A 3D georeferencer and viewer to relate landscape pictures with VGI

Timothée Produit, Jens Ingensand
G2C Institute, Department EC+G
University of Applied Sciences Western Switzerland
Yverdon-les-Bains, Switzerland
timothee.produit@heig-vd.ch

Abstract

Collections of historic images are currently being scanned to stop their degradation and to enable a numeric exploitation. These collections, having a high cultural, artistic and scientific value, are shared by public and private owners on photosharing platforms.

Among these historic collections, landscape images have a non-exploited geographic value. It is not exploited because the images do not have a georeference (geographic information cannot be extracted from the picture) and are poorly labeled which makes searching in a database inefficient. On the other hand, geographic labels (location names, land cover contours) are recorded in VGI databases, but cannot be related to the images without an accurate image georeferencing.

Our idea is to offer volunteers the possibility to create a relationship between shared landscape images and geographic databases. Namely, we are developing a 3D picture georeferencer in a virtual globe. The accurate camera orientation computed with the help of volunteers enables the automatic image labeling with VGI data and the creation of historic geographic layers from the images. Subsequently, the 3D georeferenced images are the source of a cross-temporal VGI comparison: the current state of the landscape, as recorded in the geographic databases, overlaid on top of historic pictures.

1 Introduction

Computers still have a limited capacity when it comes to interpreting images. Scientists and the industry are working hard to improve the computers' ability to label objects (e.g. detect and tag a cat) or recognize the location of a picture. Indeed, these labels or tags are mandatory to search images in databases by location, category or image content. Currently, tagging is generally performed by the user and the geotag is recorded by GPS-enabled cameras.

Our ongoing project specifically focuses on collection of historic landscape photographs. As standard remote sensing images, such as satellite or aerial imagery, these pictures record a state of the landscape. However, poor georeferencing and labeling prevents their use in GI Science. Our project is based on the paradox that precise labels are stored in VGI databases (for instance toponyms and land cover classes) but they are not used to label landscape images. Inversely, images shared by public and institutions record a detailed state of the landscape, but this information can't be inserted in a GIS. Indeed, to relate a picture with geographic data, the picture must be accurately oriented, i.e. the camera position, direction and field of view must be provided. These parameters are required to i) project 3D vector layers in the image and to annotate pictures accordingly ii) to project image information in the map [9].

This problem is well known by the owners of image archives. Indeed, historic photographs have a high cultural and scientific value but they require costly georeferencing and labeling before they can be used for further research or be published on the internet. As a result, images are generally not geotagged at all or simply with a rough point location or location name. In this field, volunteers are central since they can complete image collections by adding their own images but they can also perform the georeferencing and tagging tasks rapidly, reliably and at a low price [11].

3D georeferencing has many impacts. First, it gives the opportunity to automatically tag the images with VGI labels (such as name of places and land cover classes). Second, it enables the creation of historical geographic layers: vector objects outlined by volunteers in the historic image (such as land cover) are projected in the map. Finally, volunteers can complete the tags with their personal knowledge (name of places or anecdotes) which are georeferenced in the same way.

In this paper, we will discuss the state of the art regarding the georeferencing of images. Second, we will present the VGI data required by our platform. Namely, it imports images coming from various collections and OpenStreetMap (OSM) layers. Third, we will present the prototypes of our interfaces for the georeferencing of images by volunteers, the digitization of image content and the visualization of images.

2 State of the art

The task of georeferencing a single image is discussed in three ways in the literature. First, recent research aims at finding the most probable region of a picture without resorting to a priori knowledge. Second, if a more accurate orientation is required, photogrammetric methods are based on 3D reconstruction of a collection of overlapping images. Finally, user-supervised methods are applied when a trustful and accurate orientation must be computed and the photogrammetric methods cannot be exploited.

Recent availability of on-line image databases have enabled researchers to develop algorithms to recognize a rough picture location based on visual features [14]. These algorithms are trained with geotaged images to infer a geotag to non-georeferenced images. In our context, there is a lack of a comparable georeferenced set of historical images that can be used to adapt and “train” a georeferencing algorithm. Moreover, these algorithms only provide a rough location which does not fulfill our requirement. Finally, they are not sufficiently mature to be implemented in published applications.

