

Improving Ancient Roman Coin Classification by Fusing Exemplar-Based Classification and Legend Recognition

Sebastian Zambanini, Albert Kavelar, and Martin Kampel

Computer Vision Lab, Vienna University of Technology, Austria
{zamba,kavelar,kampel}@caa.tuwien.ac.at

Abstract. In this paper we present an image-based classification method for ancient Roman Republican coins that uses multiple sources of information. Exemplar-based classification, which estimates the coins' visual similarity by means of a dense correspondence field, and lexicon-based legend recognition are unified to a common classification approach. Classification scores from both coin sides are further integrated to an overall score determining the final classification decision. Experiments carried out on a dataset of 60 different classes comprising 464 coin images show that the combination of methods leads to higher classification rate than using them separately.

1 Introduction

Numismatics deals with all forms of money and other payment media and its contemporary and historic aspects [9]. Along with studying use and production of historic money, cataloging or classifying coins is one of its major aspects. Coin hoards possibly comprise thousands of coins and therefore the classification of each individual coin is a time-consuming and cumbersome task, which up until now has to be carried out manually [9]. This is where numismatists could largely benefit from a fully-automated coin classification method. When the coins are registered in digital archives, their front (*obverse*) and back sides (*reverse*) are photographed using a digital camera and stored in a database along with textual descriptions. Hence, we suggest an image-based classification method since it does not require additional digitization steps which would cause extra work and costs and therefore would reduce the chances of such a system to be accepted by numismatists.

Several image-based coin classification systems for modern coins were developed in the past (e.g., Dagobert [16] and Coin-O-Matic [15]). However, the development of a fully-automated image-based coin classification system for ancient coins is still subject of ongoing research. One of the major reasons for this is that, as opposed to their modern counterparts, ancient coins show little inter-class variability and large intra-class variability (see Fig. 1) and thus pose different challenges to computer vision. The intra-class variability is related to the manufacturing process of ancient coins. They were cast or struck using manually engraved dies. The longer a die was used, the more worn down it became



Fig. 1. Examples of Roman Republican coins. (a) Low inter class-variability: three coins of different classes that show the same basic motive and only differ in the depicted legend (highlighted in red). (b) High intra-class variability: three coins of the same class.

and consequently the quality of the struck coins decreased over time. Hence, a die could only be used for a limited number of coins before they had to be replaced by a new one, which introduced additional variations. Moreover, the dies were not always centered and the relative alignment of obverse and reverse die on the *flan* (the raw piece of metal the coin is made of) shows inconsistencies.

In the past, promising approaches for recognizing ancient Roman coins including image matching [20] and the recognition of the coin legends [12] have been proposed. The former measures the visual similarity of coin images by the energy needed to establish a dense correspondence between them. This approach has the advantage that it is more independent on the number of images available from the same class, as no offline machine learning is performed. However, it also suffers from a low inter-class variability as different coin classes possibly depict the same central motive (e.g., a chariot or a god portrait) [5]. In such hard cases only minor sources of information like the legend can be exploited for differentiation [5].

In this paper, we present a way of fusing these methods which helps to improve the robustness of image-based coin classification. Legend recognition is based on a vocabulary which maps the words to coin classes and thus adds a different kind of background knowledge to the pure visual similarity of coins estimated from the dense correspondence. We also increase the level of robustness by exploiting both coin sides for classification.

The remainder of the paper is organized as follows: Section 2 reviews the state of the art in coin classification. Section 3 explains the proposed image matching and legend recognition algorithms and how they are combined. In Section 4 the experimental setup and test results are presented. Finally, Section 5 concludes this work.

