Journal of Cultural Heritage 12 (2011) 134–139

Available online at ScienceDirect www.sciencedirect.com

Elsevier Masson France EM|consulte www.em-consulte.com

Original article

# Data matrix (DM) codes: A technological process for the management of the archaeological record

Jorge Martínez-Moreno [a], Paloma González Marcén [a], Rafael Mora Torcal [a,b]

[a] Centre d'Estudis del Patrimoni Arqueològic de la Prehistoria, Facultat de Lletres, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain
[b] Program ICREA-Academia

ARTICLE INFO

*Article history:*
Received 17 June 2010
Accepted 8 October 2010
Available online 21 December 2010

*Keywords:*
DM codes
Artefact labelling
Digital managing
Fieldwork application

ABSTRACT

This article presents a new method for labeling archaeological material, based on the use of data matrix (DM) codes. The information that identifies an artefact (site name, level or archaeological unit and consecutive number) is coded on very small labels ($3 \times 3$ mm and $4 \times 4$ mm). This information is captured by a laser reader, which inputs it directly into a computer database. The system has been successfully applied to the pilot study presented here. Its use greatly improves provenance information and management of the archaeological record, and results in the more accurate processing of artefacts in fieldwork routines, laboratory activities and museum storage.



## 1. Introduction

Recording material from sites is an essential task in archaeological work [1,2]. The need to identify each artefact individually (be it lithic, bone, pottery or any other material) is standard practice for prehistoric sites, and is gradually being extended to sites of more recent periods. One advantage of having identifiers for each archaeological item is that by using techniques such as artefact refitting in the horizontal and/or vertical scale [3], the resolution and internal coherence of archaeological contexts can be analyzed, and inferences about formation processes or activity areas can be made.

The identification of each object and its manual marking involves a considerable investment of time and work in the inventory of an assemblage, especially when it consists of thousands of pieces. Errors habitually occur in the course of this manual process which, although difficult to quantify, must affect any subsequent research [4].

In order to limit problems of this kind, we propose a new system of recording using digital systems, specifically data matrix (DM) codes. Their application is illustrated by a pilot study undertaken at the middle Palaeolithic site of la Roca dels Bous (Lleida, Spain).

## 2. Recording archaeological material

Contextualization of the archaeological record through the recording of absolute coordinates of lithic artefacts, pottery, bones, and in general, any cultural or natural indicator recovered during an excavation, is standard practice in fieldwork. The increasingly common use of high-precision topographic instruments allows the three-dimensional positioning of artefacts and other types of remains. The advantages of this procedure have been described in various publications [5–8].

Along with three-dimensional positioning of objects [9], visual techniques help define the characteristics and geometry of the basic unit of analysis, in our case, the archaeological level or unit [8,10,11]. Such fieldwork methodology provides an object-based approach to the identification of significant assemblages, which improves the widely used concept of *archaeological context* [7,8].

Therefore, although the basic unit of analysis is the archaeological level, the contextual position and attributes associated with artefacts should not be underestimated as potential sources of information [12]. Labelling, that is, the application of codes, which identify each individual artefact, is a fundamental step in this research strategy, since the accurate definition of an archaeological level depends to a great extent on the proper identification of the elements that it comprises [3,10,13].

Codification of the site name provenance unit and inventory number supplies a complex network of data from which contextual information about the artefact can be recovered. Its temporal, cultural or taphonomic attributes and implications can be evaluated and eventually possibly even be refuted by subsequent studies [14].

*E-mail addresses:* jorge.martinez@uab.es (J. Martínez-Moreno), paloma.gonzalez@uab.es (P. González Marcén), rafael.mora@uab.es (R. Mora Torcal).

1296-2074/$ – see front matter © 2010 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.culher.2010.10.001

The lack or loss of this codification deprives us of the basic information that could lead to the formation of new hypotheses or to evaluate the validity of old inferences. If the archaeological item lacks an individual label, the artefact irretrievably loses much of its explanatory potential.

## 3. Data matrix codes

In recent years, the application of digital identification systems such as bar codes has been tested in fieldwork [15]. In this article, we present a similar method but based on the use of DM codes to label archaeological material (Fig. 1). DM codes were developed in the 1980s and are currently used for quality management and control of industrial products. They include QR-codes, Aztec codes, Semacode, idcDMatrix, icQRC and icPDF [16]. DM codes were developed by the RVSI Acuity Cimatrix Inc. corporation (Nashua, US) and are protected by the public domain standard ISO//EC16022 [17], which guarantees their free access and use.

