

# Coarse-to-Fine Correspondence Search for Classifying Ancient Coins

Sebastian Zambanini and Martin Kampel

Computer Vision Lab, Vienna University of Technology, Austria

**Abstract.** In this paper, we build upon the idea of using robust dense correspondence estimation for exemplar-based image classification and adapt it to the problem of ancient coin classification. We thus account for the lack of available training data and demonstrate that the matching costs are a powerful dissimilarity metric to establish coin classification for training set sizes of one or two images per class. This is accomplished by using a flexible dense correspondence search which is highly insensitive to local spatial differences between coins of the same class and different coin rotations between images. Additionally, we introduce a coarse-to-fine classification scheme to decrease runtime which would be otherwise linear to the number of classes in the training set. For evaluation, a new dataset representing 60 coin classes of the Roman Republican period is used. The proposed system achieves a classification rate of 83.3% and a runtime improvement of 93% through the coarse-to-fine classification.

## 1 Introduction

In ancient times, coins were the usual monetary items and thus everyday objects like they are today [1]. However, nowadays ancient coins are also considered as pieces of art which reflect the individualism of the engravers who manually cut the dies used for minting the coins [2]. Roman coins, for instance, often depict portraits of gods and emperors or historical events, in a similar manner as sculptures or paintings from this era do [2]. Fundamental work of coin experts is the classification of coins according to standard reference books since this provides additional information such as accurate dating, political background or minting place. However, classifying ancient coins is a highly complex task that requires years of experience in the entire field of numismatics [1]. As a substantial part of numismatic coin analysis, coin classification can be supported and facilitated by an automatic image-based system.

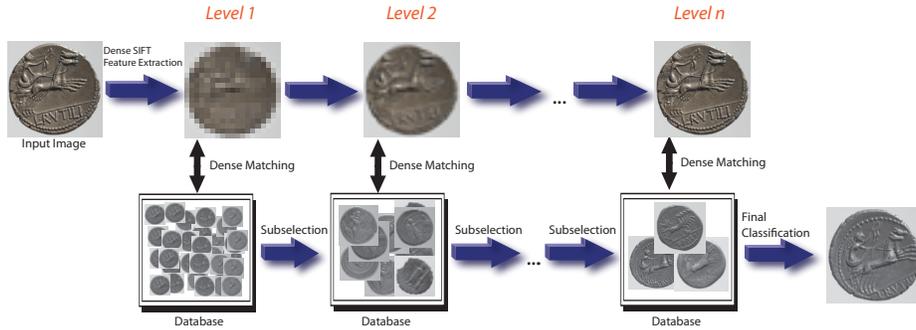
The difficulty of ancient coin classification arises from the high number of types (e.g. 550 types are defined for the Roman Republican period [2]) as well as from the high level of intra-class variability. The complexity is additionally increased when the discriminability between the classes is low, as can be seen in Fig. 1 where three coin classes with two samples each are shown. Please note the high global visual similarity between the classes. Nonetheless, one can also see local variations within a class as well as missing parts due to abrasions over the centuries. Apart from abrasions, local spatial variations of features within



**Fig. 1.** Six coins from the Roman Republican period where (a)-(c) each represents a different class

a class also stem from the fact that different dies from different die engravers were used for minting the coins. For instance, the legends are located differently on the two coins shown in Fig. 1(c). Thus, in contrast to modern coins, image features of ancient coins from the same class can not be aligned and compared by means of a global image transformation.

