

2D/3D Semantic Annotation of Spatialized Images for the Documentation and Analysis of Cultural Heritage

A. Manuel^{1,2}, P. Véron² and L. De Luca¹

¹MAP (UMR CNRS/MCC 3495), Marseille, France

²LSIS (UMR CNRS 7296), Arts et Métiers ParisTech, Aix-en-Provence, France

Abstract

The introduction of digital technologies in the documentation methods for cultural heritage has helped to design new tools for the acquisition and management of information collected for multidisciplinary studies. These tools are beginning to emerge as the preferred media for describing, analyzing and understanding the objects of study. Existing solutions for semantic annotation on images, on 3D models or with 2D/3D hybrid methods still reveal themselves today insufficient to tackle the complex problem of annotating heritage artifacts. In this field, the semantic description of the studied objects must be able to rely on a rich and structured representation by on one hand making explicit the morphological complexity of the object and on the other hand by reflecting all aspects conveyed by the acquisitions of scientific imaging.

This paper introduces an approach for conducting semantic annotations on 2D images (photography, scientific imaging ...) while facilitating the annotation work with an automatic propagation of these annotations between other correlated representations (2D or 3D) of the object. It is based on a spatial referencing method aiming at the establishment of a continuous 2D/3D projective relationship. The goal of the approach is to define an informative continuum between all phases of observation and description processes ranging from the acquisition of images and spatial data up to the building of semantically-enriched 3D representations. The idea is to insert semantics at all phases of 2D/3D data processing while ensuring a continuous correlation of annotations from a spatial, temporal and morphological point of view.

Categories and Subject Descriptors (according to ACM CCS): I.3.3 [Computer Graphics]: Picture/Image Generation—Digitizing and scanning - I.4.0 [Computer Graphics]: Image Processing and Computer Vision—Image processing software - I.4.6 [Computer Graphics]: Image Processing and Computer Vision—Segmentation

1. Introduction

Today, digital technologies provide to cultural heritage (CH) documentation experts new tools for the acquisition, the management and the dissemination of information and knowledge. These tools are beginning to assert themselves as interesting medium for the analysis, the description and the understanding of CH.

Methods for 3D digitization have known important evolutions thanks to the development of 3D laser scanning and geometry processing. The 3D model appears today as relevant medium for the documentation as it provides detailed and accurate geometric representations of the studied objects. However 3D does not respond yet to the everyday observation and analysis needs of the CH researchers and experts.

2D images have an important place in the documentation process of CH objects. Today they are easy to produce and they testify the state of an object at a given time. It is an effective tool for collecting, observing, analyzing, recording and disseminating information about heritage objects. Indeed, given the huge diffusion of digital cameras, photographs reveal themselves as an essential support to document the state of a historic object. They can be used

firstly as annotations medium and secondly to analyze the architectural forms and to characterize the surfaces state. Recent advances in photogrammetry introduce automatic reconstruction methods directly from images [FRN15].

2D images used for analysis are not limited to "standard" photographs, that is to say images reflecting the color of the object as it is in reality. In the last years, the scientific imaging applied to CH analysis has explored records of other types of radiation, such as infrared or ultraviolet. This help to bring out information that is not visible to the naked eye.

All the used images have not always the same geometric projection system: perspective projection for photographs (or infrared acquisition), orthogonal projection for orthophotos and spherical projection for TOF laser scanning acquisition.

During the CH object documentation process, the use of a graphical support or another one depends on the observation and analysis objectives. These different images are used in parallel in order to highlight different aspects. So these images are used for documentation as a basis for extracting information such as color, conservation state, geometric descriptors extraction, measurement, or

temporal transformation. As a consequence, an important issue today relies on the study of the conditions through which various 2D and 3D graphical representations can support the documentation process of heritage objects from a semantic and geometric point of view.

2. Semantic Annotation Methods

Whether it is images or 3D models, these representations can be enriched by a semantic annotation process. The principle of the annotation is based on linking a resource (partially or entirely) with additional information by using an information structure that can be defined in 4 models [AZP12] (tags, attributes, relationships and ontology). This aims at establishing advanced research.

Adding a semantic content to images is based on several methods. The simplest and most commons are manual methods. The user himself defines the terms to associate to the image or to image parts with simple tags [Mar] or ontologies [STS*06]. Automatic methods are based on image processing using segmentation, feature extraction, feature selection, classification and finally an annotation model, which can be either probabilistic, by classification, parametric or by using graphs [PS14]. There are also semi-automatic methods [LW06] that use the treatment process of automatic methods to offer a user a list of terms that have to be validated or not.