These codes have a two-dimensional structure consisting of square cells that store numeric and/or alphanumeric data. The quantity of information they can contain depends on their size, but is sufficient to identify an artefact. In our study, this information corresponds to the site name, archaeological context or level from which the artefact has been recovered, and inventory number. These identifiers enable reconstruction of the spatial positioning of the recovered artefact once its context has been excavated.

DM codes offer two basic advantages over bar-codes [15]. Barcode labels are large, often cannot be stuck directly on to the object and have to be stored with the object in a bag. Because DM codes are so small, they can be attached directly both on to a bag *and* to the surface of the object, thus reducing the possibility of loss or errors while handling objects during analysis, or mixing identification codes on containers.

Depending on its size, the DM code stores a series of numeric and/or alphanumeric characters (Table 1). DM codes also support advanced coding systems that allow all information to be recovered from the code even if part of it is damaged.

## 4. Artefact labelling

An essential concern when labelling archaeological objects is reliability, an attribute that depends on the durability or survival over time of the identifier. The use-life of labels depends on many variables, such as ink quality, application method (nib, stylographic pen with Indian or indelible ink), protection given to the handwritten label, and the conditions under which artefacts are kept in storage or museum collections. The unreliability of traditional manual marking techniques, although not directly evaluated in relation to archaeological objects, has led to their gradual substitution in industrial and commercial international organizations, first by bar codes and then also by DM codes [18]. Indeed, as a consequence of the lack of a reliable marking procedure, a significant part of museum work consists of the relabeling of assemblages in order to prevent information loss, especially in the case of relevant collections.

**Table 1**
DM code sizes and their maximum numeric and alphanumeric capacity for data storage.

| | Data capacity | |
|---|---|---|
| Code size (mm) | Numerics | Alphanumerics |
| 3 × 3 | 10 | 6 |
| 3.5 × 3.5 | 16 | 10 |
| 5 × 5 | 34 | 24 |

DM codes do not last indefinitely either, but the material used for the label, the type of printing, the ink and the products used for fixing the code, suggest that these digital identifiers are an improvement over traditional methods of labelling [19].

The labels on which the codes are printed are made of polypropylene, a plastic produced by the company Manter [20]. Because of polypropylene's technical characteristics, labels are resistant, well suited to digital printing and remain stable in extremes ranges of temperature and humidity. They are printed with a desktop digital thermal wax printer with a resolution of 300 dpi, which provides adequate resolution and an excellent ratio between quality and product price. The labels can be designed and printed according to the size and number of DM codes that they include.We have used Paraloid B72 to attach the code onto the artefact. This acrylic resin is widely used with excellent results in the restoration and conservation of archaeological materials [21,22]. Paraloid provides fixation, transparency, resistance to discolouration and chemical stability for the labels [23]. Another useful property is its rapid adhesion; because this acrylic resin is dissolved


**Fig. 1.** Label with DM code: top row 2.5 × 2.5 mm DM code capable of storing the catalogue number code; middle row 3 × 3 mm DM code that contains the level and catalogue number codes; bottom row 5 × 5 mm DM code with the series of codes – site, level and catalogue number.


**Fig. 2.** Examples of the application of DM code to different archaeological materials. The size of the DM code is selected on the basis of artefact size. For special samples, it can be applied directly on to the container, in this case, a tube with seed samples. Graphic scale 0.5 cm.

in volatile organic compounds (such as acetone or alcohol) before its application, it impregnates and dries immediately [24]. Similarly, it can be removed easily by applying solvents that do not affect the archaeological material.These attributes are essential as the DM code label is fixed between two coats of acrylic resin that are applied to the piece, the first directly on to the material to produce an adhesive surface, and the second covering the label to ensure that the identifier does not become detached from the artefact. This procedure protects the code against the effects of time. Although we have no reliable information about the label's life and how long it takes to deteriorate, such a procedure increases its conservation (Fig. 2). A similar procedure for DM code application is suggested by NASA for small pieces marked with ink jet, using, in this case, acrylic lacquer to coat the code [18].

