handle the massive amount of photographs and allow access to all participating establishments, the project is intended to be developed using an online context (i.e. a web application). As a whole, its focus goes to the indexing, interlinking, and visualization of these historical collections (i.e. the lastest, being the center of the Ph.D.). Archivists and researchers in humanities are the user targets of the final prototype. Specifically, this dissertation research will study the rendering process to enable a **spatio-temporal, immersive** and **interactive navigation** of the 3D environment that is enriched with old photographs (see Figure 1).



**Figure 1:** Mockup of an immersive co-visualization of an image from Archives Nationales/Fonds LAPIE on a WebGL application.

## 2. Problem Statement

During the years many large collections of historical datasets (e.g. postcards, engravings, paintings, street level or aerial photographs,

Institution	Collection	Type	Context	Point of View	Date	No. Images	Example
National Archives	Carto MRU	image	topographical	oblique	1948-1970	13562	
	Lapie		commercial	oblique with medium-low altitude	1955-1965	29140	
Niéphore Niépce Museum	Combier	image	postcard	randomly	1949-1974	10000	
	Bouquet		army	vertical	1914-1918	49	
	Bron			vertical and oblique	1924-1926	130	
	Falba		aerial	vertical	1951-1960	550	
	Henrad			oblique	1930-1970	1100	
	Franjus			oblique	1960-1990	877	
IGN	Stereopolis	image + lidar (pointcloud)	mobile mapping	terrestrial, panoramic and stereo	2007-2016	8280000	
	BD TOPO	geometry	3D vectorial description	vertical	1998-2005	-	

**Table 1:** Overall description of the historical image collections of the ANR ALEGORIA project covering the whole French territory.

etc.) have been assembled with the use of capturing devices. The digital era brought with it new practices to allow the access and archiving of these pictures (i.e. even remotely). To examine them, the most widespread and known approach is the use of a photo library for the listing and displaying of the photographs. A simple 2D grid is used for the navigation and handled through different metadata (e.g. keywords, categories, dates, etc.). Nevertheless, this traditional technique does not present sufficient context to give users an immersive feeling or for them to fully appreciate the spatial relationship between images (i.e. in terms of locations, dates or any other semantic aspect). Therefore, we propose to facilitate the discovery and continuous navigation within these collections through a spatio-temporal approach (i.e. exploration in a 3D environment).

It is necessary to have in mind that datasets of historical images are a special input. They are likely to present sparsely sampled results as a consequence of: (i) a diversity of existing sources (e.g. old photographs, postcards, acquisitions of mobile cartography, aerial photographs, etc.); (ii) temporalities (e.g. diachronism and evolutions of the scene); (iii) conditions of acquisition (e.g. photographic device used, scanning, illumination, etc.). Accordingly, the rendering of a scene for a historical picture cannot be based on a current 3D model because of its contrast in appearance due to environmental changes over time. This indicates that the 3D model may not be existent or its 3D reconstruction may be impracticable due to insufficient data. Therefore, the objective of this Ph.D. research project is to accomplish a **real-time immersive rendering** engine that offers **interactive** and **continuous navigation in large heterogeneous collections of historical images with uncertain geolocation**. In summary, the main contributions will address the following issues:

