THE ILAC-PROJECT: SUPPORTING ANCIENT COIN CLASSIFICATION BY MEANS OF IMAGE ANALYSIS

Albert Kavelar*, Sebastian Zambanini*, Martin Kampel*, Klaus Vondrovec** and Kathrin Siegl**

*Computer Vision Lab,
Vienna University of Technology
Favoritenstr. 9/183-2, 1040 Vienna, AT
{kavelar, zamba, kampel}@caa.tuwien.ac.at
http://caa.tuwien.ac.at/cvl/research/ilac/

**Coin Cabinet,
Museum of Fine Arts
Burgring 5, 1010 Vienna, AT
{klaus.vondrovec, kathrin.siegl}@khm.at

KEY WORDS: ancient coins, computer vision, numismatics, optical character recognition, image matching

ABSTRACT:

This paper presents the ILAC project, which aims at the development of an automated image-based classification system for ancient Roman Republican coins. The benefits of such a system are manifold: operating at the suture between computer vision and numismatics, ILAC can reduce the day-to-day workload of numismatists by assisting them in classification tasks and providing a preselection of suitable coin classes. This is especially helpful for large coin hoard findings comprising several thousands of coins. Furthermore, this system could be implemented in an online platform for hobby numismatists, allowing them to access background information about their coin collection by simply uploading a photo of obverse and reverse for the coin of interest. ILAC explores different computer vision techniques and their combinations for the use of image-based coin recognition. Some of these methods, such as image matching, use the entire coin image in the classification process, while symbol or legend recognition exploit certain characteristics of the coin imagery. An overview of the methods explored so far and the respective experiments is given as well as an outlook on the next steps of the project.

1 INTRODUCTION

ILAC stands for *I*mage-based c*L*assification of *A*ncient *C*oins and is an interdisciplinary project that links the field of Computer Vision with Numismatics and aims at generating a selection of image-based classification methods for ancient Roman Republican coins. The competence in the field of Computer Vision is provided by the Computer Vision Lab of the Vienna University of Technology, whereas the Department of Coins and Medals at the *Kunsthistorisches Museum Wien* (KHM, Museum of Fine Arts, Vienna) contributes the numismatic know-how. The benefit is mutual: Computer Vision aids numismatists in the task of classifying coins, whereas the field of Numismatics provides a large number of new and challenging data for fundamental research in Computer Vision. While the project's primary focus lies on basic research in the field of Computer Vision, the long-run goal is to create an application which assists numismatists in their everyday work. Thus, the project bridges the gap between fundamental scientific and application-oriented research.

Despite the existence of classification systems for modern coins, no fully-automated image-based classification system for ancient Roman Republican coins has yet been researched successfully. This results from various fundamental differences between ancient coins and their modern counterparts: Modern coins of the same class look identical because they are machine-made and undergo quality assurance while ancient coins were struck manually. For this reason and due to wear from use and environmental influences, such as chemicals in the soil, every ancient coin is unique. Thus, ancient coins show a high intra-class and a low inter-class variability (see Fig. 1), which facilitates the recognition of a specific coin specimen but impedes its classification.

In the course of this project, various techniques and concepts of Computer Vision are evaluated for the application of image-based coin classification. The sub-fields investigated reach from content-based image retrieval over symbol recognition to optical character recognition. The project, originally set for three years, has already reached its final year. Besides primarily dealing with the development of best-practices for coin image acquisition and the creation of a large-scale image database comprising over 4100 Roman Republican coins, a registration-based coin classification technique and an initial prototype of a coin legend recognition system were developed in the first year.

In the second year, we concentrated on the research of illumination-invariant local image features that are robust against highlights and shadows introduced by the metallic relief surface of ancient coins, which can change drastically depending on the lights angle of incidence. The proposed image feature is based on Gabor filters and outperforms state-of-the-art image descriptors in scenarios of varying illumination conditions (Zambanini and Kampel, 2013b). Furthermore, the refinement of the legend recognition algorithm based on local image descriptors and the adoption of methods first introduced in the field of scene text recognition were pursued in the second year. The legend recognition uses SIFT features for the description of the individual characters and combines them to meaningful words of a given lexicon via pictorial structures.

