Classifying Ancient Coins by Local Feature Matching and Pairwise Geometric Consistency Evaluation

Sebastian Zambanini, Albert Kavelar, Martin Kampel
Computer Vision Lab, Vienna University of Technology, Favoritenstrasse 9/1832, A-1040 Vienna, Austria
Email: {zamba,kavelar,kampel}@caa.tuwien.ac.at

***Abstract*—Classification of ancient coins is a substantial part of numismatic research which needs a large amount of expert knowledge due to the high number of classes to be considered. In this paper we propose an automatic image-based classification method for ancient coins to support this time-consuming and difficult process. We demonstrate that previously proposed learning-based methods suffer from the practical conditions of this problem: a high number of classes, limited number of training samples per class and complex intra-class variations. As a solution we propose a similarity metric based on feature correspondence which is designed to be robust against the possible intra-class coin variations like degraded parts, non-rigid deformations and illumination-induced appearance changes. The similarity metric is used in an exemplar-based ancient coin classification scheme which shows to outperform previously proposed methods for ancient coin recognition. Experiments are conducted on a dataset of 60 Roman Republican coin classes where the presented method achieves classification rates ranging from 72.7% for the case of one training sample per class up to 97.2% when nine training samples per class are used.**

(a) Crawford 410/3 (b) Crawford 442/1

(c) Crawford 346/1 (d) Crawford 393/1

Fig. 1: Reverse side of Roman Republican coins from four different coin classes.

I. Introduction

The aim of this paper is to automatically classify ancient Roman coins from single query images of the reverse (back) sides. Our motivation stems from the challenging nature of this task, as even for trained experts this is a time-consuming process, which can be speeded up and supported by an automatic image-based classification [1]. Additionally, coin recognition can be utilized for detecting illegal coin trade on the internet by checking the images against a repository of stolen ones [2].

As illustrated in Fig. 1, the difficulty of ancient coin recognition is caused by two main factors: the high number of classes with low inter-class and high intra-class variations as well as the suboptimal conditions of the coins. For instance, for the gold and silver coins of the Roman Republican age 550 main types with over 1000 subtypes are defined [3]. The coins of Fig. 1(a)-(b) all show the goddess Calliope and the two classes can only be differentiated by the coin legend. On the other hand, non-rigid intra-class deformations due to the non-industrial manufacturing can be spotted within the classes (e.g., Calliope's hand and the legend structure in Fig. 1(b)). The conditions of the coins lead to further challenges: missing and worn parts due to the coins' age (Fig. 1(c)) as well as appearance variations due to lighting variations and the metallic relief-like structures of the coins (e.g., the two coins in Fig. 1(d) are illuminated from opposite directions).

In this work we build upon correspondence-based methods for image classification [4]–[8]. These methods have shown superior performance in recognition scenarios with non-rigid intra-class deformations and low number of training samples like face recognition [6], [8]. This is also the case for recognizing ancient coin classes with different levels of rarity: in the *Museum of Fine Arts* in *Vienna* around 3900 coins of the Roman Republican age are available, but for only 237 of the 515 classes more than three coins are available [9]. Correspondence-based methods estimate image similarity by means of optimizing a cost function with first-order (local feature similarity) and second-order (regularization) constraints. Hence, the matching costs are robust against image clutter and non-rigid deformations, and this methodology has been successfully applied to exemplar-based coin classification [9].

In our method, instead of regularizing the matching process by geometric constraints, we perform a data-driven first-order matching and use geometric constraints afterwards to reason about the geometric plausibility of the correspondences found. This improves the reliability of the similarity measure while reducing the computation time. The improved reliability comes from the introduced potential of using stronger constraints with higher computational complexity, as the constraints have to be evaluated only once for the given correspondence configuration. Second, in contrast to optimization-based approaches, the "freedom" of data-driven matching contributes to a statistically more meaningful way of using the matching costs as dissimilarity measure: the geometric plausibility of the matched features will be higher for similar coins than for dissimilar coins, as statistically more correspondences are correct. In contrast, in optimization-based approaches the correspondence search is highly forced by the geometric constraints in case of local appearance ambiguities, which consequently reduces the similarity metric's gap between similar and dissimilar image pairs, and hence the discriminative power.