In this paper we present a method for automatically estimating the visual similarity between two coin images and show how this visual similarity can be used in a coarse-to-fine scheme for the ancient coin classification task. The classification does not rely on machine learning techniques and offline training which eases database extensions as new classes do not involve a re-training of the database. The idea was recently introduced in [3] and we extend this approach towards a rotation-insensitive coarse-to-fine matching, yielding a significant classification speed-up. The main motivation for excluding machine learning techniques from our method is that this way we are less dependent on the availability of a large and representative set of training images. The training of classifiers which provide both a sufficient discriminability for the hundreds of different classes and an adequate representation of the possible variability within a coin class is hindered mainly by the low number of available training samples. Our database of Roman Republican coins from the *Museum of Fine Arts in Vienna* is one of the biggest in the world and comprises around 3900 coins. These coins represent 515 different types but for only 237 of them more than three pieces are available. Therefore, instead of trying to cope with the high intra-class variability in a heavy offline classifier training phase we tackle the problem online in the matching stage. We use SIFT flow [4] for this task, but introduce a coarse-to-fine classification scheme which provides a significant decrease of computational time



**Fig. 2.** Schematic illustrating our coarse-to-fine coin classification procedure. Given an input image, a dense set of SIFT features is extracted and matched against the database at the coarsest level. A defined amount of most similar coin images is selected and forwarded to the matching step of the next finer level. This process is continued until the finest level  $n$  is reached where the final classification decision is made

needed for matching against a coin database without losing the discriminative power of SIFT flow for classification. An illustration of our method is shown in Fig. 2.

The work presented in this paper contributes to the research of image-based ancient coin classification, a quite new application field in the area of computer vision. Older methods dedicated to modern-day coins [5–8] have already been shown to be inappropriate for ancient coins [9]. The first approach especially designed for ancient coins was presented by Kampel and Zaharieva [10] with a classification rate of  $\sim 90\%$ , however by evaluating only three coin types. Their approach is based on matching sparse SIFT keypoints between coin images. Similarity of coin images is then estimated by simply counting the number of matching keypoints while ignoring the geometric configuration of the keypoints. A method based on offline learning was recently proposed by Arandjelović [11]. For feature description also SIFT keypoint detection is used, but geometry is introduced by calculating directional histograms at the keypoint positions. The improvement of adding geometry is shown in the experiments where the proposed method outperforms the bag-of-words approach with a classification rate of 57.2% against 2.4% on a dataset containing 65 Roman Imperial coin classes. This indicates that geometry is an important aspect for ancient coin recognition that we account for by using SIFT flow. This way, we do not depend on the availability of a large set of training images to generalize the intra-class variation: while in [11] between 9 and 160 samples per class were used for training, we show results on a training database of only 1 or 2 samples per class.

The rest of this paper is organized as follows. In Section 2 our SIFT flow based method for coin classification is presented. Results on a dataset representing 60 types of Roman Republican coins are reported in Section 3. The paper is finally concluded in Section 4.

## 2 Coin Classification Methodology

Our method is based on SIFT flow [4], a method for computing dense pixel-to-pixel correspondences between two images. SIFT flow works by minimizing an energy function which can be exploited to estimate the visual similarity of the images. In a classification scenario, the energy function values between a query image and all class image samples are determined to assign the query image to the class sample with minimum energy. Using this classification scheme, SIFT flow has shown superior results for scene and face classification in scenarios with a low number of available training samples [4]. The motivation to use SIFT flow for coin classification is further rooted in its ability to cope with large image variations in the form they appear within classes of ancient coins. However, we describe some modifications to account for differences of coin location, coin scale and coin rotation. Additionally, we present a coarse-to-fine classification scheme for runtime improvement.

### 2.1 Insensitivity to Coin Location and Coin Scale by Coin Segmentation

As a dense set of SIFT features with fixed scale has to be computed for SIFT flow computation, we normalize all images to a standard dimension of  $150 \times 150$  in order to account for scale differences between the coin images. Normalization is achieved by segmenting the coins in the images using a shape-adaptive thresholding approach [12]. The method applies a range and entropy filter to the image which is assumed to provide higher responses at coin regions than on background regions. For the final segmentation mask an optimal threshold value is found by minimizing an objective function describing the circularity of the binary thresholding result.