The goal for the final year is to enrich the method-mix by researching a coin image recognition technique, to further improve and evaluate the methods researched so far and to ultimately combine all methods to a fully-fledged coin recognition framework, which allows to perform an image-based classification of ancient Roman Republican coins.

Figure 1: Examples of Roman Republican coins. (a) Low inter-class variability. Three coins of different classes that only differ in their legend. (b) High intra-class variability. Three coins of the same class that show little similarity.

The remainder of this paper is organized as follows: Section 2 reviews the state of the art in coin recognition. Section 3 provides a general overview of the project. Section 4 gives an outline of the various image-based coin recognition methods that have been implemented so far. Finally, Section 5 summarizes the project's progress and points out the next steps to be taken.

2 RELATED WORK

The first to apply computer vision methods to coin recognition were (Fukumi et al., 1991). They presented a system for separating 500 yen from 500 won coins. The coin image is split into ring segments, for which the sum of gray levels is computed. The sums are used to form a feature vector that is classified by a multi-layered neural network. *Dagobert* is a powerful coin recognition system for modern coins which can tell more than 600 different coins apart and was introduced by (Nölle et al., 2003). Their coin classification system uses binarized edge images of input coin images that are compared to images stored in a master database. A nearest neighbor classifier assigns the input image to the class resulting in the closest match. However, their system is not solely image-based but also uses two additional sensors to capture the coin's thickness and diameter. Bremananth et al. (Bremananth et al., 2005) proposed a two-step approach for the classification of 1-, 2- and 5-rupee coins. Initially, the location numeral on the coin is found using template matching. The region containing the numeral is extracted and Gabor features are computed for the rotationally invariant character recognition. Van der Maaten and Poon (van der Maaten and Poon, 2006) introduced COIN-O-MATIC, another coin classification system for modern coins which works similar to Dagobert. An edge image is computed from the input image and split into segments using a log-polar grid. This grid allows to compute a histogram capturing the angle and distance distribution of edge pixels, which is classified by a nearest neighbor classifier. The coin classification system proposed by Reisert et al. (Reisert et al., 2006) transforms the coin image to polar coordinates after successfully segmenting the coin from the background via a generalized Hough transform (Ballard, 1981). Gradients are computed for the image and quantized into

different orientation bins. For each orientation bin a binary image is calculated and classified using a nearest neighbor method.

One of the first end-to-end coin recognition techniques for ancient coins was presented by Kampel and Zaharieva (Kampel and Zaharieva, 2008). In their work, various interest point detectors, local image descriptors and their combinations are tested for their applicability to image-based coin recognition. A nearest neighbor classifier assigns the input image to one of three classes. Arandjelovic (Arandjelovic, 2010) introduced an automated coin attribution system for Roman imperial coins. The central contribution of his work is a novel feature type called *locally-biased directional histogram* that captures geometric relationships between interest points determined with the Difference-of-Gaussian detector (Lowe, 2004). Huber-Mörk et al. (Huber-Mörk et al., 2011) presented an extension to the coin identification of Kampel and Zaharieva (Kampel and Zaharieva, 2008). A preselection step which analyzes the contour of the coin is introduced and allows for an efficient pruning of mismatching coins. Hence, the computationally more complex matching using local descriptors only needs to be performed for a smaller number of coins. The contour is analyzed by casting equally spaced rays from the coin's center of gravity and intersecting them with the coin border. The distances of the intersection points to the border of a hypothetical perfect circle fit to the actual coin contour are measured and form a feature vector. Arandjelovic presented another coin classification system exploiting coin legends (Arandjelovic, 2012). His system is tailored to Roman imperial coins and assumes that the legends are running along the coin border. This assumption allows a transformation of the image to polar coordinates, where the legends are aligned horizontally. Thus, the image descriptor does not need to be rotationally invariant; consequently, HoG features (Dalal and Triggs, 2005) can be used for describing the individual characters. The legend is detected by shifting a sliding window along the horizontal coin border line and computing HoG features for every window location. These character probabilities are combined to words using dynamic programming. The legend word recognition enables preselecting a subset of all possible classes for which SIFT-based matching is performed.
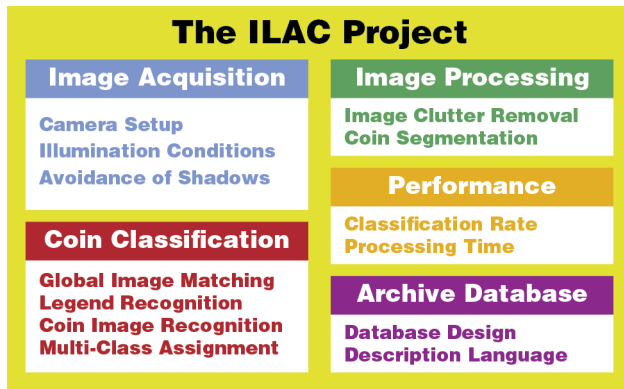
Figure 2: Overview of the ILAC project.

