

2018

Crowdsourcing the georeferencing of historical pictures

Timothée Produit

University of Applied Sciences and Arts Western Switzerland

Nicolas Blanc

University of Applied Sciences and Arts, Western Switzerland

Sarah Composto

HEIG-VD / INSIT, University of Applied Sciences Western Switzerland

Jens Ingensand

University of Applied Sciences and Arts, Western Switzerland

Loic Furhoff

University of Applied Sciences and Arts, Western Switzerland

Follow this and additional works at: <https://scholarworks.umass.edu/foss4g>

 Part of the [Digital Humanities Commons](#), [Environmental Engineering Commons](#), and the [Geographic Information Sciences Commons](#)

Recommended Citation

Produit, Timothée; Blanc, Nicolas; Composto, Sarah; Ingensand, Jens; and Furhoff, Loic (2018) "Crowdsourcing the georeferencing of historical pictures," *Free and Open Source Software for Geospatial (FOSS4G) Conference Proceedings*: Vol. 18 , Article 6.

DOI: <https://doi.org/10.7275/awa4-4w48>

Available at: <https://scholarworks.umass.edu/foss4g/vol18/iss1/6>

This Paper is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Free and Open Source Software for Geospatial (FOSS4G) Conference Proceedings by an authorized editor of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.

Crowdsourcing the Georeferencing of Historical Pictures

Timothée Produit^{a,*}, Nicolas Blanc^a, Loïc Fürhoff^b, Sarah Composto^a, Jens Ingensand^a

*University of Applied Sciences Western Switzerland,
School of Management and Engineering Vaud,*

^a *Territorial Engineering Institute,*
^b *Media Engineering Institute,
CH-1401 Yverdon-les-Bains,
Switzerland*

Abstract: Countless organizations own archive pictures which are not used due to the difficulty of localizing them. Only the participation of volunteers can help to identify the image content. Through an accessible web interface called *smapshot*, citizens are able to georeference historical images without any prior specific competence. To avoid the volunteers to face the complexity of 3D digitization and navigation, we implemented a state of the art photogrammetry camera orientation algorithm in the platform. Currently several thousand images from all around Switzerland have been georeferenced by the crowd. At this time *smapshot* is mainly used by the population for the visualization experience and the nostalgia behind it. The ambition is to support scientists to analyze landscape evolution.

*Corresponding author

Email address: timothee.produit@heig-vd.ch (Timothée Produit)

1. Introduction

Despite their value for scientists and the general public, archive photos suffer from a low visibility. In our project we focus on historical cityscape and landscape images. These images are complementary to regular remote sensing images such as satellite and photogrammetric images. They often have a better spatial and temporal resolution and a more natural viewpoint.

One problem is that historical photos usually have poor metadata especially in terms of geospatial metadata. We can consider two types of geospatial metadata for these images. The information about their orientation (X, Y, Z, camera rotations and field of view) and the information about their content (place names). In addition, images are scattered in several archives (e.g. personal, libraries) and various websites. Typically, search engines provide incomplete results due to poor geographic metadata. An accurate georeferencing can provide a solution to these issues.

The state of the art for the georeferencing of images is the Structure-from-motion (SfM) which uses key-points detected in overlapping images to reconstruct the geometry of the camera orientations and a 3D model of the scene (Hartmann et al., 2016). In the context of non-photogrammetric historical images the small overlap and the seasonal and temporal variation make key-point detection and SfM difficult in general. Computer vision and machine learning researches are growing fast in the field of image recognition. However, the problem of exact georeferencing of photos at the scale of a region or an entire country remains a very complex task (Baatz et al., 2012).

The manual georeferencing of historical images is a working concept. It was applied with software such as the pic2map plugin for QGIS (Produit et al., 2012) and the WSL-MPT (Bozzini et al., 2012). The natural evolution of this concept is its implementation in a web platform to provide a georeferencing and visualization solution. The idea of a virtual globe for the visualization and the georeferencing of images was discussed in the discontinued project Viewfinder (University of Southern California California, 2008).

