

Automatic Coin Classification by Image Matching

Sebastian Zambanini and Martin Kampel

Computer Vision Lab, Vienna University of Technology, Austria

Abstract

This paper presents an automatic image-based ancient coin classification method that adopts the recently proposed SIFT flow method in order to assess the similarity of coin images. Our system does not rely on pattern classification as discriminative feature extraction and classification becomes very difficult for large coin databases. This is mainly caused by the specific challenges that ancient coins pose to a classification method based on 2D images. In this paper we highlight these challenges and argue to use SIFT flow image matching. Our classification system is applied to an image database containing 24 classes of early Roman Republican coinage and achieves a classification rate of 74% on the coins' reverse side. This is a significant improvement over an earlier proposed coin matching method based on interest point matching which only achieves 33% on the same dataset.

Categories and Subject Descriptors (according to ACM CCS): I.4.9 [Image Processing and Computer Vision]: Applications—

1. Introduction

Determining the type of an ancient coin needs a lot of numismatic experience and is in general a time consuming task, especially when it comes to the classification of huge coin hoards [Gri75]. In this paper we address the problem of automatic coin classification from 2D images. Such a methodology has the potential to act as a supporting tool for numismatists that enables a much faster processing of coins. In the long run, an automatic image based coin classification system could be of use for the whole numismatic community (e.g. researchers at museums as well as private collectors), allowing for instance the automatic classification of self-uploaded coin images in an online webtool.

From a computer vision point of view, ancient coins prove to be a challenging type of data for image classification due to their high level of degradation and variability. As ancient coins are textureless objects with a 3D relief, the appearance of a coin in a 2D image is strongly influenced by illumination conditions. The loss of 3D information and the lighting variations that appear in 2D coin imaging impede coin classification as the algorithms have to be made preferably insensitive to the resulting effects. Obviously, acquisition techniques which are able to capture the 3D characteristics of the coin like active stereo coin scanning [ZSHK10] or Polynomial Texture Mapping of coins [MVSL05] provide more suitable data for the classification task. However, 3D acqui-

sitions are more laborious and expensive than traditional 2D imaging and comprehensive 3D databases of ancient coins are nowadays not available. Therefore, in this work only 2D image data is addressed.

Conventionally, image classification is achieved by a supervised machine learning procedure [Bis07]: a set of features is extracted from the images and passed to previously trained classifiers that map the features to the defined classes. Following this approach, powerful machine learning techniques can be exploited that allow a robust classification for high-dimensional complex feature spaces. However, the performance of these methods depends on the discriminative power of the extracted features. Finding and extracting such features with high discriminative power is difficult for ancient coins because of their high level of degradation and variability. Another drawback of using classifiers concerns the expandability of the database: classifiers have to be retrained every time new coin classes are included in the database which impedes database extensions.

For these reasons, we propose a classifier-free procedure for coin classification based on image matching. The idea is to measure the similarity between a coin image and all coin images in the training set (the database) and to finally choose the class with the highest image similarity. The measurement for image similarity is thereby derived from a SIFT flow based image matching [LYT11]. SIFT flow aligns im-

ages by minimizing an energy function defined over a dense grid of SIFT features [Low04]. We argue that this method is well suited for coin images as it allows a spatially coherent matching with local variations that is robust to image clutter. As coins from the same class show a similar spatial arrangement of local features, matching of these images is assumed to produce a lower energy than matching images from two different classes.

The remainder of this paper is structured as follows: an overview of related work in coin classification is given in Section 2. In Section 3 the specific challenges of image based ancient coin classification are outlined in detail. A description of the proposed coin classification methodology based on SIFT flow is given in Section 4. Experiments on a dataset of Roman Republican coins are reported in Section 5. Conclusions are finally drawn in Section 6.

2. Related Work

Related work in image based classification of ancient coins is scarce so far. Recent research approaches for coin classification focused mainly on the classification of present day coins [HRM*05, NPR*03, RRB06, vdMP06a]. However, the differences between present day and ancient coins exposed to be too large to effectively apply such methods to ancient coins [vdMP06b]. Due to the abrasions over the centuries and the non-industrial manufacturing, ancient coins naturally exhibit a larger variation within a class. This impedes the coin classification problem to a large degree since a classification method has to determine the features common to class while ignoring the coin-specific ones. On the other side, this fact favors the automatic identification of ancient coins. For instance, the shape of the coin border has shown to be a very discriminative feature to identify individual coins [HMZZK10].