Figure 3: Obverse and reverse of a Roman Republican coin with its individual parts highlighted. This example's obverse shows the crowned head of the godess Liberta while its reverse depicts a quadriga as main coin image.

3 PROJECT OVERVIEW

The overall aim of the ILAC project is to provide a selection of image-based classification methods for ancient Roman Republican coins. Fig. 2 gives an overview of the different aspects this project comprises. Since coins have a three-dimensional surface structure, illumination conditions have a strong influence on the coin's appearance in photographs taken for digitization. In order to provide an optimal starting point for an image-based classification, best practices for coin image acquisition were developed in an initial step. These practices suggest an ideal camera and lighting setup which allows for the optimal visibility of the coin relief and ensures that coins of the same class are illuminated in the same way. The acquired images are stored in a database along with additional textual information of the coin. This information comprises a textual description of the depicted imagery and legend, the coin's weight and diameter and the coin type according to the categorization provided by Crawford (Crawford, 1974).

Obverse (front) and *reverse* (back) of a coin are captured individually; however, not only is the coin depicted in the images but also rulers and cardboard signs providing handwritten information for size and type reference respectively. Thus, before image-based classification methods can be applied to these pictures, a coin segmentation robust against image clutter has to be performed (Zambanini and Kampel, 2009). After a successful segmentation, the image can be further processed by image-based classification methods.

In the classification step, discriminative features need to be extracted from the segmented coin images. The exploration and evaluation of such features constitutes a major aspect of the ILAC project. There are several features that show a high potential for a reliable differentiation between the various coin classes. Besides the computation of local image features to gain an overall description of the coin, the textual inscriptions and coin images (see Fig. 3) are highly discriminative features, for which individual recognition methods were implemented. A combination of this methods is expected to increase the classification performance while reducing the computation time, since certain methods can be computed faster than others and can thus be used for initial pruning. The different methods that have been developed so far are described in detail in Section 4.

Another goal of the ILAC project is to provide image-based classification with reasonable response times, that is, significantly faster than existing methods. This may require finding a trade-off between classification accuracy and processing performance. Finally, in order to verify the classification result achieved by the implemented method, ground truth needs to be provided and queried. Therefore, a database has to be designed with respect to these requirements.

3.1 Coin Data

The ILAC project focuses exclusively on the gold and silver coinage of the Roman Republic (ca. BC $280/225 - 27$). The decision to work on Roman Republican coins is grounded on two main reasons. First, Crawford (Crawford, 1974) provides a standard reference book which gives a comprehensive overview of the different coin types of this period; his classification system is still accurate. This relates to the fact that the structure of the coin-issues is almost linear, since the magistrates responsible for minting were replaced every year, contrary to other monetary fields of comparable size and importance. Crawford lists 550 distinct coin reference numbers, many of which comprise various denominations and slight variations resulting in additional subtypes. Counting all the different variants, the number of actual types lies at approximately 2000. The differences between the individual types can often be found in the legends or tiny control and minting marks. Thus, methods dedicated to the recognition of those highly discriminative features were developed.

