An Information System for Medieval Archaeology Based on Photogrammetry and Archaeological Database: The *Shawbak Castle Project*

Pierre Drap, Djamal Merad, Julien Seinturier, Jean-Marc Boï, Daniela Peloso[1], Guido Vannini, Michele Nucciotti, and Elisa Pruno[2]

[1] LSIS umr CNRS 6168, Centre National de la Recherche Scientifique, Marseille, France
`name.surname@univ-amu.fr`

[2] Dipartimento di Studi Storici e Geografici dell'Università degli Studi di Firenze, Florence, Italy
`{guido.vannini,nucciotti,elisa.pruno}@unifi.it`

Abstract. The paper presents an interdisciplinary project which is a work in progress towards a 3D Geographical Information System (GIS) dedicated to Cultural Heritage with a specific focus application on the Castle of Shawbak, one of the best preserved rural medieval settlements in the entire Middle East). The Shawbak archaeological project is a specific and integrated project between medieval archaeological research and computer vision done thanks to a long cooperation between University of Florence and CNRS, LSIS, Marseille. Focusing mainly on stratigraphical analysis of upstanding structures we provide archaeologists with two-step pipeline. First a survey process using photogrammetry, both in a traditional way with additional annotations and using the most advanced technique to obtain dense maps and then a tool for statistical analysis. Two main applications are presented here, stratigraphy analysis with Harris matrix computed on the fly from the 3D viewer and statistical tools, clustering operation on ashlar in order to show new relationships between the measured artifacts.

All these developments are written in Java within Arpenteur framework[4].

Keywords: *Light* Archaeology, Photogrammetry, upstanding structures, database, Harris matrix, chronotypology, cluster analysis.

1 Introduction

Our research is an interdisciplinary research on tangible Cultural Heritage concerning the Medieval Archeological studies in the Petra Valley settlements, especially aimed at Shawbak Castle, which is focused on the documentation and presentation of archaeological data (CH content). The main goal of this interdisciplinary research is the archaeological analysis of the site and its interpretative design, especially based on all the data collected also with the comparison of the Petra Valley medieval settlements.

M. Ioannides et al. (Eds.): EuroMed 2012, LNCS 7616, pp. 119–128, 2012.

The interdisciplinary research is characterized by common efforts to the general digital data acquisition and particularly 2D and 3D data capture methodologies and data processing. Our first aim was the realization of the 3D GIS.

To achieve these goals we merge two approaches coming from two different scientific fields. From one side an archaeological analysis with a database elaboration (the archaeologist' team use since long years a database designed for this application named PetraData [5]. This database is populated by all the archaeological observations and use concepts which are also used by the survey process, such as CF (Corpo di Fabbrica, i.e. Building Unit, according to Brogiolo [2]) or USM (standing for Unità Stratigrafica Muraria, i.e. Wall Stratigraphic Unit).