Our solution is to provide to volunteers a platform¹ for the georeferencing of large collections of historical images. The key issues are:

1. Provide a user-friendly interface for the computation of the camera orientation by the general public
2. Generate geographic metadata: image footprints (the polygon enclosing the area visible in the image) and visible place names
3. Represent the images on a map
4. Represent the images in a virtual globe with comparison functionality
5. Provide an efficient geographic image search engine.

¹https://gitlab.com/produitt/smashot_FOSS4G/blob/master/README.md
<https://smashot.heig-vd.ch>

2. 3D Crowdsourcing platform

2.1. Georeferencing

The georeferencing of images is usually provided in photogrammetric software dedicated to professional users. The main challenge in our project is to deliver a user-friendly georeferencing functionality which hides the complexity of the collinearity equation and that is rapidly handled by volunteers.

First the volunteer has to indicate an approximate camera position and azimuth on a 2D map (fig. 1). With these parameters, we initialize a virtual globe. The volunteer is then able to click on similar locations in the image and in the virtual globe (see fig. 2). These locations, also defined as Ground Control Points (GCP) are thereafter used as 2D-3D correspondences in a camera orientation algorithm that computes camera location, rotation and field of view.

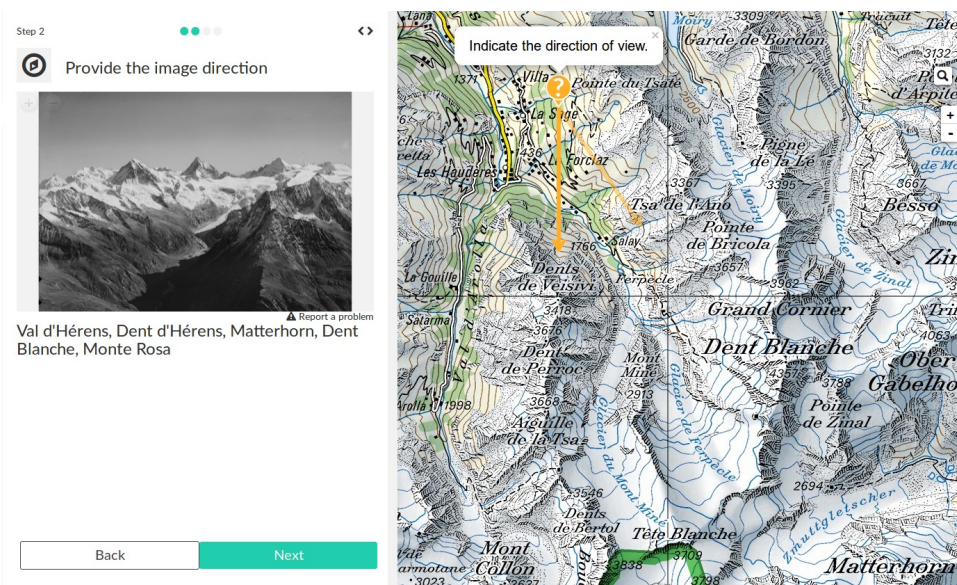


Figure 1: The two first steps of the georeferencing process. The user moves the marker to provide a rough location and indicate the image azimuth with the arrow.

We implemented some procedures in order to partially avoid wrong georeferencing (deliberate or not). First, only four correspondences are mathematically required to compute the camera orientation: In our platform, the georeferencing can be validated only if at least six correspondences are provided. This redundancy allows us to compute an error on each correspondence. If the maximum error exceeds a limit value, the georeferencing can not be validated and we ask the volunteer to improve the correspondences. Eventually, the images are validated by the archivist team. The dashboard provides the team with quality indicators (number of correspondences, the maximal error, surface of the image covered by the correspondences) and the volunteer's username. The validation team can visualize the image, check its alignment with the virtual globe and decide whether they validate the georeferencing, improve or reject it. In the last two scenarios, the validator can send the reason for the rejection or improvement to the volunteer. Hence the volunteer has some feedback on his work and can improve his skills.