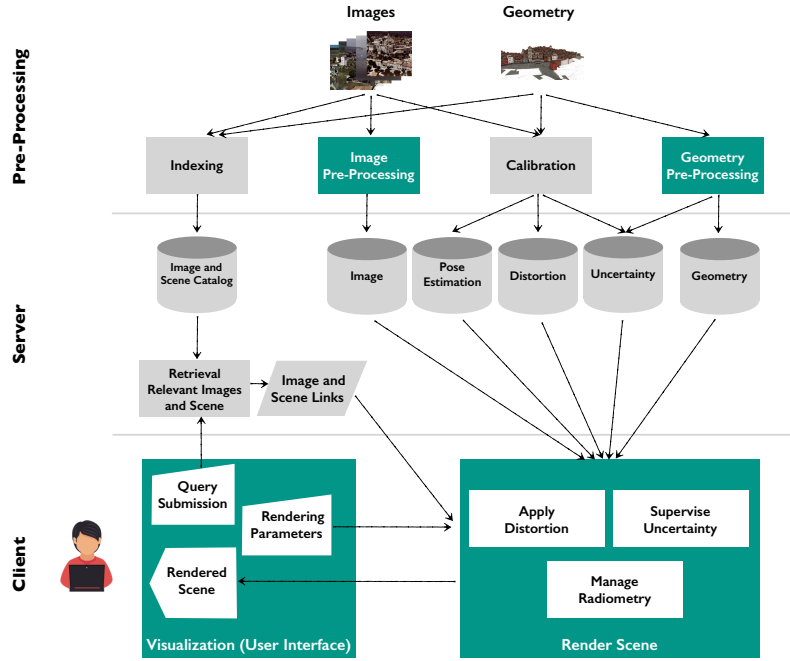
- A photographic process may exhibit small to severe **distortion**. Rendering methods need to reproject the input image taking into account these distortions (i.e. changing the image to fit the scene). Historical pictures specifically should not be deformed due to their cultural value. To leave the image untouched, a new visualization can be reached by applying the distortion on the surrounding scene instead.

- **Uncertainties** (e.g. on relative or absolute poses, on distortions, on 3D geometries, on view and texture cameras, etc.) may hinder an accurate co-registration and calibrations between photographs and the 3D scene. To avoid the resulting misalignments, uncertainty needs to be taken into account during the rendering process.
- Color between one photograph and the other may not be directly analogous. Heterogeneous **radiometry** (e.g. manipulation of pixel values) can be applied to homogenize, accentuate or highlight the pictures or certain areas inside the scene.
- The rendering environment must be interactive for the users. The **visualization**, of one image or the complete massive dataset, needs to be easily navigated and presented. A human-computer interaction approach may be followed for the user inputs and queries.

### 3. Related Work

On the context of cultural heritage, it does exist some advanced work for the exploration and browsing of photographs. The ANR project Poeme [poe] concentrates on an immersive 3D interactive system but its objective is museographic devices. Historypin [his] is an existing online tool for street view of photographs/postcards. Users are allowed to integrate their old photographs and geolocalize them manually in a consistent location using Google Street View. However, this application does not provide distortion or proper alignment with today's environment. Smapshot [sma] follows the same idea of using users input for the georeferencing of images in Swiss territory. Although the system allows smooth navigation through the photographs, the final scene only projects the selected image without any mix between overlapping images. Photocloud [BBT\*13] offers a real-time interactive navigation in large datasets. This application allows a hybrid visualization of 3D models along with thousands of calibrated photographs using a thumbnail bar for user navigation. Nevertheless, its main focus is in the management of large collections. The method does not use any sophisticated rendering technique to project the images in the 3D scene.

To our knowledge, **navigation** through photographs with **heterogeneous spatial sampling** and very **wide time scale** in



**Figure 2:** Visual depiction of project scope highlighted in color: (i) distortion; (ii) uncertainty; (iii) radiometry; (iv) visualization.

a real 3D environment has been little explored. One particular possibility is to analyze it through image-based rendering IBR methods [LH96, GGSC96], which exploits a relatively heterogeneous and dense set of pictures to obtain synthesized novel views from the morphing and blending of the input images [HRDB16]. One of the main advantages of this type of rendering is the fact that it may cope with little or no geometrical information [BBM\*01]. The novel view may be directed displayed and as the synthesized viewpoint departs from the viewpoint of the input photo, its result degrades gracefully [GAF\*10]. Additionally, these models can be extended to take other aspects into account (e.g. uncertainty [EDDM\*08, Bre14] or heterogeneous datasets with point clouds and simplified meshes [DB16]).