3D models can also be annotated. Here the idea is to associate the annotation to parts of 3D models, that is to say, points, surfaces, areas or objects in the scene ([3DS] [RASFO7] [MR10]).

[SSS06] and [Bus10] tried to combine the 2D information and the 3D information. Both systems rely on the use of a spatial relationship between the images and a 3D representation which allows introducing a link between 2D and 3D and thus enabling automatic propagation of annotations. These works show that the annotation process could be clearly improved by connecting images with a 3D representation of the object.

3. Approach

The aim of our work is to develop a hybrid approach for the semantic annotation of heterogeneous images. The main issue relies on the problem of propagating annotations between different representations of an object.

Our approach is based on images annotation and the use of a sufficiently dense spatial reference in order to define the semantic description as the common characteristic between the 3D representations of the object geometry and the various images of the acquisition. This requires introducing a complete methodology of 2D/3D semantic annotation based on the concept of spatial referencing. Our methodology thus includes:

- the registration of different images around a common geometric reference (represented by an implicit dense point cloud),
- the definition of a 2D/3D projective relation
- a semi-automated annotation propagation system
- the development of 2D and 3D analysis tools

These different aspects are explained in next sections.

The objective is to define an informative continuum at all phases of the documentation process from the acquisition of images and

spatial data to semantically enriched 3D representations by integrating the different aspects useful for the CH object analysis (multi-scale, multi-layers, multi-temporal etc).

The main advantage of this approach is that it takes into account the annotation within an overall process. Each of its concerns is based on works from the state of the art which are enriched by specific contributions.

4. Common Spatial Reference

The problem here is to establish the conditions under which the various 2D and 3D representations can be linked in the same spatial reference. The aim is to spatially link various iconographic sources in relation with a 3D representation of the object. This spatial relation is thereafter used in our method to create a projective relation between the two representations.

4.1. Point Cloud and Images Registration

The characteristics defining an image alignment in space depends on how this image was obtained and the image projection type. The position of a digital image around a 3D model has not the same characteristics as a spherical image or an orthoimage.

The CH survey nowadays includes the acquisition of point cloud by photogrammetric processing or by laser scanning.

In the case of photogrammetric processing, digital images alignment is directly computed during the process. Indeed the process evaluates the orientation and the calibration data of images and uses these results to generate the point cloud. Some of photogrammetric tools also include orthoimages generation and thus their position characteristics. The main advantage in this method is that the alignment between images and point cloud is totally consistent.

Spherical images alignment around a point cloud coming from a laser scanner survey is implicitly recorded if these images are also recorded by the laser device.

However the other alignment scenarios (spherical image/photogrammetry, digital or ortho-images/laser, photogrammetry/laser etc) could be processed with different methods of spatial referencing. For now, we mainly focused on the use of a photogrammetric tool (MicMac [PDDL11]) with digital and ortho-images but the other alignment scenarios are under consideration.

4.2. New Images Registration

As seen in the introduction, 2D sensors can acquire images containing information that is different from the real color of the object. The infrared, ultraviolet or thermographic images are as well useful for the analysis. To be integrated into our system, it is necessary that they can be oriented around the point cloud.

With the photogrammetric tool, these images cannot be directly used in the steps for the point cloud generation. However they can at least be oriented among images used in this generation. In this case, the ability to automatically orientate an image among other ones is based on the problem of automatically finding enough common tie points between images.

Our approach considers thus two alignment strategies with these images. If it is possible to automatically find enough common tie point with the already-oriented images, the new image can be

aligned directly among the other images with the photogrammetric tool. This is generally the case for infrared or ultraviolet images. On the other hand, if it is not possible to find these common points (for example, with thermographic images), the image cannot be integrated directly in the images base. Here we propose a method where the images integration requires another image which can be directly oriented in the already-oriented images base and whose orientation is known relative to the image that cannot be oriented. This supposes to have specific acquisition protocols.

So, these alignment strategies allows us to register different scientific imaging images in the same spatial reference and relative to a point cloud.

4.3. Projective Relation

Starting from these results, registration data and point cloud, our methodology includes the establishment of a projective relation between images and point cloud. We consequently consider separately each images projection types.

For digital images, we interpret the orientation and the calibration data to define a transformation of a 3D point to a pixel position in the image. The calibration data allows taking into account the optical distortion aspect of the acquisition device. Thus correspondences between 3D points and pixels of the image are obtained.

For orthoimages, the image orientation consists in identifying the position of the image surface in space. From this surface, we orthogonally project 3D points on pixels to get the 2D/3D correspondences.

For spherical images, the orientation consists of the definition of a point, position of the sphere center in space, and at least one direction. The projection is made from these spherical data and also gives the 2D/3D correspondences.