The remainder of this paper is organized as follows. In Section II we give an overview of related work in image-based ancient coin classification. Our proposed methodology is explained in Section III and empirically evaluated and compared to the state-of-the-art in Section IV. Conclusions are finally drawn in Section V.

II. Related Work

Due to the aforementioned challenging characteristics of ancient coins, methods for present-day coins have shown to be inappropriate for ancient coins [10]. The first approach dedicated to ancient coins was presented by Kampel and Zaharieva [11]. They define coin similarity by the number of matched SIFT features [12] and achieve 90% classification rate, although using only three coin classes in the experiments. We extended this approach of local feature matching by establishing dense matching costs similar to SIFT flow [8]. The method [9] was incorporated into an exemplar-based coarse-to-fine classification scheme and achieved a classification rate of around 83% on 60 classes of Roman Republican coins.

Learning-based methods for ancient coin classification have been proposed by Anwar et al. [13] and Arandjelović [14]. Both methods also rely on local SIFT features which are quantized into a fixed vocabulary of visual words. In [13] the image is tiled into spatial regions and the concatenated single histograms of visual words of each region are used as image feature. This approach is rather used for a coarse-grained classification of coins based on the reverse-side symbols than fine-grained classification and achieved a classification rate of up to 90% for eight common Roman Republican symbols. Arandjelović's method [14] exploits the spatial configuration of visual words in a different way: locally-biased directional histograms (LBDHs) are introduced for encoding the distribution of visual words around a detected keypoint in eight directions relative to its canonical orientation. The LBDH features are then again subject to vocabulary creation and the histogram of LBDH words serves as final image feature. This method achieves a classification rate of around 57% on 65 classes of the Roman Imperial age. Learning-based methods have also been exploited to support coin classification by means of legend recognition [15], [16].

III. Local Feature Matching and Geometric Consistency Evaluation

The core of our proposed exemplar-based coin classification methodology is to estimate the similarity of two coin images robustly against scale differences, illumination conditions, image background and non-rigid deformations. Robustness against scale differences and image background is achieved by segmenting the coin region in the image (Section III-A). Robustness against illumination conditions is accomplished by extracting illumination-insensitive local features for matching (Section III-B). Coin similarity insensitive to non-rigid deformations is finally enabled by first-order matching followed by an evaluation of the geometric consistency of the correspondences (Section III-C).

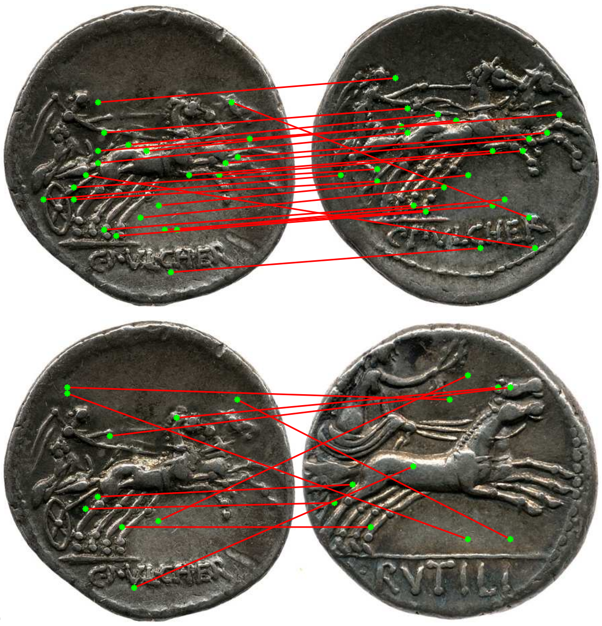
A. Segmentation and Scale Normalization

As a preprocessing step, we perform coin segmentation by means of a shape-adaptive method [17] in order to mask out local features outside of the coin region for further processing. The second achievement of the segmentation step is the ability of scale-normalizing the images. By selecting the coin area from the images and resizing it to a standard size (150×150 in our case) we are able to compute local features at a constant scale and do not have to sacrifice a certain amount of discriminative power and reliability by scale-invariant feature detection [12].