2 Related Work

Fukumi et al. [8] were the first to use computer vision methods for coin classification. Their system can distinguish 500 yen from 500 won coins. After separating the coin from the background, it is divided into several ring segments by a log-polar grid; this separation ensures rotational invariance. For each segment, the sum of gray levels is computed and fed to a multilayered neural network. Nölle et al. [16] presented a powerful coin recognition system called Dagobert, which discerns 600 different coin types. Their method computes a binarized edge image from the coin input image which is compared to master images stored in a database. The coin is assigned to the class which gives the closest match using nearest neighbor classification. To reduce computation time, a pre-selection based on the measurements of two additional sensors, which capture the coin's diameter and thickness, is carried out beforehand. Bremananth et al. [4] proposed a method for the recognition of 1-, 2- and 5-rupee coins. After successful edge detection, the location of the numerals depicted on the coin is found via template matching. Next, Gabor features are computed for the subimage containing the numeral, which are then classified by a back propagation network. The COIN-O-MATIC system suggested by van der Maaten and Poon [15] relies on edge information, similar to the Dagobert approach. After computing an edge image from a coin input image, a histogram based on a log-polar grid, which captures the angle and distance distribution of the edge pixels, is calculated and classified using nearest neighbour classification. Finally, Reisert et al. [17] introduced another classification system for modern coins. Their approach segments the coin from the background using a generalized Hough transform [3]. Next, the segmented image is transformed to polar coordinates and gradients are computed. The gradients are quantized into different orientational bins and for each of which a binary image is constructed. The binary images serve as feature vectors and are classified using a nearest neighbor classification.

Kampel and Zaharieva were among the first to present an end-to-end recognition workflow for ancient coins [11]. They experimented with various interest point detectors and local image descriptors in order to determine the best combination. Recognition is performed by finding the nearest neighbour using the Euclidean distance. While Kampel and Zaharieva only considered three different coin types in their experiments, Arandjelovic [1] works with 65 classes. He introduced a new type of feature called *locally-biased directional histogram*, which captures geometric relationships between interest points found via the Difference-of-Gaussian detector [14]. Huber-Mörk et al. [10] extended the work of Kampel and Zaharieva [11] to coin identification by introducing an additional

preselection step which exploits the coin’s contour as a characteristic feature for single coin specimens. The contour is analyzed by intersecting rays cast from the coin’s center of gravity with the coin border. The distances between the intersection points and a hypothetical perfect circle fit to the coin contour are measured and used to build a descriptor, which can be computed quickly and allows for an efficient preselection. Thus, the introduction of this step allows fast pruning and increases the recognition rate while the computation time of the coin recognition algorithm can be reduced. Arandjelovic recently introduced another coin classification method which exploits coin legends [2]. He focuses on Roman Imperial denarii, which have legends running along the coin border. This *a priori* assumption about the coin layout allows to align the legends horizontally by transforming the coin images to polar coordinates. The individual letters are then detected by a sliding window for which HoG-like features [6] are computed. This results in a likelihood for each letter and sliding window location, which is then combined to legend words using dynamic programming. The recognized legend is used to select a subset of matching coin types for which SIFT-based matching is performed in order to determine the final class.

Like in the work of Arandjelovic [2], we combine legend recognition with image matching for an improved classification. However, we are focusing on Roman Republican coins instead of Roman Imperial coins. Thus, legends in our case can be arbitrarily located on the coin and are not restricted to be circularly placed at the coin border [5]. Furthermore, instead of doing legend recognition on one side and image matching on the other, we apply both methods to both sides to achieve a more reliable and robust classification.

3 Methodology

The method we propose combines image matching with legend recognition in order to improve the classification rates the methods achieve separately. The general architecture of the fused methodology is depicted in Fig. 2. It takes two input parameters: an image showing the obverse of a coin and one of the corresponding reverse. For each of these images, image matching and legend recognition compute their separate probabilities $P_{s,c}^M$ and $P_{s,c}^L$ for each coin side $s \in \{1, 2\}$ and class $c \in C$. These values are then combined to an overall probability for each class, which finally allows the determination of the coin class to which the input image belongs.

3.1 Coin Classification by Image Matching

Coin classification by image matching [20] has the advantage that there is no need for a large amount of training data as no machine learning techniques with an offline straining stage are involved. Instead, classification is done by means of reference images per class whose visual similarity to an input image is estimated. This methodology has the benefit that we can cope better with the partially limited amount of available training data for ancient Roman coins [20].

