Computer vision and machine learning for archaeology

Laurens van der Maaten[1], Paul Boon[1], Guus Lange[2],
Hans Paijmans[1], and Eric Postma[1]

[1] MICC-IKAT, Maastricht University
Maastricht, The Netherlands
l.vandermaaten @micc.unimaas.nl
[2] Dutch State Service for Archaeology
Amersfoort, The Netherlands

Abstract
Until now, computer vision and machine learning techniques barely contributed to the archaeological domain. The use of these techniques can support archaeologists in their assessment and classification of archaeological finds. The paper illustrates the use of computer vision techniques for archaeology with two examples: (1) a content-based image retrieval system for historical glass and (2) an automatic system for medieval coin classification. The content-based image retrieval system automatically finds artifact drawings in a reference collection that are most similar to a photograph or drawing of an excavated historical glass. The similarity measurements are based on the outer shape contours of the artifacts. The system can speed up the process of classifying historical glass, and make it more objective and controllable. The coin classification system will be trained on a collection of Dutch early-medieval coins. For this system, we present preliminary results on modern coin data.

1 Introduction
Developments in computer science and statistics have provided archaeologists with a wide range of tools. These tools include tools for GIS, 3D modeling, predictive modeling, visualization, simulations, remote sensing, resource management, etc. However, the archaeological domain lacks tools that support the archaeologist in one of his main tasks, viz. the classification of artifacts. Since classification of artifacts is usually performed by visual inspection, techniques from computer vision and machine learning can be applied in order to develop such tools. Computer vision is the field concerned with the development of techniques that allow computers to evaluate and analyze images or sequences of images (i.e., video). Important tasks in computer vision include image segmentation, object detection, and object classification. Overviews of current techniques for computer vision can be found in (Forsyth and Ponce 2003; Gonzalez and Woods 2002). Machine learning is the field that is concerned with the recognition of patterns in statistical data. When the statistical data is labeled (i.e. all classifications for the data are specified), the use of machine learning techniques allows for the automatic classification of unlabeled data. Overviews of techniques for machine learning can be found in (Duda, Hart, and Stork 2001; Mitchell 1997).
The paper illustrates the contribution of computer vision and machine learning techniques to classification of artifacts with two examples: (1) a content-based image retrieval system for historical glass and (2) an automatic system for coin classification. These systems are being implemented within the RICH project, a project that aims at the application of image analysis techniques to archaeology.
The outline of the remainder of the paper is as follows. In section 2, we give an overview of work that applies computer vision techniques to the archaeological domain. Section 3 presents the system for content-based image retrieval of historical glass. In section 4, we present work on an automatic system for classification of coins. The performance of the two systems, as well as the contribution of such systems to archaeology is discussed in section 5. In section 6, we conclude that our work illustrates the potential of machine learning and image analysis techniques for the domain of archaeology.

2 Related work
Until now, the number of studies in which computer vision and machine learning techniques are applied in the archaeological domain is limited. This section provides a (non-exhaustive) concise of the work done until now.
In (da Gama Leitão and Stolfi 2002), a system is presented for the automatic reassembly of two-dimensional fragmented objects such as ceramic sherds. The system is tested on a small amount of artificially created sherds. Although the results presented in the paper are promising, it is not likely that the

system is applicable to large scale fragment reconstruction problems such as the Forma Urbis Romae problem. Recently, Stanford University started a project aiming at solving the Forma Urbis Romae reconstruction problem using computer techniques (Koller et al. 2005).
In (Kampel and Melero 2003), a system is presented that reconstructs a virtual 3D vessel using profile information of a ceramic fragment. Two different approaches are proposed: (1) an approach based on Hough-features and (2) an approach based on genetic algorithms. The Hough-based approach is capable of automatic classification of the fragment. Although the results look promising, no classification performances are presented for the system.
A system that allows for automatic classification of ceramics profiles is presented in (Karasik et al. 2004). The system is based on analysis of the curvature function of the ceramics profile. Although no classification performances are reported, the authors show that the system captures increasing complexity of ceramics profiles over time.