First promising results on the classification of ancient coins were presented in [KZ08]. Similar to our work in their paper a classifier-free approach based on SIFT matching is used for coin classification. Image similarity is measured by counting the number of matched SIFT features between two images. Compared to our image similarity based on SIFT flow, this methodology provides a much simpler and thus less robust measure of coin similarity, as only a sparse set of interest points is matched without considering their spatial relationships (this is also demonstrated in our experiments in Section 5). The authors report a classification rate of $\sim 90\%$, however achieved on a small dataset containing only three coin classes. Recently, a promising method based on feature classification has been proposed by Arandjelović [Ara10]. In his work the so-called *locally-biased directional kernel feature* (LBD) is proposed as a feature that encodes the spatial context of local interest points. For coin classification, frequencies of a quantized set of LBD features are computed and classified using a previously trained Gaussian mixture model. A classification rate of $\sim 57\%$ on a testset containing



Figure 1: Reverse side of different coin specimens from the Roman period. Coins in the same row denote coins of the same class.

65 classes is reported which shows that spatial-context features like LBD are generally promising as they offer a powerful representation able to capture the class-specific coin appearance. However, in this work, we aim at a classifier-free approach for coin classification that can be easily extended to large databases.

3. Specific Challenges of Image Based Ancient Coin Classification

For an image based coin classification method, substantial differences between present day and ancient coins have to be taken into account. All of these differences make the classification of ancient coins a harder problem. The main difference between present day and ancient coins lies in the much higher intra-class variability that comes both from the age as well as the non-industrial manufacturing of the coins. On the other hand, inter-class variability can be comparatively low. This is exemplified in Figure 1 where two classes of Roman silver denarii are shown, each class represented by three reverse coin images in the upper and lower row, respectively. Please note that both coins show the same motive, a naked, bearded warrior in a chariot, indicating the low inter-class variability for these two classes. The only obvious difference between the two classes is the inscription *SCAVRI* under the chariot. The possible degree of change due to abrasions within a class (i.e. the high intra-class variability) can be spotted in Figure 1(b) and (c). For instance, the chariot's wheel is completely or partly missing on these two specimens.

High intra-class variability can also result from the manual manufacturing of the coins. In ancient times, coin dies were made manually and used only for a limited amount of coins. Therefore, die variations and thus finally coin variations as the one shown in Figure 2 can occur. Here a detail of two Roman denarii depicting the goddess Roma are shown. The green arrows illustrate that facial features have a differ-

ent arrangement due to the different dies used for striking these coin specimens.



Figure 2: The obverse side of two coins from the same class. Due to different dies used for striking, the spatial relationship of facial features is different on the two coins.

Another problem are lighting variations that significantly affect the surface appearance and hence make coin comparison difficult. Ancient coins were made of metal and have a higher relief than present day coins, which can lead to strong shading and reflection variations under different lighting conditions. An example for this kind of variation is given in Figure 3, where two images of the same coin specimen under different lighting directions are shown. A detailed look at the man's upper body reveals how the local appearance is affected by the lighting direction.

Please note that the challenge of lighting variations is neglected in the presented SIFT flow based method. SIFT descriptors are generally not invariant to lighting variations. However, the underlying descriptor of SIFT flow can be easily changed from SIFT to a more suitable descriptor. We leave this open for future research.

4. Coin Classification Methodology

SIFT flow was recently introduced by Liu et al. [LYT11] as a dense image matching method that finds nearest neighbors for each pixel in the presence of large image variations. Despite SIFT flow being able to maintain spatial discontinuities in the matching, it includes a smoothness constraint in order to prefer smooth alignment results. These properties motivate its use for ancient coin classification as it can cope with the high intra-class variability arising from abrasions and manual manufacturing. Thus, it is insensitive to local spatial variations as those shown in Figure 2. Missing parts do not affect the matching as long as the major part of the coin can be matched correctly.