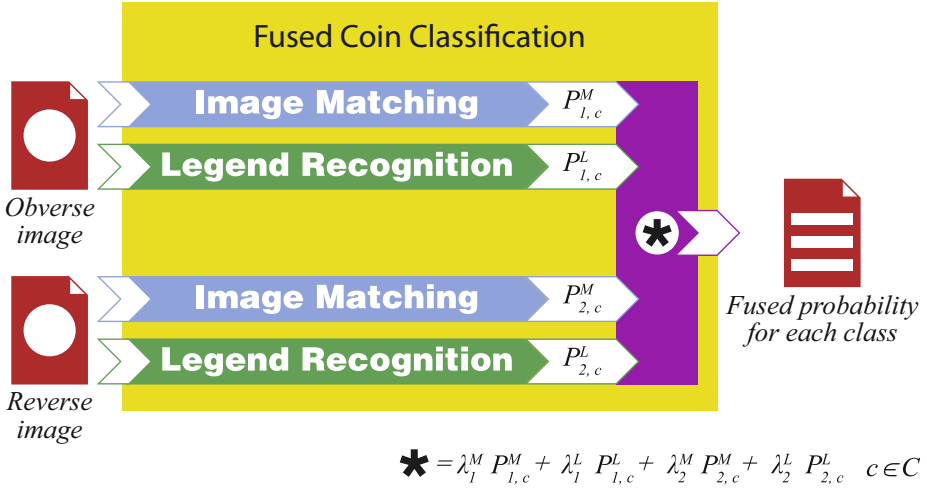


Fig. 2. The fused coin classification workflow

The challenge of high intra-class variance is thereby overcome by using a flexible model for dense correspondence that is able to handle the spatial variations of local structures between coins of the same class and is robust to image clutter.

Similar to the SIFT flow method proposed in [13], we use a dense field of SIFT features [14] to compute pixel-to-pixel correspondences between two images. Each SIFT feature provides a rotation-invariant description of its local neighborhood by means of gradient orientation distributions and thus the costs of matching two pixels can be computed by their Euclidean distance. This can be described in form of an energy term $E(\mathbf{w})$ with an additional regularization term to favor smoothness of the result:

$$E(\mathbf{w}) = \sum_{\mathbf{p}} \|S_1(\mathbf{p}) - S_2(\mathbf{p} + \mathbf{w}(\mathbf{p}))\|_2 \tag{1}$$

$$+ \sum_{(\mathbf{p}, \mathbf{q}) \in \Omega} \min(\alpha|u(\mathbf{p}) - u(\mathbf{q})|, d) + \min(\alpha|v(\mathbf{p}) - v(\mathbf{q})|, d) \tag{2}$$

where S_1 and S_2 are dense fields of SIFT features computed at every image point \mathbf{p} of the two images to be compared, $\mathbf{w}(\mathbf{p}) = (u(\mathbf{p}), v(\mathbf{p}))$ describes the correspondences, Ω contains all four-connected pixel pairs and α and d control the influence of the regularization term.

As in the SIFT flow method [13], a dual-layer belief propagation [18] is used along with a coarse-to-fine matching scheme to minimize $E(\mathbf{w})$ and determine the optimal image matching. The resulting energy value is taken as measure of dissimilarity between the two coin images. Therefore, for a given input image class scores $\Psi_{s,c}^M$ are determined by comparing it to single reference images of all classes in the database.