### 2.2 Insensitivity to Coin Rotation and Local Spatial Variations by SIFT Flow Image Matching

SIFT flow is based on SIFT features [13] which provide a rotation-invariant description of the local neighborhood by means of gradient orientation distributions. The SIFT features are computed densely over the image, resulting in the so called *SIFT image s*. The pixel-to-pixel correspondences between two SIFT images  $s_1$  and  $s_2$  are represented as a field of flow vectors  $\mathbf{w}(\mathbf{p}) = (u(\mathbf{p}), v(\mathbf{p}))$  at grid coordinates  $\mathbf{p} = (x, y)$ . The optimal correspondences are found by minimizing the following energy function of  $\mathbf{w}$ :

$$E(\mathbf{w}) = \sum_{\mathbf{p}} |s_1(\mathbf{p}) - s_2(\mathbf{p} + \mathbf{w}(\mathbf{p}))| \quad (1)$$

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

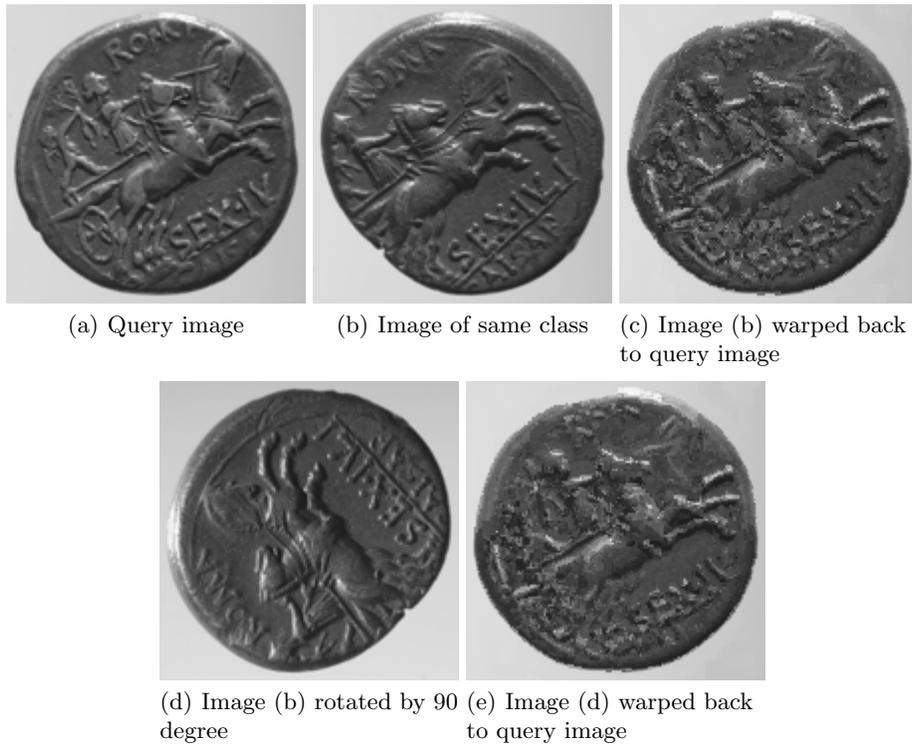
where  $\psi$  contains all four-connected pixel pairs. The energy function is composed of two terms, the *data term* (1) and the *smoothness term* (2), and the influence of the smoothness term is controlled by the parameters  $\alpha$  and  $d$ . Please note that in our case we do not consider a *small displacement term* in contrast to the original SIFT flow energy function defined in [4]. The reason lies in possible coin rotation differences between the images which demand to allow large pixel displacements in the correspondence search without additional costs. Therefore, without a small displacement term a rotation between image pairs only affects the correspondence search by producing a slightly larger energy in the smoothness term. We quantitatively prove in the experiments in Section 3.2 that the influence of the smoothness term in such cases is negligible and that classification performance is not affected by coin rotation differences. Examples of correspondences found by using the presented energy function can be seen in Fig. 3. Here the query image of Fig. 3(a) is matched against an image of a coin from the same class (Fig. 3(b)), which produces the correspondences visualized in Fig. 3(c) by warping back the image to the query image. We can see that reasonable correspondences have been found despite the variations between the two coins. If we compute SIFT flow for a rotated version of the coin (Fig. 3(d)), the result is almost identical (Fig. 3(e)).

