Identification of ancient coins based on fusion of shape and local features

Article *in* Machine Vision and Applications · November 2011

DOI: 10.1007/s00138-010-0283-y · Source: DBLP

4 authors:

Reinhold Huber-Mörk
Oesterreichische Nationalbank
93 PUBLICATIONS **570** CITATIONS
SEE PROFILE

Sebastian Zambanini
TU Wien
63 PUBLICATIONS **460** CITATIONS
SEE PROFILE

Maia Rohm
TU Wien
55 PUBLICATIONS **338** CITATIONS
SEE PROFILE

Martin Kampel
TU Wien
220 PUBLICATIONS **1,698** CITATIONS
SEE PROFILE

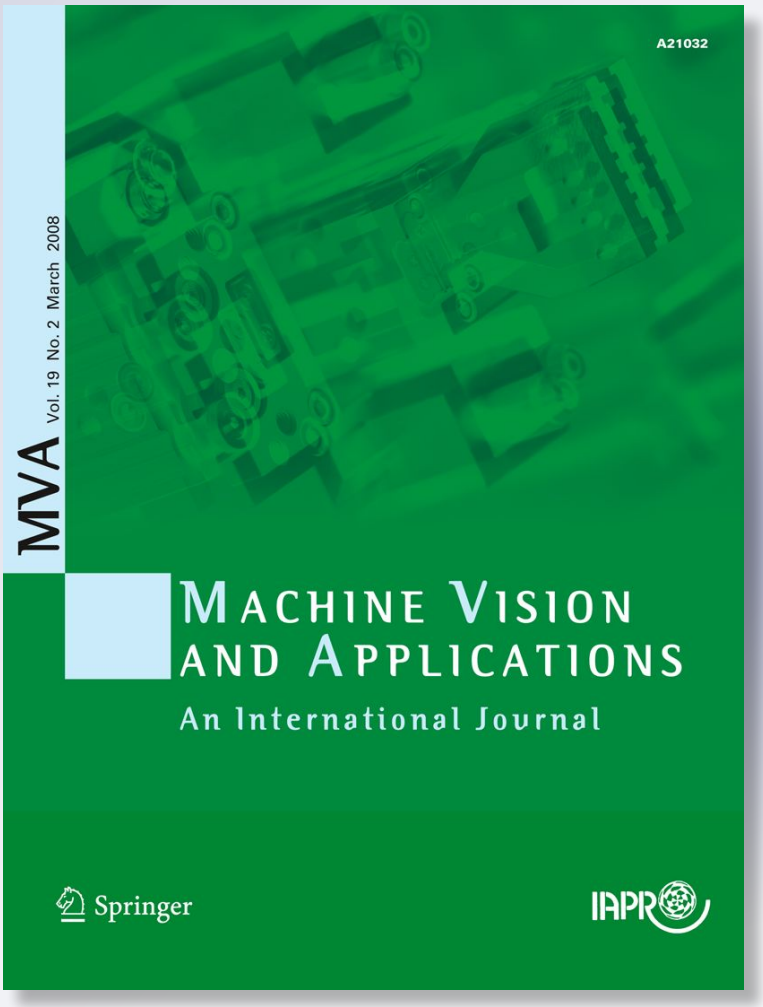
Identification of ancient coins based on fusion of shape and local features

Reinhold Huber-Mörk, Sebastian Zambanini, Maia Zaharieva & Martin Kampel

Machine Vision and Applications

ISSN 0932-8092
Volume 22
Number 6

Machine Vision and Applications (2011)
22:983-994
DOI 10.1007/s00138-010-0283-y



Springer

Machine Vision and Applications (2011) 22:983–994
DOI 10.1007/s00138-010-0283-y

ORIGINAL PAPER

Identification of ancient coins based on fusion of shape and local features

Reinhold Huber-Mörk · Sebastian Zambanini · Maia Zaharieva · Martin Kampel

Received: 4 August 2009 / Revised: 23 April 2010 / Accepted: 20 June 2010 / Published online: 11 July 2010

Abstract We present a vision-based approach to ancient coins' identification. The approach is a two-stage procedure. In the first stage an invariant shape description of the coin edge is computed and matching based on shape is performed. The second stage uses preselection by the first stage in order to refine the matching using local descriptors. Results for different descriptors and coin sides are combined using naive Bayesian fusion. Identification rates on a comprehensive data set of 2400 images of ancient coins are on the order of magnitude of 99%.