In all cases, this projection provides a precise correspondence between the 3D points and pixels of images. So whatever the type of image projection is, we create a common relation between the images pixels and 3D points. To easily store this relation, we index the common point cloud. Then, each pixel of images is connected with an index corresponding to the projected 3D point.

5. Hybrid Annotation Process

This spatialization and the 2D/3D link show that it is possible to structure the 3D data in relation with 2D data. Thanks to the projective relation we introduce, each pixel of images is linked directly with data contained in the image (color, gradient etc) and with the data of the point cloud (3D point coordinates, normals etc). Thus this projective relation allows developing new interaction scenarios between images and point clouds.

5.1. Selection Tool

To annotate an image, the first problem is to select interesting areas on the image. This can be done completely manually but some 2D or 3D semi-automated selection tools already exist and can be improved by the use of the 2D/3D link.

5.1.1. Geometry-based Magicwand

The commonly used "magic wand" tool selects pixels based on color and a threshold. The idea of this tool relies on a simple principle: a reference pixel and a tolerance define a set of colors around

the reference pixel color. Each image pixel, whose color is in the set, is selected.

As the projective relation links pixels with different types of data, our method explores this principle on other data types. For example with normals, a reference point is defined on the image, and by the projective relation defined a reference normal. The tolerance is defined here by a maximum angle to the reference normal. All the points of the image satisfying the criteria may be extracted or only contiguous points around the reference point (Figure 1).

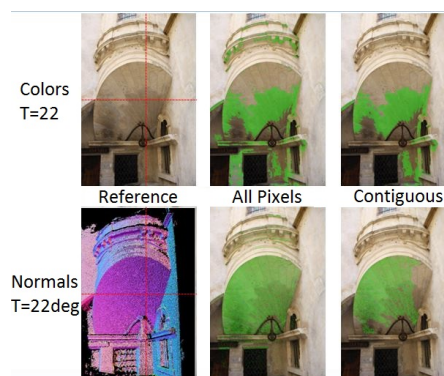


Figure 1: Magic Wand on different data

5.1.2. Images and Point Cloud Segmentation

Rather than defining a tolerance, it is possible to prior subdivide the image using segmentation methods according to various criteria. There are many methods of segmentation on images or on point clouds. Our approach uses 2D or 3D segmentation results in a self-expanding selection method. By comparing the selected area and the segmentation, we extended the selection to the covered part of the segmentation in the image view (Figure 2).

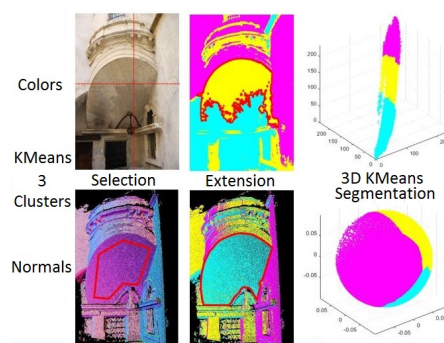


Figure 2: Segmentation selection with the KMeans segmentation method

5.2. 2Dto3Dto2D Annotations Propagation

Although these two tools allow to quickly select image areas, the specificity of heritage survey supposes to acquire large amounts of images of all types for the analysis and the documentation of the object. So even if these tools can facilitate the selection of areas on images, it is not conceivable to annotate all images one by one.

The developed approach therefore incorporates our previous

work about the automatic propagation of annotations that exploits this projective relation. The detailed process has already been described in [MDLV14] (Figure 3). With the point cloud indexation, our process now only needs to make comparison on indexes and no more on coordinates triplets. The indexation table of point cloud helps converting the annotation indexes into a point cloud. Thus with the projective relation, we create annotations that are represented both by regions on images of different types and a 3D region which contains important geometric information.

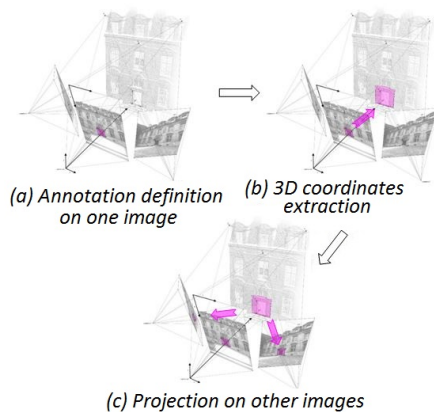


Figure 3: Propagation Steps

6. Analysis Tool

The analysis and the documentation of CH do not rest only on annotating. 2D or 3D analysis tools are also used.

6.1. Structuring Multi-Layer Annotations

Our method structures annotations by using the layer concept. The principle is that a layer represents a specific aspect and therefore can contain several regions relative to this aspect. Thus regions of a layer do not overlap. The use of such a structure allows us introducing two main methods for the information extraction.