Secondly, the ILAC project is carried out in cooperation with the Department of Coins and Medals of the *Kunsthistorisches Museum Wien* (KHM, Museum of Fine Arts Vienna), which houses one of the world's largest collections of Roman Republican gold and silver coinage. It comprises 5800 pieces, out of which 4100 are relevant to this project. This cooperation allows experimentation with image acquisition setups and provides firsthand access to numismatic and historic knowledge.

3.2 Image Acquisition

In the course of this project, obverse and reverse of the 4100 relevant Roman Republican coins of the KHM were digitized using a DSLR camera. In the process of digitization, coins of the same type were aligned consistently with respect to the light source. Coins were always aligned so that depicted heads face the key light. All coins were captured under the exact same lighting conditions to minimize variations introduced by highlights and shadows. Moreover, the lighting was placed to maximize the legibility of the imagery.

4 IMAGE-BASED COIN CLASSIFICATION

ILAC is an image-based classification system combining various computer vision techniques that focus on different aspects of the coins. Global image matching considers the entire coin and tries

to match an input image to sample images of every class using densely computed SIFT descriptors. Since legends and depicted local coin images are highly discriminative, techniques dedicated to their recognition were developed. The proposed legend recognition method computes scores for a set of densely sampled candidate character locations and combines them to meaningful lexicon words using dynamic programming. Finally, a coin image recognition providing a coarse-grained classification based on the depicted images using *Bags of Visual Words* (BoVW) was researched. Moreover, a fruitful combination of the two former methods was implemented and showed an increased classification accuracy.

4.1 Global Image Matching

Global image matching has the advantage that it does not rely on machine learning techniques, and thus, no training data is required. This is a considerable advantage when working with ancient coins, because museums usually prefer to have a few examples of many different types to provide their visitors a broader overview. Thus, only a few images per type are available, which impedes the use of machine learning techniques. The missing training is compensated by the flexible matching model for dense correspondence that can handle the spatial variations of local structures between coins of the same class. The matching algorithm is reminiscent of the SIFT flow method (Liu et al., 2011). For every pixel in the input image, SIFT features are computed and form a dense field of features. This allows the computation of pixel-to-pixel correspondences between two images. The Euclidean distance between two SIFT features is considered as their matching costs. The matching of the entire SIFT field can be described as an energy term, which compares every SIFT descriptor of one field with the descriptors in the respective pixel neighborhood of another field. Minimizing this term yields the final matching score for the two images. However, minimizing this function for large-scale images is a complex operation. In order to accelerate the classification process, a coarse-to-fine approach was proposed (Zambanini and Kampel, 2013a). That is, SIFT flow matching is performed in multiple steps with each subsequent step operating on images in a higher resolution than the previous one. For each step k, the acceptance rate defining how many images are passed on to be inspected at a higher resolution in the classification stage can be tweaked with a parameter λ_k ranging from 1% to 100%.

The experiments were carried out on 180 images of the reverse side of coins belonging to 60 different classes. Each class comprises 3 different images and the coarse-to-fine granularity was set to four steps. Moreover, two different test/training set configurations were used; the first option uses only one reference image while the other configuration uses two reference images per class. The best result of 83.3% classification rate is achieved when two reference images per class are used and 70% of all images are rejected in the first subselection step as well as 50% in the second and third step. The same classification rate can be achieved when no subselection is performed during matching, but takes approx. 8 times longer to compute.