## 5. Reading and recovering the information

Manual labelling is a repetitive process and as such is subject to a percentage of errors that is difficult to quantify, but which varies depending on factors such as the experience and motivation of the person who is doing the labelling, or the existence of controls that evaluate the legibility of the labels. Marking by hand is not uniform and is conditioned by imponderables such as the porosity, irregularity, texture and, a very important variable in our case, size of the artefact. These factors, although not easy to evaluate and quantify, affect the quality of the record made on the artefact. However, it is not unusual for a manual mark to become illegible after a few years, resulting in the irreversible loss of the artefact's contextual information.

The use of DM codes solves some of these difficulties since the printed labels are not subject to the problems described for marking by hand, and fixing them with adhesive ensures their survival.

One technical question relates to the way information contained in the code is captured. The most common method is to use a laser scanner, which transfers the data directly from the code to the database (Fig. 2). This is not the only method as codes can be captured with other digital devices such as video and photographic cameras, or even a mobile phone camera, many of which include a reading utility that can be downloaded from the Internet [25].

Another advantage of these codes is that they can be read irrespective of their position on the artefact, or that of the scanner in relation to it. This is not the case with barcodes, which have to be precisely positioned in front of the scanner to be read correctly [15].

## 6. Required accessories and devices

The devices needed to use this system are a code scanner, a printer (Fig. 3), consumables, software to list out the codes and adhesive. There is currently a wide range of two-dimensional laser scanners and digital printers in the market that can be used to print

**Fig. 3.** Accessories used to apply the DM code: (a) desktop digital thermal wax printer (real size 240 × 200 × 200 mm); (b) laser reader for reading two-dimensional codes.

polypropylene labels. The labels and the acrylic resin are easy to obtain from companies that specialise in the distribution of products used in restoration, conservation and storage of archaeological material. Some have websites with catalogues giving technical information about the products, prices and ordering that can be consulted online [24].

Along with the wide availability of the necessary accessories for DM marking, their relative low cost is a further advantage of this system. Although the exact amount of necessary economic investment depends on the different brands and suppliers, in our pilot study the resultant cost of each printed label and the proportional use of a code scanner has been calculated to be around 0.01 € ().

However, specifically-designed software is required to generate the labels and to introduce the selected attributes to be coded. In our case, this was achieved in collaboration with the Spanish company *Internet Web Serveis* (IWS). The selected information fields are: site identifier, archaeological level and catalogue sequence number. These attributes are coded in alphanumeric sequence and are ordered hierarchically according to their greater to lesser significance, the most important being the sequence number, which has a maximum length of six digits (Fig. 1). This number can never be repeated within the same archaeological level. A second field identifies the level and the final field, the site code. Each of these latter fields has two digits. The combination of these fields produces a 10-digit sequence, sufficient to catalogue any collection regardless of size as, by combining figures and letters within the sequence, these alphanumeric chains exponentially multiply the number of possible combinations (Table 2).

Depending on the size of the DM code selected, a sequence of data of variable length can be input. Thus, a 3 × 3 mm code can store an alphanumeric chain of 10 digits. If smaller codes are used, the fields that need to be incorporated into the label must be selected (Table 1).

In order to minimize the limitations of the non human readable character of DM codes, different labels are printed: smaller ones only with the DM code to be fixed directly on to the archaeological

**Table 2**
Example of a DM code codification table using a 3 × 3 mm label with six alphanumeric digits (which contains the catalogue number), while the 5 × 5 mm label capable of storing 24 alphanumeric digits contains the catalogue number, level and site.

| Identify | Digits in the code | Length of the code |
|---|---|---|
| Catalogue number | 012345 | 6 |
| Level or unit | 01 | 2 |
| Site | 01 | 2 |

items following the procedure described above, and larger ones with both the DM code and the alphanumeric chain which are attached to the specific bag in which the item is placed during the excavation and later for definitive storage. In this way, it is possible to combine the need for direct manipulation of archaeological material for research purposes without the risk of losing contextual information with the implementation of coded and uncoded marking systems used for the storage and archiving of archaeological material.

## 7. The application of data matrix codes: the example of Roca dels Bous

The procedure followed in recent fieldwork at the Middle Palaeolithic site of la Roca dels Bous is an example of how these technological resources can be integrated into excavation and recording routines.

Excavations at Roca dels Bous have been ongoing since 2001 and are part of a project that aims to analyze human settlement patterns in the southern slopes of the Pyrenees over the last 50,000 years [10,13]. From the beginning of fieldwork, a systematic methodology, using a total station, has been applied in the recovery and recording of archaeological material, including *microdebitage* and remains of less than 1.5 cm in length [26].

