

ALEGORIA: Joint Multimodal Search and Spatial Navigation into the Geographic Iconographic Heritage

Florent Geniet
LASTIG, Univ. G. Eiffel, IGN ENSG, France
florent.geniet@ign.fr

Valérie Gouet-Brunet
LASTIG, Univ. G. Eiffel, IGN ENSG, France
valerie.gouet@ign.fr

Mathieu Brédif
LASTIG, Univ. G. Eiffel, IGN ENSG, France
mathieu.bredif@ign.fr

ABSTRACT

In this article, we present two online platforms developed for the structuring and valorization of old geographical iconographic collections: a multimodal search engine for their indexing, retrieval and interlinking, and a 3D navigation platform for their visualization in spatial context. In particular, we show how the joint use of these functionalities, guided by geolocation, brings structure and knowledge to the manipulated collections. In the demonstrator, they consist of 54,000 oblique aerial photographs from several French providers (national archives, a museum and a mapping agency).

CCS CONCEPTS

• **Information systems** → **Retrieval models and ranking**; • **Human-centered computing** → **Visualization techniques**; • **Applied computing** → *Arts and humanities*.

KEYWORDS

image retrieval, HCI, geolocation, heritage, digital humanities

ACM Reference Format:

Florent Geniet, Valérie Gouet-Brunet, and Mathieu Brédif. 2022. ALEGORIA: Joint Multimodal Search and Spatial Navigation into the Geographic Iconographic Heritage. In *Proceedings of the 30th ACM International Conference on Multimedia (MM '22)*, October 10–14, 2022, Lisboa, Portugal. ACM, New York, NY, USA, 3 pages. <https://doi.org/10.1145/3503161.3547746>

1 INTRODUCTION

The ALEGORIA project [10] aims at facilitating the promotion of iconographic institutional collections describing the French territory, from the interwar period up to nowadays. The collections are of variable sizes, between thousands and hundred thousand elements, which are oblique, vertical aerial or terrestrial photographs. Contrary to the well structured regular surveys in satellite imagery, the promotion of those collections remains confidential and scattered. They are generally spread among various institutions, heterogeneously documented and weakly georeferenced, while they represent a rich heritage, little known by the general public and exploited in a way forced by their main users (researchers, institutions and local authorities), on site at the library or via standard online digital libraries. In order to improve their structuring, we have implemented two web-based platforms dedicated to the promotion of this heritage: a multimodal and large-scale search engine, coupling

search by content and by metadata within collections (section 2); a visualization engine for their immersive restitution in their spatial context, allowing spatio-temporal navigation and interaction in the 3D environment (section 3). In section 4, we highlight the interactions between these platforms to improve the content geolocalization, for their better structuring and restitution in context, while section 5 gives an insight into several use cases.

Collections. We managed more than 90,000 iconographic contents in ALEGORIA, coming from several institutions (French national archives, French mapping agency and Niépce Museum), which offer a bird's eye view of the French territory from the 20's to the 70's (links to the [description](#) and [spatial distribution](#) of the collections). In the demo, a subset of 54,000 views (called DB-54k in the following) is indexed in the search engine; they share some common locations, 2/3 of them are associated with metadata and various levels of geolocation (no information up to a 6-DoF pose).

2 MULTIMODAL SEARCH ENGINE

In order to structure the collections, and to access them more easily, we have developed a web-based multimodal search engine, which indexes both the content of the images and the metadata associated when available. As they come from several suppliers having different missions, the metadata were harmonized and extended according to the RiC-O Ontology [7]. The content-based part of the indexing had to deal with image descriptors tuned to the contents considered. We have studied generalizable and robust deep descriptors for the problem of low-data, heterogeneous image retrieval, as well as post-processings to improve retrieval, such as geometric verification and query expansion [5]. These modalities are made available to query one or several collections together, by combining metadata and image content search (through the query-by-example paradigm and various image descriptors). Fig. 1 shows the web client and an example of content-based retrieval. The search engine is client-server based, with a distributed server architecture that makes it scalable in terms of database size and descriptor complexity. We use the distributed OS DC/OS from Mesos, where every service is a docker container, that communicates with other services via an internal network, and DC/OS deals with instance distribution.