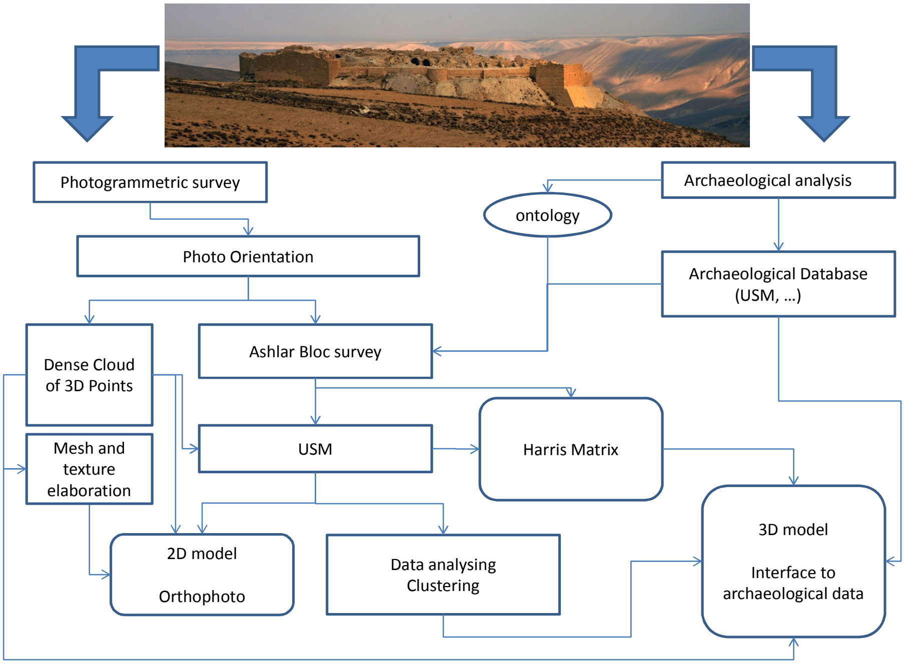


Fig. 1. Synoptic schema of the interdisciplinary project, with detail of archaeology and computer science cooperation and a view on the Shawbak castle

And the second approach is photogrammetry and computer science in order to produce 3D or 2D models linked with geometrical data analysis driven by archaeological knowledge.

The photogrammetric survey is organized in two distinct ways. After the orientation phase, the photographs are used to produce 3D models with dense cloud points generation in an automatic process and also to measure and represent ashlar bloc, at this stage the survey is done by an expert archaeologist [4; 9].

The ashlar blocs are the atomic element chosen for photogrammetric measure, the link with the archaeological database allows connecting it to the archaeological structure (USM) and giving us a good approach of the positive USM [9].

Every USM is a building homogeneous action which is representative of different building moment. In other words to identify and record the different USM in upstanding structures allows archaeologists to define a relative chronology of the building by the relative position of every USM.

Once the ashlar blocs are measured and linked to the archaeological database (at least by their own USM) we have two important links between them: temporal relations, given by the Harris relationships in the database and structural relations given by the dimensions and position of each bloc or USM.

We have developed a 3D viewer able to visualize all data used, photogrammetric models, with photographs and the 2D/3D points, ashlar blocs, dense cloud of 3D points, USM and clustering analysis and finally the USM graph used in the Harris Matrix.

The main idea of such 3D viewer is not to produce 'nice' image of the site but to drive the archaeological study offering a control panel to all the produced data.

2 The Archaeological Project: Aims and Methodology

The archaeological mission of the Florence University has been working for many years in a stratigraphic analysis project aimed at analyzing the forms of Crusader settlements in the territories of the Lordship of Transjordan in the twelfth century. The research started to document the characters of the first phase of Crusader occupation of the Holy Land where they are better preserved and stratigraphically clear, i.e. in Transjordan. The fortifications in this area, in fact, were completely and permanently abandoned by the Crusaders after the battle of Hattin (1187) - where the army the Latin Kingdom of Jerusalem was defeated by Salah al-Din - except for specific exceptions.

Gradually, the program of investigations, directed by the results *in itinere* of the research, has turned its attention to one of the most important fortified site of the entire region, Shawbak, as a 'stratigraphic observatory', after the identification and study of the surprising role of Petra [10]. Shawbak is to be considered an extraordinary archaeological-monumental area that holds the possibility to be fully explored. It is characterized by a very readable archaeological stratification and it is a consistent material archive spanning over at least 1600 years, from Roman- Byzantine age (structures belonging to these periods have been recognized thanks to recent archaeological investigation [10]) through Crusader-Ayyubid, Mameluk and Ottoman periods [14].

The research methodology employed since the beginning has been the Light Archeology that means an integrated system of landscape archaeology, environmental archaeology, upstanding structures stratigraphy, excavation samples. All these research fields are strictly connected in an informatics system that covers the phases of collection, recover and analysis of all the archaeological data. In this kind of research the upstanding structures stratigraphy is very important because it allows to know in relatively short time many living phases of the site (and also of the entire region). It is also a very important knowledge-key for the restoration program, which is the object of an agreement of Italian-Jordanian scientific and cultural cooperation between the Florence University and the Department of Antiquities of Jordan.

