INTERDISCIPLINARY SCIENCE REVIEWS, Vol. 34 No. 2–3, 2009, 189–205

Digital Support for Archaeology

Paul Boon, Laurens van der Maaten, Hans Paijmans and Eric Postma

Tilburg Centre for Creative Computing, Tilburg University, The Netherlands

Guus Lange

Dutch State Service for Archaeology, Built Monuments, and Cultural Landscape, Amersfoort, The Netherlands

We describe an interdisciplinary approach in which computer scientists develop techniques to support archaeology. In the Reading Images for the Cultural Heritage (RICH) project, a variety of methods have been developed to support archaeologists in the visualization, categorization, and characterization of archaeological objects, such as medieval glass, coins, ceramics, and seeds. The methods are based on image processing and machine learning algorithms that are tailored to the task at hand. We describe the algorithms and illustrate their application on archaeological datasets. The virtues and pitfalls of the interdisciplinary approach to archaeology are discussed.

KEYWORDS Archaeology, Digitization, Information retrieval, Visualization, Artificial intelligence

Introduction

The ultimate task of specialists in archaeological studies is to assign meaning to new discoveries. The interpretation of new discoveries is possible because objects and features are thought to resemble each other more when they are closely related to each other in a space–time window. The newly discovered objects and features and their contexts are thereto compared with those found earlier and described in the literature. Paramount is the accessibility to this collective memory or knowledge base, which is composed of the relevant literature and databases.

The accessibility of data, information, and knowledge is a major issue in archaeology for two main reasons. First, the accessibility to information is limited because of the dispersion of archaeological expertise over a multitude of private archaeological firms and institutes. Clearly, traditional manners of

DOI 10.1179/174327909X441108

communication are stretched to their limits. Second, the implicit nature of much of the expert knowledge (so-called tacit knowledge) makes it hard to translate into formalized descriptions such as if-then rules, which are amenable to processing by a computer. Each excavation site is unique in the sense that the geographical context differs from site to site as are the diversity and constellations of objects encountered. As a result of the context-dependency of archaeological finds and the fact that observations in the field cannot be repeated, the scientific discipline of archaeology is highly dependent on individualistic observations and judgements. The main challenge is to incorporate the individual, implicit knowledge of archaeologists into objective and transparent descriptions, taxonomies, or theories, without losing specific contextual information. Intelligent digital techniques may support archaeologists in meeting the challenge.

The successful implementation of intelligent digital techniques in the archaeological domain requires close interaction between computer scientists and archaeological experts. This interaction was facilitated by The Netherlands Organisation for Scientific Research (NWO) within the Continuous Access to the Cultural Heritage (CATCH) programme. The Reading Images for the Cultural Heritage (RICH) project is part of CATCH, and aims at meeting the challenge by developing intelligent tools to support archaeologists. Within the RICH project, the emphasis is on the use of image processing and machine learning techniques for reasons that will be explained in the following section. In addition, machine learning tools for text mining and information retrieval are developed in close conjunction with another CATCH project called Mining for Information in Text from the Cultural Heritage (MITCH), which is described elsewhere in this issue. The numerous cooperations between experts from various domains have led to the development of a variety of products and ideas, which makes the RICH project a good example of how the interdisciplinary cooperation of domain experts can foster innovation. In this contribution, we describe the tools that were developed within the RICH project, and we reflect on the virtues of the cross-fertilization between the archaeological and the computer science domain.

The outline of the contribution is as follows. We describe the institutional embedding of the RICH research. In addition, we provide an outline of the three main subprojects executed: (1) the digitization of ceramic fabric, (2) the development of techniques for the automatic classification and visualization of archaeological objects (i.e., coins and ceramic profiles), and (3) the development of software for information retrieval from archaeological reports. These three subprojects are described in detail in the subsequent sections and finally, we discuss the virtues of the interdisciplinary cooperation within the RICH project.

Embedding of the project

The RICH project is executed on-site at the Dutch National Service for Archaeology, Cultural Landscape and Built Monuments (RACM) in

Amersfoort, The Netherlands. The RACM has a coordinating role in protecting the national heritage in the physical public space.

The RICH project is integrated with existing initiatives of the RACM that provide the infrastructure for information exchange and the safeguarding of knowledge: the Knowledge Infrastructure for Cultural History (KiCH) and the National Reference Collection (NRc). KiCH is basically a GIS-based[1] web service for professional end-users to provide access to dispersed cultural historical information sources provided by a growing number of different governmental institutions and other partners. The NRc is an information infrastructure, primarily for specialists, to exchange knowledge on, and define standard terminologies for, cultural heritage sub domains. The NRc is being developed with the ultimate aim to enhance the quality, transparency, sustainable accessibility and reusability of the shared knowledge base. The principal contents of the NRc are richly illustrated standard vocabularies and thesauri on material culture categories (Lange 2004; Lange and Drenth 2006). This information structure is congruent with the traditional archaeological method of using archetypes put in a chronological (chronotypology) or other order (type series) as a measure for all other objects to be positioned along and giving them meaningful labels (Adams and Adams 1991). The NRc serves as the interface allowing meaningful communication between authors and users of the information (persons or computers), who may operate geographical and chronological separated spaces. KiCH and NRc are complementary and may evolve towards a single overarching project.