B. Feature Extraction and First-Order Matching

In our work, we extract local features from the images at positions $\mathbf{p}_i = (x_i, y_i)$ on a regular grid with pixel interval $\Delta\mathbf{p} = (\Delta x, \Delta y)$. We use this dense sampling scheme as we found the repeatability of keypoint detection on ancient coins to be unsatisfying. Furthermore, dense sampling gives us more features and thus a statistically more valuable estimation of the quality of feature correspondences. In contrast to other works [9], [11], [13], [15], we also do not build upon SIFT features but use the recently proposed LIDRIC features [18]. These features use locally normalized oriented even Gabor filter responses and have shown to outperform SIFT and other descriptors under illumination changes. As can be observed in Fig. 1, illumination changes can induce effects like highlights and edge polarity changes due to opposite lighting directions which disturb the image gradients and thus the SIFT descriptor. LIDRIC is more robust against these effects by using oriented even Gabor filters. We use eight oriented filters at a single fixed scale. In contrast to the original descriptor [18], we adapt the local normalization of filter responses in order to make the descriptor more appropriate for the problem at hand. Instead of dividing each filter response by the pure L2-norm of all eight responses (denoted as $\|F\|$) for normalization, we take the power $\|F\|^c$ with $c > 1$ of it before division. This reduces the relative influence of the highest responses in the image which likely arise from highlights on the metallic surface of the coin. Finally, by performing the same spatial pooling with 4×4 squared cells as in SIFT [12], we end up with a 128-dimensional descriptor $\mathbf{d}_i$ for the image point $\mathbf{p}_i$.

After the local descriptors $\mathbf{d}'_i \in \mathcal{D}'$ and $\mathbf{d}''_j \in \mathcal{D}''$ have been extracted from the two images I' and I'', we aim to find robust matchings between them. An option would be to accept only nearest neighbors with a certain distance to their second nearest neighbors as proposed by Lowe [12], but in our case a one-to-one symmetric search [19] turned out to be the better choice. Two features $\mathbf{d}'_i$ and $\mathbf{d}''_j$ are matched only if $\mathbf{d}''_j$ is the nearest neighbor of $\mathbf{d}'_i$ in $\mathcal{D}''$ and $\mathbf{d}'_i$ is in turn the nearest neighbor of $\mathbf{d}''_j$ in $\mathcal{D}'$. The indices of the descriptors in $\mathcal{D}'$ with a match in $\mathcal{D}''$ are stored in the set $\mathcal{M}$ and the function $\phi(i)$ relates the indices of $\mathcal{D}'$ to the corresponding indices of $\mathcal{D}''$, i.e. $\phi(i) = j$ if $\mathbf{d}'_i$ corresponds to $\mathbf{d}''_j$. Fig. 2 shows the result of the one-to-one symmetric correspondence search for



(a)

(b)

Fig. 2: Result of one-to-one symmetric matching of a coin image with (a) an image from the same class and (b) an image from a different class. Only random 10% of the overall correspondences are shown for better illustration.

two coins from the same class and two coins from different classes.

C. Similarity Estimation from First-Order Correspondences

The basic assumption of our approach is that by first-order matching more correct correspondences can be found for coins from the same class than for coins from different classes. By examining the matching results of similar and dissimilar coins as shown in Fig. 2 we are able to identify three key observations that lead to the definition of our final similarity measure:

1) Number of Correspondences: The number of matched features is likely to be higher for similar coins than for dissimilar ones. We use this property for a similarity score by

$$\Theta_n = \frac{|\mathcal{M}|}{\min(|\mathcal{D}'|\,,|\mathcal{D}''|)} \tag{1}$$

2) Displacement of Corresponding Feature Points: The displacement of correct correspondences is low which can be used as a dissimilarity score by

$$\Theta_d = \frac{1}{|\mathcal{M}|}\sum_{i\in\mathcal{M}} \|\mathbf{p}'_i - \mathbf{p}''_{\phi(i)}\| \tag{2}$$

with $\|\cdot\|$ being the L2-norm.