4.2 Legend Recognition

The legend recognition method (Kavelar et al., 2012) uses object recognition techniques rather than standard OCR methods, since OCR relies on successful binarization for the separation of text and background. Coin legends have the same color as the rest of the coin, thus intensity changes result only from the coin surface relief structure. Therefore, effective binarization is not possible. Instead, the appearance of the individual letters occurring in coin legends is taught to a *Support Vector Machine* (SVM) using

	$n = 5$	$n = 10$	$n = 20$	$n = 30$
Best, initially	31.7%	21.7%	17.8%	13.9%
Top 3, initially	62.2%	39.4%	21.7%	18.3%
Best, re-scored	53.3%	42.8%	34.4%	28.9%
Top 3, re-scored	81.1%	66.1%	57.8%	51.7%

Table 1: Legend recognition results. Top 3 indicates that the correct word is among the three most probable words found.

a single SIFT descriptor (Lowe, 2004) (see Fig. 4 (a)). The legend words considered comprise 18 different characters which the SVM has to distinguish. The training process uses 50 100×100 pixel-sized, manually segmented images per class. The recognition process works as illustrated in Fig. 4 (b). The input image is first scaled down to a standardized size of 348×348 pixels to ensure an approximately equal font size for all legends. This spares the computation of SIFT features in various scales. Next, in the keypoint extraction step, regions of interest are detected with an entropy filter. From this regions, a list of candidate character locations (CCLs) is densely sampled and passed to the character recognition step, which computes a SIFT descriptor for each of the CCLs and tests it against all 18 SVMs to receive 18 scores for every CCL. Hence, this step leads to a dense likelihood map for every letter. In the final word recognition step, pictorial structures (Felzenszwalb and Huttenlocher, 2005) are used to generate word hypotheses for all legend words provided via a lexicon. To increase the confidence in the hypotheses, they are re-scored by computing SIFT descriptors for fixed orientations based on the CCLs in the individual hypothesis. Finally, words that received scores above a certain threshold are rejected. Thus, for an input image and a lexicon containing the possible legend words, the legend recognition algorithm outputs a list of legend words ordered by their scores. The classification rate depends on the lexicon size n used and whether the word hypothesis are re-scored or not. Table 1 gives an overview of the results achieved for 180 coin images. *Best* indicates the cases where the word with the lowest score (i.e., the best match) equals the ground truth while *Top-3* indicates that the correct word is among the three best matching words detected.

4.3 Coin Image Recognition

Just like the legend, the coin image provides a highly discriminative feature. For Roman Republican coins, the coin image, such as a she-wolf or a dolphin, is depicted on the reverse side, while the obverse usually shows the head of a god. Since multiple coin types share the same coin images, a fine-grained classification cannot be performed based solely on the coin image. However, it can serve as a preselection step, which prunes the set of classes. The proposed algorithm (Anwar et al., 2013) is based on a *Bag of Visual Words* (BoVW) algorithm. For an input image, SIFT features are densely extracted at a constant pixel stride. The computed features are quantized using k-means clustering; the number of clusters k determines the size of the visual vocabulary. To describe a novel image, the computed SIFT features are mapped to their closest visual word in Euclidean space. Finally, a histogram based on the number of features assigned to the individual words is constructed and describes the depicted symbol. This process, however, does not consider spatial relations of the individual parts of a symbol. Thus, various tiling patterns capable of capturing spatial relationships (rectangular, circular, log-polar) were evaluated. The best classification rate of 90.6% is achieved with a vocabulary size of 100, a pixel stride of 5 and rectangular tiling.
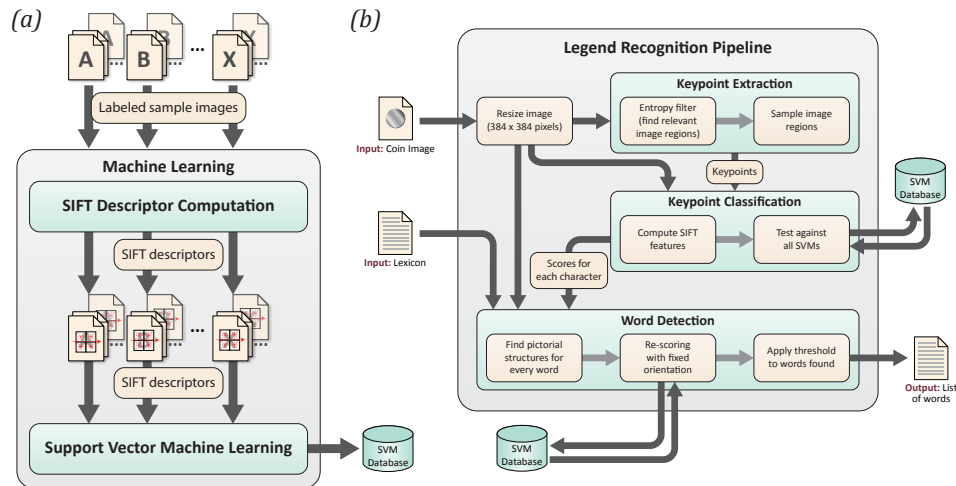
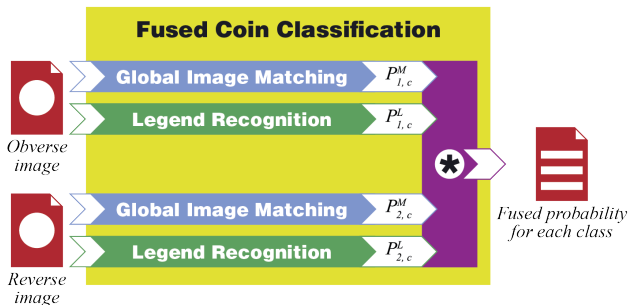Figure 4: (a) Training of SVMs. (b) Legend Recognition Pipeline.