The results of the RICH project are essential contributions in the development towards the realization of the NRc and the reuse of information. RICH provides archaeologists, and in their wake other cultural heritage specialists, with digital tools to access, analyse, and enrich the shared knowledgebase by increasing the efficacy and efficiency of access, reinforcing its infrastructure, and improving the quality and consistency of the information stored.

The RICH project emphasizes the development of software based on image processing and machine learning algorithms. Both types of algorithms are well suited to meet the challenge of archaeology for the following two reasons. First, the emphasis on image processing in the RICH project agrees well with the predominance of visual analysis in the field. Second, the tacit nature of archaeological knowledge can only be acquired by a computer using computerized learning methods that learn from examples, rather than by explicit descriptions in terms of if-then statements. More concretely, an archaeological expert will have considerable difficulty in making explicit all variables that the expert takes into account to assign a certain object to a certain class, whereas assigning the class label to a find is straightforward. Machine learning algorithms do not require explicit descriptions and can learn a model by means of labelled, partially labelled, or even unlabelled examples. Hence, the use of machine learning seems highly appropriate for the archaeological domain.

Three main subprojects that illustrate the contribution of RICH and the use of image processing and machine learning are: (1) the digitization of ceramic fabric, (2) the development of techniques for the automatic classification and visualization of archaeological objects, and (3) the development of software for information retrieval from archaeological reports. The digitization subproject illustrates the importance of interdisciplinary collaboration. The nature of ceramic fabric poses special requirements on its digitization. The digital techniques for the subproject on classification and visualization of archaeological objects were motivated by the desire of archaeologists to classify objects in a consistent manner and, more importantly, to visually assess the variables underlying the classification. Finally, the development of software for the subproject on semi-automatic information retrieval from archaeological reports was motivated by the desire to come to grips with the heterogeneity of descriptions in the documents in the Dutch archaeological domain. In the following three sections, each of the subprojects is discussed in detail.

Digitization of ceramic fabric

Fragments of ceramic artefacts are commonly found at archaeological excavations. The analysis of their structure and composition is of pivotal importance to reconstructing the production methods and distribution of the pottery in the past. In particular, the visual analysis of fractured surfaces reveals important information on the source and composition of the fragment. Typically, the fragments are composed of a variety of tempering materials in addition to clay in order to facilitate the production process or to enhance the quality of the final ceramic product. The combination of particular types of material gives an indication of age and origin. Figure 1 shows an illustration of a fresh break of a sherd with tempering material clearly visible. The analysis of fractured sherds can be performed even if the fragments are too small to reveal the shape of the original pottery (the analysis of shape profiles is discussed in *Ceramic profiles* section).

The aim of the digitization subproject is to collect digital images of fractures to realise a large database to support research and evaluation. We collaborated with fabrics expert Gert van Oortmerssen of the Laboratory for Conservation & Material Studies of the Groningen Institute of Archaeology, who provided us with ceramic material and his expert knowledge. The interdisciplinary collaboration resulted in a non-standard digitization procedure that relies on a three-dimensional scanner, rather than on a two-dimensional scanner (e.g., a single photograph). In their visual examination of a fractured surface, archaeologists exploit the three-dimensional surface structure by moving the sherd subtly. The movements reveal the intricate physical texture of the fracture to the observer. Standard digitization of a fracture as a two-dimensional image would hamper the examination because important visual cues, such as transparency, lustre, and roughness would be lost. The interdisciplinary cooperation prompted us to develop a new method for the three-dimensional capture of fractured surfaces, rather than elaborating on the traditional two-dimensional photographs or scans. The basic idea of our method is to digitally represent a surface by means of a sequence of

FIGURE 1 Fresh fracture of a ceramic fragment revealing the fabric of the sherd.

images that, when shown in succession, forms a movie showing the movement of the surface, closely following the examination process in real life.

In our image sequences, we decided to define a single axis of rotation, which can be played using the browser's standard support for movie playing. The suggestion of reality is further improved by allowing an online viewer to virtually rotate the object by dragging with the mouse on the image. Examples of such movies can be seen at the RICH website.[2] Although other techniques for digitizing fabric could be used, the rotation sequences are preferable, because the images are photorealistic by definition and already useful without any filtering or interpolation.

In our digitization subproject, we furthered our interdisciplinary cooperation with Sylvia Pont of the Physics of Man, Human Perception group of Utrecht University. She suggested to employ the Bidirectional Texture Function (BTF) to quantify the surface of the fracture in our digital analyses (Dana *et al.* 1999). The BTF describes how the observed texture changes as a function of the angle of view and the angle of illumination owing to effects of shading, shadowing, occlusions and inter-reflections. Our shared interest in developing equipment for measuring the BTF gave rise to the joint development of the digitization system in which both rotation and image acquisition are computer-controlled. This enables keeping constant the illumination and viewpoint for the digitization of all fragments in the collection. The results obtained with an early prototype of the system yielded encouraging results (Boon *et al.* 2007). Recently, we completed the development of the BTF acquisition device. Figure 2 shows the device during acquisition of a sherd. The fabric movies are acquired by keeping the camera and lamp fixed while the object is rotated. Two arcs are attached to the apparatus; one for the camera and one for holding the lamp. The arc holding the lamp can also be rotated around the central rotation platform. Combinations of different camera and lamp positions allow for the acquisition of the BTF, which in turn provides a quantitative representation of the perceptual properties of the fractured surface. The acquisition system will be used to digitize fabric reference collections. When sufficient data are acquired, we will apply machine learning techniques (as described in the next section) to the digitized fragments. The techniques and procedures we have developed for acquisition and visualization of ceramic fabric can also be applied to other archaeological materials and objects, such as flint and coins. Clearly, the interdisciplinary collaboration was immaterial to the successful completion of the digitization subproject.