3 Content-based image retrieval for historical glass

Currently, classification of historical glass artifacts is performed manually by an archaeological expert. Generally, the expert attempts to find already classified artifacts that are perceptually similar to the unclassified artifact. In order to effectively find such artifacts, usually, the expert searches through a reference collection. A reference collection is a collection of reference artifacts, which is usually published as a set of formalized descriptions together with line drawings of the artifacts. An important Dutch reference collection for historical glass is described in (Kottman 1999). Manual comparison of historical glass artifacts with artifacts from a reference collection is a time-consuming process. Additionally, the identification of artifacts in this way is a highly intuitive and uncontrollable process.
In order to partially overcome these drawbacks, we propose the use of a content-based image retrieval system by the archaeological expert. Content-based image retrieval is the retrieval of images that are perceptually similar to a query image. In order to retrieve similar images, a content-based image retrieval system employs a similarity measure based on certain image features.
We developed a content-based image retrieval system that compares photographs of artifacts with drawings from a reference collection. When an archaeologist finds a historical glass in the soil, he can photograph the artifact, and present the artifact photograph as a query to the system. Subsequently, a number of objects from the reference collection are presented that are perceptually similar to the photographed historical glass. The expert can now easily make a classification decision on the historical glass, by comparing the found artifact with the perceptually similar objects from the reference collection that are presented by the system. In this way, classification of historical glass by experts can be improved and speeded up.
The value of a content-based image retrieval system for the classification of historical glass is highly determined by the way it finds images that are perceptually similar, i.e., by the choice of the similarity measure. This section briefly describes the similarity measure that is used in our system. Our choice of the similarity measure is based on the historical glass dataset that is described in (Kottman 1999). In section 3.1, we give a brief description of the historical glass dataset. Section 3.2 describes the similarity measure that is used in our system. In section 3.3, the results and the contribution of the system to archaeology are discussed.

3.1 Historical glass dataset

The historical glass reference collection is described in (Kottman 1999). It consists of 314 glass artifacts[1], which all date from the period 1400-1915. The artifacts are described by drawings and their corresponding classifications. Figure 1 shows an example of an artifact drawing from the reference collection. The left part of the drawing represents the profile of the artifact, whereas the right part of the drawing shows the surface of the artifact. The drawings in the reference collection are based on selected artifacts. We photographed the selected artifacts in a way that is commonly applied for the presentation of archaeological artifacts. An example of a photograph is shown in Figure 2. The artifact in Figure 2 corresponds to the artifact presented in Figure 1. The photographs allow for evaluation of the content-based image retrieval system.
Usually, an artifact drawing looks different than an artifact photograph. The drawings in the reference collection are based on reconstructions of archaeologically complete objects by fitting sherds and

[1] The collection is shown on http://www.referentiecollectie.nl/rich/

fragments. Missing information is indicated by a dashed line. Furthermore, colour information and non-relevant texture information is eliminated in the drawings. The texture information that is present in the drawings is an abstract representation of the texture on the artifact. Therefore, the visual information in the drawings that is most accurate for use in our content-based image retrieval system is the information on the outer shape of the artifact. I.e., content-based image retrieval in historical glass reference collections relies on similarity measures based on outer shape features.

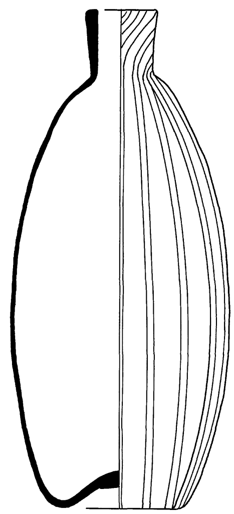

Figure 1: Drawing of a historical glass artifact.