### 2.3 Runtime Reduction by Coarse-To-Fine Classification

A disadvantage of using SIFT flow for example-based classification is that the runtime is linear to the amount of images in the database. However, SIFT flow itself uses a coarse-to-fine matching scheme for speed-up and better matching results, i.e. correspondences are propagated and re-estimated from coarser to finer levels. We utilize this scheme by selecting only the most similar coin classes at each level for further processing and thus subsequently reduce the amount of possible target coin classes. This way, the computational effort of the whole classification process is reduced as the more costly computations at finer levels have to be conducted only on a subset of coin classes.

More formally, for each sift image  $s$ ,  $n$  pyramid levels  $s^{(k)}$  are constructed where  $s^{(n)} = s$  and  $s^{(k-1)}$  is downsampled from  $s^{(k)}$  by a factor of 2. If we denote the set of coin target classes by  $C$  and the SIFT flow energy obtained at level  $k$  by  $E^{(k)}$ , classification of a query image  $s$  is achieved in the following manner:

1. For all levels  $k$ ,  $k = 1 \dots n$ 
  - (a) Compute SIFT flow energies  $E^{(k)}$  between  $s^{(k)}$  and all SIFT images of level  $k$  of classes  $C$ .
  - (b) For each class in  $C$ , compute the average energy  $\bar{E}^{(k)}$  for all its SIFT images in the database.
  - (c) Sort all energies  $\bar{E}^{(k)}$  and reduce  $C$  by selecting only a percentage  $\lambda^{(k)}$  of  $C$  with lowest energy.
2. Finally, take the class with lowest energy.



**Fig. 3.** Results of SIFT flow applied to images of the same class

### 3 Experiments

For the experiments we use a set of 60 classes of Roman Republican coins [2], where each class is represented by three coin images of the reverse side. Sample images of all classes are shown in Fig. 4. We report both classification and runtime performance on our dataset for different subselection values  $\lambda^{(k)}$ . Additionally, we compare the performance when one or two database images are available per class. As most of the coins show the same rotation in our database, the insensitivity against coin rotations is addressed in an individual experiment where the coin images are artificially rotated. Throughout all experiments, we used the same empirically determined parameters for SIFT flow matching: dense SIFT features were computed for a local neighborhood of  $12 \times 12$  pixels, the number  $n$  of pyramid levels was set to 4, and the parameters controlling the influence of the smoothness term were set to  $\alpha = 12$  and  $d = 1200$ .

#### 3.1 Classification Results

In each classification run, one of the 180 coin images served as query image and one or two of the remaining images per class served as training images.