In areas densely photographed (e.g. tourist hotspots) a 3D model can be recomputed from shared images with Structure-from-Motion (SfM) [6, 12]. The relative 3D model is georeferenced with the GPS tags recorded in the images and optionally with Ground Control Points (GCP) digitized by an operator. This technology is applied in Mapillary which uses a collection of images recorded by a volunteer to reconstruct a 3D model of a street. Such a 3D model can be the reference for a georeferencing algorithm [10, 5]. Unfortunately, these approaches can not be used in this particular project. Obviously, historic images do not have a GPS tag. Moreover, landscape images, and especially historic landscape images have no or only a small overlap and are of varying quality. It limits the possibility to detect similar point features in multiple images, which is the initial step of a SfM algorithm.

Hence, applicative researchers have developed alternative supervised georeferencing methods. They are often applied on sets of historic images which contain valuable information for various field of research [13]. An inefficient but straightforward method is the “rephotography”. The user goes in the field and identifies the picture location by comparing the image with the real world. A new image, recording the current state of the landscape is shot and compared with the historic image. Although this method is time-consuming and not very accurate, it is widely spread in the scientific community [13]. The second method, having the favor of scientists aware of Geographic Information Systems (GIS), is the mono-photogrammetry. The picture is georeferenced with Ground Control Points (GCP): similar point features clicked in the picture (2D) and in the reference map (in 3D, the altitude is extracted from a digital elevation model (DEM)) [1, 7]. These 2D-3D correspondences are the inputs of a camera orientation (or 2D-3D pose estimation) algorithm which computes the unknown parameters of the camera.

The insertion of images in a 3D environment is proposed in Google Earth (function Photo Overlay). However, with the proposed navigation functionality, the recognition of the image location, direction and field of view is tedious and inaccurate. A more accurate method to insert an image in a 3D environment is the digitization of GCP inside the 3D model. This method is similar to the monophotogrammetry method where the reference is a 3D environment rather than a 2D map, the main difference is that the 3D environment provides a natural navigation and viewpoint and ease the digitization of GCP. It was proposed for an online application in the never achieved project View Finder. The same principle was successfully integrated as a plugin for QGIS [8].

Hence, accurate 3D georeferencing is generally proposed to professional users, familiar with geographic information and maps, although it does not require particular skills. On the contrary, we believe that a volunteer knowing the region represented in a picture is able to rapidly identify its initial location in the virtual 3D environment and to recognize and click on stable features (GCP). The georeferencing by volunteers has been mainly implemented using 2D maps [2].

For instance, HistoryPin.com or WhatWasThere.com give the possibility to volunteers to share historic images and “pin” it on the map or orient it more precisely in Google Street View. The images can be commented by the community. Nevertheless, a point location is too rough to enable the interaction of the image content with other geographic data and the georeferencing in Google Street View limits the georeferencing to urban images shot at the ground level from a street recorded by the firm.

3 Data

Our platform has as input database of images. It is connected to web-sharing collections and gives the possibility for volunteers to upload their own images. The second input is the VGI vector layers used to annotate the images.

3.1 Input image databases

The web is bursting with public pictures. With the rise of connected cameras, millions of images are shared everyday. The most valuable images in term of geographic content are landscape images. Among them, historic images have a very high value. The artistic and cultural value of such collections is proved by the numerous exposition and books dedicated to historic photographs. However, they also have a scientific value. Indeed, they record a past state of a landscape going back to 1900. They may provide information about landscape evolution: both man-made (urbanization, deforestation...) and natural (glacier evolution) or document natural disaster (flood, landslide). Generally, the low and various quality of the images and their low overlap limit the possibility to use (semi-) automatic georeferencing methods (except in the specific scenario of photogrammetric collections). Currently, historical images stored in archives and libraries are being scanned. Digitization has two goals. First, it is a solution to stop the degradation of the physical support. Second, it is a gateway to the publication of the archive on the web to reach a large and diverse public. However, the publication requires a tidy and costly labeling. Hence, the access to the image database through the web is limited to a small subset of collections which are properly labeled. Finally, a large amount of historic images belongs to private owners. It is necessary to implement a photo-sharing platform which gives the impulse for their scanning and sharing and prevent the likely lost of these private collections.

3.2 Databases of geographic layers

In our project, we are particularly interested in two VGI layers. First, the name of places databases which link a name with a spatial coordinate. For instance, Geonames.org contains more than 10 millions locations which can be corrected, edited and completed by the public. Second, the database of land cover classes of OpenStreetMap describes the land use of a region. These vector layers are projected in the image:

- To tag an image with the name of places visible in the picture,
- To tag regions of the picture according to the land use,
- To compare the current land cover with the state recorded by the historic image.