3.2 Similarity measure

Many similarity measures based on outer shape are proposed in literature. We selected and implemented five shape similarity measures: similarity measures based on (1) curvature scale spaces (Mokhtarian, Abbasi, and Kittler 1996), (2) shape contexts (Belongie, Malik, and Puzicha 2001; Mori, Belongie, and Malik 2005), (3) turning functions (Tanase and Veltkamp 2005), (4) Hausdorff distance (Huttenlocher, Klanderman, and Rucklidge 1993), and (5) moment invariants (Hu 1962). Preliminary experiments revealed similarity measures based on shape contexts to be best-performing. This is due to the coarse nature of shape contexts, which makes them more robust to distortions caused by 3D-rotations, broken artifacts, noise, etc. Therefore, we focus on the shape context similarity measure in this paper.

The computation of shape context similarity measures consists of three steps: (1) the pre-processing of the image, (2) the computation of shape context descriptors, and (3) the computation of the similarity measure. First, a shape profile has to be computed from the artifact photograph. This is necessary in order to compute outer shape features of the artifact. The pre-processing process consists of five steps. First, a Canny edge detector (Canny 1986) is applied on the artifact photographs. Second, in order to connect unconnected edges, a morphological dilation operation (Serra 1982) on the resulting image is performed. Third, a NOT-operation and a bucket fill are performed on the resulting image, assuming that the upper-left pixel of the image is not part of the artifact. In this way, we obtain an image in which the background is black, and the artifact is white. However, edges caused by shadows or table edges still remain in the image. Therefore, a morphological erosion operator (Serra 1982) is applied fourth, in order to remove these edges. An example of the resulting shape profile is shown in Figure 3. Fifth, a Sobel edge detector is applied on the resulting image, in order to obtain the final shape representation.

Figure 2: Photograph of historical glass artifact.

Once the outer shape profile is extracted from the photograph, shape contexts are computed from the obtained shape contours. Shape contexts are global shape descriptors first introduced in (Belongie, Malik, and Puzicha 2001). In a shape context representation, a shape is represented by a number of points that is sampled from the boundary of the shape contour. The points are described as shape context descriptors. Shape context descriptors describe the distance and angle of a point to all other points in a discretized log-polar space. By means of this description, a set of shape context descriptors (i.e., a shape context) contains global information about the shape.

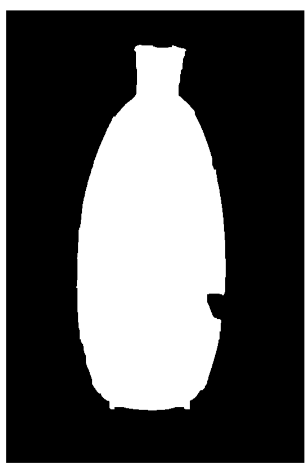

Figure 3: Shape profile of historical glass artifact.

After the computation of the shape context descriptors, a similarity measure is computed. The similarity measure is based on the dissimilarity between two shape contexts. In order to compute the dissimilarity between two shapes, Belongie et al. first compute the Euclidean distance between all shape context descriptors in the two shapes. Subsequently, they apply a Hungarian matching algorithm on the distance matrix. In this matching algorithm, the optimal matching between the shape context descriptors of two shapes is computed. The costs of this matching provide a measure for the dissimilarity of the two shapes. The bending energy of the thin plate spline warping describing the warping between both shapes is added to the costs, in order to further enhance the shape matching (Belongie, Malik, and Puzicha 2001). Additionally, Belongie et al. add image appearance costs in order to further improve their results.

However, we found this approach to be to time-consuming for application in a content-based image retrieval system. Therefore, we propose the use of a simple k-nearest neighbor matching algorithm for the matching of shape descriptors instead of the Hungarian matching algorithm. In order to be able to use a k-nearest neighbor matching algorithm, the shape context descriptors should be ordered. Furthermore, the computation of the similarity measure should be invariant to changes in the starting point of the shape contour. Starting point invariance can be achieved by computing the distance between the first shape context and *n* shifted versions of the second shape context (where *n* is the number of shape context descriptors in the shape context). Subsequently, the minimum distance is set as the final similarity measure. This yields a computational complexity of $O(n)$ (where n is the number of shape context descriptors), whereas the original Hungarian matching algorithm has a computational complexity of $O(n^3)$. In order to further reduce computation time, we eliminate the use of thin plate splines and image appearance costs.