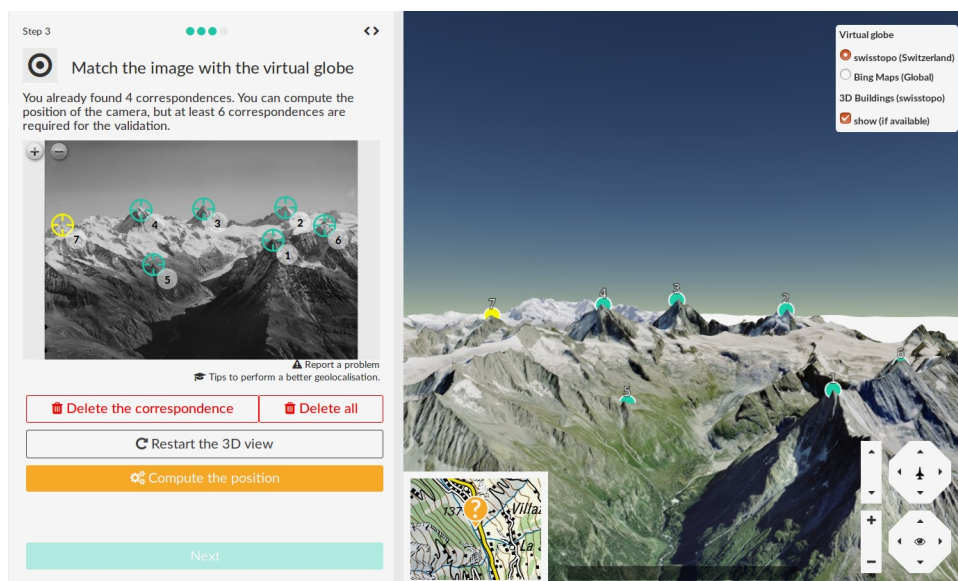


Figure 2: 2D-3D correspondences clicked by the volunteer in the image and in the virtual globe.

2.2. Geographic metadata

We deliver several type of metadata, the camera position and orientation, but also the image viewshed and the list of place names visible in the image. In the context of images, the viewshed is the visible area in the image (fig. 3). To the best of our knowledge, there is no open source software which is able to take into account the camera parameters (especially the camera roll) to generate a viewshed. In our platform, the computation of the viewsheds has three steps. First, the 3D coordinates of the image border are computed on-the-fly by the user's browser using the virtual globe. Second, the 360° viewshed of the camera position is computed with the GIS SAGA and a DEM with a resolution of 200m. Finally, the intersection of these two polygons generates the image footprint. The intersection of the footprint with the spatial database of the Swiss place names (Swiss Federal Office of Topography: *swissNames3D*) provides geotags that can be linked to the image. The geographic metadata can be downloaded by the archivists to improve their database.