Automatic classification and visualization

Within the RICH project, we developed a wide variety of classification and visualization techniques based on image-recognition and machine-learning algorithms. In this section, we describe our techniques for the classification of coins and ceramic profiles. For the automatic classification of coins, the main result is expressed numerically as a classification performance on unseen coins. The sometimes ambiguous results obtained were found to be due to inconsistent labelling and the lack formal descriptions of the critical features of the medieval coins under consideration. Extensive interdisciplinary discussions with archaeologists and coin experts revealed that ground truth is hard to find in almost all archaeological data. This led to the realization that the standard methodology of machine learning, in which a perfectly labelled training set is used for creating a model from the data, does not apply to most of the archaeological domain. We therefore adopted an alternative machine-learning methodology in which the results were visualized, rather than quantified, so that archaeological experts could employ our tools in a semi-automatic manner. As an illustration, we describe the visualization of the similarity structure of profiles. The next two subsections describe the classification (coins) and visualization (profiles) approaches and their results.

Coin recognition

The aim of coin recognition is to identify the properties (such as currency or authority) of the depicted coin, or to find coins with a similar stamp in a reference collection of coins. The aim of a coin recognition system is to

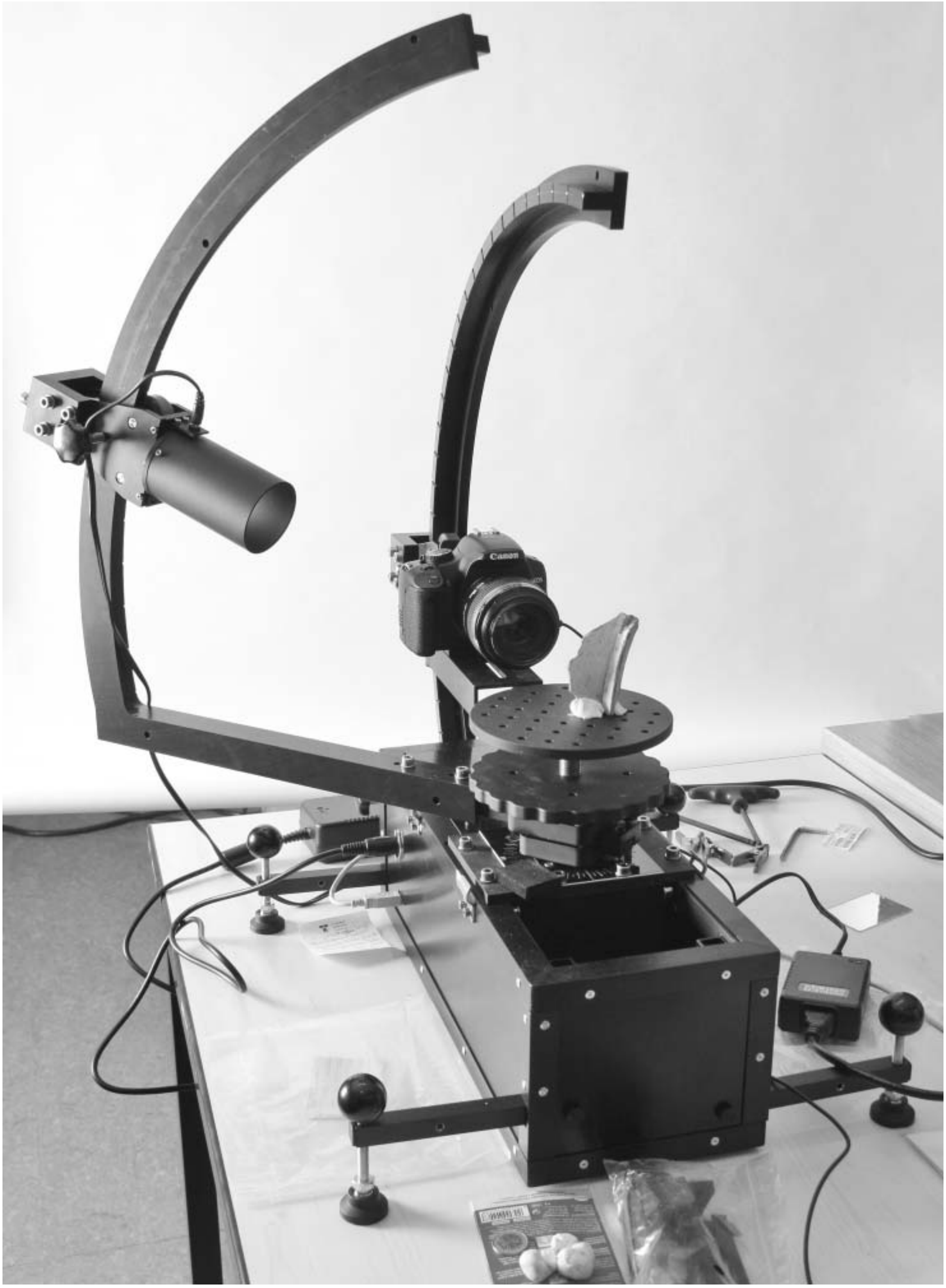

FIGURE 2 The acquisition device rotating a Roman potshard.

compute a similarity measure between coins that is invariant to rotations of the coins, because the orientation of the coin on the photograph is unknown beforehand. The similarity measure can be used in classification or retrieval tasks using a nearest neighbour scheme.

