Vassilis Tsioukas[*], Alexandra Koussoulakou[**], Maria Pazarli[***], Nopi Ploutoglou[****],
Miltiadis Daniil[*****], Ioanna Stergiopoulou[******]

Scanning or digitizing in libraries? A test on the efficiency of dedicated book-scanning devices in digitizing bound atlases and maps.

Keywords: Atlas digitizing; map scanning; book scanning; 3D digitizing; historical maps; bound maps.

Summary: The digital documentation of books in a library, mostly using dedicated book scanners, is a means of modern storage of their textual and pictorial content and heritage for future use. The primary task for this purpose is to convert book-pages in digital images, so that the books' original use can be recreated digitally. Within this context books such as bound atlases and maps constitute a particular case: their content is more than textual or pictorial -as is the case with plain books- because maps are a form of graphic representation characterized by a significant property: the geometrical reference to geographic space, in any degree of approximation. Consequently, it is of particular importance that the metric qualities inherent in atlas bound maps are taken in consideration when going digital. In the map case the result of scanning requires a really numerical quality, which is the quite literal meaning of the proper digital output. The aim of the study presented is to test the results of different means used for such proper digital conversions. Maps in an old atlas from the collection of rare books at the Koventareios Municipal Library of Kozani, in Greece, are used for that purpose. The atlas dates back to the early 18th century and is authored by the well-known German geographer, cartographer and atlas publisher Johann Baptist Homann (1664-1724). The maps are converted to digital form by means of a dedicated book-scanning device, familiar to the map librarians, and by means of a high precision large-format scanner as well as by means of 3D digital camera. The results are analyzed and evaluated by comparing optimally the relevant images with respect to physical measurements of features (i.e. on the maps) with their image counterparts.

Introduction

Digital technology offers a great potential for research and education in cultural heritage. For cartographic heritage in particular, the possibility of browsing and examining unique and rare objects outside the controlled environment of archives, libraries and museums is unprecedented. An indicative index of the activity in the digital "domain" of cultural and cartographic heritage is the amount of funding and research projects in Information and Communication Technologies worldwide and in the EU (Tsioukas and Daniil, 2009).

But digital technology, like any technology, is a tool that has to be used with forethought. Digitization of maps, in particular, might seem a trivial task, but in reality it proves to be more demanding than initially anticipated.

---

[*] Associate Professor, Aristotle University of Thessaloniki [tsioukas@topo.auth.gr]

[**] Associate Professor, Aristotle University of Thessaloniki [kusulaku@topo.auth.gr]

[***] PhD. Candidate, Archaeologist, National Map Library [pazarli@maplibrary.gr]

[****] PhD. Candidate, Surveyor Engineer, National Map Library [ploutoglou@maplibrary.gr]

[*****] Dr., Surveyor Engineer [mdaniil2010@topo.auth.gr]

[******] Librarian, Head of Koventareios Municipal Library of Kozani [iostergiop@kozlib.gr]

Our concern in this study is the digitization of cartographic material found in libraries in the form of atlases and bound maps. What is the best practice to treat such documents when going digital? Are they to be regarded as mere books –and consequently treated so when it comes to their digital conversion- or are they of a different nature?

Converting into digital form: the case of atlases and bound maps

The digital documentation of books in a libary is mostly done by using dedicated book scanners. The primary task for this purpose is to convert bookpages in digital images: in this way the books' content and heritage is stored for dissemination and further use. Atlases and bound maps constitute a particular case of books (or "books"); their form signifies a use which is actually different from that of their content. Speaking in terms of documents existing in plain books, one could argue that documents in books can be divided in *textual* and *pictorial*. In atlases, on the other hand, an additional form of document is prevailing and that is maps, i.e. *cartographic* documents. Maps are a form of graphic representation of the milieu, characterized by a unique and significant property: their geometrical reference to geographic space, in any degree of approximation. Consequently, it is of capital importance that the metric qualities inherent in atlas bound maps are taken in consideration when converting to digital.

In the case of books therefore, scanning *textual* and *pictorial* documents should not be the same as scanning *cartographic* documents: in the map case the result of scanning requires a really numerical quality, which is the quite literal meaning of the proper digital output. So we could differentiate between *scanning* and *digitizing* by arguing that in digitizing primarily *metric* (but also colour) properties are of main concern.