3) Geometrical Consistency of Correspondences: Pairs of correct correspondences will not drastically change their relative position to each other. Hence, given two points $\mathbf{p}'_i$ and $\mathbf{p}'_j$ and their corresponding points $\mathbf{p}''_{\phi(i)}$ and $\mathbf{p}''_{\phi(j)}$, the vector

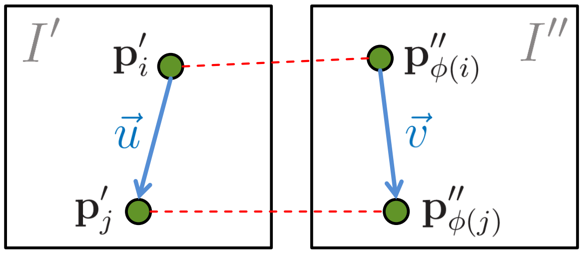


Fig. 3: The geometric plausibility of the correspondences of the points $\mathbf{p}'_i$ and $\mathbf{p}'_j$ in I' with the points $\mathbf{p}''_{\phi(i)}$ and $\mathbf{p}''_{\phi(j)}$ in I'' is assessed by the comparing the vectors $\vec{u}$ and $\vec{v}$ in terms of length (Eq. 4) and orientation (Eq. 5).

$\vec{u} = \overline{\mathbf{p}'_i\mathbf{p}'_j}$ will be similar to the vector $\vec{v} = \overline{\mathbf{p}''_{\phi(i)}\mathbf{p}''_{\phi(j)}}$, as illustrated in Fig. 3. We compute their difference by

$$\begin{aligned}\Omega_{(\mathbf{p}'_i,\mathbf{p}'_j,\mathbf{p}''_{\phi(i)},\mathbf{p}''_{\phi(j)})} =&\lambda\cdot\eta_{(\mathbf{p}'_i,\mathbf{p}'_j,\mathbf{p}''_{\phi(i)},\mathbf{p}''_{\phi(j)})} \;+\\ &(1-\lambda)\cdot\alpha_{(\mathbf{p}'_i,\mathbf{p}'_j,\mathbf{p}''_{\phi(i)},\mathbf{p}''_{\phi(j)})}\end{aligned} \tag{3}$$

$$\eta_{(\mathbf{p}'_i,\mathbf{p}'_j,\mathbf{p}''_{\phi(i)},\mathbf{p}''_{\phi(j)})} = \frac{\left|\|\mathbf{p}'_i-\mathbf{p}'_j\| - \|\mathbf{p}''_{\phi(i)}-\mathbf{p}''_{\phi(j)}\|\right|}{\|\mathbf{p}'_i-\mathbf{p}'_j\| + \|\mathbf{p}''_{\phi(i)}-\mathbf{p}''_{\phi(j)}\|} \tag{4}$$

$$\begin{aligned}&\alpha_{(\mathbf{p}'_i,\mathbf{p}'_j,\mathbf{p}''_{\phi(i)},\mathbf{p}''_{\phi(j)})} =\\ &\frac{1}{\pi}\arccos\left(\frac{\mathbf{p}'_i-\mathbf{p}'_j}{\|\mathbf{p}'_i-\mathbf{p}'_j\|}\cdot\frac{\mathbf{p}''_{\phi(i)}-\mathbf{p}''_{\phi(j)}}{\|\mathbf{p}''_{\phi(i)}-\mathbf{p}''_{\phi(j)}\|}\right)\end{aligned} \tag{5}$$