#### 4. Research Approach and Current Progress

This research work is focused on the efficient exploration and easy browsing of the photograph collections. Here it is proposed to carry out this study in four main phases: **(i) distortion;** **(ii) uncertainty;** **(iii) radiometry;** **(iv) visualization.** To achieve the final project prototype, additional modules will be developed by the ALEGORIA project contributors (see Figure 2). To optimize the rendering process and have real-time interaction a pre-processing step is likely to be necessary, both on the images and 3D geometries. Already existing methods (e.g. thresholding, filtering, edge enhancement, etc.) and self-created ones will be applied. For uncalibrated images (i.e. historical photographs are normally only scanned, camera information is not known) a pose estimation is going to be required. To manage queries from the user, context-based indexing and retrieval will be proposed.

### 4.1. Distortion

A traditional graphics pipeline commonly uses a pinhole camera model, which is natively pure projective (i.e. distortion-free). Nevertheless, real-world cameras have a diversity of effects (e.g. distortion, chromatic aberration, depth of field, etc.) that may arise during the photographic process. Therefore, to render an input image (i.e. reprojection of the picture), it is necessary to fit the real-world data into the pinhole model. Focusing solely on image distortion, this can be achieved following two different paths: (i) through the image; (ii) through the scene.

The first method corrects particularly the photograph. This means that the distortion is applied to the image for it to fit the displayed scene. On the other hand, the second approach extrapolates the distortion of the picture to the whole synthesized view. The image will then be displayed without any correction of its distortion because the distortion will be applied to the entire scene instead. This second scheme will be followed to achieve a non-perspective distortion that is sought in the current historical datasets. As the general public, archivists and researchers in humanities (i.e. which are the user target of the research prototype) are used to look at the unaltered rectangular version of the originally distorted image, rather than undistorted versions. The method of [LD09] is able to achieve aliasing free and real-time non-perspective rendering of a dynamic scene. As their approach requires geometry shaders, it is not directly applicable in our web context. Furthermore, the distortion is only defined by the calibration process within the distorted image boundaries. An extrapolation method for the distortion will thus have to be proposed.

## 4.2. Uncertainty

Rendering artifacts may arise when imprecise data (i.e. calibration, geometry, distortion or camera placement) is used on an image-based rendering process. This is caused by the misalignment between images and geometry. In order to decrease this issue, uncertainty must be taken into account. Extending the work of [EDDM\*08, GAF\*10, Bre14], the following uncertainties will be studied:

**Calibration:** even if it is not possible to know an absolute position during pose estimation, relative orientation are likely to be estimated with better accuracy. The calibration of the camera synthesizing the view can be judged as stochastic, just like those of the pictures to be visualized. The target visualization will be then defined as the expectation of the rendering of these stochastic calibrations. This method should make it possible to take benefit of uncertain but non-independent calibrations, viewing them as blocks of images finely calibrated in relative but grossly in absolute.

**Camera:** the viewing camera is capable of taking the position of the image. For this, it is necessary that the image-based rendering technique is only affected by uncertainties relative to the viewing camera. This will allow the possibility to render an uncertain image from its own view and obtain an unmodified image.

## 4.3. Radiometry

Once different images to be co-visualized are put in place geometrically, they may have very different radiometries (i.e. contrast, color or pixel value) due to differences in shooting and digitalization (e.g. lighting, camera, film, etc.), even without any change in the scene. These discrepancies may be desirable (e.g. emphasizing or differentiating a photograph in relation to its context) or not (e.g. when the perception of the individual photographic object is not sought). Here it will be studied real-time context alteration with the possibility of pre-treatment of the images for both cases. Coloring techniques, gradient-based IBR [BHST17], histogram transfer [TC17] and soft z-buffer with alpha normalization [GAF\*10] are going to be considered. In the end, the objective is to allow continuous rendering from the original radiometry of one or more images to a merged picture rendered through the masking of the radiometric heterogeneities of these images.

## 4.4. Visualization

The immersive rendering system needs to be interactive through the browsing of the large datasets (i.e. comprising of thousands of pictures), with a continuous, integrated and seamless transition through the images. Based on the [BBT\*13], the visualization of the photographs could be represented using carousels and framelets. This may be continued by exploring the clustering of the images for a broader overview of the samples (i.e. clustering the pictures when viewing on the top or a map and displaying the carousels and framelets for a street view). The objective is to reach a human interactive technique in which different experienced users will be able to benefit from its effectiveness.