4.1. SIFT Flow Image Matching

SIFT flow is based on the SIFT (Scale Invariant Feature Transform) descriptor [Low04]. This descriptor is computed for every pixel in an image by first dividing each pixel's local



Figure 3: Two images of the same coin, illuminated from different directions. A high variation of specular reflections can be seen due to the surface characteristics of the coin.

neighborhood into a 4×4 cell array. The gradient orientations of each cell are quantized into 8 bins, generating a 128-dimensional vector for every pixel, the so called *SIFT image* \mathbf{S} . To find an image matching, corresponding SIFT features between two SIFT images \mathbf{S}_1 and \mathbf{S}_2 have to be determined for each pixel location, represented as a field of flow vectors $\mathbf{w}(\mathbf{p}) = (u(\mathbf{p}), v(\mathbf{p}))$ at grid coordinates $\mathbf{p} = (x, y)$. This is achieved by minimizing the following energy function on \mathbf{w} :

$$E(\mathbf{w}) = \sum_{\mathbf{p}} \min(|\mathbf{S}_1(\mathbf{p}) - \mathbf{S}_2(\mathbf{p} + \mathbf{w}(\mathbf{p}))|, t) \quad (1)$$

$$+ \sum_{\mathbf{p}} \gamma(|u(\mathbf{p})| + |v(\mathbf{p})|) \quad (2)$$

$$+ \sum_{(\mathbf{p}, \mathbf{q}) \in \epsilon} \min(\alpha|u(\mathbf{p}) - u(\mathbf{q})|, d) + \min(\alpha|v(\mathbf{p}) - v(\mathbf{q})|, d) \quad (3)$$

where ϵ contains all four-connected pixel pairs. The energy function is composed of three terms, the *data term* (1), *small displacement term* (2) and *smoothness term* (3). These three terms force the algorithm to produce smooth alignments (i.e. flow vectors of adjacent pixels are similar) with preferably small flow vectors. t, γ, α and d are parameters that control the influence of the various constraints.

In order to minimize the energy function and to obtain an optimal image matching, a dual-layer belief propa-

gation [SKH08] is used. Additionally, a coarse-to-fine matching scheme is applied for speed-up and better matching results. A detailed explanation of the energy minimization is given in [LYT11].

4.2. Coin Classification Using SIFT Flow Image Matching

The adoption of the SIFT flow algorithm for coin classification relies on the following idea: matching two coin images of the same class will likely produce a lower energy $E(\mathbf{w})$ than matching coin images from different classes, since a smooth matching can be more likely found in the former case. An example for this is shown in Figure 4. Matching the test coin image with a coin image from the same class produces a reasonable result, as can be seen in Figure 4(c), where the result of warping the image back to the test image using the SIFT flow vectors is shown. In contrast, matching the test coin image with a coin image of a different class produces an unsuitable result and thus a higher energy. As a consequence, we achieve coin classification by matching the coin image with all coin images in the database and finally choosing the class of the image with lowest energy.

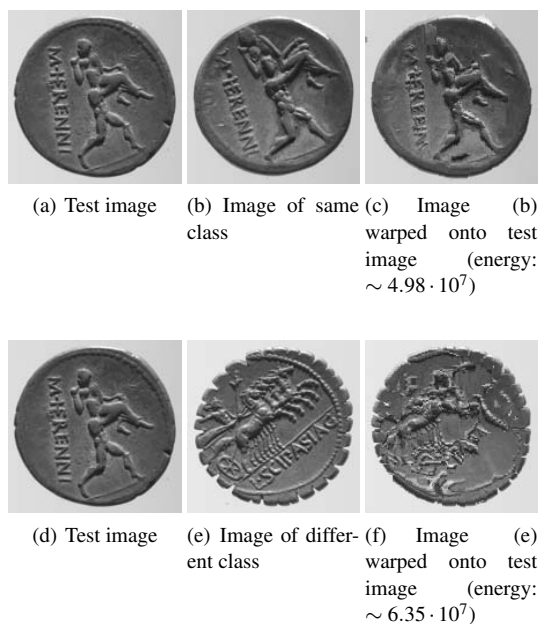


Figure 4: SIFT flow applied to images of the same class (top) and images of different classes (bottom).

In detail, the following steps are conducted in our coin classification system (outlined also in Figure 5):

1. **Coin Segmentation:** In order to remove background clutter such as rulers in the image, the coins have to be