Fig. 4. Sample images of all 60 classes of the evaluation database

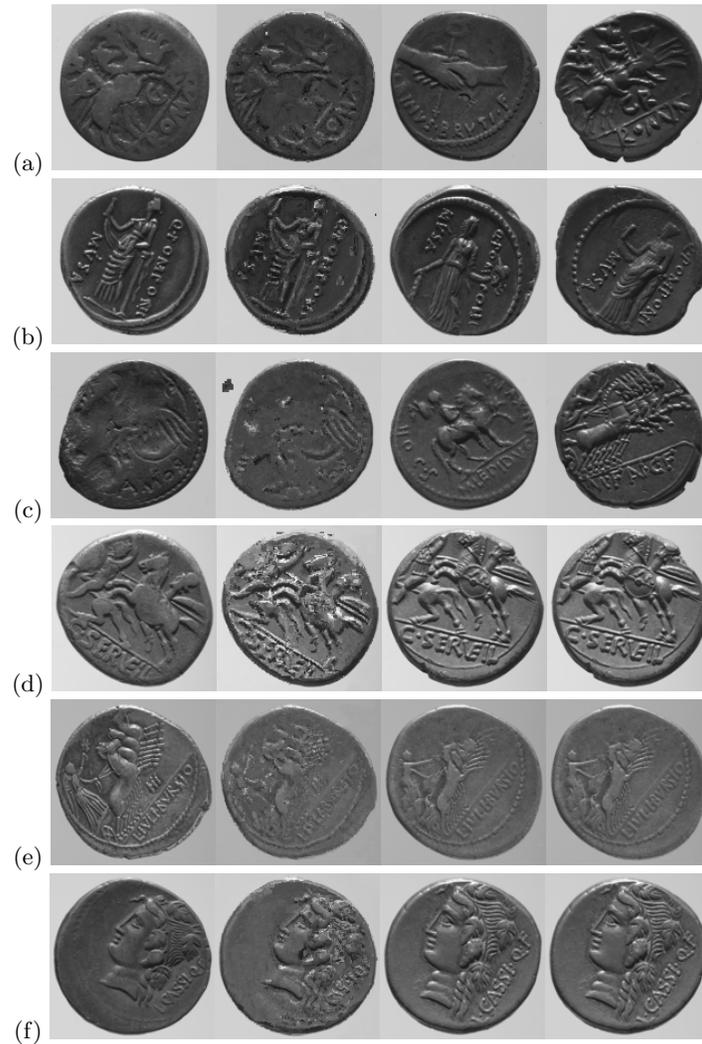
**Table 1.** Classification results

Training set size	$\lambda^{(1)}$	$\lambda^{(2)}$	$\lambda^{(3)}$	Correct classifications	Classification rate	Average classification time
1	100%	100%	100%	257/360	71.4%	235.8s
1	1%	100%	100%	220/360	61.1%	7.1s
1	100%	1%	100%	234/360	65.0%	21.2s
1	100%	100%	1%	257/360	71.4%	70.7s
1	30%	50%	50%	258/360	71.7%	32.5s
1	10%	50%	50%	249/360	69.2%	16.5s
2	100%	100%	100%	150/180	83.3%	471.6s
2	1%	100%	100%	127/180	70.6%	14.1s
2	100%	1%	100%	133/180	73.9%	42.4s
2	100%	100%	1%	141/180	78.3%	141.5
2	30%	50%	50%	150/180	83.3%	65.0s
2	10%	50%	50%	149/180	82.8%	32.9s

This leads to 180 (two training images per class) or 360 classification runs (one training image per class). For runtime evaluation, we measured the average runtime of computing the SIFT flow between two coin images by using the C++ implementation provided by the authors on a standard machine with a quad-core 2.70 GHz processor. The resulting average SIFT flow matching time was 3.93s, where around 3%, 6%, 21% and 70% are needed for the first, second, third and fourth level, respectively. In Table 1 classification results for both training set sizes as well as various values of  $\lambda^{(k)}$  are shown. Runtimes are indicated as the time for classifying one coin against our database of 60 classes, without considering feature extraction of the query image. Subselection parameters of  $\lambda^{(1)} = \lambda^{(2)} = \lambda^{(3)} = 100\%$  mean that no subselection is performed. Subselection parameters of  $\lambda^{(1,2,3)} = 1\%$  mean that only the energies of the first, second or third level, respectively, are used for classification.

One can see that, without subselection, over 70% of the images can be classified correctly with only one training image per class available. Adding a second training image brings a performance improvement of about 7 – 12%. Based on the results on this dataset, a reasonable choice for the subselection parameters is  $\lambda^{(1)} = 10\%$  and  $\lambda^{(2)} = \lambda^{(3)} = 50\%$ . The classification rate is very close to the case without subselection (−2.2% for a training set size of 1 and −0.5% for a training set size of 2, respectively), whereas the runtime improvement is around 93%.

