



HAL
open science

Time transcendance, metadata and future utilization in 3d models of point clouds for heritage elements

Alvaro Rodriguez Miranda, José Manuel Valle Melon, Ane Lopetegui
Galarraga

► To cite this version:

Alvaro Rodriguez Miranda, José Manuel Valle Melon, Ane Lopetegui Galarraga. Time transcendance, metadata and future utilization in 3d models of point clouds for heritage elements. Virtual Retrospect 2007, Robert Vergnieux, Nov 2007, Pessac, France. pp.115-123. hal-01770106

HAL Id: hal-01770106

<https://hal.science/hal-01770106v1>

Submitted on 18 Apr 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



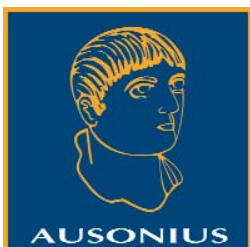
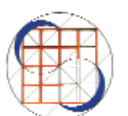
Vergnien R. et Delevoie C., éd. (2008),
Actes du Colloque Virtual Retrospect 2007,
Archéovision 3, Editions Ausonius, Bordeaux

Tiré-à-part des Actes du colloque Virtual Retrospect 2007

Pessac (France) 14, 15 et 16 novembre 2007



J.M. Valle Melon, A. Lopetegi Galarraga, A.R. Rodriguez Miranda,
*Time transcendence, metadata and future utilization in 3d
models of point clouds for heritage elements* pp.115-123



Conditions d'utilisation :
l'utilisation du contenu de ces pages est limitée à un usage
personnel et non commercial.
Tout autre utilisation est soumise à une autorisation préalable.
Contact : virtual.retrospect@archeovision.cnrs.fr

<http://archeovision.cnrs.fr>



Time transcendence, metadata and future utilization in 3d models of point clouds for heritage elements

Álvaro Rodríguez Miranda alvaro_rodriguez@ehu.es

José Manuel Valle Melón jm.valle@ehu.es

Ane Lopetegui Galarraga ane.lopetegui@ehu.es

Laboratorio de Documentación Geométrica del Patrimonio
Universidad del País Vasco-Euskal Herriko Unibertsitatea (Espagne)
<http://www.ldgp.es>

Résumé : Le patrimoine n'est pas une réalité statique, bien au contraire, il subit des modifications et transformations qui, à la limite, vont jusqu'à sa complète disparition ; en conséquence, l'information à propos du patrimoine doit être considérée aussi comme du patrimoine, ce qui veut dire qu'elle doit être entretenue pour son utilisation par les générations suivantes. Non seulement, les futurs utilisateurs doivent être capables d'accéder à l'information, mais aussi ils auront besoin de la comprendre et même, prévoir un nouvel usage différent de celui qui fut l'origine de son acquisition. Étant donnée que, de nos jours, la presque totalité de l'information est numérique, nous sommes en train de traiter avec des concepts comme transcendance temporelle, métadonnées et ré-utilisation. Dans ce texte, on fera un tour d'horizon de ces concepts et leurs applications pratiques à un type de données dont l'utilisation est très étendue dans la documentation géométrique du patrimoine : les nuages de points.

Mots-clés : Métadonnées, ré-utilisation, transcendance temporelle, héritage, nuage de points

Abstract : Heritage is not static, on the contrary, it is subject to continuous modifications and transformations which, at worst, go as far as to their complete disappearance; for this reason, information referring to heritage must be considered in itself like heritage, which implies that it must be maintained able to be used for the future generations. Not only future users have to be able to access to the information, but also they will need to understand it and even it should be foreseen that the new use could be different from the one which caused its creation. Bearing in mind that, nowadays, almost all the information is digital, we will be dealing with concepts such as time transcendence, metadata and re-use. In this text we will give a general overview of these concepts and their practical application to a specific kind of data widely used in geometric documentation : the point clouds.

Keywords : metadata, re-use, time transcendence, heritage, point clouds

INTRODUCTION

The tangible cultural heritage, specially the built one, undergoes continuous processes of modification, restoration, change of use, and even disappearance. Due to these circumstances the original data and final representations of the mentioned elements, such as descriptions, photographs, engravings, plans or measurements, become heritage elements in many times.