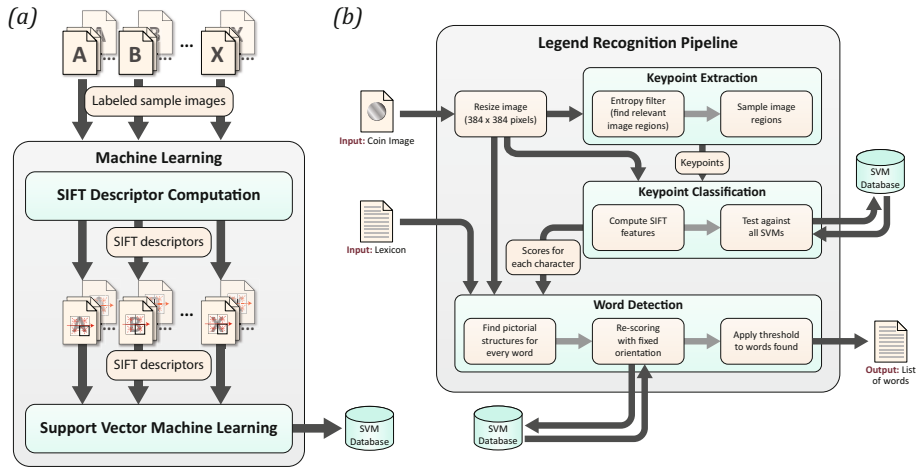


Fig. 3. Overview of the legend recognition methodology [12]. (a) Training of SVMs. (b) Legend recognition pipeline.

3.2 Coin Classification by Legend Recognition

The overall architecture of the legend recognition [12] method is depicted in Fig. 3. The method is classification-based and therefore requires an initial training step before the actual recognition can be performed.

Training. In order to teach the appearance of the individual legend letters to our system, we again use SIFT features [14] for letter representation (see Fig. 3(a)). The training set comprises 50 100×100 pixel-sized images per character class. The legends considered can be represented by 18 different capital letters. Since structures resembling the letter 'I' often occur in the depicted imagery, this letter is not trained. Due to high variations in the illumination conditions and heavy degradations related to wear and chemicals in the soil, some letters are barely visible and the respective edges in the image cannot be used for robust keypoint detection. Thus, only one descriptor which is centered and scaled according to the letter size is computed and used in the Support Vector Machine (SVM) training. Because both rotationally invariant and fixed SIFT descriptors are required in the recognition step, two different databases are created.

Legend Recognition. The legend recognition step is depicted in Fig. 3(b). Its input parameters are a coin image and a lexicon containing all legend words the image should be scanned for. Since the legend recognition results are more reliable for longer words, we do not consider words having less than five letters. After an initial downscaling of the image to a standard size of 384×384 pixel, the image is passed to the *Keypoint Detection* step where flat image regions are identified using an entropy filter and ignored for further processing. Next, a densely sampled grid is constructed in the remaining image regions which represents the *candidate character locations* (CCL) [19]. These locations serve as

input for the *Keypoint Classification* step, in which a SIFT descriptor is computed for each of the CCLs. Every descriptor is tested against the 18 SVMs trained for rotationally invariant SIFT features and yields a score telling how likely it is to encounter a letter of the respective class at this location. Thus, every CCL is now associated with 18 scores. In the *Word Detection* step, the optimal configuration for every word of the input lexicon is found within the image using pictorial structures [7]. Finally, all words found are rescored. Since the individual character locations are known for a specific word hypothesis, the respective character orientations are known as well and can be used to compute SIFT descriptors aligned accordingly. This increases the confidence in the hypothesis since orientationally fixed SIFT descriptors lead to a better character recognition rate [12]. In the end, the output of the legend recognition pipeline is a list of words ordered by their scores. The respective coin classes are retrieved via a lookup table which maps the legend words to coin types, resulting in the class scores $\Psi_{s,c}^L$.

3.3 Combination of Image Matching and Legend Recognition

Both image matching and legend recognition provide class scores $\Psi_{s,c}^M$ and $\Psi_{s,c}^L$, respectively, which indicate the dissimilarity of the input coin's side $s \in \{1, 2\}$ to class $c \in C$. In order to combine the individual class scores, we first transform them to the same range. More specifically, the class scores are transformed to pseudo probabilities $P_{s,c}^M$ and $P_{s,c}^L$ in such a way that the class score with lowest dissimilarity is mapped to the highest probability and $\sum_{c \in C} P_{s,c} = 1$. Furthermore, we define a parameter δ that describes the factor between the class with highest probability and the class with lowest probability, i.e., the class with lowest class dissimilarity has a probability δ times higher than the class with highest dissimilarity:

$$P_{s,c} = \frac{1}{\tilde{\Psi}_{s,c} \cdot \sum_{c \in C} 1/\tilde{\Psi}_{s,c}} \quad (3)$$

where

$$\tilde{\Psi}_{s,c} = \frac{\Psi_{s,c} - \min_{c \in C}(\Psi_{s,c})}{\max_{c \in C}(\Psi_{s,c}) - \min_{c \in C}(\Psi_{s,c})}(\delta - 1) + 1. \quad (4)$$