The coin recognition system

In coin recognition, most visual information that is of relevance to the class of the coin is contained in the stamp of the coin. From a computer vision perspective, the contours of the stamp of a coin are characterized by sudden changes in the pixel values in the coin image (so-called edges). These edges can be localized by inspection of the derivative or gradient of the coin image. The coin image gradient has a *magnitude*, which measures the strength of the change in pixel values, and an *orientation*, which measures the direction of the change in pixel values. As a result of the low contrast in many coin images, the gradient magnitude is often very noisy, and thus not very suitable for such an analysis. In contrast, gradient orientations are not sensitive to low contrast in coin images (Reisert *et al.* 2007). The robustness of gradient orientations to low-contrast images led us to use it as a basis for our coin recognition system.

Our approach consisted of three stages, which are described in detail by van der Maaten and Boon (2006). In the first stage, the gradient orientations of the coin images are computed. For each location in the coin, the gradient orientation represents the direction of change of pixel values. By definition, the gradient orientation at a location on the contour of the coin stamp is at right angles to the direction of the contour. In the second stage, the pairwise similarities between gradient orientation images are computed. The similarity between two gradient orientation images is computed by counting the number of identical orientations in the two gradient orientation images. The resulting count is an indication for the similarity of the corresponding two coin stamps. In the third stage, classification of the coins is performed based on the pairwise similarities. The classification of a coin face is performed using a 1-nearest neighbour classifier (see, e.g., Duda *et al.* 2001) that uses the similarity measure developed in the previous stage. The reason for the use of a 1-nearest neighbour classifier is that, in many historical coin datasets, the number of duplicate coins is limited. Generally, images of both coin faces are available, and thus a combination of the classification of the two coins is required. If the two coin faces of a coin are classified differently, we reject both classifications and classify the coin as unknown. Note that the pairwise similarities can also be used as input into a visualization technique such as the technique discussed in the *Visualization* section.

Coin recognition experiments

In the evaluation of the performance of our approach to coin recognition, we performed experiments on two datasets: (1) a modern coin dataset and (2) a historical coin dataset. We describe the set-up of our experiments and their results separately below.

The modern coin dataset contains approximately 30,000 coin images, corresponding to 15,000 coins.[3] The dataset is divided into a fixed training set of 20,000 coins, and a fixed test set of 10,000 coins. The training set contains 2268 different coin faces, corresponding to 692 coin classes. The test set contains 4 per cent coins that are not in the training set, and that should be classified as unknown.

The archaeological coin dataset is a dataset that was provided by Arent Pol of the Dutch Money and Bank Museum. The dataset contains 4822 coins from the Merovingen dynasty that were produced and used between the fifth and the eighth century in Northern Europe.

Classification of archaeological coins is more difficult than the classification of modern coins, for instance, because the variation in coins from the same origin (stamp) may be larger than variations between specimen from different origin. Moreover, obtaining ground truth for archaeological coin data is generally impossible as we lack the knowledge of the composition of the original parent population. The class labels of the historical coins are only available for a small portion of the coins. In addition, the class labels are known to contain errors. Hence, we decided not to use the labelling of the coins in our experiments.

For our experiments on the archaeological coin dataset, we selected 50 historical coins of which we know that the dataset contains one or more coins with similar coin faces. A larger test set would have been desirable, but cannot be constructed because the number of duplicate coins in the archaeological coin dataset is limited. The performance of the system is evaluated by human comparison of the 50 coins with the best matching coins that were identified by the system. The human subjects evaluate whether a query coin and the best matching prototype as identified by the system (i.e., the query-match pair) have perceptually similar coin faces. We report on the number of times in which the human subjects evaluate a query-match pair as perceptually similar.

Results

In Table 1, we present the results of the evaluation of our approach to coin classification on the modern coin dataset. We present the results of template matching based on gradient magnitude images. The results in Table 1 reveal that our approach is capable of correctly classifying a large percentage of the coins, while only making a low percentage of misclassifications (taking into account that 4 per cent of the coins in the test set were not in the training set).