Keywords Coin identification · Ancient coins · Shape matching · Local features · Fusion

This work was partly supported by the European Union under grant FP6-SSP5-044450. However, this paper reflects only the authors' views and the European Community is not liable for any use that may be made of the information contained herein.

R. Huber-Mörk (✉)
Business Unit High Performance Image Processing, Department Safety & Security, Austrian Institute of Technology GmbH, 2444 Seibersdorf, Austria
e-mail: reinhold.huber-moerk@ait.ac.at

S. Zambanini · M. Kampel
Computer Vision Lab, Institute of Computer-Aided Automation, Vienna University of Technology, Favoritenstr. 9/183-2, 1040 Vienna, Austria
e-mail: zamba@caa.tuwien.ac.at

M. Kampel
e-mail: kampel@caa.tuwien.ac.at

M. Zaharieva
Interactive Media Systems Group, Institute of Software Technology and Interactive Systems, Vienna University of Technology, Favoritenstr. 9-11/188/2, 1040 Vienna, Austria
e-mail: zaharieva@ims.tuwien.ac.at

1 Introduction

Illegal trade and theft of coins is a major part of the illegal antiques market. There is an annual estimate of half a million to a million coins sold in the North American market. Those coins were primarily excavated in Eastern Europe and 80% of all coins are undocumented [5]. In order to facilitate prevention and repression of illicit trade of stolen coins, technologies aimed at allowing permanent identification and traceability of coins are a matter of interest. Every individual coin has signs, caused by minting techniques for pre-industrial ones or by use-wear for more recent ones, that make it unique and recognizable to an expert's eye. This enables the traceability of pre-industrial coins based on visual inspection. Once passive security systems have been defeated and coins have been stolen, fighting against their illicit trade must allow for the identification of stolen items among the incredibly vast number of those being offered for sale by antiquity traders, at auctions or on the Internet.

Methods for computer-aided identification of ancient coins become significant for the protection of cultural heritage and associated conservation. In [16] the potentials of image analysis for a systematic study of coinage and the scope of image comparison for die studies are highlighted. Traditionally, coin identification is performed by manual search in coin publications and auctions catalogues. The purpose of existing digital coin collections mainly focuses on internet services and inventory. Nowadays, there are large digital coin collections available, e.g., the collection of coins and medals at the Fitzwilliam Museum, Cambridge, UK, contains more than 40,000 digital coin descriptions. Details on numismatic photography using digital cameras were recently published by Goodman [12], detailed treatment of coin collections is given in [14]. For coin classification, the textual description, if present, already provides relevant information.

For the identification of a given coin, its image is even more informative, as the appearance of an ancient coin is often unique, e.g., due to variations in the hammering process, die, mint signs, shape, scratches, wearing, etc. The uniqueness in the appearance of coins results from variations in the coin blank material and application of the tools in minting, as well as from wear of the coin. Therefore, for numismatists the shape of the coin edge is regarded to be an important feature to characterize a coin.

In this paper we propose a two-stage procedure for the identification of ancient coins in images. The first stage performs a shape matching of the coin edge and serves as a preselection step for the second stage. In the second stage local features are matched between the query image and images contained in the preselection set. The outputs of both stages are finally fused in order to obtain a more reliable decision. A similar procedure was suggested in an earlier paper [20] where the determination of preselection size was done ad-hoc. A more systematic way to adjust the preselection size is presented here. Furthermore, this work focuses on a systematic treatment of fusion of different sources of information i.e., features derived from shape and local descriptors, and extension to fusion of images of reverse and obverse sides of a coin.

The paper is organized as follows. Section 2 presents recent work on coin recognition, especially modern coins, and recent object recognition methods, especially methods based on shape and local descriptors. The discussion of ancient coin identification is contained in Sect. 3 and our approach is detailed in Sect. 4. The investigated data sets are introduced in Sect. 5 and achieved results are presented and discussed in Sect. 6. Finally, in Sect. 7 conclusions and outlook for further research are drawn.