In Fig. 5 we show some of our classification results where Fig. 5(a)-(c) depict incorrect classifications and Fig. 5(d)-(f) depict correct classifications. We see that strong abrasions, like in Figure 5(a), as well as the low inter-class variability, like in Figure 5(b), still pose a problem to the system, since the SIFT flow energy becomes less reliable under such conditions. However, also the examples shown in Fig. 5(d)-(f) represent strong abrasions and variations between the images which can be dissolved by SIFT flow. Figure 5(c) demonstrates the general limits of image-based ancient coin classification. The query image represents a misprint,

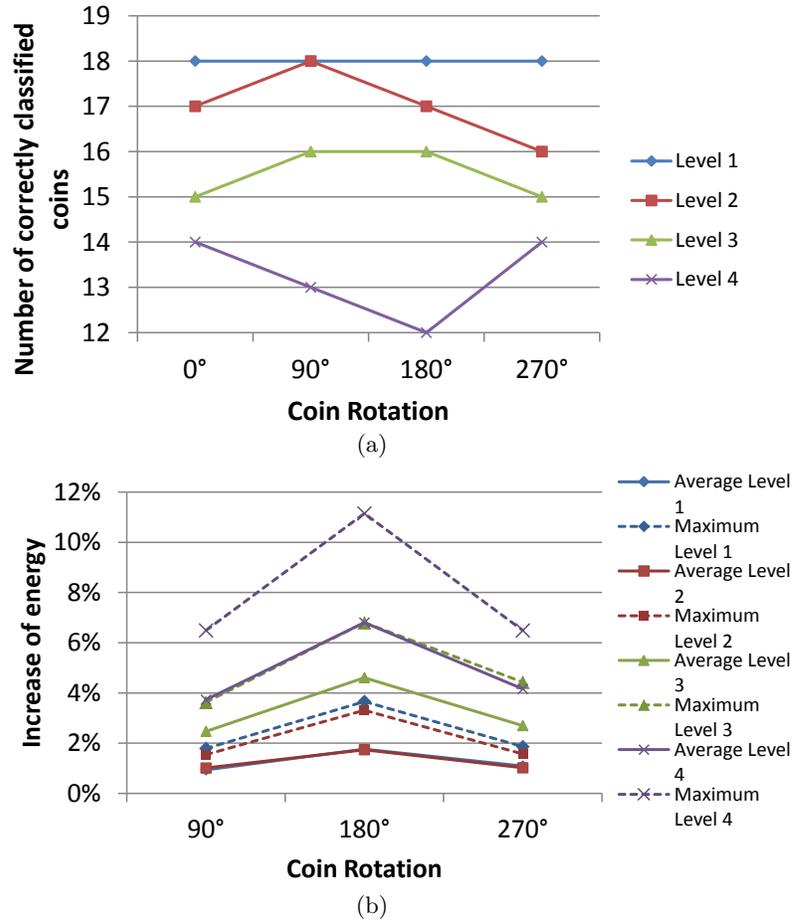


**Fig. 5.** Six classification results on our dataset. From left to right: query image; most similar image found in the database warped back using the SIFT flow correspondences; original most similar image found; correct most similar image depicting a coin of the same class

which makes it impossible even for human experts to accurately classify the coin if only this coin side is available for examination.

### 3.2 Insensitivity to Coin Rotations

In order to assess the sensitivity of our SIFT flow matching to coin rotation differences, we randomly took a query and a training image from 20 coin classes



**Fig. 6.** Results of evaluating sensitivity to coin rotations. (a) Number of correct classifications; (b) Average (solid lines) and maximum increase (dashed lines) of SIFT flow energy between images of the same class

and simulated different coin rotations by rotating the query image in 90 degree steps. Figure 6(a) shows the classification results of all four runs for the four levels of SIFT flow matching. In Fig. 6(b) the average increase of energy due to the additional costs in the smoothness term are plotted. We see that at a coarser level the energy values are more sensitive to coin rotations, thus producing a decrease of classification performance and a higher relative increase of the energy value. Nevertheless, by using a coarse-to-fine classification with subselection parameters  $\lambda^{(1)} = 10\%$ ,  $\lambda^{(2)} = \lambda^{(3)} = 50\%$ , 18 out of 20 classes can be classified correctly for all coin rotation differences. This shows that, although the method is in theory not invariant to coin rotation differences, a high degree of insensitivity is given.