The geometric documentation registers the form, dimensions and layout. Considering that, nowadays, almost all the used data carriers are digital, transcendence mechanisms will have to be settled down in order to pass on the information that conforms it to the future generations. On the other hand, this information have to be understood by them, that is to say, they will have to know the data type, its characteristics, organization, etc.

In the following sections, in which this text is articulated, these concepts and their practical application to the documentation projects are reviewed. Although the exposed ideas are applicable to any documentation technique, special attention will be paid to the register of heritage elements by means of point clouds due to the great development and spread it presently enjoys.

Specific procedures allowing the conservation and recovery of the data in the future will be exposed, we will focus the transcendence on two aspects : firstly, the physical transcendence, it means that we will be able to read the carrier in which the data is stored; secondly, the transcendence of the information itself, implying that we will dispose of a software able to understand the stored data.

Once the mechanisms have been established to guarantee the data will be consulted in the future, it is necessary to decide what information will be stored, two possibilities will be taken into account :

- If, in the future, the geometric documentation will be used with the same purpose for which it has been generated, only the final results must be stored; however, they have to include all the complementary data for their complete definition in form of metadata.
- If it is anticipated that this documentation can be used, at least partially, with different aims, it will be also necessary to store the original data with their corresponding metadata and the workflow.

TRANSCENDENCE

Data transcendence has been a patent problem for years, aggravated by the rapid evolution of the hardware and software. The expected usable life for most data carriers is not longer than 5 years [Webb, p. 113] that is the reason to lay down measures if we want to overcome this limit.

A common way to circumvent this eventuality has consisted of making massive copies when migrating to a new hardware, but file formats also stop being understood by new releases of the software, so it is necessary to update them or, at worst, keep in operation a series of obsolete hardware.

- In order to avoid the obsolescence of data carriers, the most important thing is not to store data in a static way but to put them into a cyclic sequence of copies which enables the existence of one, at least, in a current carrier. Actually, this is the less demanding process and external companies can offer it like the web-hosting services provided they guarantee the back-up copies and the restoration of the data in case of failure.
- Proprietary formats do not ensure their access in the future so it is better to use open formats, if possible easy to read (plain text), well documented and whose documentation is available to the public, preferably with international standards (ISO, DIN, UNE, ...). The more it is accepted and widely used today, the more it will be likely to find a software able to read it in the future, nonetheless, if we did not have a proper software, the easier to read and the better documented, the easier it will be translated into a current format.

Going on about the formats, it is interesting to distinguish between work files and files aimed at the storage. During the development of the project we use the more suitable formats according to the available software, at this moment, the choice between open or proprietary formats will be defined by the policies and the resources of the company; but what is inadmissible is to issue the final outputs and digital reports only in proprietary formats because that puts in jeopardy the future of the data by making compulsory the ownership of that software by the user.

In the particular case of the point clouds, a simple text file with the data arranged in columns, together with a brief description indicating the meaning of each column is a very efficient system to ensure the transfer through the time. The description of the columns is necessary for the user to

understand the meaning of the data and it will be a part of the metadata, of which we will return later.

Although the resulting files are of a big size, we must be reluctant to compress the data because we often relapse into the proprietary systems and we lost the legibility.

RE-USE OF THE GEOMETRIC DOCUMENTATION

Points clouds, in the same way that photogrammetry, are a mass register method for geometric data. Cartographic products obtained from these systems can be defined as "indirect" because, during the register phase, it is registered in a systematic way all the possible information, later, at the processing phase, we will extract this one which is important for the aimed cartographic output. On the contrary there are the "direct" methods, such as surveying, in which it is only registered the information which is going to appear in the final model and plans.

In direct methods, the aim of the documentation is clearly defined from the beginning because it conditions the data to capture, but in the indirect ones that is not compulsory, which permit the re-using of the information, that is to say, the same laser scans or photogrammetric pair can be used today to analyse several features and, in some years, to study completely different ones not considered in the original planning. Re-use is a main concept due to the fact that, in many cases, the object is not longer accessible and it is not possible to go back in order to made a new documentation (table 1).