Labeling has two consequences. First, information is added in the picture during visualization: for instance summits, villages, rivers are named or the current contour of a village or glacier can be compared with the historic state recorded in the image. Second, images are tagged with the visible locations and land cover classes. These tags enable search forms based on visible features, for instance complex requests such as “pictures where Zermatt, the Matterhorn and a glacier are visible”.

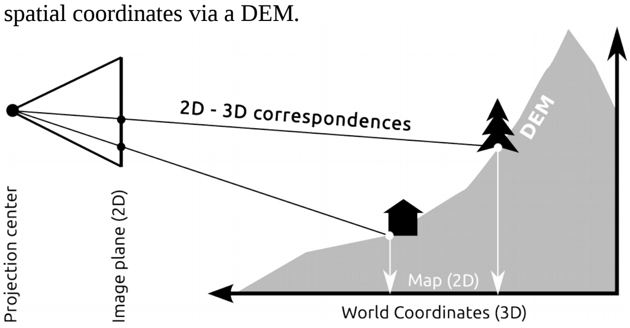
4 Results

The functions which relate pictures with geographic information and transform pictures in geographic information require the development of dedicated web-apps:

- **A 3D georeferencing platform:** Similarly to 2D georeferencing, the volunteer digitizes Ground Control Points (GCP) to perform the image orientation.
- **A geographic image viewer:** Based on the relation between a picture and the map via a DEM (see Figure 1), this platform enables the visualisation of geographic data in the image.
- **A historic geographic database:** The images, their orientation and the geographic information digitized by the volunteers are stored in a geographic database.
- **A database visualization interface:** This interface provides a geographical search form to rapidly select a subset of pictures and visualize them in a virtual globe.

These apps are discussed in the next sections.

Figure 1: Perspective geometry. Image pixels are related to the spatial coordinates via a DEM.



4.1 Image orientation (3D georeferencing platform)

An accurate 3D georeferencing of landscape images enables their accurate labeling and the possibility to extract geographic layers from the pictures. The georeferencing is therefore essential to fully exploit the geographic potential of a landscape image. Nevertheless, the georeferencing of an entire collection is beyond the human and financial resources of the images owners. Hence, the exploitation of volunteers is a low cost, fast and (hopefully) an accurate method to improve an image collection with geographic tags and innovative visualization interface.

We implemented a three step georeferencing method. First, the virtual environment is initialized at the geotag location. Indeed, most historic images are at least related with a geographic name. Its correspondence in the name of places database provide an initial location for the georeferencing. Second, the volunteer navigates in the virtual environment until he reaches a viewpoint roughly similar to the image. Finally, he is asked to digitize Ground Control Point (GCPs) required to compute the accurate camera orientation. This final step is iterative: the initial orientation facilitates the digitization of more accurate GCPs.

Additionally, two methods are used to validate the quality of the camera orientation. First, the GCPs residuals give an information about the quality of georeferencing. The residuals measure the distance between the 2D GCP digitized by the volunteer and the 3D GCPs projected in the image with the computed orientation parameters. Large residuals exhibit potential GCP errors. Hence, we can set a threshold on the residuals to prevent a bad orientation. Second, the public visiting the image gallery has the possibility to report bad georeferencing.

4.2 Monoplotting to generate geographic layers (Geographic image viewer)

The camera orientation gives a direct relation between the image coordinates and the spatial coordinates. Hence, an object outlined by a volunteer in an image can be exported to a spatial database. This functionality is used to:

- Complete the name of places database: The geonames database is far from being complete. We take the opportunity of having an interested volunteer at work to complete the database.
- Create a database of historic geographic layers: OSM keeps track of the land use since the beginning of the project. Historic images can be used as a detailed source of information since 1900. These images can be orthorectified directly or the historic state of the land cover can be drawn in the images by the volunteers and inserted into the geographic database.

However, a main limitation is the inaccuracy of the monophotogrammetry (influenced by the ability of the volunteer, the image quality, the GCP distribution and the topography) [3]. Indeed, an oblique image oriented by an operator can never compete with the accuracy of an orthoimage processed with a state-of-the-art workflow.