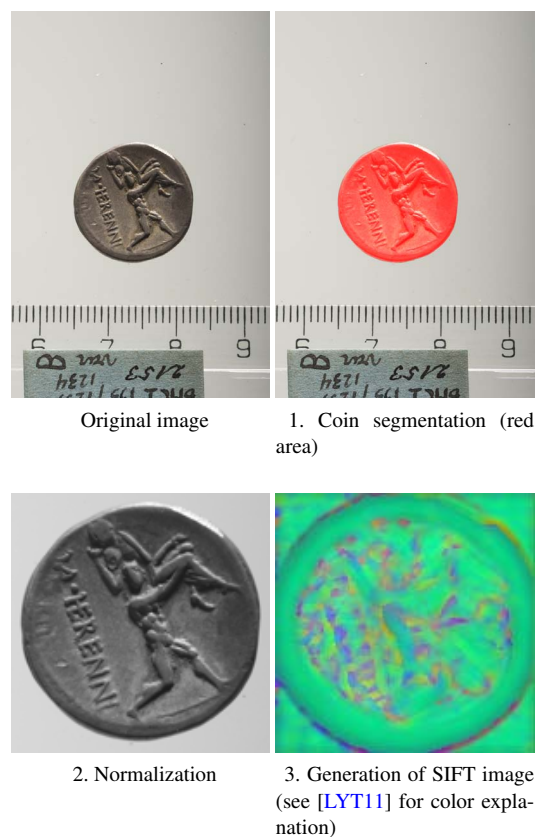


Figure 5: The first three steps of the presented coin classification system.

segmented first. For this task we use a shape-adaptive thresholding method [ZK09] that has proven to provide a robust segmentation for a variety of coin images. It uses a range and entropy filter under the assumption that the coin in the image provides more information and details than the background. The final segmentation mask is obtained by a thresholding operation where the optimal threshold is found by means of the formfactor [Rus06] of the resultant binary mask.

The output of the segmentation is a binary mask identifying the pixels belonging to the coin. Therefore, we achieve translational invariance of the classification method by the coin segmentation step.

2. **Normalization:** To additionally achieve scale invariance, the segmented coin image is normalized to a standard size of 150×150 . This is needed as we use a dense grid of SIFT features with fixed scale (i.e. fixed window size of the SIFT descriptor). Additionally, the image is converted to grayscale for SIFT descriptor computation.
3. **Generation of SIFT image:** SIFT descriptors are computed at every pixel location to produce the SIFT image.

Empirical tests have shown that a window size of 12×12 (i.e. cell size of 3×3) works best to capture the local coin details.

4. **SIFT flow matching:** Finally, SIFT images are matched to assess the visual similarity of the coins. As the coins do not necessarily have the same orientation in the images, the matching has to be done in a rotation-invariant way. Therefore, we ignore the small displacement term (2) for our application in order to allow large displacements between coin images. The SIFT features themselves are rotation-invariant, as gradient orientations are represented relative to the dominant orientation.

5. Experiments

For the experiments Roman Republican coins are used which have been kindly provided by the *Coin Cabinet of the Museum of Fine Arts, Vienna*. Coins from this era have been chosen because a big stock of material is available at the museum. There is a comprehensive standard-reference work by Crawford [Cra74] that defines the various types (classes) of coins. Crawford's work has 550 distinct reference numbers, however, currently only a small amount of coins from 131-102 B.C. is available as 2D images. Therefore, experiments are conducted on a set of 24 classes with three images each. The obverse and reverse side of a specimen from each class are shown in Figure 6.

As we have three images of each coin class, we use a 3-fold cross validation to test classification performance. The dataset is divided into three subsets, each set containing one image from each class. Three classification runs are executed whereas in each run one subset serves as testset and the remaining two serve as training set. An image from the testset is then matched with all images from the training set. The sum of SIFT flow energies of the two images of a class defines the class-energy, and thus finally the image is classified as the class with minimum class-energy. We test the single performance on the obverse side as well as on the reverse side. Obviously, information from both sides could be combined to improve classification performance. We leave this open for future research.

To compare our method to a previously proposed classifier-free method for coin classification, we furthermore apply the standard SIFT matching method proposed in [KZ08] to the database. In this method, similarity between coins is measured by the number of matched interest points, extracted at Difference-of-Gaussians extrema [Low04] and described by SIFT. The same evaluation procedure as for the SIFT flow matching is applied.

The overall results are listed in Table 1. Our proposed SIFT flow method clearly outperforms the method of [KZ08]. As the SIFT descriptor itself is potentially noisy on ancient coins due to the lighting problem, the simple matching of SIFT interest points is more vulnerable than SIFT flow

matching. SIFT flow introduces an additional constraint for a spatially meaningful matching with local variations. Therefore, SIFT flow provides a more robust matching and coin similarity measure.