Nevertheless, maps in atlases are commonly considered as books and usually scanned by standard dedicated book scanning machines, as normal book pages (i.e. as if they were texts). This induces deformation altering the geometry (and often the colour) of the content. How this is done is demonstrated by the comparison carried out and described in the following.

When converting maps (and images, in general) to digital form one has to deal with two main issues: the content of the maps and its description; it should be stated at this point that this work is concerned strictly with the former issue.

A case study : an early 18th century atlas from the Municipal Library of Kozani, Greece

The aim of the study presented here is to test the results of different means used for the properly digital conversions, as described in the previous section. For the purposes of the study, maps in an old atlas from the collection of rare books at the Koventareios Municipal Library of Kozani, in Greece, are used. The two-volume atlas dates back to the early 18th century and is authored by the well-known German geographer, cartographer and atlas publisher Johann Baptist Homann (1664-1724). It has a size of 65 by 51 cm (open) and the second volume which was used here contains 85 maps of approximate dimensions 56 by 46 cm. It has survived in a rather bad shape, exhibiting alterations due to age and atmospheric conditions, which actually affect the natural curvature of the book pages, when open (Fig. 1).

To accommodate for the different shape of the book pages in different parts of the document, three different maps from different parts of the atlas were selected for digitization (Fig. 2). For the purposes of the comparison two different ways for scanning the maps were followed. First the maps were converted to digital form by means of a dedicated book-scanning device, familiar to the map librarians, also available at the Koventareios Library (a *Bookeye 3* scanner). Then the digitization of the maps was done by means of a high precision large-format scanner (a *Cruse CS-220ST*). The images resulting from these scans were compared to physical measurements (i.e. measurements on the maps themselves) which were used as the “ground truth”.

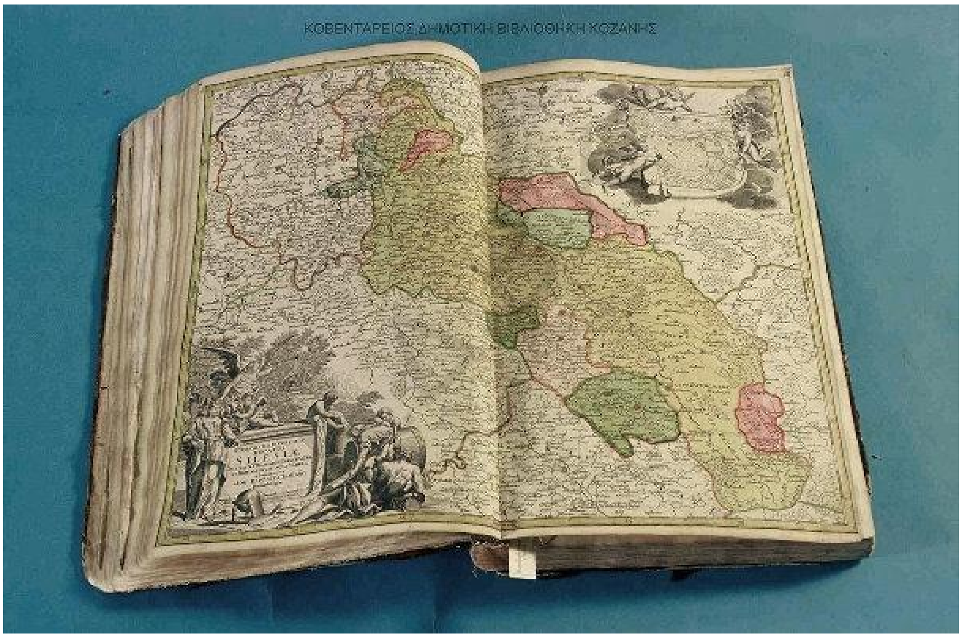

Figure 1: The early 18th century Homann atlas at the Koventareios Municipal Library of Kozani in Greece (image source: www.kozlib.gr).

Figure 2: The atlas maps used for the comparison test.

The approach followed: scanning, 2D digitizing, 3D modelling

The deformations induced by the digitization process and affecting the digital images of the pages of the atlas, fall into two main categories: deformations caused by the scanning device itself and deformations due to the form of the pages i.e. their curved and irregular surface. Dealing with the deformations of the second category is a rather complicated issue and requires a more sophisticated approach; it is considered at the second step of this study (*3d digitizing and* *modelling*). Of our initial concern at this point is the proper device for the digitization of maps in atlases and books; therefore common sense drives us to the comparison of flat (i.e. 2D) areas as the first step to carry out.