2.3. 2D image representation on the map

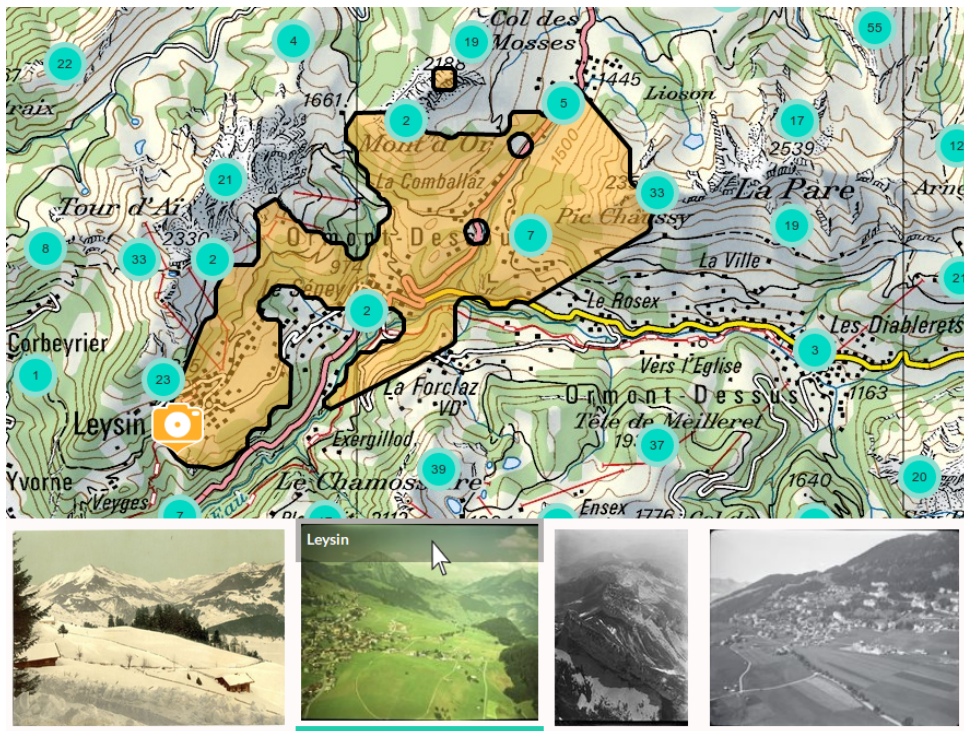


Figure 3: An example viewshed for one picture, the green circles are the clusters of images locations.

With several thousand images to be shown, the map could quickly appear overloaded. There are mainly two alternatives for the representation of large quantities of points on a map:

- Showing the points density with a raster heat map
- Point clustering

Both graphical representations are usually computed on-the-fly at each zoom level.

We used both approaches in two contexts. We opted for the clustering to keep the information about the location of the georeferenced images. On the map, each point depicts a camera location, but we also wanted to show the image content. Hence, moving the mouse over an image or a single point shows the image viewshed (fig. 3). A heat map representation is used for the images which are not georeferenced. These images are linked to several potential locations, the heat map allows us to represent the georeferencing fuzziness. The goal is therefore to indicate regions where images are available.

2.4. 3D image representation in the virtual globe

A central feature of our platform is the virtual globe. The virtual globe is based on the web services of the Swiss Federal Office of Topography Swisstopo (orthoimage, DEM, 3D buildings). We opted for the technology that Swisstopo uses for their own portal (<https://map.geo.admin.ch>), namely Cesium (<https://cesiumjs.org/>). In order to show 2D images in the virtual globe, we generated 3D files in the gltf format using the camera orientation, position and focal length.