2 Related work

For modern coins, i.e., machine struck coins, judging systems using electromechanical devices are commonly based on measuring weight, diameter, thickness, permeability and conductivity [4], oscillating electromagnetic field characteristics [34], and photo- and piezoelectric properties [41]. Machine-minted coin classification using image processing is described by Fukumi et al. [8] uses a neural network approach to discriminate between 500 Won and 500 Yen coins. Nölle et al. [35] and Fürst et al. [9] described the *Dagobert* coin recognition and sorting system for high volumes of coins and a large number of currencies. The approach by Huber et al. [17] used the same mechanical setup including a coin-reversing unit as described in [9] and the classification method is based on matching features derived from multiple edge-based eigenspaces. A special acquisition device for coins employing colored illumination from various angles was suggested by Hossfeld et al. [15]. Recently, methods based on matching gradient directions [37] and color, shape and wavelet features [49] were suggested.

A number of coin authentication methods for modern coins employing optical means are to be found in patents. Hibari and Arikawa [13] describe a system by which both sides of a coin are first acquired by cameras, then features are derived from binarized images, and finally results of each side are combined with a magnetic sensor measurement. Another approach, also based on binarization followed by area measurement and comparison of coin center and center of gravity, is suggested by Onodera and Sugata [36]. Tsuji and Takahashi [45] analyze one side of a coin by transformation of its image into polar coordinates and matching of profiles taken along angle direction.

For hand struck coins, some approaches for classification appeared recently [48,51]. However, in contrast to object classification, object identification relies on those unique features which distinguish a given object from all other members of the same class. First results on identification of ancient coins were reported in [18] where the combination of shape and local descriptors to capture the unique characteristics of the coin shape and die information was suggested. Figure 1 demonstrates the challenges of the identification process. At the first glance, all three different coins bear the same characteristics (e.g., nose, mouth, hair or helmet pattern). However, all three coins are produced by different dies which is clearly visible on the respective helmet part of the coin.

For ancient coin recognition recently SIFT [28] features were used [51]. The authors presented evaluation on a small set of 350 coin images of three different coin types and achieved 84.24% classification and 76.41% identification

Fig. 1 Example for ancient coins produced from different dies

rate. In contrast to existing approaches, we aim at an extended evaluation of the identification efficiency of top performing local features on a data set of 2400 images of the same coin type. The application of local features in computer vision is manifold ranging from object [7] and texture recognition [25] to robot localization [33], symmetry detection [29], wide baseline stereo matching [46], and object class recognition [31]. In spite of their success and generality, these approaches are limited by the distinctiveness of the features and the difficulty of appropriate matching [7]. For a comprehensive survey and evaluation on the performance of local features in the context of their repeatability in the presence of rotation, scale, illumination, blur and viewpoint changes please refer to [32].

Recently, two algorithms have been applied to shape matching of ancient coins: a shape context description and a robust correlation algorithm [51]. Ancient coins are in general not of a perfect circular shape. From a numismatic point of view, the shape of a coin is a very specific feature. Thus, the shape serves as a first clue in the process of coin identification and discrimination. We are especially interested in the shape described by the edge of a coin given by the set of pixel positions sampled along it. The comparison of objects characterized by their shape representations is termed shape matching [50].

A comparison of shape descriptors based on curvature-scale space, wavelets, visual parts, Zernike moments, multilayer eigenvectors and directed acyclic graphs is given in [24]. In this study, curvature-scale space and visual parts outperformed the other approaches. Recently, further improvements were reported using a hierarchical likelihood cut-off scheme [30], using the shape context approach [2] to compare shapes based on the earth movers distance [27], and using hierarchical deformable shapes, the so-called shape tree [6]. We will present a method called deviation from circular shape matching (DCSM), i.e., the description of the shape border as the deviation from a circular shape, and compare it to the state-of-the-art shape context method. As DCSM makes use of background knowledge on coin shapes, e.g., coin shapes are close to circles, it is less suited for matching of general shapes but it performs significantly better on ancient coins.

3 Image-based ancient coin identification

The suggested coin identification method involves the steps of segmentation, shape and local descriptor extraction, matching of features and fusion of results.

3.1 Coin image segmentation

The separation of an object of interest from background is commonly termed segmentation. Due to textured background, presence of other objects in the image, inhomogeneous or poor illumination and low contrast, straightforward methods based on global image intensity thresholding tend to fail.