2D measurements and comparisons

For the purposes of 2D comparisons, linear distances on the flat part of the maps' pages were chosen carefully, along the longitudinal (i.e. horizontal) and latitudinal (i.e. vertical) direction, approximately. The lines chosen are defined by characteristic points of detail, easily and clearly recognizable on the maps. The physical measurements (i.e. on the maps) were carried out with a high precision ruler and with careful handling of the document.

The same linear distances were then measured on the digital images of the respective two scans i.e. the book-scanning device (*B* for short here) and the large-format scanner (*C*, for short). The results of the comparisons are summarized in Fig. 3, where one can see the order of magnitude of the differences from the physical measurements (i.e. from the real values of the linear distances, as measured on the maps). On the upper part of each page, the values (in green) refer to the images scanned by *C*, the values below (in purple) refer to the images scanned by *B*. The values actually show the magnitude of the deformations in each case: *C* introduces minor deformations, while the values of the deformations induced by *B* are remarkably larger.

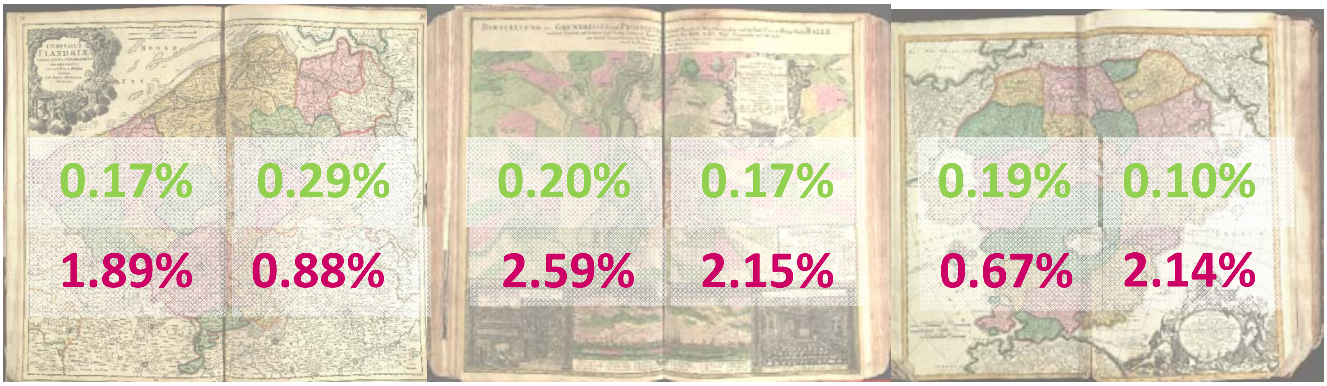

Figure 3: The deformations of linear features for each page, due to scanning: the values (%) show order of magnitude of the difference from true values. The figures above (in green) refer to the output from the high precision scanner, the figures below (in purple) refer to the output from the book scanning device (see text for further explanations).

But what does a value of a difference of e.g. 2.59% actually mean? (see Fig. 3: the left page of the map in the middle, B scan below). For a distance of e.g. 35cm on the map this results in a difference of 0.9 cm from the correct value (or 9mm, that is almost 1cm). A magnitude of 9mm on a map is far from being negligible: it might represent the length of a city block on the ground (e.g. in a large scale map) or the diameter of a town (in a medium scale map)!

3D digitizing of small objects - 3D modelling

Physical measurements on the pages of atlases or bound maps are not always possible, due to their curved and often ondulated surface (commonly found in old, historical maps). This, however, could be overcome if a digital, virtual surface of the pages is generated, with the help of 3D digitizing and 3D modelling techniques. In the following this subject will be presented from the point of view of its use in atlases and bound maps.

The 3D digitization is often used to capture the full geometry of objects with a particular high cultural and therefore and economic value. Items that are usually 3D recorded are archaeological artifacts and objects digitized to be reproduced in replicas and sold (e.g. museum exhibits). There are also many cases where objects are recorded in 3D in order to be examined and studied virtually from scholars through their digital models. The ease of electronic reception and downloading of the 3D models and the use of special software to focus on specific areas and extract measurements makes the 3D recording procedure very useful.

The most common techniques used to digitize objects is through structured light configuration where a non-measuring laser beam or even a video projector image is used to project known patterns on the object. The way that these light patterns deform when striking the objects' surfaces allows vision systems to calculate the depth and surface information of the objects in the scene (*URL1*).