A simple documentation project made from indirect methods will consist of a set of raw data (registers) from which the outputs have been obtained. A new user, who wants to re-use this data in the future with a different aim, will begin heading towards the outputs, she or he will analyse their features (metadata) in order to see whether they can be used directly, otherwise, the original data and corresponding metadata will be analysed to check if, at least, a new set of outputs could be obtained from them. If even the registers did not give the required information, they would have to go back to the object (fig. 1).

In any case, it is easily deduced that, without the existence of the metadata, the re-use is not possible because we cannot know if the features of the results or registers fit the necessities required by the new aims.

A more complete case will be analysed now : the "Portada de los Hierros" of the Cathedral of Valencia (Spain). In order to focus briefly the project, we can say that it is a 35 meters high façade which has been documented by means of about 150 laser scans (30 millions of points per scan), the final product is a model of the surface with a point every 5 mm.

The following workflow describes the structure of the work; it shows the field processes in red, the office ones in blue and the outputs in red (fig. 2).

In this case, the original data are the raw point clouds (1) plus the surveying data about coordinates of the benchmarks

	Direct methods	Indirect methods
Recorded data	Only the important one for the pursued aim.	Massive, all the visible data within the established resolution.
Possibility of re-use with different aims	Very limited, data only fit the original purpose.	Yes.
Documents to store	Final outputs with their features (metadata).	Registers and outputs, both with their features (metadata).
Necessity for documenting processes	Yes, in order to guarantee the features of the outputs.	Yes, in order to guarantee the features of the outputs and to permit re-using.
Examples	Surveying, GPS.	Photogrammetry, laser scanner.

Table 1. Differences between direct and indirect methods for geometric documentation.

and sketches permitting their identification in the scans. The accuracy of the measures will be part of the metadata. Scans are oriented (alignment) putting all them in the same coordinate system (2) by identifying the benchmarks in each point cloud (fig. 3).

Point clouds go through an edition process in which it is removed all the information not related with the object to document, in this case, the façade (3) (fig. 4).

The next step consists in isolating each element on the façade : columns, statues, reliefs, etc (4). For each one, all the available information is added and it is processed in order to generate a single model with uniform resolution (a point every 5 mm). Then, we add a new homogeneous colour by means of artificial shading (fig. 5).

Finally, from this model, the products such as coloured thematic models (5), cross-sections (6), orthoimages (7) or three-dimensional line drawing (8) are obtained (fig. 6).

As it can be seen, each processing step removes or modifies part of the input data (those not interesting for the aim of the current project) and it obtains a refined product, closer to the aimed objective.

The problem in re-use is that we do not know what will be valuable in the future, hence, it is advisable to store the intermediate steps, allowing in that way the possibility of recovering data just before the moment in which they were erased, that is, when they were more refined.

The following table shows a series of hypothetical re-uses not foreseen in the initial project but useful for new researches. The table also shows in which level the project should be resumed (table 2).

Some of these mentioned uses are illustrated by the following image, as it can be seen that, apart from the façade, it contains information about a craftwork market. All this information will be removed during the edition process since it is not interesting for the actual aim, however, in the future, it can be most interesting for ethnographic purposes or scaffolding systems (fig. 7).

In case we supply with different intermediate products, it is compulsory that the users also have at their disposal the workflow which permit them to know the relationships between the elements, since the workflow is the key which arranges the levels of data processing and allows the organised transition through the processes.

On the other hand, each process (transforming the point cloud to the following level) has to be conveniently documented. Here are the parameters which have to be considered at any rate :

- Process identification : a descriptive name.
- Process aim : what the process intends to do.
- Input data : registers with their technical features, outputs from previous process and complementary information (sensor calibration parameters, sketch with the location of the benchmarks, ...).

Aim	Level from which the project should be resumed
Obtaining three-dimensional lines of different elements.	Classified model
Producing a new thematic cartography by following different criteria.	Classified model
Generating a new model with different spatial resolution or colour.	Edited point clouds
Re-classifying the façade into different elements.	Edited point clouds
Study of scaffolding systems at the beginning of the 21 st century	Oriented point clouds
Ethnographic studies, such as traditional pottery.	Oriented point clouds
Study of the accuracy of algorithms for orientating point clouds.	Raw point clouds

Table 2. Examples of re-use and the level of the workflow from which we should resume.