In the 2010 campaign, we had the necessary infrastructure in the field laboratory to develop the DM system, including laser printer and readers, labelling software and consumables (labels).

As a DM code label is assigned to be attached to each coordinated artefact, and as there is no electricity at the site to run printers, the first step in a daily routine is to print out enough DM code labels to cover the expected needs of a day's fieldwork. Devices such as the total station or PDAs are powered by batteries that are recharged daily in the laboratory.

During the excavation, the procedure is relatively simple. When a series of artefacts is collected, an individual label is assigned to each of them. During the recovery of the artefact, the laser theodolite registers its coordinates and gives the item a unique and unrepeatable identifier that contains a sequence of data consisting of the level and catalogue sequence number. This alphanumeric sequence is the same as the one printed on the DM code label. Once the artefact's position has been recorded, it is bagged individually with its corresponding label. Up to this point, the procedure is quite similar to the traditional method, differing only in the use of polypropylene labels bearing printed DM codes instead of manually produced labels (or those previously printed in the laboratory) (Fig. 4a).

But a complementary procedure has been developed for use with the DM codification beyond numerical identification and to ensure the code assignation to each recovered object. For this purpose, we use a personal digital assistant (PDA) as a notebook to capture quantitative data and attributes that define an archaeological object, be it a lithic artefact, bone, charcoal or seed. Since PDAs are programmed to contain the same sequence of alphanumeric code that identifies each object, it is unlikely that an error will occur in the assignation of IDs when describing the object [26]. This procedure ensures that the match between

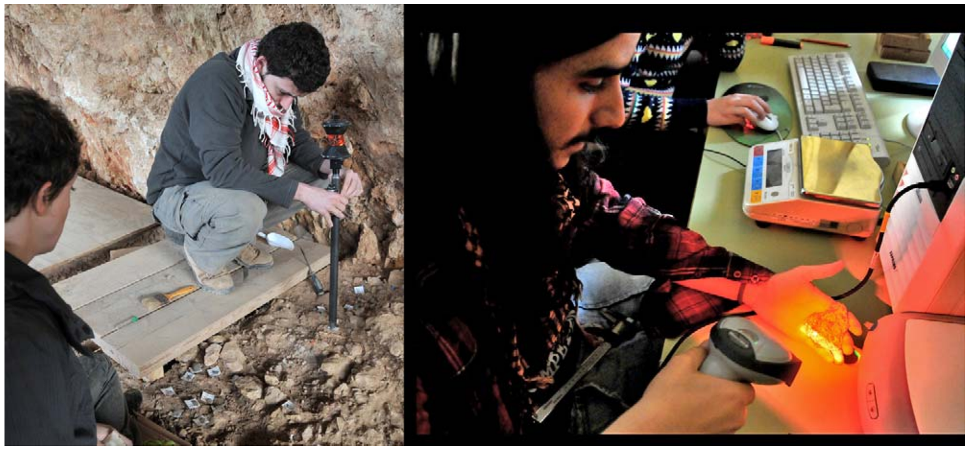

**Fig. 4.** (a) recording of archaeological material using a total station, in which individual labels have been assigned to each artefact; (b) reading an artefact to which a DM code has been applied with a laser scanner connected to a PC.

object and code has been visually identified and verified, reducing the possibility of labels being mixed up. Furthermore, the result of the combined use of DM codes and PDAs during excavation is to unambiguously ascribe relevant contextual field and object properties to each item for later laboratory analysis and study.

The PDAs store different types of information depending on the particular software developed for the project. For example, in the case of stone tools qualitative parameters recorded include raw material (chert, quartzite, quartz), category (core, flake, retouched pieces), while quantitative data include orientation and slope of the axis of the pieces in the archaeological and sedimentary context [27]. This information can be processed or reduced depending on the particular characteristics of each excavation and the software used. In the case of a Palaeolithic excavation, there is no point in creating fields to capture attributes that define ceramic, metal or vitreous materials, but if necessary, PDA fields can be modified accordingly.

Subsequently, in the laboratory, the data from the laser theodolite are downloaded onto a computer, and introduced into the database on a daily basis. The information in the PDA is downloaded directly to the database in the same way, thus the coordinates captured in the excavation combine the spatial position with the contextual data.