3 Photogrammetric Survey

Archaeological excavations are always irreversibly destructive, so it is important to accompany them with detailed documentation reflecting the accumulated knowledge of the excavation site. But not only the excavation needs an accurate documentation, also the upstanding structures research [2]. This documentation is usually iconographic and textual. Graphical representations of archaeological upstanding structures such as drawings, sketches, watercolors, photographs, topography, and photogrammetry are indispensable for such documentation and are an intrinsic part of an archaeological survey. However, as pointed out by Olivier Buchsenschutz in the introduction to the symposium Images and archaeological surveys, in Arles, France, in 2007 [8], even a very precise drawing only retains certain observations that support a demonstration, just as a speech retains only some arguments, but this choice is not usually explicit. This somewhat lays the foundation of this work: a survey is both a metrics document and an interpretation by archaeologist.

The photogrammetric approach is here to solve two kinds of problem: the first one is to build 3D or 2D representation of the site, or part of the site in order to have a representation of the site which can be used as an interface to the large amount of textual and iconographic data collected and computed by archaeologists. The second

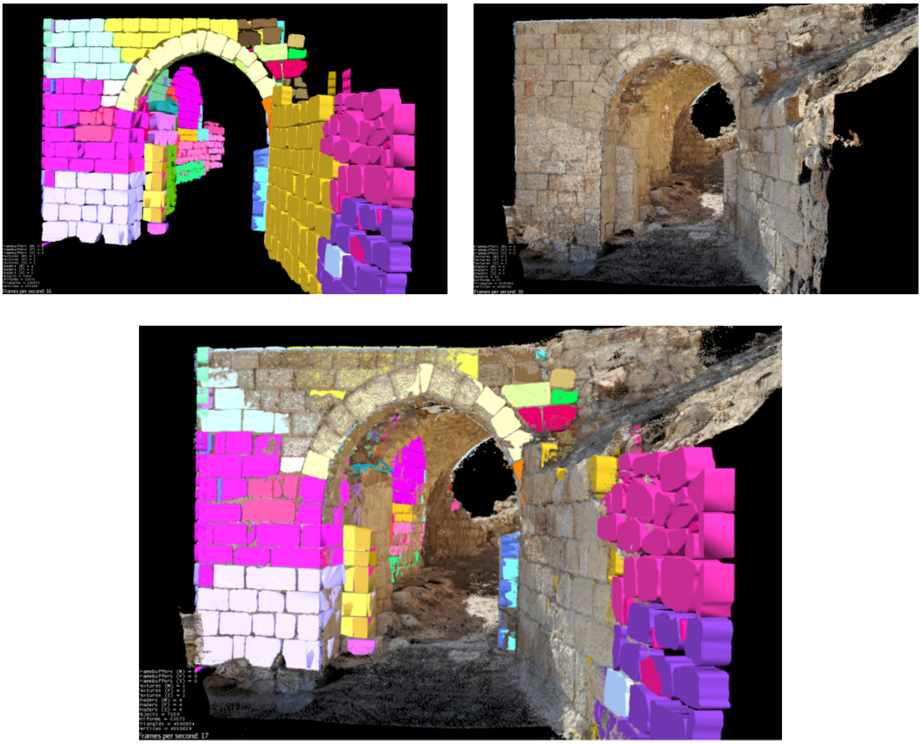


Fig. 2. Merging the two photogrammetric approaches in a single, interactive, 3D viewer

one is the necessity to collect metric data on identified artifact in order to be able to make dimensional analysis, clustering and others statistical computation.

The first use of photogrammetry is the ashlar bloc measurement. This is done with a specific method, already published[4; 9]. Users can make a 3D measurement using one single photograph, without altering precision of the result. This method allows the user to concentrate on the archaeological aspects of the survey with less attention to the photogrammetric one. In this approach the user, which is an expert of the domain, here the archaeologist of the middle age can take benefit of a manual process to add annotations to the measured points. This semantic tags can be used to complete the survey (for example to obtain a complete bloc by extrude using an arbitrary depth given by the archaeologist or simply adding archaeological data to the bloc, elaboration tools, stone quality, etc..)[9].