Type of metadata	Fricher (2003)	Extension
General (compulsory for any file or document)	- Catalogue metadata	- Name of the schema
- Catalogue metadata		
Specific (optional, according to the characteristics of the data to describe)	- Commentary metadata	
- Bibliographical metadata		
	- Commentary metadata	
- Bibliographical metadata		
- Additional parameters		
Other (e.g. preservation metadata)		They will be appended like additional parameters until a new specific group is defined.

Table 3. Metadata schema presented by Fricher and proposed extension.

- Actions performed by the process (description and mathematical background).
- Outputs and their technical features.

METADATA

In this section we will focus on selecting the metadata to include with point clouds and the way to do it.

Roughly speaking, we can indicate three different category of metadata :

- General metadata (e.g. Dublin Core, ISO 19115, ...) applicable to whatever file or type of information (photograph, report, plan, multimedia, etc.) such as title, subject, data, author, ...
- Specific metadata, depending on the subject we can add some metadata related with heritage, informatics, history, surveying, ...
- Other, like preservation metadata. Both previous groups correspond to the metadata of the document when it was delivered for the producer, nevertheless, it is possible to add new metadata where we track the storage of the file, for instance the different carriers, backup copies, migrations, integrity indicators, etc.

If we analyse the schema presented by Fricher [Fricher 2004], we will see that it includes the Dublin Core as general metadata (called catalogue metadata in the reported source) and two specific groups for virtual models (commentary metadata and bibliographical metadata)

We propose to adjoin two more groups to this schema :

- Name of the schema : in first place, the name of the schema should be indicated, this name can refer to an international standard (ISO_???) or a collection made by a producer (My_schema v1.0). Moreover, a mean of accessing to the description of this schema have to be supplied, this way, the user will understand the meaning of each value, this access could be a link to a web document, but it should be taken into account that the average life of a web page does not last too much (between 44 days and 2 years, [Webb]), therefore, if we

want to guarantee a lasting access we should indicate the location of a physical copy (paper), for this purpose, an official registry will be a suitable place to deposit the copy since it ensures public access, integrity and rights.

- Additional parameters : no matter how complete our schema can seem, users will be always trying to append some unexpected information, hence, it is interesting to make our schema flexible, for instance, by means of a block of additional parameters where the user can define their own metadata. In this case, we propose couples of records indicating, on the one hand, what is stored and, on the other hand, its value. These additional parameters are the source for new groups of specific metadata.

Returning to our case study, the points clouds of the Cathedral of Valencia, at the beginning we thought of including the metadata at the heading of the file with the coordinates but we considered that this solution would made more difficult to read them in the future, then, we decided to include them in a separate file, like the following one (table 4).

CONCLUSIONS

Documents about the geometry of the heritage become heritage in itself, therefore, they have to be preserved and for this purpose transcendence criteria have to be established.

It is essential to distinguish the file formats used while we are processing the data and the ones suitable for storing the final outputs, which will be designed for being read easily.

When storing data with the aim of re-using, the workflow is the key element which structures the project. In this case, in addition to the registers and the results, the process have to be also documented and the additional information (sketch, calibration reports, ...) have to be included. If we do it like that, a future user will be able to redo the entire project and to obtain the same results from our original data (bibliographical metadata can be very useful for this purpose, since they can be used to link the documents).