3.3 Results

In order to evaluate the performance of our content-based image retrieval system, we presented the photographs as described in subsection 3.1 as queries to the system. Figure 4 shows a number of typical queries[2]. Although the results in many cases look promising, our system suffers from one major drawback. Glass artifacts are often distorted by missing parts. The missing parts of the glass affect its outer shape in a significant way. Therefore, our system does not perform well on severely damaged glass artifacts. A solution to this problem is the manual adjustment of the shape profile by the user.

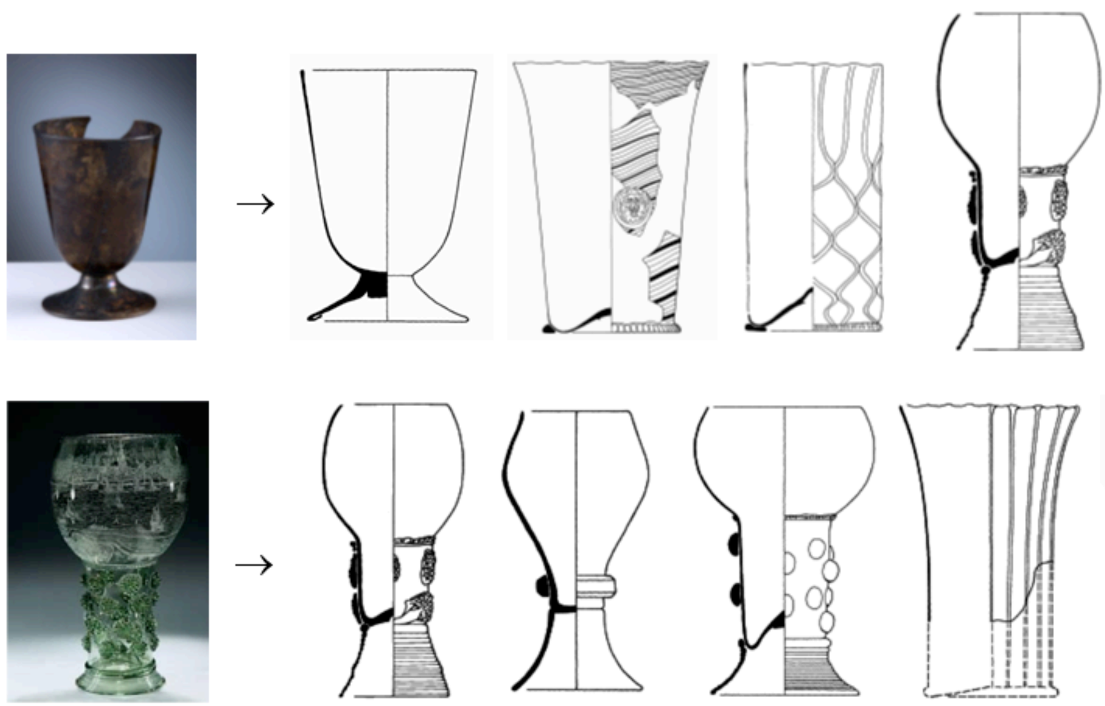

Figure 4: Typical results of a query.

Although our system does not perform well on highly degraded artifacts, we still think it forms an important contribution to archaeology. Even when the retrieved artifacts are not perfect matches, the system provides an entry into the reference collection website, allowing the archaeologist to classify artifacts faster and with a lower risk of errors. In addition, our system can be used for automatic shape analysis of entire reference collections. By performing the shape analysis on the entire reference collection, it is possible to present perceptually similar artifacts for every artifact in the reference collection[3]. This can

[2] The system is available for testing from http://www.referentiecollectie.nl/rich/
[3] This is done for the historical glass collection on http://www.referentiecollectie.nl/rich/

be beneficial to projects such as NRc (Lange 2004) and eRC, by providing a new way of navigation through reference collection websites.