In Table 2, we present the results of our experiments on the archaeological coin dataset. The table shows the percentage of query-match pairs that the human subjects evaluated as perceptually similar. In addition, we report the corresponding standard deviation, as well as the minimum and maximum values. Figure 3 shows 10 archaeological coin images and their corresponding best matches as identified by the system. In Figure 3, the upper coins are the query coins and the lower coins the best matches as identified by the system. The upper two rows depict five correct query-match pairs, whereas the lower two rows depict five incorrect query-match pairs. The results reveal that our approach is to some extent capable of dealing with distortions as a result of wear and small morphological variations of ancient coins. Hitherto, the performance of our system on archaeological coin data is limited by the sensitivity to incorrect estimations of the centre of the coin stamp, and by the

TABLE 1

GENERALIZATION PERFORMANCE OF THE APPROACH FOR THE MODERN COIN DATASET

Gradient	Correct (%)	Unknown (%)	Incorrect (%)
Magnitude	38.58	21.88	39.54
Orientation	92.92	6.70	0.38

TABLE 2

EXPERIMENTAL RESULTS ON THE ARCHAEOLOGICAL COIN DATASET

Gradient	Query-match sim. (%)	Standard dev. (%)	Min. (%)	Max. (%)
Magnitude	20.60	8.80	12.00	30.00
Orientation	43.60	10.19	26.00	60.00

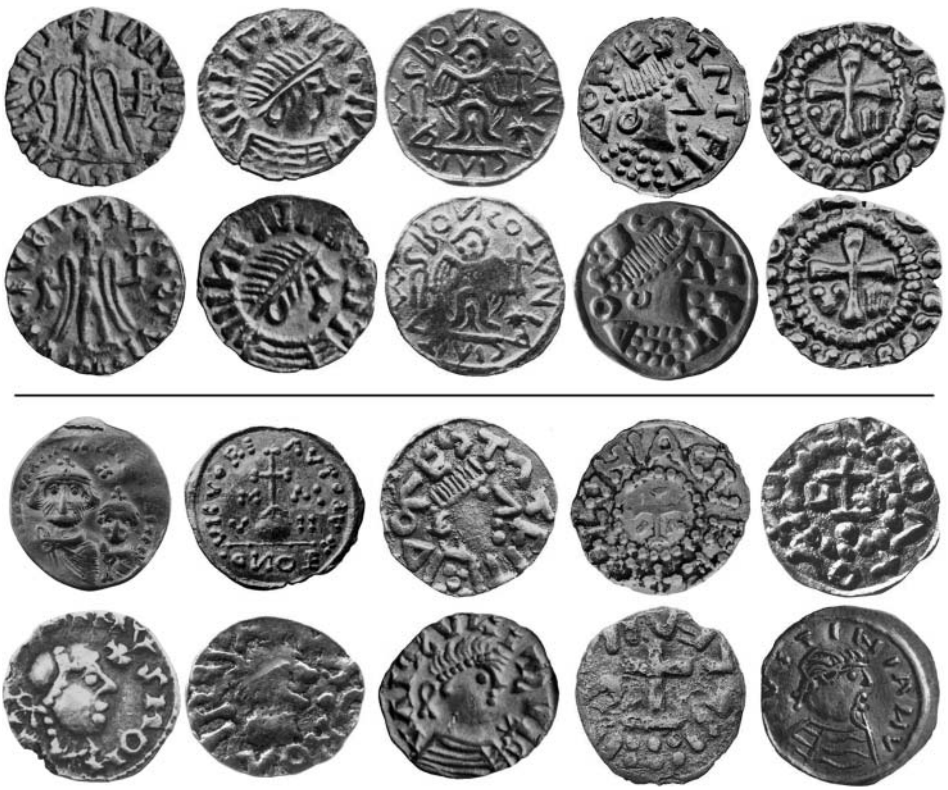

FIGURE 3 Query-match pairs from the historical coin experiment.

inability to deal with large morphological shape variations in archaeological coin stamps.

Ceramics profiles

In archaeology, pottery is typically documented using profile drawings that show a cross-section of the sherd or the complete pot. The shape is an important criterion for identifying the type and hence the age and origin of the pottery The assessment of shapes by humans is subjective and error-prone, which is why we propose to use techniques for shape matching to perform objective measurements on the similarity of ceramics profiles. Below, we describe our approach to the assessment of ceramics profiles, which consists of two main stages: (1) shape matching using shape contexts and (2) visualization using t-Distributed Stochastic Neighbour Embedding (t-SNE).

Ceramic shape matching

In the literature, various schemes have been proposed for the comparison of shapes (Mokhtarian *et al.* 1996; Kim *et al.* 2000; Ricard *et al.* 2005; Zhang and Lu 2003; Grigorescu and Petkov 2003). After preliminary experiments with various shape matching schemes, we found that a matching scheme based on shape contexts was best tailored to the task of computing similarities between ceramics profiles. The shape matching scheme used is described below.

Shape contexts represent a shape contour by means of a collection of points that are sampled from the shape contour (Belongie *et al.* 2001). The sampling of points from the shape contour is performed using a method that selects points as uniformly as possible over the shape contour. The points that are sampled from the contour are represented by means of so-called shape context descriptors. Shape context descriptors represent statistical information for a single point on a shape contour in a small histogram. A shape context descriptor is computed for each point on the shape contour, and the complete set of shape context descriptors (a so-called shape context) for a single shape can be used to characterize a shape.

The shape contexts of two shapes can be compared using an elaborate scheme that attempts to determine how one shape should be warped in order to be transformed to the other shape. The 'strength' of this transformation (the so-called energy) is diagnostic for the similarity between the two shapes: the larger the transformation, the more dissimilar the two shapes. As in the coin recognition system, the resulting similarity measure can be employed in a 1-nearest neighbour classifier or in the visualization technique we discuss below.