Other ways to create models of small objects is through *stereoscopic photogrammetric* methods (Tsioukas and Daniil, 2008) and the use of *structure from motion techniques* found in computer vision applications (*URL2*). Both are using recognition-matching processes that operate on a large number of points on the object that are necessary to define the object’s outer surface. They require image capturing of the object using a conventional (but properly calibrated) optical digital sensor from multiple locations. After the geometry of the camera shots is calculated, the reconstruction of the light rays able to calculate the 3D coordinates of a large number of points on the object leads in the next stage to the creation of the digital surface model of the object.

Both techniques (structured light and photogrammetric-motion) are based on visual recognition and reconstruction of geometry and have led to the development of commercial solutions for capturing objects. Their cost starts from few hundreds of dollars (David 3D Scanner, *URL3*) to hundreds of thousands (Arius 3D, *URL4*). The most appropriate solution will depend mainly on the objects' size and the required accuracy of recording. In the case of historical maps a solution that we suggest would be a combination of

1) a scanning device for the high-fidelity reproduction of colour information
2) a low-cost 3D technique to extract the geometry capable of defining in low resolution the object's relief.

Our proposal suggests the additional use of software that will be able to exploit both the colour information and the extracted geometric content of the 3D recording of the historical map. The software uses the image representation of the accurate (colour) digital copy of the map and extracts precise measurements using the digital surface model created in whatever technique is available to each user. Thus without distorting the photographic imagery but exploiting the generated 3D model, users can derive correct measurements on the map. The only prerequisites are:

1) The generation of 3D digital model of the surface map
2) The registration (using automated procedures) of the image on the digital surface model that is able to define for every point of the map chromatic and additional geometric (in 3D) information.

In order to examine the validity of our theory that claims to provide precise measurements on the deformed images of historical maps, a case study was implemented.

Case study

In our case study a regular grid of dots are placed on a paper that is deformed to follow the shape of an internal book centerfold (Fig. 4). The grid has been recorded using *Photomodeler* software and the DSM (Digital Surface Model) and ortho image were generated.

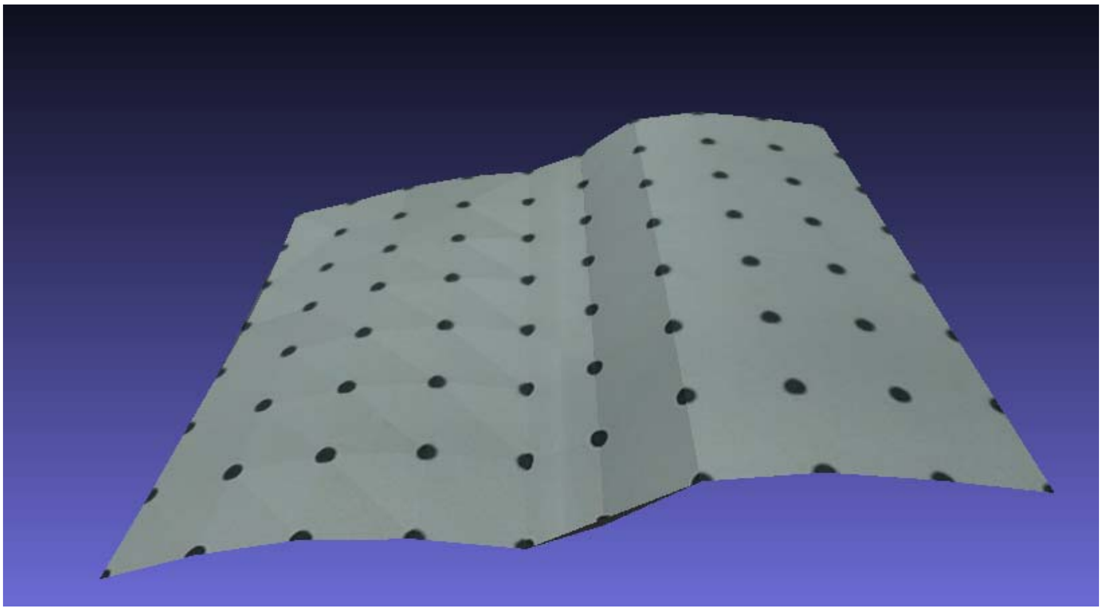

Figure 4: The DSM of deformed grid recorded using *Photomodeler* software.