4 Automatic coin classification

Coins are commonly found archaeological artifacts. For instance, the collection at the Dutch Money and Bank Museum contains over 200,000 historical coins. However, since the coins are usually stored in saves, these collections are generally not accessible to the public. Therefore, a great effort is done to provide access to the coins digitally, e.g. in NUMIS[4]. However, the current presentation of collections is not well suited for helping non-experts to classify coins. Classification of coins requires a great deal of expert knowledge. Therefore, a system that aids the classification of coins can be an important asset to medieval coin collections managers and users.
In other work (van der Maaten and Postma 2006), we presented a number of features for automatic classification of coins. We briefly discuss these features and some preliminary results in subsection 4.1. Currently, we are working on the development of a similar system for early-medieval coins from the Merovingen-dynasty. An example of these coins is shown in Figure 5.

Figure 5: Example of a coin.

4.1 Modern coin classification

Until now, work on automatic coin classification is limited. Some attempts for classification systems for modern coins can be found in (Adameck, Hossfeld, and Eich 2003; Davidsson 1996; Fukumi et al. 1992; Passeraub et al. 1997). Generally, these systems are limited by performance or require specific devices such as proximity sensors. Vision-based systems with high classification performances are presented in (Nölle et al. 2003; Huber et al. 2005). A drawback of the systems by Nölle et al. and by Huber et al. is that they highly rely on additional sensor data on coin area and thickness.
In (van der Maaten and Postma 2006), we presented preliminary results for image features based on contour and texture information. Contour information is information contained in the stamp of the coin. We represent contour information from the coin by means of multi-scale edge-based statistical features. Edges are obtained by applying a Sobel edge-detection on the coin images. The best-performing edge-based statistical features measure the joint angle-distance distribution of edge pixels over the coin, and represent this information in a multi-scale histogram (van der Maaten and Boon 2006). Texture information is information on the texture of the coin. Since coins usually do not have a texture, texture features might seem uninteresting in coin classification. However, modern coins often contain very detailed pictures. Therefore, these pictures can be considered as the texture of the coin. Texture information can be represented using approaches based on the Gabor wavelet (Lee 1996) or the Daubechies D4 wavelet. For technical details we refer to (van der Maaten and Postma 2006; van der Maaten and Boon 2006), they are outside the scope of this paper.
We trained our system on a set of modern European coins that were collected after the introduction of the euro. The dataset contains 692 different coin types with 2,270 different coin faces. Our system classifies approximately 78% of the coins in the test set correctly (van der Maaten and Boon 2006). Usually, misclassifications are due to very dirty coins or due to unknown coins.

---

[4] See http://83.149.77.24:8080/numis/ for more information

5 Discussion

In section 3 and 4, we presented two examples of image analysis systems that can be applied on archaeological data. The two systems illustrate the possibilities of computer vision and machine learning techniques. Application of these techniques in the archaeological domain has two major contributions to the field. First, it allows for more objective and more controllable classification of archaeological artifacts. Furthermore, it can speed up the classification process. By using a wireless internet connection, archaeologists can semi-automatically perform important classifications in the field. Hereby, use of computer vision and machine learning techniques can even influence excavation decisions. Second, the use of computer vision and machine learning techniques can give a broader public access to archaeological knowledge, both by providing automatic classification systems to non-experts, and by allowing for new presentation methods for on-line archaeological collections. In addition, computer vision and machine learning techniques allow archaeologists to develop new typologies or to evaluate old ones. For example, applying visualization techniques such as MDS on the shape analysis of the entire historical glass collection leads to similarity maps such as the one shown in Figure 6. By manually creating clusters in these maps, the archaeologist can easily create a new typology. This process can even be automated using techniques for unsupervised learning.