The second use of photogrammetry is the intensive production of cloud of 3D points with color information. As we are working since a long time on this site we have a huge set of oriented photographs and it was possible to use part of them in order to generate a dense cloud of 3D point (see Figure 2) with the method proposed by Furukawa and Ponce [12]. This approach, which looks like highly accurate and quite exhaustive, is, at the first level, really poor from the semantic point of view. Of course we can manage a huge quantity of 3D point but we lose all the knowledge carried out by the original photogrammetric survey. The main problem is how to combine from one side the high quantity of information, (3D points, color information for each point) and from the other side the indispensable archaeological knowledge.

We developed a hybrid approach, based on both photogrammetry and 3D visualization in order to use the data usually produced by archaeologist. Even in a 3D context archaeologist uses mainly 2D documents such as photos or orthophoto and USM are designed onto these documents. This approach allows us to produce orthophoto with specific request to the database and also Harris Matrix from the 3D model (for example partial Harris graph from an USM identified in the 3D model).

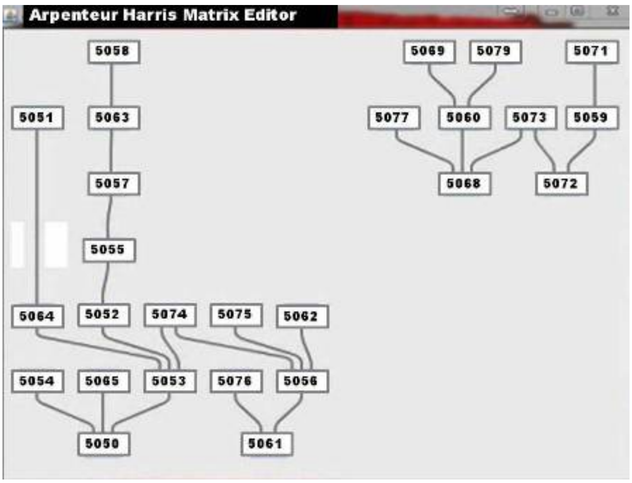
The 3D viewer, which support both dense cloud of 3D points and artifact visualization, allows also bidirectional interaction and can be used to visualize database request or statistical result as clustering [15]. We can see on Figure 2 the ashlar, measured by an archaeologist with annotations which able a link toward the archaeological database, colored according to their own USM. In the same viewer the dense cloud of 3D points is displayed. As we show in a precedent paper we can broadcast the semantic information toward the 3D points and change, for example their color according to the USM [15].

4 Harris Matrix and 3D Models

Following the most important concept of the stratigraphical archaeology, the stratigraphic unit, we try to document directly on the survey the main characters of each of them. From the publication of the first edition of E. Harris's book Principles of archaeological stratigraphy, many archaeologists follow the idea that all archaeological sites, to a greater or lesser degree, are stratified and that for this reason it is necessary to know the main principles of the archaeological stratigraphy to obtain all the possible information. In fact all forms of stratigraphic unit are the result of natural or human action of deposition or erosion (i.e. in the last case of construction or destruction).

Concerning the stratigraphy of upstanding structures, stratigraphic unit is the result of a human action of construction. Its position on the entire wall and the physic and stratigraphical relationships with the others stratigraphic units are necessary to detect the relative chronological sequence of the entire building (and, comparing all the structures, the relative chronological sequence of the site). But the concept of interface is also very relevant. According to Harris "There are thus two main types of interface: those which are the surfaces of strata and those which are only surfaces, formed by the removal of existing stratification." [1 p. 54]. These interfaces create their own surfaces and area, they have stratigraphic relationships with other units of stratification and their own boundary and surface contours. So it is very important first of all to detect this kind of features and then to survey the position and the physical and stratigraphical relationships with the others stratigraphic units.