In situations where explicit knowledge on the properties of objects is available this knowledge can be used to steer segmentation parameters. For example, the compactness measure was used to find an intensity threshold in images showing circular spot welds [38]. Similarly, ancient coins were localized by thresholding the local intensity range, i.e., the difference between maximum and minimum gray-level in a local window [52].

Typically, the shape of modern coins is circular, whereas ancient coins deviate from this shape but still stay close to a circular outline. Therefore, we employ a measure of compactness c_t related to a threshold t defined as

$$c_t = 4\pi A_t / P_t^2, \tag{1}$$

where A_t is the area of the region covered by the coin and P_t is the perimeter of the coin, respectively. A_t and P_t are obtained by connected components analysis [44] applied to the binary image which is derived from thresholding the intensity range image. Figure 2a shows an intensity image of a coin, Fig. 2b is the corresponding intensity range image and Fig. 2c–g shows thresholded images for different selections of t along with calculated values for compactness c_t. The image thresholded at the optimal level t_{opt} with highest compactness is given in Fig. 2h. Figure 3 shows a typical plot of the relationship between compactness and gray-level range threshold, where the sudden decrease of the compactness measure is due to segmentation of the coin into several small regions, e.g., compare to Fig. 2g. Although Fig. 2c, d look similar, the low value for segmentation in Fig. 2c is due to the noisy border which results in a larger perimeter P_t and much smaller compactness c_t in Eq. 1.

3.2 Coin shape description and matching

The approach of shape comparison is based on a description of the difference between the shape of a coin and the shape of a circle. Therefore, the suggested approach is called deviation from circular shape matching (DCSM). In order to represent the coin shape, a border tracing on the binary image resulting from segmentation is performed. A list of border pixels is obtained and is resampled to l samples using equidistantly spaced intervals with respect to the arc length. Figure 4a–c shows this operation.

A one-dimensional descriptor, i.e., a curve describing the border, is obtained from fitting the coin edge to a circle and unrolling the polar distances between sample points and fitted circle into a vector. The center $s_c = (x_c, y_c)$ of the fitted circle is derived from the center of gravity and the radius r is the mean distance between the center and all sample points

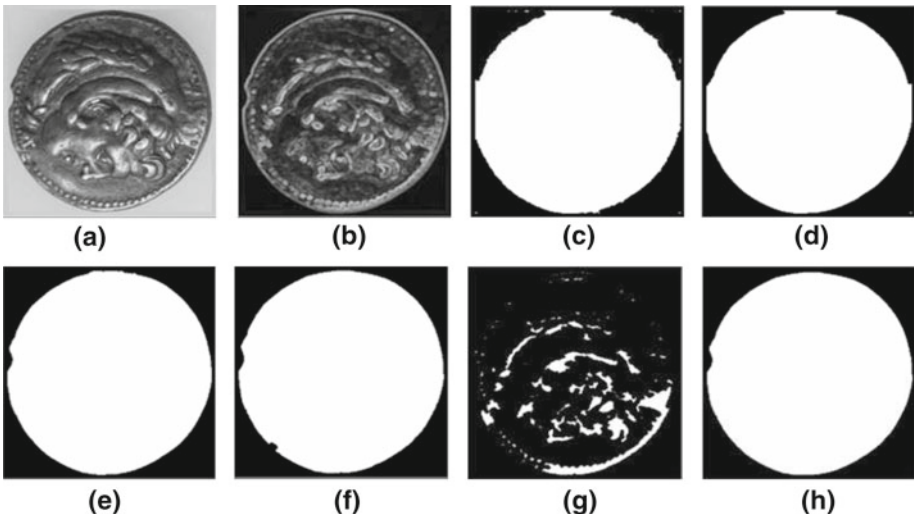


Fig. 2 Image of a coin, intensity range image and different binary images with corresponding threshold and compactness. **a** Intensity image. **b** Intensity range image. **c** Binary image: $t = 5,\ c_5 = 0.096$. **d** Binary image: $t = 25,\ c_{25} = 0.868$. **e** Binary image: $t = 45,\ c_{45} = 0.879$. **f** Binary image: $t = 65,\ c_{65} = 0.859$. **g** Binary image: $t = 85,\ c_{85} = 0.018$. **h** Binary image: $t_{\text{opt}} = 49,\ c_{49} = 0.888$