For the final class decision we apply a weighted sum rule, since it allows for a flexible combination which accounts for the unequal recognition accuracies of the individual votes. Thus, a coin is finally classified to class c^* with

$$c^* = \arg \max_{c \in C} (\lambda_1^M P_{1,c}^M + \lambda_2^M P_{2,c}^M + \lambda_1^L P_{1,c}^L + \lambda_2^L P_{2,c}^L) \quad (5)$$

4 Experiments

The experiments were conducted on images of 232 coins from 60 classes. All coins are from the Roman Republican age and have been provided by the *Museum of Fine Arts* in *Vienna*. Class membership is defined by Crawford's standard

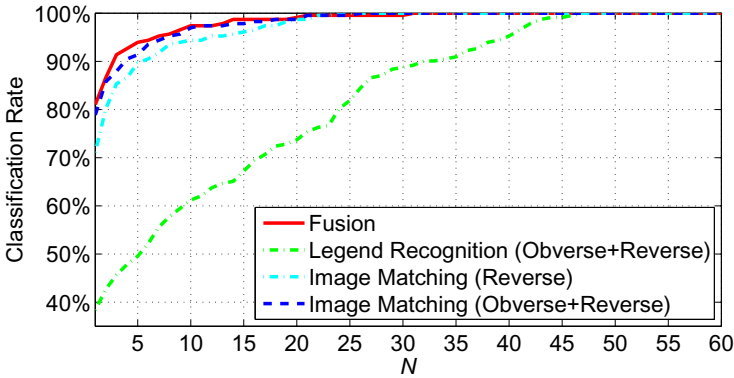


Fig. 4. Classification accuracy of legend recognition, image matching and our proposed fusion method for the top N classes

reference book [5]. For each coin, an image of obverse and reverse is present in the dataset; thus, an overall of 464 images was used. The lexicon of obverse and reverse words contains 21 and 41 words, respectively, since some classes have words in common and not all words present on the coins can be used: legend recognition is not robust enough for words with less than five letters because falsely classified letters have a stronger influence on the overall score for shorter words. The parameters for classifier combination have been empirically chosen as $\delta = 10$, $\lambda_1^M = \lambda_2^M = 0.4$ and $\lambda_1^L = \lambda_2^L = 0.3$. The weights of the legend recognition are lower due to the general worse performance of legend recognition compared to image matching.

In Fig. 4 the results for legend recognition, image matching and our proposed fusion method are compared. The plot shows the respective percentage of coins where the correct class is within the top N probabilities. First of all, it can be seen that combining both coin sides supports the classification, as for image matching only the classification rate is generally higher compared to the case where only the reverse side is considered (78.9% vs. 71.1% for $N = 1$). Legend recognition is more error-prone than image matching due to the challenging nature of the problem: in contrast to image matching only small coin parts are analyzed and thus this process is more affected by noise and the variation of the image data. Nonetheless, it is clearly shown that in combination with image matching legend recognition improves the classification rate, as for $N = 1$ the classification rate is increased from 78.9% to 81.0%. The highest gap between image matching only and the fused result is spotted at $N = 3$, where an improvement from 87.9% to 91.4% can be achieved.

The benefit of using legend information is also demonstrated by the two classification results shown in Fig. 5. By using image matching only, the coins are classified to the wrong class (Fig. 5(b)), since the reference images (Fig. 5(c)) are in a sub-optimal condition and the strong abrasions on the coin impede to establish the true correspondences. However, with the support of legend recognition the final decision is changed to the correct class.