All these stratigraphical relationships are formalized in the Harris matrix, or graph, the most convenient system to collect and to place in a chronological order all the events that occur in a site. Currently we can generate the Harrix graph by a simple mouse click into the 3D model and we can also see the corresponding 3D representation of an USM from the graph, the next step will be to make possible to visualize on the 3D model also the stratigraphical relationship and the Harris graph.



Although there are some softwares which elaborate automatically the Harris matrix from the data collected, it is very interesting to have an integrated system from the survey to the matrix.

Fig. 3. Harris Matrix produced choosing an USM from the 3D viewer (see Fig. 2). Each USM represented here is a link toward the 3D viewer.

5 Statistical Analysis and Classification

A very important element of the Shawbak Castle Project, from the point of view of archaeological readings, is the possibility offered by this context to study the wall constructive technology of the twelfth century thanks to the short life of the Crusader living phases in this area (less than a century). Once this phase has been well determined it will be possible to create a fossil-guide to the entire region. In order to achieve best results, to be inserted in the stratigraphic column, it can be very useful to improve the research with other studies such as mensiochronological one, assisted also by the cluster analysis. These statistical analyses recently have been applied to the study of the upstanding structures in a Spanish project which includes the analysis of Vitoria-Gasteiz city [11].

The mensiochronological researches have been almost successfully applied to stone buildings [6] and they could usefully be tested in the Shawbak Castle Project, to be able to consolidate the masonry chronotypology that we are preparing. A tool like our 3D survey linked to the archaeological data is very useful to this purpose because at the same time of the survey we have all the necessary steps for each subsequent statistical processing and also they are inserted in the stratigraphic diagram. Compared to measurements taken directly on the field it has a total coverage of the existing surveyed building (which is almost impossible to perform on the field for several reasons, including above all the lack of time, and also, often, the position not physically accessible for certain walls or pieces of them). But we need more time to improve and to test these kind of statistical analysis, which are very interesting because they allow us “to see” different kind of clusters in the pattern of the constructive elements potentially useful to analyze, for example, some different technological choice and so on.

In our previous work [15], we used a Principal Component Analysis (PCA) so as to study the whole measures' information of each bloc. We also highlighted the existing links between each bloc's measures and the corresponding USM. In the following works, we first applied the same procedure on the rest of the building (1191 blocs which are classified over 44 USM); then we used the results of the PCA in order to build clusters in a unsupervised way. We kept 5 variables features (*Heigth, Length, X, Y, Z*) and we implemented a hierarchical ascendant classification method. This algorithm principle is to gather points according to a proximity parameter. In these kinds of methods, there are two important steps: the choice of the dissimilarity measure between blocs and the choice of the cluster index corresponding to the between-class distance. For these two measures, many solutions can be found. For our

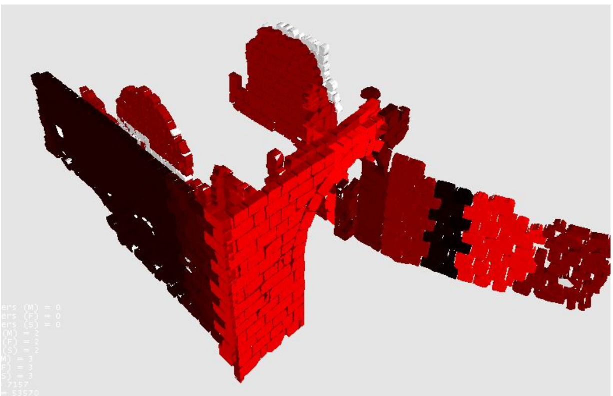


Fig. 4. Ashlar clustering visualization

application, we chose the euclidean distance for the between class distance and Ward's method so as to estimate the cluster index between classes.

The results are illustrated on Fig. 4. We can see similarities between these clusters and the classification made by archeologists with USM (Fig. 2). However, we also found that this classification is more influenced by the positional parameters (*X, Y, Z*). In order to alleviate this effect, we applied a non-linear classification method.