Another observation from the results is that classification rates are higher on the reverse sides of the coin. This is caused by the typical composition of Roman coins from the investigated period. Customarily, obverse sides show the heads of gods or emperors. In the given evaluation dataset the obverse side of 15 of the 24 classes depict the goddess Roma (see Figure 6). Reverse sides depict certain scenes and thus have a higher inter-class variability. However, the given dataset contains coins from an early stage of Roman Republican coinage and therefore this difference is less pronounced since 16 of the 24 classes show chariots (a common motive [Cra74]).

	SIFT Flow Matching	SIFT Matching
Obverse side	63.9%	25.0%
Reverse side	73.6%	33.3%
Total	68.8%	29.2%

Table 1: Classification rates of the proposed SIFT flow matching and standard SIFT matching [KZ08] on obverse and reverse coin sides.

In Figure 7 some misclassifications are presented to note the limitations of the presented method in its current form. One limitation is the unsatisfactory handling of lighting variations due to the use of the SIFT descriptor. This is evidently shown in Figure 7(a): the test image and training image of class 314/1 have a different orientation and relief height, thus different lighting effects appear on their surface (please note the stronger and larger shadows in the training image). The second main source of error is the low variability between two coin classes. For instance, the classes 282/1 and 282/4 both show chariots and differ from each other only by the inscription. In the presence of strong abrasions on the coin as shown in the training image of class 282/1 (Figure 7(b)), the SIFT flow method prefers a matching with coins of similar condition, in this case of class 282/4. A similar situation can be seen in Figure 7(c) where the test image shows strong abrasions. In this case large image regions are not matchable by the SIFT flow method and finally a misclassification is produced.

6. Conclusions

In this paper we presented a classifier-free method for ancient coin classification on 2D images. The method is based on image matching and therefore avoids the need for a feature extraction and classification step, which is difficult to achieve for ancient coin data. We use the flexible SIFT flow method which is well qualified for ancient coins as it takes account of the spatial variations that arise from the manual

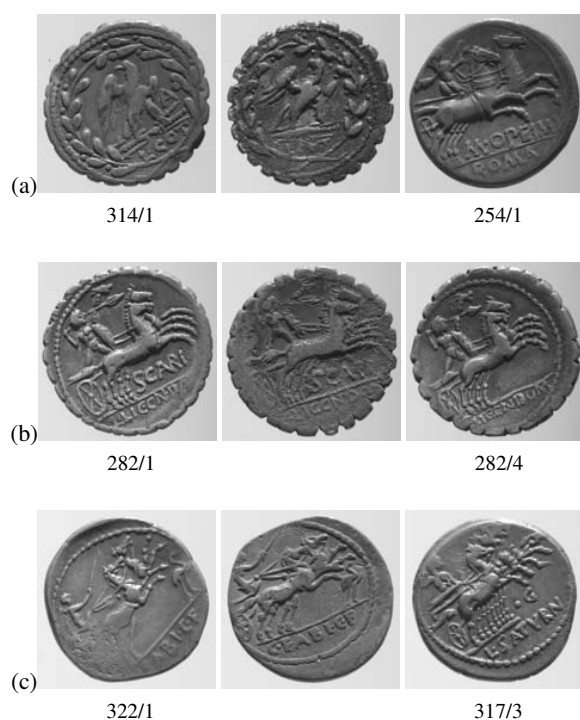


Figure 7: Three selected misclassifications of the presented SIFT flow method, from left to right: test image, coin image of correct class, coin image of class selected by proposed method.

manufacturing of the coins. The experimental results show the potential of the method for ancient coin classification with a classification rate of 73.6% on the reverse sides of 24 coin classes. These results are achieved on a database including all the main challenges of ancient coin classification, that are lighting variations, abrasions and high intra-class variabilities in combination with low inter-class variabilities. Nevertheless, we plan to validate the results on a larger database in the future. This can be easily achieved as the method does not rely on machine learning methods and thus needs no offline training phase.

One drawback of the method is the long computational time for SIFT flow. Currently SIFT flow matching of two images takes about 5-6 seconds on a standard machine and thus matching a test image with all images in the database is a time-consuming task. One possible solution to decrease computational time is to use a hierarchical coarse-to-fine classification scheme as already used within SIFT flow: on a coarse level the test image is matched with all images in the database, but on a finer level only the most similar coins are matched further. Such a preselection could also be achieved by using low-level global image features that are fast to com-

pute but are informative enough to initially reject a certain amount of classes.