After the artefacts have been cleaned using non-aggressive methods, a base coat of paraloid B72 is applied on which to affix the DM code. Each label contains codes of different sizes and those on each row contain the same information. Code size is chosen to suit the characteristics of the artefact, and although the size of the object is important, criteria used depends on the particular needs of the project (Fig. 2). Since there are several codes on each row, the one that is best centred on the label sheet is selected (Fig. 1). The code is removed from the sheet, attached to the first coat of paraloid, and then sealed by a second coat of paraloid to protect it. Because the labels are small, the most suitable part of the artefact's surface (e.g. a flat or smooth area) is selected on which to correctly affix the label.

As paraloid B72 dries very quickly, the piece can immediately be placed near a computer so that the DM can be identified by the laser reader, recognised by and entered in the database (Fig. 4b). At this point, other data relating to the item but not recorded during excavation can also be entered on the database: e.g. weight, dimensions, information relating to state of preservation, taphonomic alterations or other relevant data. This information enables rapid completion of the daily inventory of materials recovered during fieldwork, so that their subsequent detailed study can proceed.

## 8. Discussion

The entire procedure, in which the DM code marking technique plays a central role, has been tested both in the laboratory and in the field, and to date nearly 10,000 lithic artefacts and bone remains have been provenanced without any major problem. Errors involve less than a dozen objects, which is less than 0,2% information loss. The present pilot study demonstrates that use of this system is viable in terms of parameters such as its speed, reliability and its compatibility with fieldwork routines. Furthermore, the current evaluation has found that these digital systems are not subject to the habitual errors that commonly occur in manual labelling, and which lead to the irreversible loss of contextual information.

The use of paraloid B72 as an adhesive will certainly preserve the labelling for longer than regular adhesives, although its use needs to be monitored with particular studies to determine label life. However, we would suggest that uncertainty about the duration of labels is also true of traditional systems.

Another relevant question is whether changing from traditional systems of labelling to digital methods such as DM codes is cost effective. This technological process represents a significant improvement by overcoming the inherent deficiencies of manual labelling. Moreover, although implementing such a digital system involves an economic investment in the purchase of laser scanners and printers, today, the cost of these devices is not excessive.

The results of this pilot study suggest that the ease and reliability of digitally labelling artefacts offers numerous advantages that more than compensate for the financial cost of the process. In our case, it is especially relevant that the size of codes can be used with very small artefacts, so that analytical procedures which required the handling and mixing of objects can be carried on without risk of losing contextual information. This is not the case for bar-codes that need larger labels.

Moreover, DM codes can be used as part of a wider procedure related to artefact recovery, recording, analysis and storage. With the help of the appropriate instruments (laser theodolite, PDAs) and specifically designed software, it is possible to link, on a more reliable basis than traditional procedures, objects with their contextual information, both for their study and permanent storage.

From this viewpoint, DM codes are a major advance, and should be considered an alternative, not only for archaeological fieldwork, but also for the provenance and management of collections held in museums and archaeological storage facilities.

## Conflicts of interest statement

The authors declare no conflict of interest.

## Acknowledgements

We would like to thank the two anonymous reviewers and Eleonora Stella for their valuable comments, Andreu Ibañez from IWS and Andreu Rivadulla for their support. The fieldwork project in Roca dels Bous has been supported by the Servei d'Arqueologia i Paleontologia-Generalitat de Catalunya. Roca dels Bous fieldwork is part of the project *Human settlement during the Upper Pleistocene and Holocene in the South-eastern Pyrenees* funded by the Spanish Ministry of Education and Science (HUM2007-60317/HIST, HAR2010-15002). This is a contribution to research group 2009SGR-0729.