Fig. 5. Two classification results comparing the use of image matching only and the proposed fusion method

5 Conclusion

In this paper, a novel way for the classification of ancient Roman Republican coins is proposed that combines image matching with legend recognition to improve the classification rates the methods achieve individually. The presented experiments carried out on 464 images prove the effectiveness of the fused approach as it outperforms both methods when applied separately. The proposed method is more flexible than state-of-the-art coin classification systems for ancient coins as it does not rely on a specific coin layout and legend alignment. However, for future research both methods, but especially the legend recognition, have to be further improved to provide a stronger benefit of combining both methods.

Acknowledgment. This research has been supported by the Austrian Science Fund (FWF) under the grant TRP140-N23-2010 (ILAC).

References

1. Arandjelovic, O.: Automatic attribution of ancient roman imperial coins. In: Conference on Computer Vision and Pattern Recognition, pp. 1728–1734 (2010)
2. Arandjelović, O.: Reading ancient coins: Automatically identifying denarii using obverse legend seeded retrieval. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part IV. LNCS, vol. 7575, pp. 317–330. Springer, Heidelberg (2012)
3. Ballard, D.H.: Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition* 13, 111–122 (1981)
4. Bremananth, R., Balaji, B., Sankari, M., Chitra, A.: A new approach to coin recognition using neural pattern analysis. In: IEEE INDICON 2005, pp. 366–370 (2005)

5. Crawford, M.H.: Roman republican coinage. Cambridge University Press (1974)
6. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: Conference on Computer Vision and Pattern Recognition, pp. 886–893 (2005)
7. Felzenszwalb, P.F., Huttenlocher, D.P.: Pictorial structures for object recognition. *International Journal of Computer Vision* 61, 55–79 (2005)
8. Fukumi, M., Omatu, S., Takeda, F., Kosaka, T.: 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 (1991)
9. Grierson, P.: Numismatics. Oxford University Press (1975)
10. Huber-Mörk, R., Zambanini, S., Zaharieva, M., Kampel, M.: Identification of ancient coins based on fusion of shape and local features. *Machine Vision Applications* 22, 983–994 (2011)
11. Kampel, M., Zaharieva, M.: Recognizing ancient coins based on local features. In: Bebis, G., et al. (eds.) ISVC 2008, Part I. LNCS, vol. 5358, pp. 11–22. Springer, Heidelberg (2008)
12. Kavelar, A., Zambanini, S., Kampel, M.: Word detection applied to images of ancient roman coins. In: International Conference on Virtual Systems and Multimedia, pp. 577–580 (2012)
13. Liu, C., Yuen, J., Torralba, A.: Sift flow: Dense correspondence across scenes and its applications. *IEEE Pattern Analysis and Machine Intelligence* 33(5), 978–994 (2011)
14. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* 60(2), 91–110 (2004)
15. van der Maaten, L.J., Poon, P.: Coin-o-matic: A fast system for reliable coin classification. In: Muscle CIS Coin Competition Workshop, pp. 7–18 (2006)
16. Nölle, M., Penz, H., Rubik, M., Mayer, K., Holländer, I., Granec, R.: Dagobert - a new coin recognition and sorting system. In: 7th International Conference on Digital Image Computing - Techniques and Applications, pp. 329–338 (2003)
17. Reisert, M., Ronneberger, O., Burkhardt, H.: An efficient gradient based registration technique for coin recognition. In: Muscle CIS Coin Competition Workshop, pp. 19–31 (2006)
18. Shekhovtsov, A., Kovtun, I., Hlaváca, V.: Efficient mrf deformation model for non-rigid image matching. *Computer Vision and Image Understanding* 112(1), 91–99 (2008)
19. Wang, K., Belongie, S.: Word spotting in the wild. In: Daniilidis, K., Maragos, P., Paragios, N. (eds.) ECCV 2010, Part I. LNCS, vol. 6311, pp. 591–604. Springer, Heidelberg (2010)
20. Zambanini, S., Kampel, M.: Coarse-to-fine correspondence search for classifying ancient coins. In: 2nd ACCV Workshop on e-Heritage, pp. 25–36 (2013)