In general, we see such a hybrid system that combines feature-based classification and image matching as a promising research direction. This way, the best of both worlds can be achieved: feature-based classification like in [Ara10] enables a fast classification once the classifier has been trained, but the feature representation might lose its discriminative power when hundreds of coin classes are contained in the database. In contrast, coin classification based on image matching like the one presented in this paper does not transform a coin image into a less informative feature representation and is therefore able to provide an in-depth analysis of coin similarity. However, this ability comes at the cost of high computational effort.

The main research challenge for the future will be the investigation of image features that are invariant to the lighting variations discussed in Section 3. We assume that more appropriate image representations or features will provide a significant step towards a more powerful coin classification method.

Acknowledgements

This research has been supported by the Austrian Science Fund (FWF) under the grant TRP140-N23-2010 (ILAC). The authors would like to thank Dr. Klaus Vondrovec and Kathrin Siegl from the Museum of Fine Arts, Vienna, for expert assistance and for providing the test images.

References

- [Ara10] ARANDJELOVIC O.: Automatic attribution of ancient roman imperial coins. In *Conference on Computer Vision and Pattern Recognition* (2010), pp. 1728–1734.
- [Bis07] BISHOP C. M.: *Pattern Recognition and Machine Learning*, 1 ed. Springer, 2007.
- [Cra74] CRAWFORD M. H.: *Roman Republican Coinage, 2 vols.* Cambridge University Press, 1974.
- [Gri75] GRIERSON P.: *Numismatics.* Oxford University Press, 1975.
- [HMZZK10] HUBER-MÖRK R., ZAMBANINI S., ZAHARIEVA M., KAMPEL M.: Identification of ancient coins based on fusion of shape and local features. *Machine Vision and Applications* (2010), 1–12.
- [HRM*05] HUBER R., RAMOSER H., MAYER K., PENZ H., RUBIK M.: Classification of coins using an eigenspace approach. *Pattern Recognition Letters* 26, 1 (2005), 61–75.
- [KZ08] KAMPEL M., ZAHARIEVA M.: Recognizing ancient coins based on local features. In *Proc. of International Symposium on Visual Computing* (2008), vol. I of *Lecture Notes in Computer Science*, pp. 11–22.
- [Lowe04] LOWE D. G.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60, 2 (2004), 91–110.

- [LYT11] LIU C., YUEN J., TORRALBA A.: Sift flow: Dense correspondence across scenes and its applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33, 5 (2011), 978–994.
- [MVSL05] MUDGE M., VOUTAZ J.-P., SCHROER C., LUM M.: Reflection transformation imaging and virtual representations of coins from the hospice of the grand st. bernard. In *Proc. of VAST* (2005), pp. 29–39.
- [NPR*03] NÖLLE M., PENZ H., RUBIK M., MAYER K. J., HOLLÄNDER I., GRANEC R.: Dagobert – a new coin recognition and sorting system. In *Proc. of DICTA'03* (2003), pp. 329–338.
- [RRB06] REISERT M., RONNEBERGER O., BURKHARDT H.: An efficient gradient based registration technique for coin recognition. In *Proc. of the Muscle CIS Coin Competition* (2006), pp. 19–31.
- [Rus06] RUSS J. C.: *The Image Processing Handbook*, 5th ed. CRC Press, 2006.
- [SKH08] SHEKHOVTSOV A., KOVTUN I., HLAVÁČ V.: Efficient mrf deformation model for non-rigid image matching. *Computer Vision and Image Understanding* 112, 1 (2008), 91–99.
- [vdMP06a] VAN DER MAATEN L. J., POON P.: Coin-o-matic: A fast system for reliable coin classification. In *Proc. of the Muscle CIS Coin Competition* (2006), pp. 07–18.
- [vdMP06b] VAN DER MAATEN L. J., POSTMA E. O.: Towards automatic coin classification. In *Proc. of the EVA-Vienna 2006, Vienna, Austria* (2006), pp. 19–26.
- [ZK09] ZAMBANINI S., KAMPEL M.: Robust automatic segmentation of ancient coins. In *4th International Conference on Computer Vision Theory and Applications (VISAPP'09)* (2009), vol. 2, pp. 273–276.
- [ZSHK10] ZAMBANINI S., SCHLAPKE M., HÖDLMOSE M., KAMPEL M.: 3d acquisition of historical coins and its application area in numismatics. In *Proc. SPIE* (2010), vol. 7531, p. 753108.



Figure 6: Obverse (left) and reverse side (right) of the 24 coin classes used for evaluation, listed along with Crawford reference number.