Future work should focus on further development of the discussed systems. The main challenge here is to incorporate archaeological knowledge into already existing computer vision algorithms. Incorporation of archaeological knowledge can be obtained by cooperation with archaeological experts in the development process, or by the development of semi-automatic adaptive systems. Such systems employ a so-called 'human-in-the-loop' approach in order to automatically learn archaeological knowledge over time.

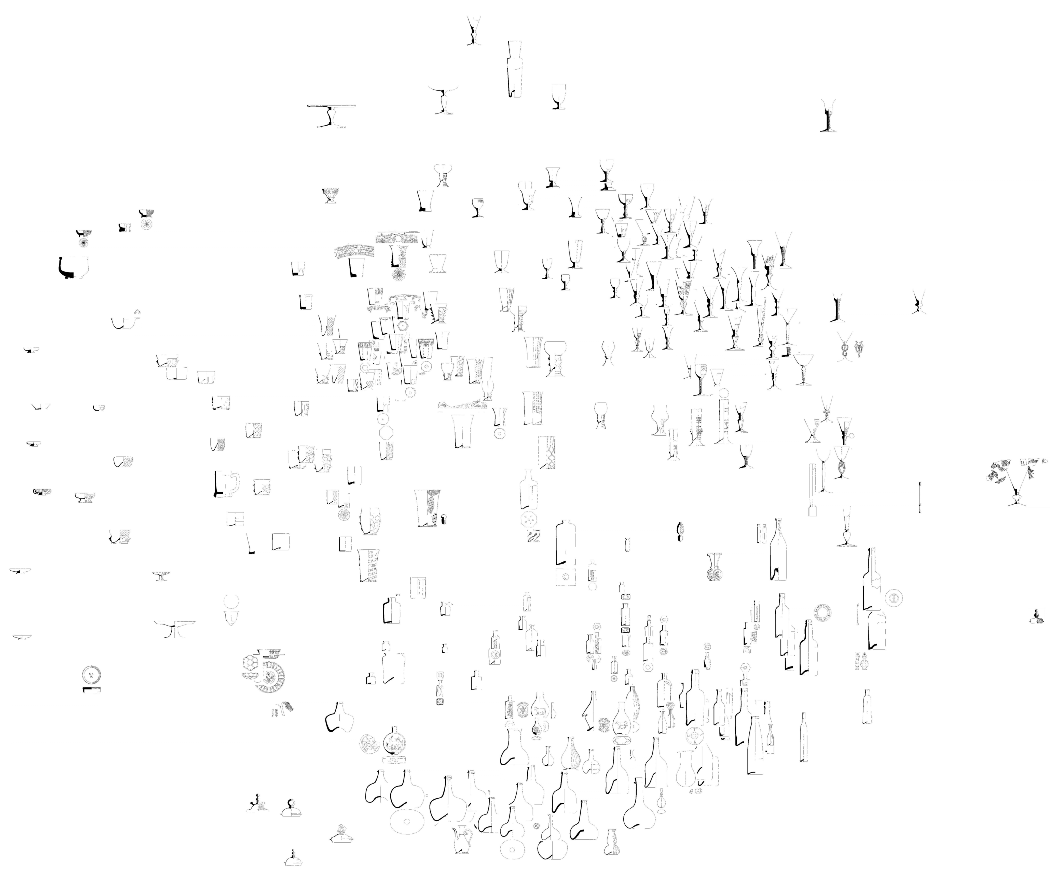

Figure 6: Historical glass similarity map.

6 Conclusions

The presented work illustrates the power of computer vision and machine learning techniques for archaeology with two examples. Although both systems are in a preliminary state, and a great deal of work has yet to be done, the advantages of the application of computer vision techniques in the archaeological domain have become clear. We have indicated various ways in which the techniques can improve archaeological processes. Additionally, we indicated directions on which further work should focus.

Acknowledgements

This work was supported by NWO/CATCH under grant 640.002.401. We thank the Dutch Royal Coin Cabinet for providing the medieval coin dataset.

References cited

Adameck, M., Hossfeld, M., and Eich, M. 2003. Three color selective stereo gradient method for fast topography recognition of metallic surfaces. Proceedings of the SPIE, 5011:128–139.