4.3 Combining VGI in a 3D interface

The images, their orientation and labels will feed a database of historic images. We use a 3D virtual environment to offer the opportunity to put every VGI source in a single viewer. Indeed, a 3D environment provides the possibility to switch from aerial to terrestrial viewpoints and to visualize standard geographic information as well as pictures (see Figure 3).

The platform we are implementing will have the following features:

- Searching of images by visible name of places, content type, spatial area,
- Navigation and visualization of images in their spatial context (in the virtual environment) and historical context (historic maps as base layers),
- Access to the Geographic viewer with a click on a picture,
- Guided visits through the space and the collection.

5 Discussion

Our prototype proves that the proposed app is a technically working concept. However, it raises some questions, regarding its implementation in a participative website:

- **Community building:** Crowdsourcing requires the building of a community. The strategy that we investigate is based on local communities: the initial

Figure 2: 3D Georeferencing platform. The volunteer digitizes correspondences in the image and the virtual globe.

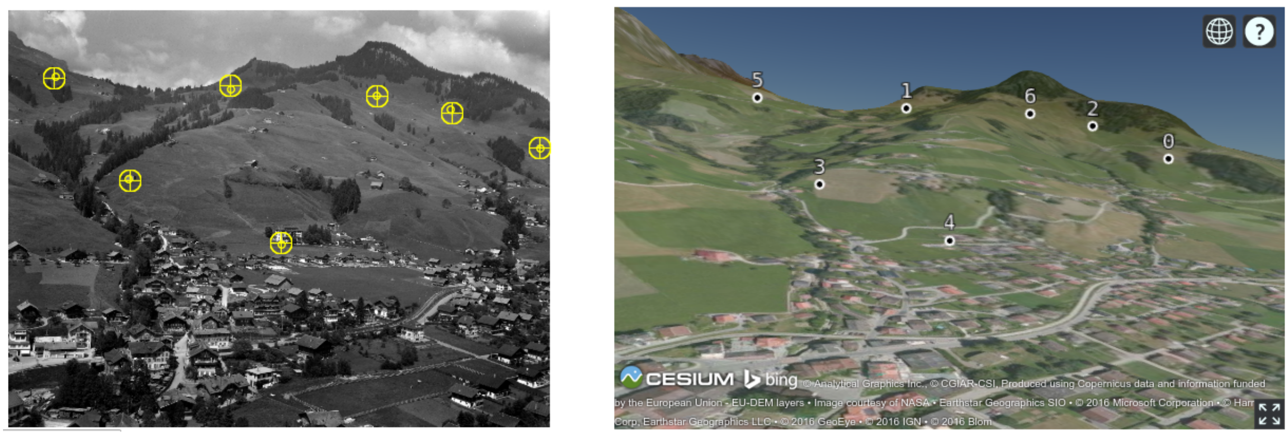
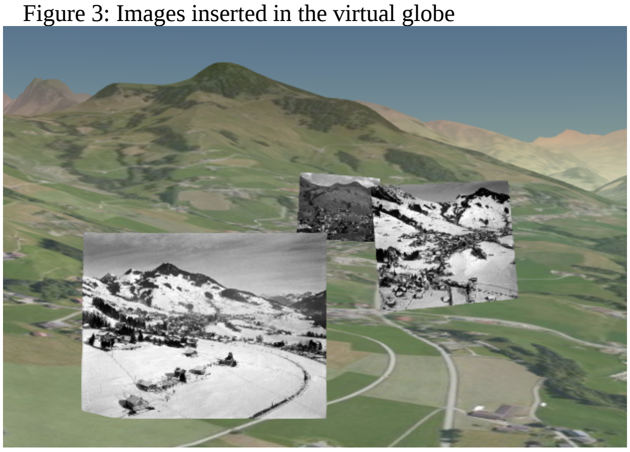


Figure 3: Images inserted in the virtual globe



set of images will cover a specific region. Hence, our media campaign (social and regular media) will target this region. We will also specifically contact local communities such as the alpine club and the heritage protection societies, which are expected to have an interest in historic images.

- **3D georeferencing and navigation in the virtual globe:** Currently, the platform is tested by operators accustomed to maps and GIS. The general public will undoubtedly encounter difficulties with the navigation in the 3D virtual environment. To limit this difficulty we will train the users with a game and provide an FAQ.
- **Image sources:** Initially, we will work with a postcard collection provided by the *Archive de la Construction Moderne* of the EPFL, as well as with images shared by the volunteers. However, the intended connection with other web-shared image databases will raise issues such as the various format of images and databases.
- **Languages:** The initial set of images will cover a french speaking area. Hence, the annotations by the volunteers will be stored in this language. However, similarly to OSM, we will implement the possibility to store geographic names in several languages.