Metadata are unavoidable to understand the project and re-use is not possible without them. Nonetheless, we should

```

# METADATA
# BLOCK 1: Name of the schema
# metadata_type: ldgp_metadata_v0.1
# metadata_type_definition: Registro de Propiedad Intelectual (España): 00/2006/4163 (2006_01_09)
# metadata_type_description:
http://www.vc.ehu.es/docarq/LDA/publicaciones/ldgp_metadata_v0_1es.pdf
# BLOCK 2: General metadata (Dublin Core)
# Title: Nubes de puntos editadas de la portada de Los Hierros de la Catedral de Valencia
# Creator: Laboratorio de Documentación Geométrica del Patrimonio (Universidad del País Vasco-
Euskal Herriko Unibertsitatea)
# Subject: portada, catedral, nube de puntos, Valencia
# Description: Nube de puntos orientada con valores de coordenadas y niveles digitales
# Publisher: Laboratorio de Documentación Geométrica del Patrimonio (Universidad del País Vasco-
Euskal Herriko Unibertsitatea)
# Contributor: none
# Date: 20070201
# Type: 3d
# Format: Ascii
# Identifier: ldgp_2007_valencia_nube_edit_a0_reja2_color.txt
# Source: Documentación Geométrica de la Portada de Los Hierros de la Catedral de Valencia
# Lenguaje: es-ES
# Relation: http://www.ldgp.es
# Coverage: 2006
# Rights: Generalitat Valenciana (Consejería de Cultura Educación y Deportes Dirección General de
Patrimonio Cultural Valenciano y Museos, Servicio de Arquitectura), Laboratorio de Documentación
Geométrica del Patrimonio (UPV-EHU)
# BLOCK N: Additional parameters
# add_par_definition(1): Estructura de la información
# add_par_description(1): Coord.X Coord.Y Coord.Z Canal_Rojo Canal_Verde Canal_Azul
# add_par_definition(2): Número de puntos
# add_par_description(2): 303313
# add_par_definition(3): Valor máximo de la coordenada X
# add_par_description(3): 473,715
# add_par_definition(4): Valor mínimo de la coordenada X
# add_par_description(4): 471,368
# add_par_definition(5): Valor máximo de la coordenada Y
# add_par_description(5): 1005,685
# add_par_definition(6): Valor mínimo de la coordenada Y
# add_par_description(6): 1002,423
# add_par_definition(7): Valor máximo de la coordenada Z
# add_par_description(7): 205,563
# add_par_definition(8): Valor mínimo de la coordenada Z
# add_par_description(8): 200,705
# add_par_definition(9): Valor máximo del canal Rojo
# add_par_description(9): 255
# add_par_definition(10): Valor mínimo del canal Rojo
# add_par_description(10): 0
# add_par_definition(11): Valor máximo del canal Verde
# add_par_description(11): 255
# add_par_definition(12): Valor mínimo del canal Verde
# add_par_description(12): 0
# add_par_definition(13): Valor máximo del canal Azul
# add_par_description(13): 255
# add_par_definition(14): Valor mínimo del canal Azul
# add_par_description(14): 0
# add_par_description(15): Tamaño de muestreo (metros)
# add_par_definition(15): 0,005
# add_par_definition(16): Precisión geométrica (metros)
# add_par_description(16): 0,01

```

Table 4.

prevent that the user receive metadata whose meaning is unknown, so, the schema will be as well at their disposal.

As we have seen, it is possible to define our own schemas, however, it will be always better if we reach to a consensus among groups of producers, like that, the number of schemas will be reduced significantly and their efficacy improved. On the other hand, this consensus would permit keeping a working group entrusted to the maintenance of the schema by collecting suggestions of the users and by publishing newer versions.

We want to finish with a reflection: after reading the title of this paper someone might have supposed that it would contain a collection of very sophisticated techniques to store and descriptions of the more recent file formats, however, the main recommendation stated is to use plain text files, even we did not present any standard to organize the information. All this is due to the fact that, above the standards, we should keep to the criteria of logic, simplicity and the ability of self-description, qualities towards which any standard have to head.

Acknowledgments

The project "Documentación Geométrica de la Portada de los Hierros de la Catedral de Valencia" was carried out at the request of the Dirección General de Patrimonio Cultural Valenciano y Museos of the Consejería de Cultura, Educación y Deporte of the Generalitat Valenciana from the Fundación Euskoiker.