There are several non-linear classification methods. We choose to use Support Vector Machines (SVM). They are not only effective, but also less sensitive to the change in the learning set in the data modeling. It has to be noticed that we find ourselves in an unfavourable situation where the number of explanatory variable features is greater that the number of each class' blocs and where the multicolinearity risk is high. We then choose supervised learning procedure. Over the 44 classes (USM), we kept 10 classes (USM) whose cardinality is greater than 30 blocs (for example: USM 5032, 5035, 5040, ...). As the classes' cardinalities remain insufficient (with an average of 40 blocs), we used a binary classifier with a cross-validation. Cross-validation consists in dividing the data into two-thirds for the learning step and on third for the testing step, and repeating the procedure *N* times.

For our calculations, we applied the C-SVC algorithm which is implemented by [13] with the variable features *Heigth*, *Length* and *Z*, with the following parameters (linear kernel, cost parameter equal to 1, weight parameter of each class equal to 1).The classification rates range from 74.64% to 97% for example for the USM 5036 and 5032 respectfully.

We wanted to test this method with just the blocs' intrinsic parameters (*Heigth*, *Length*). The results are illustrated in table 1. Each cell stands for the classification rate between the classes in the corresponding line a column. These first results seem interesting because, for all the classes, we manage to get a classification rate which next to 80% (even 90% between the USM 5326 and 20214). As a consequence, we can consider a combination of several classifiers so as to help archeologists classifying the blocs which are found near a site, thanks to these two measures.

Table 1. Classification results

USM	***5032***	***5035***	***5038***	***5040***	***5102***	***5105***	***5326***	***5328***	***20214***	***5036***
5032	100									
5035	73,57	100								
5038	64,88	55,72	100							
5040	75,38	67,61	62,41	100						
5102	51,42	**80,27**	64,66	74,28	100					
5105	76,48	68,87	63,78	57,4	75,38	100				
5326	**83,18**	59,54	**79,74**	76,21	**84,72**	66,5	100			
5328	65,02	51,11	56,81	68,57	72,32	**89,76**	60,37	100		
20214	63,61	**82,36**	71,28	63,66	66,22	71,25	**90,51**	**79,27**	100	
5036	54,28	75,13	71,33	72,08	52,85	73,23	**83,09**	73,75	59,55	100

6 Conclusion and Future Work

This project is due to a long cooperation between University of Florence, Italy and CNRS in Marseille (LSIS laboratory). We are working on a documentation system dedicated to medieval archaeology which integrates three main components:

- A survey tool based on photogrammetry and computer vision, using traditional photogrammetry where an operator is able to design specific point with annotation belonging to a specific knowledge and also, with recent advance in photogrammetry and computer vision, which can produce huge quantity of 3D points with color information but, this time, without any semantic information. In this case we develop a bridge between the two approaches in order to add semantic data on the cloud of 3D points.
- A 3D and 2D visualization tool, which allows interactivity between the stored data (the data persistence is done with both XML formalism and relational database) and the user. This tool can manage geometry computed from photogrammetric data as well as dense 3D points cloud. The stratigraphic relation stored in the archaeological database can be represented, on the fly, by and interactive graph representing the Harris Matrix.
- A statistical analysis module which allows producing new knowledge on the measured artifact (clustering on ashlar bloc for example).

All these developments are using and underlying knowledge model formalized with ontology which allows to build links between the concepts managed by archaeologist as well as those used in survey techniques.

We are now working on several aspects of this project:

- Improving the module which transfers the semantic data from the manual photogrammetric process to the dense 3D points cloud. This is done using the fact that we use the same original support to produce archaeological and 3D data: the photographs. Archaeologists can design USM and ashlar bloc onto oriented photographs.
- Improving and integrating the statistical computation module to increase correlation between ashlar blocs and other archaeological data.
- Finally, from the archeological point of view the units of stratification represent an archaeological aspect of time's cycle and this is the cause of their relevance. In fact the other important goal of the Harris's work was the invention of the Harris matrix, the most convenient system to collect and to place in a chronological order all the events that occur in a site. Nevertheless the Harris matrix don't permit to indicate the concept of the life span of the stratigraphic unit and this is a real lack concerning the upstanding structure stratigraphy [16]. So our most relevant future work will be the effort to indicate more relationships between the stratigraphic units, following also the Allen's theory [3; 7], to better put in evidence all the life-phases of an archaeological structure.