6 Conclusions

With the rise of low-cost connected cameras and photo-sharing web-sites, volunteers become image specialist. Additionally, citizens are increasingly accustomed to virtual environments that they meet in computer games or more recently in web maps. Hence, they are ideal candidates to complete the georeferencing of pictures: they are able to recognize a picture location shot in their habitual neighbourhood, rapidly identify its location in a virtual globe and digitize GCP.

We are especially interested in shared historic landscape images which are, in a way similar to standard remote sensing imagery: they record the state of a landscape in high spatial and temporal resolution. However, the main difference is their inaccurate georeferencing which limits their use for geographic interpretation.

Oriented images have a high potential in the context of VGI. First, our platform provides the link of two VGI sources (images and geographic layers) which up to now could not be related. This feature enables the visualization of cross-temporal data and the annotation of the pictures for improved indexation. Second, each pixel is related to a spatial coordinate. Hence, object digitized in the image by the volunteers feed VGI databases. This function is particularly interesting to complete the databases of place names and historic anecdotes.

In this project, we focus on historic images, because they have a high value for the public, as well as for archive managers and scientists. Moreover, historic images can federate a large community of volunteers. Specifically, we hope to reach the communities of seniors which are usually less interested in technology but have a rich historic knowledge which match the age of the images. The main challenge is therefore to provide an intuitive and user-friendly interface for the navigation in the virtual globe.

Nevertheless, the implications of our project go beyond this specific subject. Many other possibilities such as the sharing of trails documented by pictures, rapid mapping of natural events or visits of image collections in augmented reality can directly derive from the web-applications implemented in our project.

References

[1] C. Bozzini, M. Conedera, P. Krebs, A new monoplotting tool to extract georeferenced vector data and orthorectified raster data from oblique non-metric photographs, *International Journal of Heritage in the Digital Era*, pages 499–518, 2012.

[2] K. Kowal, P. Pridal, Online georeferencing for libraries: The british library implementation of georeferencer for spatial metadata enhancement and public engagement, *Journal of Map & Geography Libraries*, pages 276–289 2012.

[3] K. Kraus. *Photogrammetry: Geometry from images and laser scans.* Walter de Gruyter, 2007.

[4] S. Lee, H. Zhang, D. Crandall, and I. N. Bloomington. Predicting geo-informative attributes in large- scale image collections using convolutional neural networks. In *Winter Conference on Applications of Computer Vision*, 2015.

[5] Y. Li, N. Snavely, D. Huttenlocher, P. Fua, Worldwide pose estimation using 3D point clouds, in: *European Conferen on Computer Vision*, Firenze, 2012.

[6] R. Martin-Brualla,, D. Gallup, and S. M. Seitz. Time-Lapse Mining from Internet Photos. *ACM Transactions on Graphics*, 2015.

[7] A. Messerli, A. Grinsted, Image GeoRectification and feature tracking toolbox: ImGRAFT, *Geoscientific Instrumentation, Methods and Data Systems Discussions* pages 491–513, 2014.

[8] G. Milani. Pic2map: Intégration de photographies dans QGIS. *Master's thesis, EPFL*, 2014.

[9] T. Produit and D. Tuia. An open tool to register landscape oblique images and generate their synthetic models. In *Open Source Geospatial Research and Education Symposium*, Yverdon-les-Bains, 2012.

[10] T. Produit, D. Tuia, V. Lepetit, F. Golay, Pose estimation of web-shared landscape pictures, *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 127–134, 2014.

[11] H. Rose, Crowdsourcing: How and Why Should Libraries Do It?, *D-Lib Magazine*, 2010.

[12] C. Strecha, T. Pylvanainen, P. Fua, Dynamic and scalable large scale image recon-struction, in: *Computer Vision and Pattern Recognition*, pages 406–413, 2010.

[13] R. Webb, *Repeat Photography: Methods and Applications in the Natural Sciences*, Island Press, 2010.

[14] T. Weyand, I. Kostrikov, J. Philbin, PlaNet-Photo Geolocation with Convolutional Neural Networks, *arXiv* 2016.