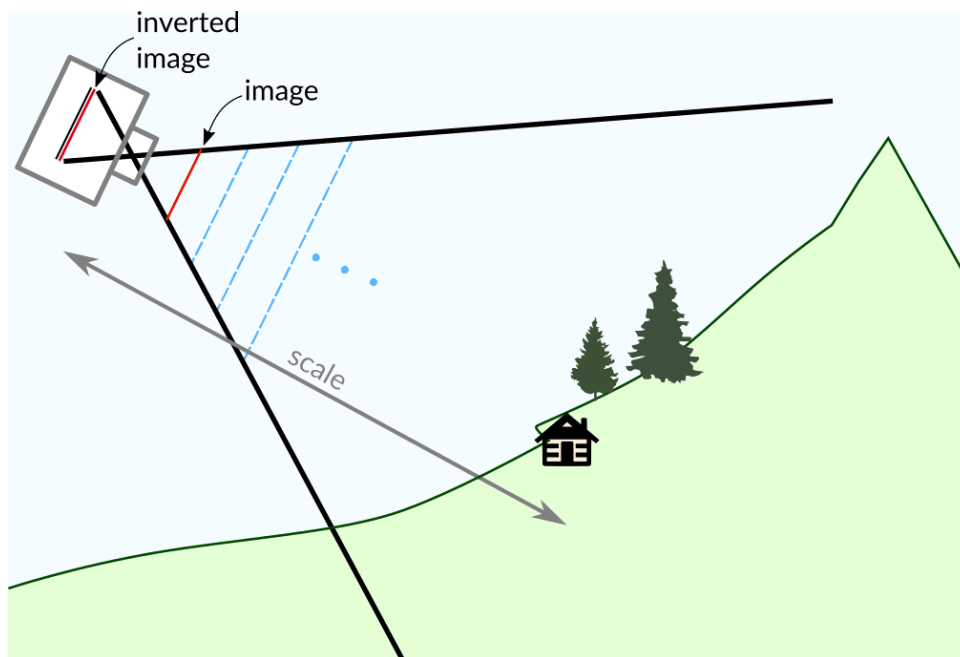


Figure 4: Image is inverted on the camera sensor. In the 3D model, scaling is done by moving the image along the optical axis.

In the real world the camera sensor is a rectangle which can be illustrated in space. Adopting the pinhole camera model, the sensor records an inverted picture of the landscape. The un-inverted image can be represented by symmetry. If we remove the scale factor, the image can be located anywhere along the camera field of view (fig. 4).

Hence, we use the computed camera orientation to initialize the virtual globe. Thereafter the image and the virtual landscape are aligned on the screen. The comparison of the current virtual landscape and the historical landscape can be done with a hide / show button as well as transparency and image scale sliders which allow the user to move the image along the optical axis of the virtual camera, thus giving the impression that the image blend in with the landscape (fig. 5).

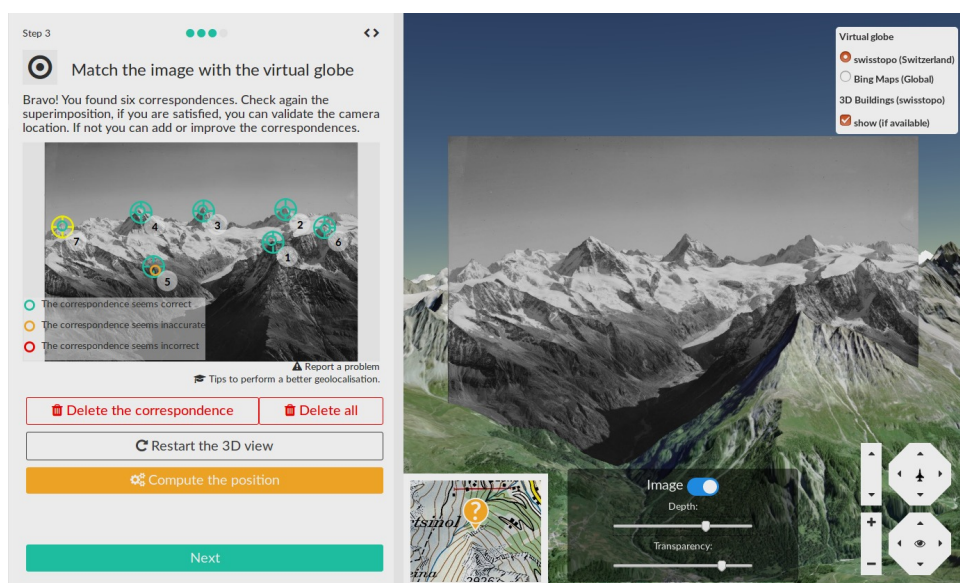


Figure 5: The image is inserted in the relief by changing the scale.

2.5. Geographic search engine

Having the camera location, the footprint and the geotags stored in a PostGIS database, efficient geographic queries can be performed to retrieve images looking at a specific location either by providing a place name or by drawing a polygon on the map. The geographic search engine is completed with a standard search field (text, collection, photographer, date).

3. Results

The prototype was published on the web on February 2017. The project was cited in mass media and volunteers were hereafter able to georeference images. Its success allowed us to contact an important Swiss archive (ETHZ library) to develop an improved and more user-friendly interface that is also more focusing on the needs of archive managers. Consequently, we added some essential features for archivists such as the tagging of images with notes and the correction and improvement of the image's title and description.