Belongie, S., Malik, J. and Puzicha, J. 2001. Shape matching and object recognition using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(24):509–522.

Canny, J. 1986. A computational approach to edge detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(6):679–698.

da Gama Leitão, H.C. and Stolfi, J. 2002. A multiscale method for the reassembly of two-dimensional fragmented objects. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(9), pages 1239–1251.

Davidsson, P. 1996. Coin classification using a novel technique for learning characteristic decision trees by controlling the degree of generalization. In Ninth International Conference on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems, pages 403–412. Gordon and Breach Science Publishers.

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

Forsyth, D.A. and Ponce, J. 2003. Computer vision: A modern approach. Prentice Hall.

Fukumi, M., Omatu, S., Takeda, F., and Kosaka, T. 1992. Rotation-invariant neural pattern recognition system with application to coin recognition. IEEE Transactions on Neural Networks, 3(2):272–279.

Gonzalez, R.C. and Woods, R.E. 2002. Digital image processing. Prentice Hall.

Hu, M.K. 1962. Visual pattern recognition by moment invariants. IRE Transactions on Information Theory, IT-8:179–187.

Huber, R., Ramoser, H., Mayer, K., Penz, H., and Rubik, M. 2005. Classification of coins using an eigenspace approach. Pattern Recognition Letters, 26(1):61–75.

Huttenlocher, D.P., Klanderman, G.A., and Rucklidge, W.A. 1993. Comparing images using the Hausdorff distance. IEEE Transactions on Pattern Analysis and Machine Intelligence, 15(9):850–863.

Kampel, M. and Melero, F.J. 2003. Virtual vessel reconstruction from a fragment's profile. In D. Arnold, A. Chalmers, and F. Niccolucci, editors, Proceedings of the 4th VAST, pages 79–88.

Karasik, A., Sharon, I., Smilansky, U., and Gilboa, A. 2004. Typology and classification of ceramics based on curvature analysis. In Computer Applications and Quantitive Methods in Archaeology 2003, pages 472–475.

Koller, D., Trimble, J., Najbjerg, T., Gelfand, N., and Levoy, M. 2005. Fragments of the city: Stanford's digital Forma Urbis Romae project. In Proceedings of the 3rd Williams Symposium on Classical Architecture.

Kottman, J. 1999. Glass. In M. Bartels, editor, Cities in sherds catalogue, volume 2, pages 939–1028.

Lange, A.G. 2004. Co-ordination of national reference collections. In Computer Applications and Quantitive Methods in Archaeology 2003.

Lee, T.S. 1996. Image representation using 2D Gabor wavelets. IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(10):959–971.

Mitchell, T.M. 1997. Machine learning. McGraw Hill.

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

Mori, G., Belongie, S., and Malik, J. 2005. Efficient shape matching using shape contexts. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(11):1832–1837.

Nölle, M., Penz, H., Rubik, M., Mayer, K., Holländer, I., and Granec, R. 2003. Dagobert - a new coin recognition and sorting system. In Proceedings of the 7th International Conference on Digital Image Computing, Techniques, and Applications, pages 329–338.

Passeraub, P.A., Besse, P.-A., de Raad, C., Dezuari, O., Quinet, F., and Popovic, R.S. 1997. Coin recognition using an inductive proximity sensor microsystem. In International Conference on Solid State Sensors and Actuators, volume 1, pages 389–392.

Serra, J. 1982. Image analysis and mathematical morphology. Academic Press.

Tanase, M., and Veltkamp, R. 2005. Part-based shape retrieval. In Proceedings of the ACM Multimedia 2005, pages 543–546.

van der Maaten, L.J.P., and Boon, P.J. 2006. COIN-O-MATIC: a fast and reliable system for coin classification. In Proceedings of the MUSCLE Coin Workshop 2006, in press.

van der Maaten, L.J.P., and Postma, E.O. 2006. Towards automatic coin classification. In Proceedings of the EVA-Vienna, in press.