## References

[1] C. Renfrew, P. Bahn, Archaeology, Theories, methods and practice, Thames and Hudson Ltd, 1991, 544.
[2] H.L. Dibble, S.P. McPherron, B. Roth, Virtual dig: a simulated archaeological excavation of a middle paleolithic site in France, Mayfield Publishing Company, California, 1999.
[3] E. Cziesla, et al., The big puzzle, International symposium on refitting stone artefacts. Studies in Modern Archaeology, Vol. 1, Bonn, Holos, 1990, 684.
[4] H.L. Dibble, T.P. Raczek, S.P. McPherron, Excavator Bias at the Site of Pech de l'Azé IV, France, Journal of Field Archaeology 30 (2005) 317–328.
[5] H.L. Dibble, S.P. McPherron, On the computerization of archaeological projects, Journal of Field Archaeology 15 (1998) 431–440.
[6] S.P. McPherron, Artifact orientations and site formation processes from total station proveniences, Journal of Archaeological Science 32 (2005) 1003–1014.
[7] J. Parcerisas, R. Mora, La estación total inteligente y sus aplicaciones en el trabajo arqueológico, in: L.A. Valdés, I. Pujuana (Eds.), Aplicaciones informáticas en Arqueología: teorías, y sitemas, Bilbao, 1995, pp. 409–417.
[8] R. Mora, J. Parcerisas, J. Martínez, Computer-based recording systems of Pleistocene deposits with large mammals, in: G. Cavarretta, P. Giola, M. Mussi, M.R. Palombo (Eds.), The World of Elephants, Roma, 2001, pp. 219–223.
[9] S.J.P. McPherron, H.L. Dibble, P. Goldberg, Z Geoarchaeology 20 (3) (2005) 243–262.
[10] J. Martínez Moreno, R. Mora, I.d.l. Torre, Methodological approach for understanding middle palaeolithic settlement dynamics at La Roca dels Bous (Noguera, Catalunya, Northeast Spain), in: N.J. Conard (Ed.), Settlement dynamics of the Middle Palaeolithic and Middle stone age, Kerns Verlag, Tubingen, 2004, pp. 393–413.
[11] H.L. Dibble, et al., Context, curation, and bias: an evaluation of the Middle Paleolithic collections of Combe-Grenal (France), Journal of Archaeological Science 36 (2009) 2540–2550.
[12] D.L. Clarke, Analytical archaeology, 2nd ed., Methuen, London, 1968, p. 526.
[13] R. Mora, J. Martínez-Moreno, J. Casanova, Abordando la noción de "variabilidad musteriense" en Roca dels Bous (Prepirineo suroriental, Lleida), Trabajos de Prehistoria 65 (2) (2008) 13–28.
[14] J.G. Bordes, La séquence aurignacienne de Caminade revisitée : l'apport des raccords d'intéret stratigraphique, Paleo 12 (2000) 387–407.
[15] H.L. Dibble, C.W. Marean, S.P. McPherron, The use of barcodes in excavation projects, The SAA Archaeological Record 7 (1) (2007) 33–38.
[16] Wikipedia. Data matrix. Available from: http://en.wikipedia.org/wiki/Data_matrix_%28computer%29; 2010.
[17] R. Stevenson. Laser marking matrix codes on pcbs printed circuit design and manufacture. Available from: http://pcdandm.com/pcdmag/mag/0512/0512stevenson.pdf; 2005.
[18] NASA. Application of data matrix identification symbols to aerospace parts using direct part marking methods/techniques. NASA Technical Handbook. NASA-HDBK-6003A. Available from: http://standards.nasa.gov/released/6000/NASA-HDBK-6003A.pdf; 2002.
[19] GS1-USA. Available from: http://www.gs1us.org/; 2010.
[20] Manter. Available from: http://www.manter.es/caste/productos/ficha_producto.php?idioma=es&id=2393&palabraClave=OPP%20IDP%20WHITE%20GLOSS%20PF-1%20CB62&tabla_productos=productos_publicados; 2009.
[21] S.P. Koob, The use of Paraloid B-72 as an adhesive: its application for archaeological ceramics and other materials, Studies in Conservation 31 (1986) 7–14.
[22] C.V. Horie, Materials for conservation, in: Organic consolidants adhesives and coatings, Elsevier, London, 2005.
[23] J. Podany, et al., Paraloid B-72 as a structural adhesive and as a barrier within structural adhesive bonds: evaluations of strength and reversibility, Journal of the American Institute for Conservation 40 (2001) 15–33.
[24] C.B. Design. Technical data sheet Paraloid® B72 fixative (SY7F). Available from: http://www.conservation-by-design.co.uk/; 2010.
[25] SMS, MiBR. Available from: http://www.2dtg.com/products/icMobile_BC_Reader.html; 2010.
[26] R. Mora, J. Martínez-Moreno, I. de la Torre, ArqueoUAB: a systematic archaeographic for the analysis of Palaeolithic sites, in: P. Melero, P. Cano, J. Revelles (Eds.), Fusion of cultures, Granada, Spain, 2010, pp. 125–128.
[27] A. Benito Calvo, et al., Sedimentological and archaeological fabrics in Palaeolithic levels of the South-eastern Pyrenees: Cova Gran and Roca dels Bous sites (Lleida, Spain), Journal of Archaeological Science 36 (2009) 2566–2577.