## 4 Conclusions

In this paper, we have presented a classifier-free system for ancient coin classification. We proposed to use image matching instead of classifier learning for ancient coin classification. The main benefit of such a methodology is that it is less dependent on the number of available training samples as similarities between coins are determined online. This is shown in our experiments where we achieved a classification rate of 71.7% on a dataset with only one training sample available per class.

The major drawback of exemplar-based coin classification is the computational effort since the expensive image matching has to be performed against all coin image samples in the database. We therefore presented a coarse-to-fine scheme that heavily reduces the time needed for classification. In our experiments the average classification time could be reduced from 471.6s to 32.9s, an improvement of about 93%. Additionally, we experimentally proofed the insensitivity of our energy function to coin rotation differences.

In general, our classification results of 83.3% on 60 classes are higher than the ones reported by [11] (57.2% on 65 classes). However, a different dataset was used in the evaluation of [11] and thus no well-founded comparison of classification performance can be presented. As a contribution to other researchers in this field, we make our dataset publicly available<sup>1</sup> which allows for quantitative comparisons of algorithms in the future.

For future research, we plan to further improve our dense correspondence methods for coin similarity estimation. We will focus on a methodology which is less sensitive to appearance variations that arise from different relief heights and lighting conditions. Although due to the efficient optimization scheme SIFT flow is able to handle a large degree of noise in the features, we assume that more adapted features and matching strategies will lead to a significant improvement. We also see potential in using a visual similarity estimation in other forms within the application field of numismatics. Visual similarity estimation can be combined with other methods like symbol or legend recognition for a more extensive classification process. It can also be used for automatic coin hoard grouping where a clustering of coins is performed based on the proposed distance metric.

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

## References

1. Grierson, P.: Numismatics. Oxford University Press (1975)
2. Crawford, M.H.: Roman Republican Coinage, 2 vols. Cambridge University Press (1974)
3. Zambanini, S., Kampel, M.: Automatic coin classification by image matching. In: VAST: International Symposium on Virtual Reality, Archaeology and Intelligent Cultural Heritage. (2011) 65–72

<sup>1</sup> The dataset is available for download at <http://www.caa.tuwien.ac.at/cvl/people/zamba/>

4. Liu, C., Yuen, J., Torralba, A.: Sift flow: Dense correspondence across scenes and its applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33** (2011) 978–994
5. Huber, R., Ramoser, H., Mayer, K., Penz, H., Rubik, M.: Classification of coins using an eigenspace approach. *Pattern Recognition Letters* **26** (2005) 61–75
6. 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) 07–18
7. 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) 329–338
8. Reiser, M., Ronneberger, O., Burkhardt, H.: An efficient gradient based registration technique for coin recognition. In: *Proc. of the Muscle CIS Coin Competition*. (2006) 19–31
9. Zaharieva, M., Kampel, M., Zambanini, S.: Image based recognition of ancient coins. In: *Proc. of the 12th International Conference on Computer Analysis of Images and Patterns*. (2007) 547–554
10. Kampel, M., Zaharieva, M.: Recognizing ancient coins based on local features. In: *Proc. of International Symposium on Visual Computing*. Volume I of *Lecture Notes in Computer Science*. (2008) 11–22
11. Arandjelovic, O.: Automatic attribution of ancient roman imperial coins. In: *Conference on Computer Vision and Pattern Recognition*. (2010) 1728–1734
12. Zambanini, S., Kampel, M.: Robust automatic segmentation of ancient coins. In: *4th International Conference on Computer Vision Theory and Applications (VISAPP'09)*. Volume 2. (2009) 273–276
13. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision* **60** (2004) 91–110