Figure 5: Overview of the fused classification method, which combines legend recognition with global image matching.

Figure 6: Classification accuracy of legend recognition, image matching and hybrid coin classification for the top N classes.

4.4 Hybrid Coin Classification

This method fuses image matching with legend recognition in order to improve the overall classification performance (Zambanini et al., 2013). This is done by computing global image matching scores and legend word scores individually for obverse and reverse of a coin; thus, four different scores, called pseudo-probabilities, are computed, as illustrated in Fig. 5. Since not all words occur on both coin sides, two different, smaller lexica can be employed. As described before, smaller lexica lead to higher recognition rates. Finally, the four pseudo-probabilities are combined to give an overall score, which results in a better classification rate suggesting that a combination of all methods will even further increase the accuracy. Fig. 6 shows the accuracy for image matching, legend recognition and the hybrid method. The x-axis shows the number of classes N within which the correct class was recognized; i.e., for $N = 5$, the correct class is found among the 5 classes with the highest computed overall probability.

5 CONCLUSIONS AND FUTURE WORK

This paper presented the ILAC project, in the course of which several image-based classification methods for ancient Roman Republican coins have been developed. Different aspects of the depicted imagery, such as the legend or the copin image, are recognized individually for an efficient preselection of classes. This way, computationally more complex classification methods considering the entire coin only need to be performed for a subset of sample classes. A first approach towards fusing the different classification m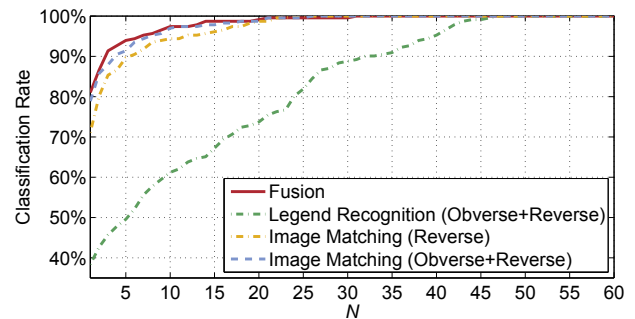ethods has already been implemented and showed promising results, which encourages a further combination of even more individual recognition techniques. Besides the improvement of the classification accuracy of the existing methods, the reduction of processing time will be tackled within the next few months. The ultimate goal is to provide a framework that generates an ordered list of classes sorted by how well they match the given input image that can be used in various applications and hides the underlying complexity from the user. Application areas range from an online web portal for the numismatic community to applications for mobile devices which allow for an image-based coin classification with smartphones.