References

1. Harris, E.C.: Principles of Archaeological Stratigraphy. Academic Press, London (1979)
2. Brogiolo, G.P.: Archeologia dell'edilizia storica. Documenti e metodi. New Press, Como (1988)
3. Allen, J.F.: Temporal reasoning and planning. In: Brachman, R.J. (ed.) Reasoning About Plans, pp. 1–67. Morgan Kaufmann Publishers Inc., San Francisco (1991)
4. Drap, P., Gaillard, G., Grussenmeyer, P., Hartmann-Virnich, A.: A stone-by-stone photogrammetric survey using architectural knowledge formalised on the ARPENTEUR Photogrammetric workstation. In: Proceedings of XIXth ISPRS Congress, vol. XXXIII, pt. 5, 8, pp. 139–145. ISPRS, Amsterdam (2000)
5. Niccolucci, F., Tonghini, C., Vannini, G., Crescioli, M.: PETRA: un sistema integrato per la gestione dei dati archeologici. Archeologia e Calcolatori XI, 49–67 (2000)
6. Nucciotti, M.: Le murature medievali di Santa Fiora (Monte Amiata - Toscana): mensiocronologia delle murature in pietra, un caso di studio. Archeologia dell'Architettura V, 65–85 (2000)
7. Accary, T., Calabretto, S.: Temporalité des corpus archéologiques. Document Numérique, Numéro spécial sur le thème Temps et Document, 111–124 (2004)
8. Buchsenschutz, O.: Images et relevés archéologiques, de la preuve à la démonstration. In: 132e Congrès National des Sociétés Historiques et Scientifiques, Arles. Les éditions du Cths (édition électronique), Pub. (2007)
9. Drap, P., Seinturier, J., Chambelland, J.C., Pruno, E.: Going To Shawbak (Jordan) And Getting The Data Back: toward a 3D GIS dedicated to medieval archaeology. In: Proceedings of 3D Arch 2009, 3D Virtual Reconstruction and Visualization of Complex Architectures Trento, Commission V, WG V/4 (2009)
10. Vannini, G., Nucciotti, M.: Un problema di visibilità archeologica: territorio, analisi "leggere" e sintesi storiche. In: Vannini, G., Nucciotti, M. (eds.) Da Petra a Shawbak, Archeologia di una Frontiera, pp. 28–31. Giunti, Firenze (2009)
11. Azkarate Garai-Olaun, A.: Archeologia dell'Architettura in Spagna. Archeologia dell'Architettura XV, 15–26 (2010)
12. Furukawa, Y., Ponce, J.: Accurate, Dense, and Robust Multiview Stereopsis. IEEE Transactions on Pattern Analysis and Machine Intelligence 32, 1362–1376 (2010)
13. Chih-Chung, C., Chih-Jen, L.: LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology (TIST) 2, 1–27 (2011)
14. Pruno, E., Sciortino, R.: La ceramica post-medievale a Shawbak (Giordania): impostazioni di metodo e prime acquisizioni della ricerca. In: XLIV Convegno Internazionale della Ceramica, pp. 39–55. Edizione Del Centro Ligure Per La Storia Della Ceramica (Ed.), Savona (2012)
15. Drap, P., Merad, D., Boï, J.M., Seinturier, J., Peloso, D., Reidinger, C., Vannini, G., Nuccioti, M., Pruno, E.: Phogrammetry for medieval archaeology, A way to represent and analyse stratigraphy. In: VSMM 2012 18th International Conference on Virtual Systems and Multimedia, Milano (2012)
16. Pruno, E., Drap, P.: Dalla stratigrafia all'archeologia teorica: il matrix di Harris e l'archeologia del tempo. In: Redi, F., Forgione, A. (eds.) VI Congresso Nazionale di Archeologia Medievale, pp. 43–45. All'Insegna del Giglio, Firenze (2012)