## 5. Status and Future Directions

I am currently in my first year of PhD and at this moment I have managed to identify and define the scope of my research project. My center on the study right now is phase one, which is the control of the image distortions in the rendering process.

## References

- [BBM\*01] BUEHLER C., BOSSE M., MCMILLAN L., GORTLER S., COHEN M.: Unstructured lumigraph rendering. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques - SIGGRAPH '01* (2001), ACM Press, pp. 425–432. doi:10.1145/383259.383309. 3
- [BBT\*13] BRIVIO P., BENEDETTI L., TARINI M., PONCHIO F., CIGNONI P., SCOPIGNO R.: PhotoCloud: Interactive remote exploration of joint 2D and 3D datasets. *IEEE Computer Graphics and Applications* 33, 2 (2013), 86–96, c3. doi:10.1109/MCG.2012.92. 2, 4
- [BHST17] BARICEVIC D., HOLLERER T., SEN P., TURK M.: User-perspective AR magic lens from gradient-based IBR and semi-dense stereo. *IEEE Transactions on Visualization and Computer Graphics* 23, 7 (2017), 1838–1851. doi:10.1109/TVCG.2016.2559483. 4
- [Bre14] BREDIF M.: Projective texturing uncertain geometry: silhouette-aware box-filtered blending using integral radial images. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences II-3* (2014), 17–23. doi:10.5194/isprsannals-II-3-17-2014. 3, 4
- [DB16] DEVAUX A., BREDIF M.: Realtime projective multi-texturing of pointclouds and meshes for a realistic street-view web navigation. In *Proceedings of the 21st International Conference on Web3D Technology - Web3D '16* (2016), ACM Press, pp. 105–108. doi:10.1145/2945292.2945311. 3
- [EDDM\*08] EISEMANN M., DE DECKER B., MAGNOR M., BEKAERT P., DE AGUIAR E., AHMED N., THEOBALT C., SELLENT A.: Floating textures. *Computer Graphics Forum* 27, 2 (2008), 409–418. doi:10.1111/j.1467-8659.2008.01138.x. 3, 4
- [GAF\*10] GOESELE M., ACKERMANN J., FUHRMANN S., HAUBOLD C., KLOWSKY R., STEEDLY D., SZELISKI R.: Ambient point clouds for view interpolation. *ACM Trans. Graph.* 29, 4 (2010), 95:1–95:6. doi:10.1145/1778765.1778832. 3, 4
- [GGSC96] GORTLER S. J., GRZESZCZUK R., SZELISKI R., COHEN M. F.: The lumigraph. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques - SIGGRAPH '96* (1996), SIGGRAPH '96, ACM, pp. 43–54. doi:10.1145/237170.237200. 3
- [his] Historypin. URL: <https://www.historypin.org/en/>. 2
- [HRDB16] HEDMAN P., RITSCHER T., DRETTAKIS G., BROSTOW G.: Scalable inside-out image-based rendering. *ACM Transactions on Graphics* 35, 6 (2016), 1–11. doi:10.1145/2980179.2982420. 3
- [LD09] LORENZ H., DOLLNER J.: Real-time piecewise perspective projections. In *Proceedings of the Fourth International Conference on Computer Graphics Theory and Applications* (2009), pp. 147–155. 3
- [LH96] LEVOY M., HANRAHAN P.: Light field rendering. In *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques - SIGGRAPH '96* (1996), ACM Press, pp. 31–42. doi:10.1145/237170.237199. 3
- [poe] ANR POEME project. URL: <http://poeme.photo.2>
- [sma] Snapshot. URL: <https://snapshot.heig-vd.ch/>. 2
- [TC17] TIAN Q.-C., COHEN L. D.: Histogram-based color transfer for image stitching. *Journal of Imaging* 3, 3 (2017), 38. doi:10.3390/jimaging3030038. 4