REFERENCES

Anwar, H., Zambanini, S. and Kampel, M., 2013. Supporting ancient coin classiffcation by image-based reverse side symbol recognition. In: CAIP 2013 - The 15th Conference on Computer Analysis of Images and Patterns.

Arandjelovic, O. D., 2010. Automatic attribution of ancient roman imperial coins. In: Proceedings of the Conference on Computer Vision and Pattern Recognition, pp. 1728–1734.

Arandjelovic, O. D., 2012. Reading ancient coins: automatically identifying denarii using obverse legend seeded retrieval. In: Proceedings of the 12th European conference on Computer Vision - Volume Part IV, pp. 317–330.

Ballard, D. H., 1981. Generalizing the hough transform to detect arbitrary shapes. Pattern Recognition 13, pp. 111–122.

Bremananth, R., Balaji, B., Sankari, M. and Chitra, A., 2005. A new approach to coin recognition using neural pattern analysis. In: Proceedings of IEEE INDICON 2005, pp. 366–370.

Crawford, M. H., 1974. Roman republican coinage. Cambridge University Press.

Dalal, N. and Triggs, B., 2005. Histograms of oriented gradients for human detection. In: Proceedings of the Conference on Computer Vision and Pattern Recognition, pp. 886–893.

Felzenszwalb, P. F. and Huttenlocher, D. P., 2005. Pictorial structures for object recognition. International Journal of Computer Vision 61, pp. 55–79.

Fukumi, M., Omatu, S., Takeda, F. and Kosaka, T., 1991. Rotation-invariant neural pattern recognition system with application to coin recognition. In: Proceedings of the International Joint Conference on Neural Networks, Vol. 2, pp. 1027–1032.

Huber-Mörk, R., Zambanini, S., Zaharieva, M. and Kampel, M., 2011. Identification of ancient coins based on fusion of shape and local features. Machine Vision Applications 22, pp. 983–994.

Kampel, M. and Zaharieva, M., 2008. Recognizing ancient coins based on local features. In: Advances in Visual Computing, Vol. 5358, Springer, pp. 11–22.

Kavelar, A., Zambanini, S. and Kampel, M., 2012. Word detection applied to images of ancient roman coins. In: International Conference on Virtual Systems and Multimedia, pp. 577–580.

Liu, C., Yuen, J. and Torralba, A., 2011. Sift flow: Dense correspondence across scenes and its applications. Pattern Analysis and Machine Intelligence 33, pp. 978–994.

Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision 60(2), pp. 91–110.

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, pp. 329–338.

Reisert, M., Ronneberger, O. and Burkhardt, H., 2006. An efficient gradient based registration technique for coin recognition. In: Proceedings of the Muscle CIS Coin Competition Workshop, pp. 19–31.

van der Maaten, L. J. and Poon, P., 2006. Coin-o-matic: A fast system for reliable coin classification. In: Proceedings of the Muscle CIS Coin Competition Workshop, pp. 7–18.

Zambanini, S. and Kampel, M., 2009. Segmentation of ancient coins based on local entropy and gray value range. In: Proceedings of the 4th International Conference on Computer Vision Theory and Applications, Vol. 1, pp. 273–276.

Zambanini, S. and Kampel, M., 2013a. Coarse-to-fine correspondence search for classifying ancient coins. In: J.-I. Park and J. Kim (eds), Computer Vision - ACCV 2012 Workshops, Lecture Notes in Computer Science, Vol. 7729, Springer Berlin Heidelberg, pp. 25–36.

Zambanini, S. and Kampel, M., 2013b. A local image descriptor robust to illumination changes. In: J.-K. Kmrinen and M. Koskela (eds), Scandinavian Conference on Image Analysis, Lecture Notes in Computer Science, Vol. 7944, Springer Berlin Heidelberg, pp. 11–21.

Zambanini, S., Kavelar, A. and Kampel, M., 2013. Improving ancient roman coin classification by fusing exemplar-based classification and legend recognition. In: International Workshop on Multimedia for Cultural Heritage.