Visualization

The shape matching stage described above (and also the matching of coins described in the coin recognition system) results in a matrix that contains the pairwise (dis)similarities between all objects in the dataset. In the visualization of these pairwise dissimilarities, we aim to construct a two-dimensional scatter plot (in which each point corresponds to a single ceramics profile) in such a way that the pairwise dissimilarities between the profiles are modelled as well as possible in terms of the pairwise distances between their corresponding points in the scatter plot. Various techniques for multidimensional scaling, such as classical scaling (Torgerson 1952) and Sammon mapping (Sammon 1969), are capable of constructing such scatter plots. We have reviewed these techniques (van der Maaten *et al.* 2009) and found them to suffer from serious shortcomings that prohibit successful visualization of real-world data. In order to alleviate these shortcomings, we opted to use a recently introduced technique called t-Distributed Stochastic Neighbour Embedding (t-SNE), which performs very well in the visualization of real-world datasets (van der Maaten and Hinton 2008).

Experiments

We performed experiments on a dataset of 996 ceramics profiles, most of which were found in the Dutch soil (the dataset also includes some African, South American, and Australian ceramics profiles). Figure 4 presents the results of our experiments. The figure reveals that our approach appropriately captures the shape variations in ceramics profiles. From the absence of clearly separated clusters in the visualization, we may conclude that there are no clearly distinguishable groups of ceramics. Such visualizations are useful to archaeologists and may be used in a semi-automatic, inductive way to meet the challenge of archaeology. More specifically, ultimately it will allow to incorporate the implicit knowledge of archaeologists into a transparent description or taxonomy of ceramic profiles.

Archaeological reports

The third subproject of the RICH project concerns the digital handling of textual information. Archaeological sources typically consist of images and text. The combined digital analysis of images and text facilitates the extraction of information from documents. As stated in the introduction, the archaeological knowledge, information, and data are distributed over a wide range of institutes. Efforts to archive archaeological reports in central digital repositories are under way. These repositories allow researchers to access reports digitally. However, searching for relevant visual and textual information from a certain period is cumbersome with plain search algorithms. Within the RICH project, information search software tailored to the archaeological domain is being developed that allows for more advanced search methodologies. The software is called ‘Open Boek’ (Open Book) and can be used by archaeologists to automatically retrieve relevant (fragments of) PDF-documents from a large repository. The development of Open Boek was prompted by a series of discussions and meetings with interdisciplinary teams composed of machine learning experts and archaeologists. Currently, Open Boek focuses on the retrieval of text only. In future versions of Open Boek, retrieval from text and images will be realised by incorporating algorithms developed in the context of the second subproject.

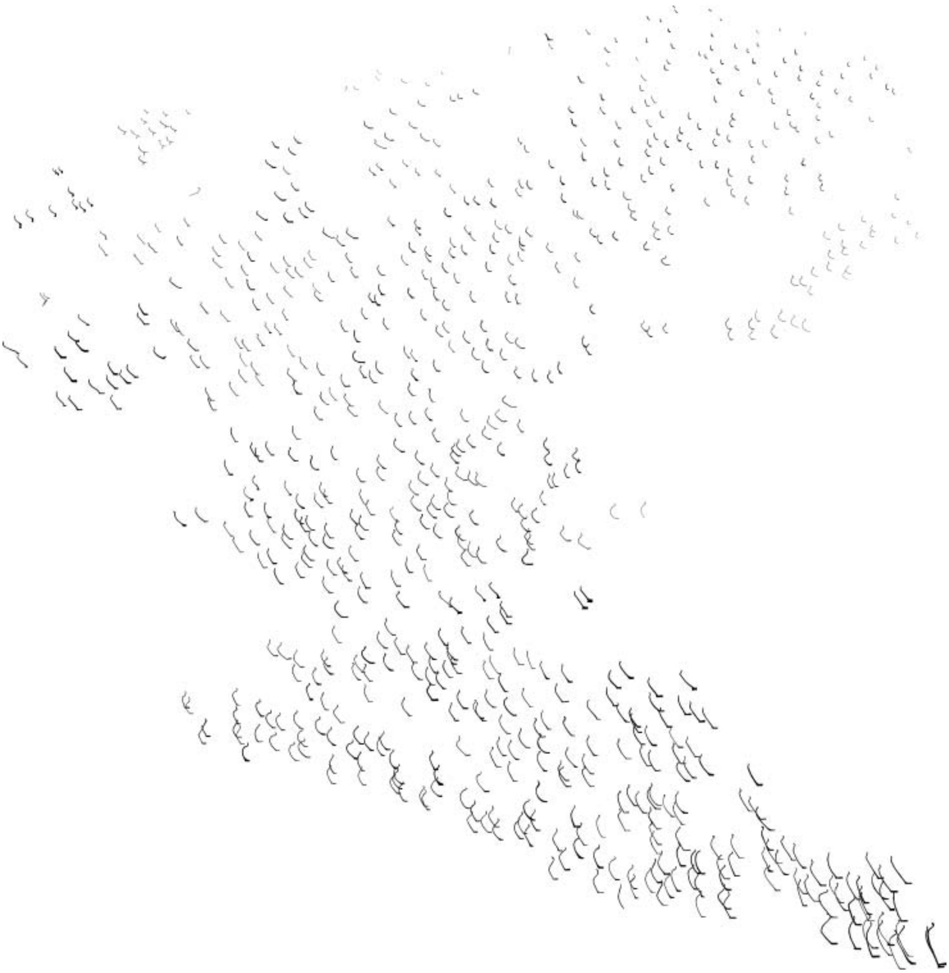

FIGURE 4 Visualization of 996 ceramics profiles using a combination of shape contexts and t-SNE. All profiles are shown at the same scale.