Currently (April 2018), the platform contains 12,300 published images among which 12,150 are georeferenced. The most active volunteer georeferenced more than 1,600 images while the three most active volunteers georeferenced more than 30% of the images. In terms of views 30% of the images have more than 5 views. The most viewed image has 340 views. However some of the images have many views due to the fact that they are difficult to georeference and that many volunteers have attempted the georeferencing. Indeed, in some situations the georeferencing of terrestrial images can be difficult :

- The computed camera position can be in the ground and therefore generates visual artifacts
- The absence of a background makes the recognition of GCP difficult

For vertical aerial images the georeferencing is not optimal since the georeferencing algorithm is optimized for oblique images.

We noticed that an accurate apriori location is crucial for the platform. If it is accurate, even a volunteer who is not familiar the area will be able to find matches with the virtual landscape. In contrast, if the apriori location is wrong or too rough, only a volunteer who recognizes the area in the image will be able to perform the geolocalization.

4. Conclusions and perspectives

The number of presently georeferenced images demonstrates that crowdsourced georeferencing of images in a virtual globe is a working concept. More generally, we show that volunteers can provide accurate 3D geographic data, however some issues of navigation and digitization in 3D have to be addressed. In our project, we use photogrammetric equations to limit the involvement of the volunteers. We still have several limitations such as the constraint of the Swiss border. Currently, the platform is only available for Switzerland due to the data sources that we use and the hard-coded coordinate system for the camera orientation computation. The extension to other parts of the globe using global data and a global coordinate system would allow us to reach a larger user community. Another difficulty will be to find appropriate 3D building datasets.

Another limitation of the project is the validation by the archivist team. It is a required feature to ensure the quality of the metadata but it is also time consuming. A validation by the crowd itself could reduce the involvement of the validation team. One possibility would be to let archivists decide whether they want to do the validation themselves or whether they want it to be done by the volunteer community.

In the future, we will have a unique database of precisely georeferenced images. Currently, the temporal component is not used at its best. We have to investigate how time data (that is associated with the images) can be used in the project as a cursor to navigate along a temporal axis. Certainly some images do not have an accurate date, but the comparison with other images and the knowledge of older volunteers can help detecting the date when a picture had been taken. This unique database can also be used to reconstruct 3D models of the landscape. As aforementioned, the variations in data that are caused by the various sensors, seasons and illumination will challenge the key-point detection. If this issue can be solved, for instance with the accurate 3D georeferencing provided by the crowd, we will be able to generate unique successions of 3D models evolving through time.

Another exciting perspective is the use of georeferenced pictures for scientific projects with the ambition to analyze landscape evolution such as the melting of glaciers, large flood deposits, landslides, rockfalls, or the evolution of build-up areas. We will also have to face the issue of a successful academic project which has to be maintained and developed with the needs of end-users in mind.

Acknowledgement The authors would like to thank the Hasler Foundation, ETHZ, Swisstopo and Migros for their support.

References

- Baatz, G., Saurer, O., Köser, K., Pollefeys, M., 2012. Large scale visual geo-localization of images in mountainous terrain. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (Eds.), *Computer Vision – ECCV 2012*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 517–530.
- Bozzini, C., Conedera, M., Krebs, P., 2012. A new monoplottting tool to extract georeferenced vector data and orthorectified raster data from oblique non-metric photographs. *International Journal of Heritage in the Digital Era* 1 (3), 499–518.
- Hartmann, W., Havlena, M., Schindler, K., 2016. Towards complete, geo-referenced 3d models from crowd-sourced amateur images. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences* 3 (3).
- Produit, T., Tuia, D., Strecha, C., Golay, F., 2012. An open tool to register landscape oblique images and generate their synthetic model. In: *Open Source Geospatial Research and Education Symposium (OGRS)*. No. EPFL-CONF-183329.
- University of Southern California California, 2008. Viewfinder: How to seamlessly flickrize google earth. URL <http://interactive.usc.edu/projects/viewfinder>