3 VISUALIZATION IN SPATIAL CONTEXT

One ambition of ALEGORIA was to show the user the visual content in its spatial environment. The proposed geovisualization platform is based on the open source web-based iTowns library [8] which leverages WebGL and OGC standards for accessing and visualizing 3D geospatial datasets [4]. By accessing geographic web services, it renders the Earth with multimodal geospatial data into a common 3D reference model, anywhere where data is available, with a high precision, at large scale. New functionalities were developed within

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

MM '22, October 10–14, 2022, Lisboa, Portugal

© 2022 Copyright held by the owner/author(s).

ACM ISBN 978-1-4503-9203-7/22/10.

<https://doi.org/10.1145/3503161.3547746>

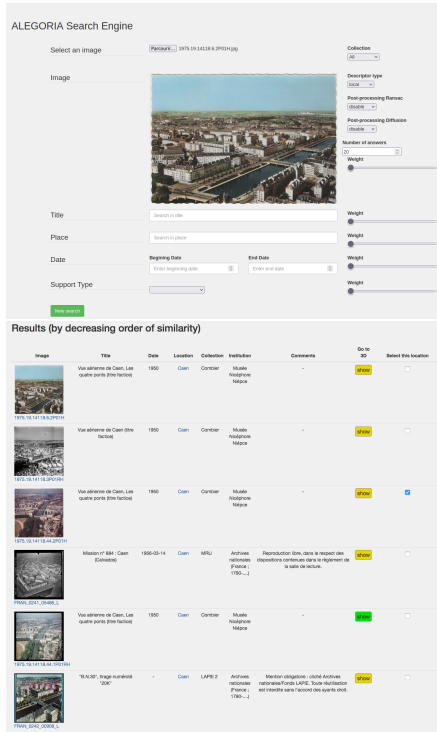


Figure 1: ALEGORIA search engine interface. The image on top was used as query by content, using local descriptors to search photographs of the same scene (Caen city). The k first results are presented with main metadata, by decreasing order of similarity; they show the same area (column "Location"), from different sources (columns "Collection" and "Institution"). Credits: Niepce Museum and French national archives.

ALEGORIA, such as navigation through time [9], free viewpoint navigation and management of massive data; see Fig. 2 and article [2] for further details. When a photographic content is associated with a 6-DoF localization (*i.e.* a 3D position and 3D orientation) and internal camera parameters (focal length, optical center, distortion...), it can be displayed in the 3D scene, as illustrated in Fig. 2 (right image). In practice, such a precise geolocation in the metadata was missing in the historical contents of DB-54k, which lead to the development of a **semi-automatic geolocalization** functionality: from an initial position in the scene, we adopted the semi-automatic approach [6], where localization is performed with 2D-3D registration with at least 4 corresponding user clicks on the overlaid image and the 3D model. In the demo, we exploit a textured terrain and a LoD1 3D model of buildings, provided by IGN on the whole French territory. When available, other models can be exploited, such as 3D point clouds (LiDAR, photogrammetry) [1].

4 JOINT NAVIGATION WITH GEOLOCATION

The two web-based platforms presented above can be used separately on the same dataset, but more interestingly, they can communicate to be exploited jointly. Here we focus on their capability of exchanging about geolocation of the image contents. When available, the visual content's geolocation can be of various types [1]: a simple textual annotation, *i.e.* a toponym usually converted in

2D position with a geocoding technique ; a 2D or 3D position provided by sensors or annotators; or a full 6-DoF camera pose with internal parameters, either available with professional systems or estimated with vision-based tools. Here, the potential geolocation of each image, whatever its type, is stored in a **georeferencing database**, this PostgreSQL/PostGIS database of point features (the camera centers) can be accessed (WFS) and updated (WFS-T) from the two platforms. When the user looks for an image in the search engine, he may want to visualize where it was taken, where was located the photographer, what the area looks like now, etc. If available, the query's geolocation can be exploited in iTowns, or, more interestingly, it can be the one associated with other similar images retrieved. According to the geolocation type, iTowns can present it more or less precisely in the 3D scene (Fig. 2), and then gives the possibility to improve the geolocation semi-automatically.

To make the link between the platforms, we use 2 buttons for each image returned in the search engine (columns "Go to 3D" and "Select this location" in Fig. 1): when clicked, the first button sends the image's ID to iTowns; it is green if a 6-DoF location is available, yellow with location's metadata only, and grey otherwise. Then the image is displayed in the scene according to the geolocation selected with the second button (the one of the image itself or, if not available, of another similar view in order to initiate a geolocalization step).