In order to explain the virtues of the Open Boek software, we focus on a specific example pertaining to the retrieval of textual information that refers to a certain historical period. Time and place are obvious dimensions in papers on archaeology and historical research. The main challenge addressed is how to retrieve documents on chronological criteria hidden in the text. A straightforward and common solution to this problem is the addition of metadata to the individual papers. The metadata, in turn, are attached to ontologies or typologies in which the time dimension of a concept is defined. However, adding metadata by humans is extremely costly and usually based on implicit knowledge. Moreover, such ontologies rapidly become unwieldy and the concepts are often contradictory. There is also the problem of granularity: a document or even a page may contain many chronological expressions that are not easily lumped into a single period.

In the Open Boek software, we adopt an alternative solution to the problem. We employ machine learning for the automatic recognition of chronological expressions in the text, either as named periods ('middle ages'), or as ordinals, cardinals, or numbers ('the fourth century', '1568–1648') and to build indexes containing these expressions in normalized form. When such indexes are built, they can be used to retrieve pages that contain references to a certain year or period.

A major problem is the categorization of the numbers in a text: when do numbers have chronological meaning (i.e., years, periods, or eras) and when do they not have this (i.e., dimensions, prices, or inventory numbers). We employed a classical and straightforward machine-learning technique called Memory-Based Learning (MBL) to assign numbers to the correct class. In the training stage of MBL, a database is created with labelled examples for a particular language. In the indexing stage, unseen examples are labelled according to their best-matching neighbours in the database. MBL methods are ubiquitous in machine learning and are all variants of the *k*-nearest neighbour algorithm (see, e.g., Duda *et al.* 2001). In Open Boek, we employ TiMBL 5.1,[4] a decision-tree-based implementation of *k*-nearest neighbour classification. TiMBL uses indexes in the instance memory extensively and therefore can handle discrete data and large numbers of various examples well.

We perform chronological indexing in three stages. In the first stage, the candidate items for classification are collected by a numeric preparser that recognizes both written (i.e., 'two', 'second', '2nd', '2-nd', 'VI') and numerical (i.e., '1', '100') items. In this stage, a list with names ('middle ages', 'iron age', 'roman period') is also consulted and the corresponding phrases are flagged as chronological phrases. All chronological phrases are labelled and stored in the database. It should be noted that the first stage is the only language-dependent stage in Open Boek. In the second stage, TiMBL is applied to the example phrases. Each (pre-parsed) phrase is matched to the stored examples in the database and labelled accordingly. The third and last stage is the normalization and creation of the index proper. This includes assignment to BC or AD, and the decision whether the expression contains a single year or is a period. The expression 'between 1200 and 1300' is clearly a period, but so is 'third century', and 'between the first century BC and the year 500'.

In this way, chronological information has become independent from the individual vernacular of the original author and with Open Boek, even higher resolution can be gained than is described in the original text. This is illustrated by the following example: assume that in a text on the famous castle 'Muiderslot', a period of '1375 to 1427' is mentioned. When a researcher for some reason is interested in what happened in the years around the turn of the fourteenth to fifteenth century, this particular treatise on the Muiderslot will be included, among many other documents in his/her catch.

The identification of useful semantic classes has been performed in collaboration between the computer scientists and archaeologists. This has led to three important additional features of Open Boek: (1) the recognition of addresses and other toponyms, (2) the possibility to generate metadata for individual documents, and (3) the automatic extraction of salient chronological and topological data, complex-types, and other archaeologically interesting properties described in the text. These features will prove of pivotal importance when handbooks and monographs are digitized. For the moment, it is of high relevance to make available the huge amount of information that is stored in the thousands of archaeological reports that are produced each year. Without Open Boek, these reports remain largely closed to the general user.

Discussion and conclusions

The three RICH subprojects described address three distinct stages of archaeological research in the digital era. The digitization can be considered to be the initial step in any digital project that deals with real-world objects. From our description of the acquisition apparatus developed for the scanning of fractured surfaces, it became clear that the interdisciplinary nature of the CATCH project allowed us to acknowledge the importance of a three-dimensional acquisition procedure. The solution can be readily adopted in archaeological practice. Information on pottery fabrics becomes more easily accessible, while strengthening transparency of decision and enhancing the efficacy and efficiency of the use of resources. Similarly, our development of automatic recognition techniques for archaeological objects (coins and ceramic profiles) was improved by arranging interdisciplinary team meetings. Our machine-learning technique presents results visually to the archaeologists, which facilitates the evaluation and detection of inconsistencies by the technique or archaeologist. New research questions will be raised and new insights will follow. In particular, the archaeologist will be able to investigate time/place relationships and dependencies. Finally, the need for the Open Boek initiative arose from a specific desire from archaeologists and historians to search in a more intelligent manner for archaeologically relevant information in a large collection of reports.