Bibliography

- Bryan, P and B. Blake (2000): *Metric Survey Specifications for English Heritage*. English Heritage (United Kingdom).
- Bryan, P (2003): *An Addendum to the Metric Survey Specifications for English Heritage – the Collection and Archiving of Point Cloud Data Obtained by Terrestrial Laser Scanning or other Methods*. English Heritage (United Kingdom).
- Frischer, B. (2004): "Mission and Recent Projects of the UCLA Cultural Virtual Reality Laboratory", *Proceedings of the conference Virtual Retrospect 2003 - Biarritz (France)*, 65-74.
- (2006): "New Directions for Cultural Virtual Reality: a Global Strategy for Archiving, Serving and Exhibiting 3D Computer Models of Cultural Heritage Sites", *Proceedings of the conference Virtual Retrospect 2005 - Biarritz (France)*, 168-175.
- Ioannides, M., A. Georgopoulos and M. Scherer (2005): *Standards in Cultural Heritage: the Missing Grammar for the Digital Documentation of the Past*. CIPA XX International Symposium, Torino (Italy).
- Mendez Rodríguez, E. (2002): *Metadatos y recuperación de la información*, Ed. Trea (Spain).
- Paquet, E. and H. L. Victor (2004): *Long-term Preservation of 3-D Cultural Heritage Data Related to Architectural Sites*. XXth Congress of the International Society for Photogrammetry and Remote Sensing. Istanbul (Turkey).
- Vergnieux, R. (2006): "Archéogrid: vers un conservatoire national des données 3D du patrimoine", *Proceedings of the conference Virtual Retrospect 2005 - Biarritz (France)*, 157-167.
- Valle Melón, J. M. (2007): *Documentación geométrica del patrimonio (propuesta conceptual y metodológica)*. Unpublished PhD thesis, Universidad de La Rioja (Spain).
- Webb, C. (2003): *Guidelines for the Preservation of Digital Heritage*. National Library of Australia (Australia).

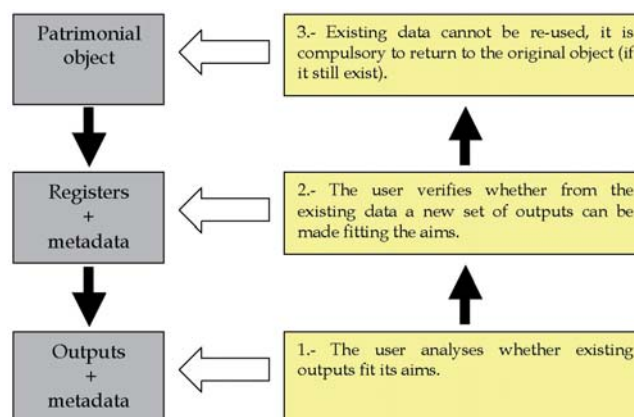


Fig. 1 Data re-use, the user checks, first of all, whether the outputs fit the current aims; otherwise, the original records will be analysed; finally, if these records are not useful either, re-use will not be possible, then, it will be compulsory to return to the patrimonial object.