Intuitively, the terms η and α measure the vector difference in terms of length and orientation, respectively, where λ serves as weighting parameter. This or a similar vector difference metric is typically used for regularization in optimization-based matching approaches [4]–[8] in order to penalize matching discontinuities and prefer smooth results. However, for computational reasons only small neighborhoods can be considered (e.g., SIFT flow uses the L1-norm of the 4-connected neighboring flow vectors [8]). In our case, these metrics have to evaluated only once for the given first order matching, which allows to use a larger neighborhood system $\mathcal{N}$ for the geometric dissimilarity score:

$$\Theta_g = \frac{1}{|\mathcal{M}|}\sum_{i\in\mathcal{M}}\frac{1}{|\mathcal{N}_i|}\sum_{j\in\mathcal{N}_i}\Omega_{(\mathbf{p}'_i,\mathbf{p}'_j,\mathbf{p}''_{\phi(i)},\mathbf{p}''_{\phi(j)})} \tag{6}$$

In general, one can define all other feature points as the neighborhood $\mathcal{N}_i$ of a given feature point, but this unnecessarily increases the computational burden without substantially improving the quality of this similarity metric. Hence, in practice it turns out to be sufficient to compare every feature point to only a small subset of feature points. In our work we have empirically chosen to compare every feature point to its neighboring features at the six distances $1\Delta\mathbf{p}$, $2\Delta\mathbf{p}$, $4\Delta\mathbf{p}$, $8\Delta\mathbf{p}$, $12\Delta\mathbf{p}$ and $16\Delta\mathbf{p}$, which on average leads to

Fig. 4: An exemplar image for each of the 60 classes in our dataset.

a comparison of a feature point with around 7.5% of the remaining feature points.

The final overall similarity score is computed from the three individual scores by

$$\Theta = (1 - g(\Theta_n; \sigma_n)) + g(\Theta_d; \sigma_d) + g(\Theta_g; \sigma_g) \qquad (7)$$

where $g(x; \sigma) = \exp(-x^2/(2\sigma^2))$ is a Gaussian membership function that transforms the individual scores to the same value range.

IV. Experiments

In this section we report the results of our proposed coin classification methodology and compare it to our previous SIFT flow based method (SF) [9] as well as to the learning-based methods using locally biased directional histograms (LBDH) [14] and bag of visual words (BOVW) [20]. Our dataset of Roman Republican coins is presented in Section IV-A. Implementation details of our and the compared methods are described in Section IV-B. In Section IV-C the results are reported and discussed.

A. Roman Republican Coin Dataset

For the experiments we use a dataset of 60 Roman Republican coin classes defined by their reference number given in [3]. For each class we collected 10 images of the reverse side as this coin side shows more inter-class variation than the obverse side and is thus the better choice for classification [3]. Different image sources are exploited to increase the diversity among the images and to mimic a more realistic scenario of coin classification under uncontrolled image acquisition conditions. Three images of each class were taken from the coin collection of the *Museum of Fine Arts, Vienna*, another three images from the collection of the *British Museum, London*[1], and the remaining four from free online ancient coin search engines[2]. As described in Section III-A, prior to classification every image is segmented and scale-normalized to 150×150 for our method and SF or 250×250 for LBDH and BOVW (as in [14]). This operation resulted in no error for all of our 600 coin images. An example image of each class is shown in Fig. 4.

We would like to note that in our dataset all images of a class are consistently oriented. Generally speaking, we found rotation differences between coin images to be an uncommon situation, which has not been encountered during our coin image search on the internet, as coins are typically imaged at a canonical orientation based on their main motive. However, as this can not be guaranteed in practice, we aim to consider rotational invariance for improved robustness in the future.

B. Implementation Details

For all methods compared suitable parameter choices were empirically determined. Local feature extraction is the first step of all methods and is accomplished either by using SIFT [12] or LIDRIC features [18] with $c = 1.3$ (see Section III-B).