From our experience in the three subprojects, it is evident to us that an approach in which computer scientists work in isolation from experts in the application domain (in this case in archaeology) is bound to fail. Technical obvious solutions often do not solve the domain-specific problems at hand. The lesson learned from the RICH project is that a successful interdisciplinary collaboration requires an intensive communication between two (or more) disciplines and preparedness to attempt to understand the 'other way of thinking'. In our specific case, computer scientists prefer to think in terms of clear definitions and of ground truths. Such thinking is warranted within the computer science domain, because computers are well defined and (theoretically) operate in a predictable manner. This contrasts with the archaeological domain where definitions and classifications are fuzzy and continuously the subject of debate or controversy. We conclude by stating that bridging both worlds is an effort in itself, but rewarding as it is to the benefit of both. Computer scientists become acquainted with a new challenging application domain and archaeologists learn how to use digital tools to work towards the goal of achieving a more objective description of archaeological knowledge. It is exactly this rich cross-fertilization that exemplifies the success of the RICH project.

Acknowledgements

The authors thank Erik Drenth and Lara Laken for many insightful discussions. Cooperation with Sylvia Pont and Gert van Oortmersen formed the foundation for the success of the digitization subproject. We thank Sylvia for supporting the development of the acquisition hardware, Gert is gratefully acknowledged for providing the ceramic material and for his expert knowledge on fabrics. We are indebted to Arent Pol (Dutch Money and Bank Museum) for providing the medieval coins dataset. We thank Vincent Mom for digitizing and annotating the collection of ceramics profiles.

Paul Boon, Laurens van der Maaten, and Hans Paijmans are supported by The Netherlands Organisation for Scientific Research (NWO), project CATCH-RICH (grant 640.000.002).

Notes

1 GIS is an abbreviation for Geographic Information System.

2 See http://www.referentiecollectie.nl/sequenceviewer

3 The dataset is publicly available for download from http://muscle.prip.tuwien.ac.at

4 TiMBL is available online from http://ilk.uvt.nl

Bibliography

Adams, W.Y., and E.W. Adams. 1991. *Archaeological typology and practical reality. A dialectical approach to artifact classification and sorting*. Cambridge, UK: Cambridge University Press.

Belongie, S., J. Malik, and J. Puzicha. 2001. Shape matching and object recognition using shape contexts. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(4): 509–22.

Boon, P.J., S.C. Pont, and G.J.M. van Oortmerssen. 2007. Acquisition and visualization of cross section surface characteristics for identification of archaeological ceramics. *Proceedings of the CAA 2007*.

Dana, K.J., B. van Ginneken, S.K. Nayar, and J. Koenderink. 1999. Reflectance and texture of real-world surfaces. *ACM Transactions on Graphics* 18(1): 1–34.

Duda, R.O., P.E. Hart, and D.G. Stork. 2001. *Pattern classification*. Wiley Interscience Inc.

Grigorescu, C., and N. Petkov. 2003. Distance sets for shape filters and shape recognition. *IEEE Transactions on Image Processing* 12(10): 1274–86.

Kim, H.-K., J.-D. Kim, D.-G Sim, and D.-Il Oh. 2000. A modified Zernike moment shape descriptor invariant to translation, rotation and scale for similarity-based image retrieval. *Proceedings of the IEEE International Conference on Multimedia and Expo*, 307–10.

Lange, A.G. 2004. The European Reference Collection. Founding the future. In *Proceedings of the International Conference on the European Electronic Reference Collection*, ed. A.G. Lange, 139–45.

Lange, A.G., and E. Drenth. 2006. The National Reference Collection. Quest for structure and content. In *Proceedings of the National Service for Archaeological Heritage in the Netherlands*, ed. R.M. van Heeringen and R.C.G.M. Lauwerier, 199–214.

Mokhtarian, F., S. Abbasi, J. Kittler. 1996. Efficient and robust retrieval by shape content through curvature scale space. *Proceedings of the International Workshop on Image Databases and Multimedia Search*, 35–42.

Reisert, M., O. Ronneberger, and H. Burkhardt. 2007. A fast and reliable coin recognition system. *Lecture Notes on Computer Science (LNCS)*, 415–24.

Ricard, J., D. Coeurjolly, and A. Baskurt. 2005. Generalizations of angular radial transform for 2D and 3D shape retrieval. *Pattern Recognition Letters* 26(14): 2174–86.

Sammon, J.W. 1969. A nonlinear mapping for data structure analysis. *IEEE Transactions on Computers* 18(5): 401–9.

Torgerson, W.S. 1952. Multidimensional scaling I: Theory and method. *Psychometrika* 17: 401–19.

van der Maaten, L.J.P., and P.J. Boon. 2006. COIN-O-MATIC: A fast and reliable system for coin classification. *Proceedings of the MUSCLE CIS Coin Recognition Competition Workshop*, 7–17.

van der Maaten, L.J.P., and G.E. Hinton. 2008. Visualizing data using t-SNE. *Journal of Machine Learning Research* 9(Nov): 2431–56.

van der Maaten, L.J.P., E.O. Postma, and H.J. van den Herik. 2009. Dimensionality reduction: A comparative review. *Journal of Machine Learning Research (submitted)*.

Zhang, D., and G. Lu. 2003. Evaluation of MPEG-7 shape descriptors against other shape descriptors. *Multimedia Systems* 9(1): 15–30.

Notes on Contributors

Correspondence to: Eric Postma, TiCC, Faculty of Humanities, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands.

Email: eric.postma@gmail.com