The analysis of heritage object implies to explore the distributions of the aspect conveyed by a layer. The distribution comparison of the different layer regions bring out the major characteristic. Layer regions can therefore provide statistical information.

Furthermore the analysis also requires comparing two regions of two different layers. The superposition study of these two regions gives an indication on the possible correlation of two aspects.

6.2. Geometric Information Extraction

The point cloud contains a significant wealth in terms of geometry. [NDL16] have shown that the point cloud can be transformed into an enriched database. This work discusses the extraction of several geometric descriptors (curvature, roughness, ambient occlusion, surfaces, volumes etc) from a point cloud. With our 2D/3D relation, these extracted descriptors are also linked with images pixels. Thus our system integrates a visualization of these 3D descriptors directly on images. They can therefore guide the user in manual selection or be exploited with the previously-presented selection tools.

7. Conclusion

This paper describes an approach for the 2D/3D semantic annotations of heterogeneous data of CH objects. The 2D/3D bijective relation we introduce is central in our methodology. Our goal is to provide a complete annotation tool to CH researchers and experts where the morphological complexity of the object can be exploited in relation with the different scientific imaging records. Our model provides thus a consistent and continuous relationship between heterogeneous data used for documentation.

The different processes of our approach are at present implemented in several modules. But the aim is to integrate them, in the future, in a collaborative web-based annotation system. This supposes consequently to consider the crowdsourcing issues in further development.

References

- [3DS] <http://3dsa.metadata.net/3dsa/>. Accessed: 2016-04-30. 2
- [AZP12] ANDREWS P., ZAIHAYEU I., PANE J.: A classification of semantic annotation systems. *Semant. web* 3, 3 (Aug. 2012), 223–248. 2
- [Bus10] BUSAYARAT C.: *La maquette numérique comme support pour la recherche visuelle d'informations patrimoniales*. PhD thesis, SMI (Sciences des Métiers de l'Ingénieur) Arts et Métiers ParisTech, 2010. 2
- [FRN15] FORLANI G., RONCELLA R., NARDINOCCHI C.: Where is photogrammetry heading to? state of the art and trends. *Rendiconti Lincei* 26, 1 (2015), 85–96. 1
- [LW06] LI J., WANG J. Z.: Real-time computerized annotation of pictures. In *Proceedings of the 14th ACM International Conference on Multimedia* (2006), pp. 911–920. 2
- [Mar] www.marqueed.com. Accessed: 2016-04-30. 2
- [MDLV14] MANUEL A., DE LUCA L., VÉRON P.: A hybrid approach for the semantic annotation of spatially oriented images. *IJHDE (International Journal of Heritage in the Digital Era)* 3, 2 (2014), 305–320. 4
- [MR10] MANFERDINI A. M., REMONDINO F.: Reality-based 3d modeling, segmentation and web-based visualization. In *EuroMed* (2010), Ioannides M., Fellner D. W., Georgopoulos A., Hadjimitsis D. G., (Eds.), vol. 6436 of *Lecture Notes in Computer Science*, Springer, pp. 110–124. 2
- [NDL16] NESPECA R., DE LUCA L.: Analysis, thematic maps and data mining from point cloud to ontology for software development. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (2016), 347–354. 4
- [PDDL11] PIERROT-DESEILLIGNY M., DE LUCA L., REMONDINO F.: Automated image-based procedures for accurate artifacts 3d modeling and orthophotographs generation. *Geoinformatics FCE CTU Journal* 6 (2011), 291–299. 2
- [PS14] PANDYA D., SHAH B.: Comparative study on automatic image annotation. *International Journal of Emerging Technology and Advanced Engineering* 4, 3 (Mar. 2014), 217–222. 2
- [RAS07] ROBBIANO F., ATTENE M., SPAGNUOLO M., FALCIDIENO B.: Part-based annotation of virtual 3d shapes. In *Proceedings of the 2007 International Conference on Cyberworlds* (Washington, DC, USA, 2007), CW'07, IEEE Computer Society, pp. 427–436. 2
- [SS06] SNAVELY N., SEITZ S., SZELISKI R.: Photo tourism: Exploring photo collections in 3d. *ACM Trans. Graph.* 25, 3 (July 2006), 835–846. 2
- [STS*06] SAATHOFF C., TIMMERMANN N., STAAB S., PETRIDIS K., ANASTASOPOULOS D., KOMPATSIARIS Y.: M-ontomat-annotizer: Linking ontologies with multimedia low-level features for automatic image annotation. In *Poster & Demo Session, ESWC 2006* (2006). 2