[1] www.britishmuseum.org/research/publications/online_research_catalogues/rrc.aspx

[2] www.acsearch.info and www.coinarchives.com

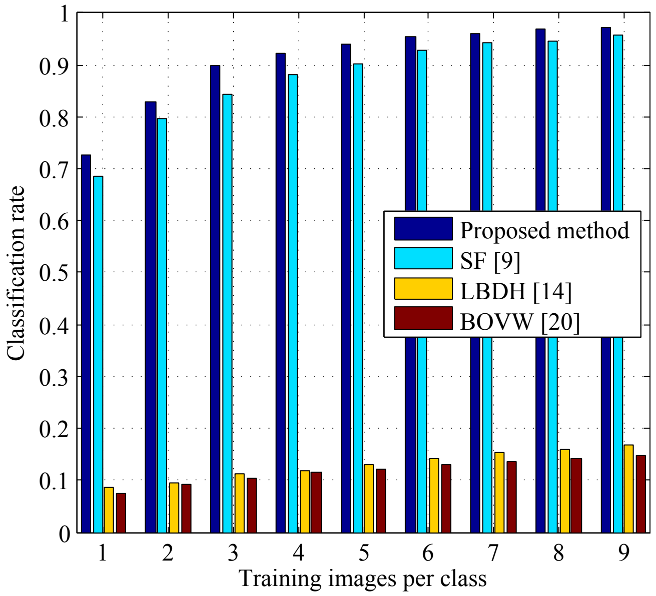


Fig. 5: Bar plot of the classification results of all methods using LIDRIC feature extraction and training set sizes of 1 to 9.

Dense sampling with $\Delta\mathbf{p} = 3$ for our method and $\Delta\mathbf{p} = 1$ for SF is performed at a feature scale of 24×24 pixels. For LBDH and BOVW the standard Difference-of-Gaussian interest point detection [12] is used. Features are only extracted from the coin region in the image provided by the initial coin segmentation step. As no rotation differences are present in our image dataset, for a fair comparison all features are extracted without rotation invariance, i.e. the canonical orientation of all features is automatically set to the same fixed value.

For our coin similarity algorithm from one-to-one symmetric correspondences we use parameter values of $\lambda = 0.7$, $\sigma_n = 0.1$, $\sigma_d = 50$ and $\sigma_g = 0.25$. For SF the SIFT flow parameters of $\alpha = 200$, $d = 20000$ and $\gamma = 12$ are used (see [8] for details). For BOVW we quantize the descriptors to 100 visual words and compute the visual word histogram as image feature. As in the original experiments [14] we use a vocabulary size of 500 for our LBDH implementation. However, we set the bandwidth R of the directional kernels to 200 instead of 1000, as this showed superior results on our data.

For the final class decision a k-nearest neighbors classifier with $k = 5$ is used for all methods compared. The distance of test and training samples is thereby determined by our proposed class similarity or the SIFT flow energy (SF). For BOVW and LBDH the Euclidean distance of the visual word or LBDH histograms is used.

C. *Classification Results and Discussion*

The goal of this paper is to achieve coin classification in scenarios with low number of training samples. Therefore, we aim to analyze the influence of the number of training samples per class to the methods' classification performances. For this purpose, we conducted multiple classification runs for each image in our dataset with increasing number of training samples n per class, i.e. $n = 1 \dots 9$. In each run, n randomly chosen images per class served as training set. This process was again repeated 10 times for each value of n and the overall classification rate out of the $60 \cdot 10 \cdot 10 = 6000$ classifications was recorded.

TABLE I: Numerical classification results of all methods using SIFT/LIDRIC feature extraction.