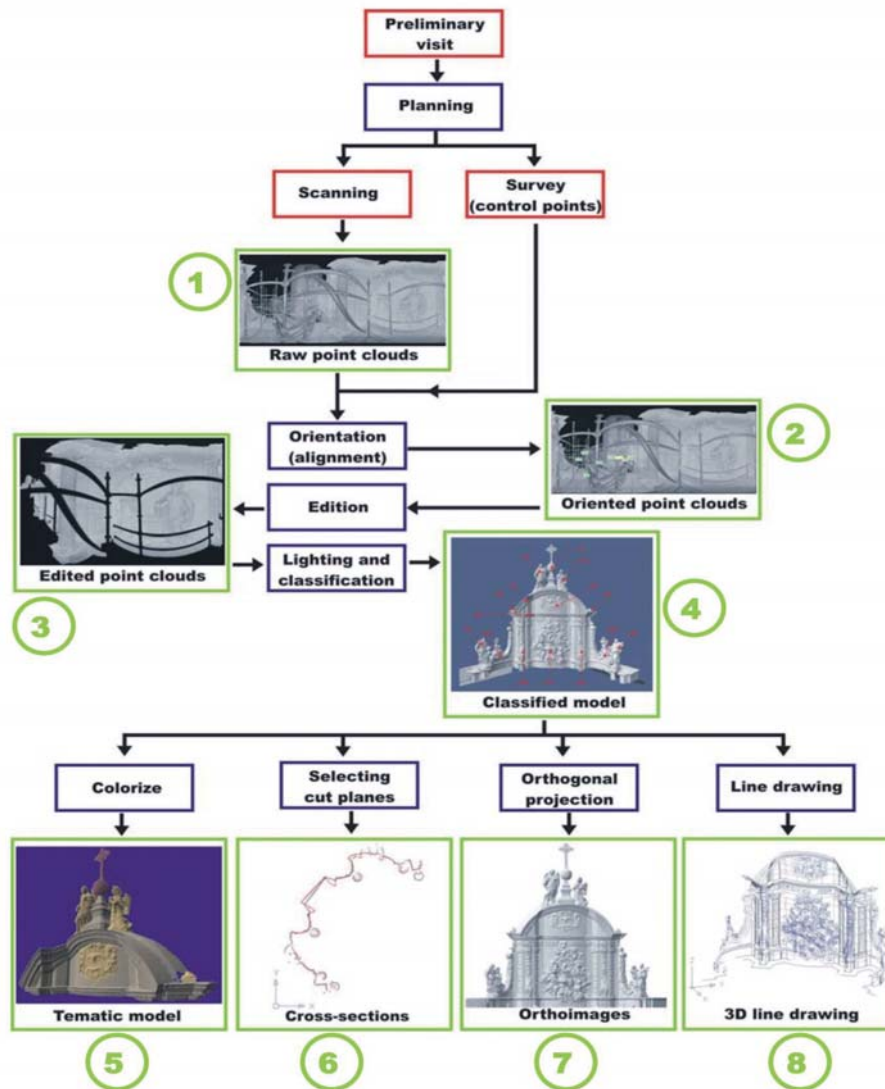


Fig. 2. Workflow of processes and obtained products.

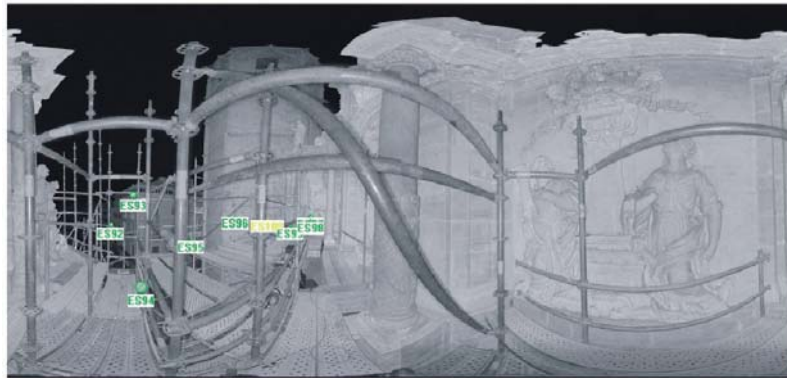
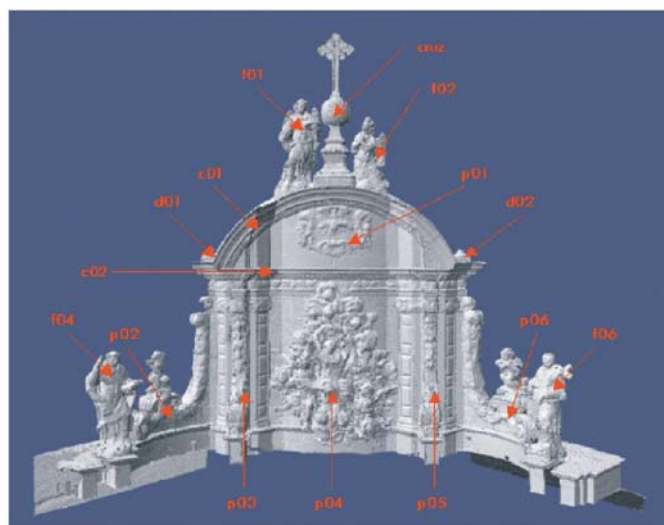


Fig. 3 Scan orientation by identifying the visible benchmarks.



Fig. 4 During the edition process, all the information not related with the main object is removed.

Fig. 5 Final model, classified in elements with a uniform resolution (every 5 mm) and with colour from artificial shading.