If we use the unregistered or even the ortho image to measure the distance between grid points incorrect measurements will be derived since the first representation has no ground truth at all, while the second is shrinking the actual distance on the grid (Fig. 5).

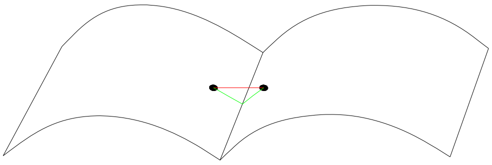

Figure 5: The actual distance between grid points is given as the sum of the green lines. The red line is giving the ortho-image distance measurement.

If we apply an algorithm that will take into account the actual slant distances of every intermediate point to its neighboring and integrate the distances for every pair of points we will unwrap digitally the deformed paper and produce the distance measured on the virtual original stretched paper. A first approach of the algorithm (Fig. 6) has been implemented using the OpenCV software library and generic image processing functions. The user can locate any pair of points on the ortho image and the software calculates the slant integrated distances for the path connecting the starting and ending location. The precision for the calculated unwrapped distance is relative to the precision of the generated relief and the grid size of the DSM.

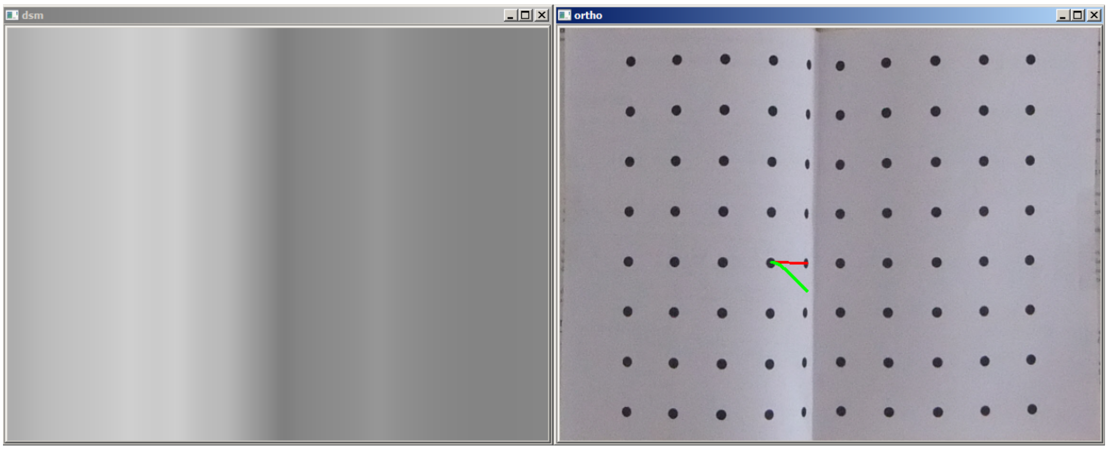

Figure 6: The software combines DSM depth image (left) and orthoimage (right). The user collects the location of the two points. The red line gives the distance on the ortho image and the green line gives their distance in slant range.

Conclusions

Although the work at its present state is still under development with respect to the overall digital treatment of atlas maps and bound maps, some definite conclusions are already drawn from its first step, regarding the suitability of scanning devices for atlases, bound maps and maps in books. From the tests carried out for this purpose it has become clear that the devices dedicated to the scanning of books in libraries are not equally appropriate for the scanning of maps in books such as atlases or bound map volumes, because these devices induce deformations unacceptable for maps, distorting their geometry and actually disregarding a fundamental characteristic of the nature of maps, the property of metric quality.

Having this characteristic in mind, the second stage of the work i.e. the development of a procedure for the digitization of maps in 3D is expected to provide sophisticated and highly accurate tools for cartographic representation and analysis, especially useful for the cases of atlases, bound maps and maps in books.

References

Tsioukas V., Daniil M., 2008: *3D digitization of historical maps*, proceedings of the 3rd International Workshop «Digital Approaches to Cartographic Heritage» organized by the ICA Working Group on Digital Technologies in Cartographic Heritage, Barcelona, 26-27/6/2008.

Tsioukas V., Daniil M., 2009: *3D digitization of historical maps, e-Perimetron, Vol. 4, No. 1, 2009 [45-52]*

URL1: *Structured light* - Wikipedia, the free encyclopaedia

URL2: *Structure from motion - Wikipedia*, the free encyclopaedia

URL3: *DAVID* 3D *Scanner*

URL4: *Arius3D*