	Training images per class		
	1	5	9
BOVW [20] - SIFT	6.3%	11.2%	13.2%
BOVW [20] - LIDRIC	7.6%	12.3%	14.8%
LBDH [14] - SIFT	6.2%	11.9%	14.3%
LBDH [14] - LIDRIC	8.8%	13.1%	16.8%
SF [9] - SIFT	48.9%	81.0%	90.5%
SF [9] - LIDRIC	68.6%	90.2%	95.8%
Proposed method - SIFT	68.0%	93.7%	97.1%
Proposed method - LIDRIC (full)	72.7%	94.1%	97.2%
Proposed method - LIDRIC (Θ_n + Θ_g)	70.5%	93.8%	97.2%
Proposed method - LIDRIC (Θ_n + Θ_d)	69.4%	92.9%	97.0%
Proposed method - LIDRIC (Θ_n)	56.0%	84.5%	91.1%

The classification results for the different training set sizes are shown in Fig. 5 for all methods with LIDRIC feature extraction. Additionally, in Table I the numerical classification results of all methods with SIFT or LIDRIC feature extraction are listed. It can be seen that the correspondence-based methods dominate the learning-based ones and that our proposed method outperforms all other methods for all training set sizes with classification rates from 72.7% ($n = 1$) to 97.2% ($n = 9$). The inclusion of spatial information provided by LBDH gives only a slight improvement over the general BOVW model and does not contribute to a performance comparable to the ones achieved by the correspondence-based methods. Due to the low number of training samples the learning-based methods are not able to sufficiently generalize over the intra-class variation. In the experiments presented in [14] LBDH achieved a classification rate of 57.2% on a 65-class problem. However, the dataset used shows a very uneven distribution of training samples among the classes which are represented by 10 up to 160 exemplars. We conjecture that the classification performance of LBDH on this dataset is mainly supported by the classes with a high number of training samples. Another reason for the low classification rate of LBDH is the erroneous interest point detection.

From the results shown in Table I it can also be concluded that LIDRIC represents a more powerful local descriptor for coin classification under uncontrolled conditions as its use improves the performance of each individual method. For our proposed method the performance is increased from 68.0% to 72.7% due to LIDRIC's lower sensitivity to illumination changes.

1) Influence of Individual Similarity Scores: As three single similarity scores are combined in our method, we are also interested in assessing their individual influence to the classification performance. It is evidently shown in Table I that all three scores have a contribution to the classification power of our method. By using only data-driven matching as similarity measure (Θ_n) and ignoring the geometric ones (Θ_d and Θ_g) only 56.0% correct classifications are achieved for $n = 1$, less than the SF method which also uses geometric information for finding the optimal correspondences (68.6%). Adding geometric information to our model either by the displacement similarity Θ_d or neighboring vector consistency Θ_g leads to classification rates that are higher than that of SF. The full model with all three terms achieves the highest classification rate of 72.7%.

2) Runtime Analysis: An important issue of exemplar-based classification is the time it takes to compare two image samples, as the query image has to be compared to all images in the database. Without feature extraction, which takes around 1s, our MATLAB implementation needs around $0.35s$ to compare two images whereas the C-implementation of SF takes around 2.2s. In practice, this means that it takes around 22s to classify a query image for our 60-class problem. However, in [9] we have shown that the classification time of exemplar-based coin classification in conjunction with feature correspondence can be reduced to one-seventh without a loss of classification accuracy by applying a hierarchical subselection scheme. We believe that the same principle can be applied to our similarity metric for speeding up the classification process, although this will be part of future research.

V. Conclusions

In this paper, we have shown that learning-based methods exhibit a poor classification performance for ancient coins when the number of training samples is low. We deal with this commonly encountered situation in ancient coin recognition by means of an exemplar-based classification methodology. The main contribution of our paper is a local correspondence-based image similarity metric that is both accurate and fast to compute. The superior classification performance of our method results also from an illumination-insensitive feature extraction that provides the needed robustness against uncontrolled image acquisition conditions, but at the same time ensures enough discriminative power to establish correct correspondences between coins without needing to guide the correspondence search by regularization.