5 USE CASES

Currently, the two platforms are designed for professionals and non specialists (historians, sociologists, archivists and museum curators) who evaluated the proposed functionalities for several scenarios [2] [3]: link by content similar images of the same location across collections, in order to share multi-source documentary information; help in structuring a collection, for instance by retrieving duplicate or derived visual contents; link similar images of the same location, across collections and times, in order to study the territory evolution; navigate into the spatial environment and exploit other contents (photographs, vector features such as statistics).

In the perspectives, we plan to study other descriptors to further improve image retrieval, especially in challenging conditions (aerial vs. ground-level, etc.) as well as improve the user experience between the two platforms for geolocalization.



Figure 2: Visualization with 2 geolocation types: left, the view, associated with toponym "Chalendrey" converted in 2D geolocation, is displayed in the 3D scene by simple overlay, before semi-automatic geolocalization; right, the view (Nanterre) is 6-DoF localized and displayed in the camera pyramid model, around which the user can turn freely. Credits: Niepce Museum and IGN.

ACKNOWLEDGMENTS

This work was supported by French ANR and DGA within ALEGORIA (grants ANR-17-CE38-0014-01 and DGA-2017-60-0048).

REFERENCES

- [1] Emile Blettery, Nelson Fernandes, and Valérie Gouet-Brunet. 2021. How to Spatialize Geographical Iconographic Heritage. In *Proceedings of the 3rd Workshop on Structuring and Understanding of Multimedia heritAge Contents (SUMAC 2021 @ ACM Multimedia 2021)*. Chengdu, China, 31–40. <https://doi.org/10.1145/3475720.3484444>
- [2] Emile Blettery, Paul Lecat, Alexandre Devaux, Valérie Gouet-Brunet, Frédéric Saly-Giocanti, Mathieu Brédif, Laetitia Delavoipiere, Sylvaine Conord, and Frédéric Moret. 2020. A Spatio-temporal Web-application for the Understanding of the Formation of the Parisian Metropolis. In *3D GeoInfo 2020 conference, ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, VI-4/W1-2020*. London, United Kingdom, 45–52. <https://doi.org/10.5194/isprs-annals-VI-4-W1-2020-45-2020>
- [3] Bénédicte Bucher, Carola Hein, Dorit Raines, and Valérie Gouet-Brunet. 2021. Towards Culture-Aware Smart and Sustainable Cities: Integrating Historical Sources in Spatial Information Infrastructures. *ISPRS International Journal of Geo-Information* 10, 9 (Sept. 2021), 1–24. <https://doi.org/10.3390/ijgi10090588>
- [4] Alexandre Devaux, Mathieu Brédif, and Nicolas Paparoditis. 2012. A web-based 3D mapping application using WebGL allowing interaction with images, point clouds and models. In *20th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems. ACM-GIS'12*, Vol. isbn: 978-1-4503-1691-0. ACM, ACM Press, 586–588. <https://doi.org/10.1145/2424321.2424422>
- [5] Dimitri Gominski, Valérie Gouet-Brunet, and Liming Chen. 2021. Connecting Images through Sources: Exploring Low-Data, Heterogeneous Instance Retrieval. *Remote Sensing* 13, 16 (Aug. 2021), 3080. <https://doi.org/10.3390/rs13163080>
- [6] Mouna Harrach, Alexandre Devaux, and Mathieu Brédif. 2019. interactive image geolocalization in an immersive web application. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLII-2/W9* (2019), 377–380. <https://doi.org/10.5194/isprs-archives-XLII-2-W9-377-2019>
- [7] Records in Contexts Ontology from the International Council on Archives. 2022. Websites: https://www.ica.org/standards/RiC/RiC-O_v0-2.html (ICA / RiC-O Ontology) and <http://data.alegoria-project.fr/def/geotheque/index-en.html> (ALEGORIA extension).
- [8] iTowns project. 2022. Website: <https://github.com/itowns/itowns>.
- [9] Evelyn Paiz-Reyes, Mathieu Brédif, and Sidonie Christophe. 2020. Geometric distortion of historical images for 3D visualization. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences V-2-2020* (2020), 649–655. <https://doi.org/10.5194/isprs-annals-V-2-2020-649-2020>
- [10] ALEGORIA project. 2018. Website: <https://www.alegoria-project.fr>.