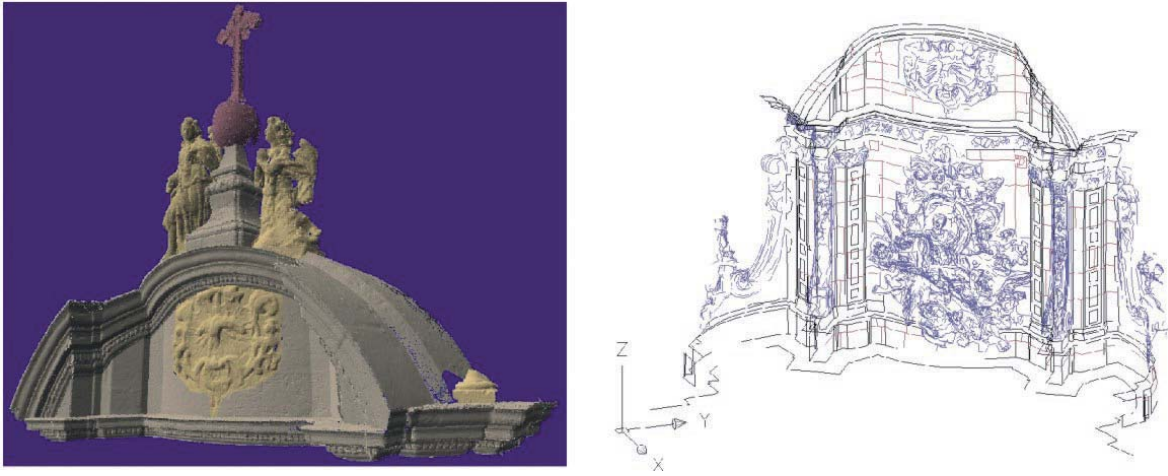


Fig. 6 Examples of stemmed products, on the left a thematic representation, on the right, a line drawing.

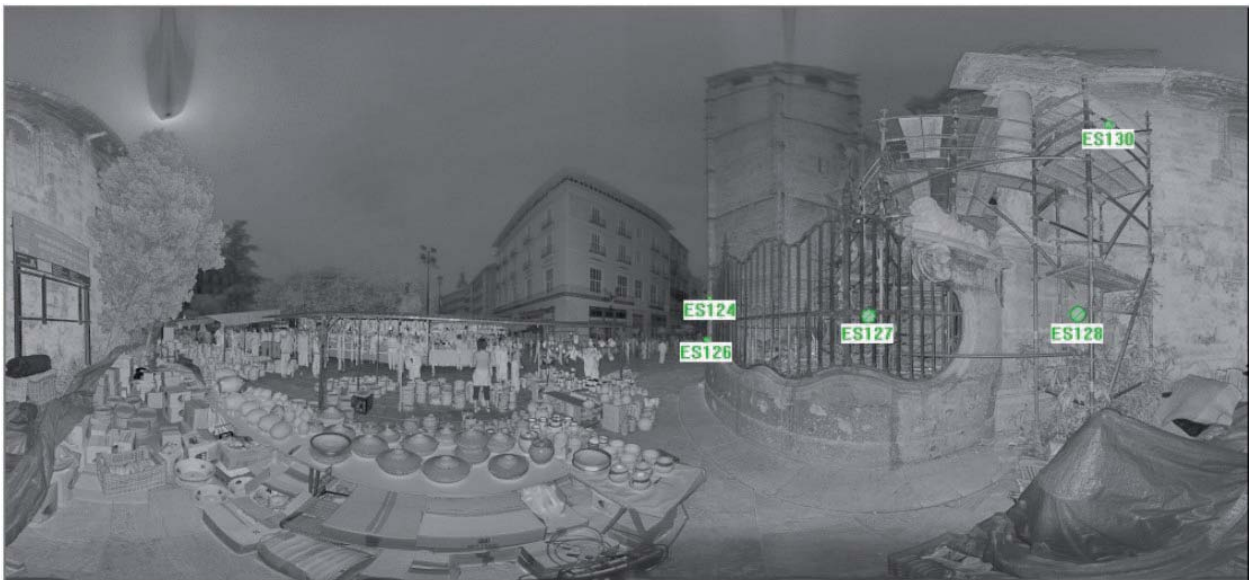


Fig. 7.- Casual record of a craftwork market.