The main practical drawback of exemplar-based classification is the long classification time which is theoretically linear to the number of samples in the dataset. However, this drawback can be mitigated by using hierarchical coarse-to-fine strategies and by using other methods like legend [16] or symbol recognition [13] for initial subselection. For future research we also plan to investigate and adapt our image similarity metric to wider class of problems as well as to extend it to other kinds of geometric variations between images. Our method allows to flexibly adapt the similarity terms to account for other required geometric invariances. For instance, rotation invariance can be achieved by using rotation-invariant local features and a rotation-invariant evaluation of correspondence consistency, e.g. by using only the distance of pairs of correspondences or by using the length and angles between triplets of correspondences.

Acknowledgement

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

References

[1] P. Grierson, *Numismatics*. Oxford University Press, 1975.

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

[3] M. H. Crawford, *Roman Republican Coinage*. Cambridge University Press, 1974.

[4] A. C. Berg, T. L. Berg, and J. Malik, "Shape matching and object recognition using low distortion correspondences," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, 2005, pp. 26–33.

[5] O. Duchenne, A. Joulin, and J. Ponce, "A graph-matching kernel for object categorization," in *Proc. of IEEE International Conference on Computer Vision*, 2011, pp. 1792–1799.

[6] A. Jorstad, D. Jacobs, and A. Trouvé, "A deformation and lighting insensitive metric for face recognition based on dense correspondences," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, 2011, pp. 2353–2360.

[7] J. Kim, C. Liu, F. Sha, and K. Grauman, "Deformable spatial pyramid matching for fast dense correspondences," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, 2013, pp. 2307–2314.

[8] C. Liu, J. Yuen, and A. Torralba, "Sift flow: Dense correspondence across scenes and its applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 5, pp. 978–994, 2011.

[9] S. Zambanini and M. Kampel, "Coarse-to-fine correspondence search for classifying ancient coins," in *Computer Vision - ACCV 2012 Workshops*, ser. Lecture Notes in Computer Science, 2013, vol. 7729, pp. 25–36.

[10] M. Zaharieva, M. Kampel, and S. Zambanini, "Image based recognition of ancient coins," in *Proc. of International Conference on Computer Analysis of Images and Patterns (CAIP)*, 2007, pp. 547–554.

[11] M. Kampel and M. Zaharieva, "Recognizing ancient coins based on local features," *Proc. of International Symposium on Visual Computing (ISVC)*, pp. 11–22, 2008.

[12] D. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[13] H. Anwar, S. Zambanini, and M. Kampel, "Supporting ancient coin classification by image-based reverse side symbol recognition," in *Proc. of International Conference on Computer Analysis of Images and Patterns (CAIP)*, 2013, pp. 17–25.

[14] O. Arandjelović, "Automatic attribution of ancient roman imperial coins," in *Proc. of IEEE Conference on Computer Vision and Pattern Recognition*, 2010, pp. 1728–1734.

[15] ——, "Reading ancient coins: automatically identifying denarii using obverse legend seeded retrieval," in *Proc. of European Conference on Computer Vision*, 2012, pp. 317–330.

[16] A. Kavelar, S. Zambanini, and M. Kampel, "Reading the legends of roman republican coins," *Journal on Computing and Cultural Heritage*, vol. 7, no. 1, pp. 5:1–5:20, 2014.

[17] S. Zambanini and M. Kampel, "Robust automatic segmentation of ancient coins." in *Proc. of International Conference on Computer Vision Theory and Applications (VISAPP)*, 2009, pp. 273–276.

[18] ——, "A local image descriptor robust to illumination changes," in *Proc. of Scandinavian Conference on Image Analysis (SCIA)*, 2013, pp. 11–21.

[19] W.-L. Zhao, C.-W. Ngo, H.-K. Tan, and X. Wu, "Near-duplicate keyframe identification with interest point matching and pattern learning," *IEEE Transactions on Multimedia*, vol. 9, no. 5, pp. 1037–1048, 2007.

[20] G. Csurka, C. Dance, L. Fan, J. Willamowski, and C. Bray, "Visual categorization with bags of keypoints," in *Proc. of ECCV Workshop on Statistical Learning in Computer Vision*, 2004, p. 22.