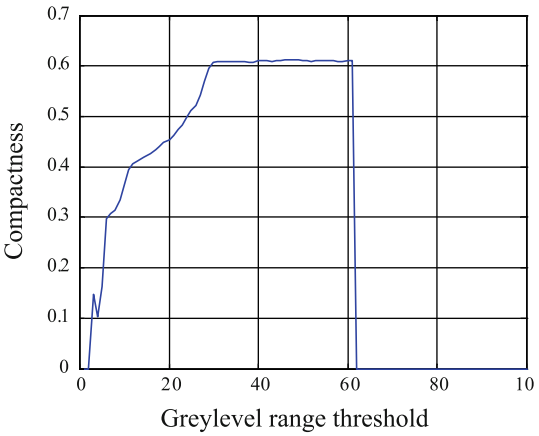


Fig. 3 Compactness versus intensity range threshold for a coin image

$s_i = (x_i, y_i)$ using

$$x_c = \frac{1}{l}\sum_{i=1,\ldots,l} x_i, \quad y_c = \frac{1}{l}\sum_{i=1,\ldots,l} y_i,$$
$$r = \frac{1}{l}\sum_{i=1,\ldots,l} \|s_i - s_c\|, \tag{2}$$

where (x_i, y_i) are the coordinates of sample point s_i and $\|\cdot\|$ denotes the L_2-norm. The 1D representation is given by $D = (d_1, \ldots, d_l)$, where

$$d_i = (\|s_i - s_c\| - r)/r, \quad i = 1, \ldots, l. \tag{3}$$

The division by r makes the representation invariant with respect to scale. Figure 4d shows the obtained 1D representation.

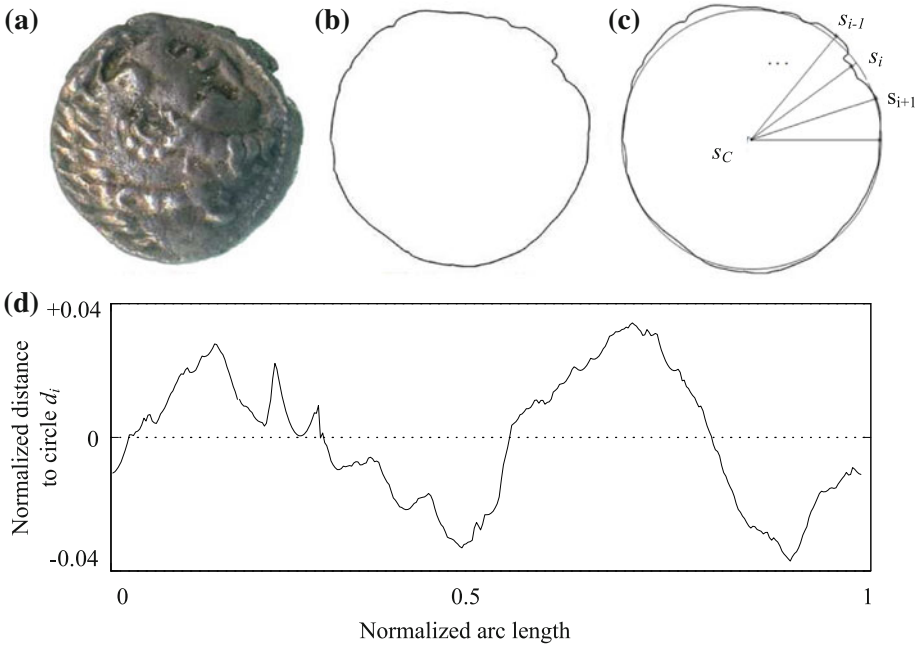


Fig. 4 Shape description: **a** ancient coin image, **b** coin edge image, **c** fitted circle and sampling along arc length, **d** normalized 1D description of coin shape

Efficient matching of shapes is performed using the fast Fourier transform (FFT) [3]. The shape descriptions of two coins are compared by a linear combination of global and local shape matching. The local matching is derived from the difference of Fourier shape descriptors, whereas the correlation coefficient between the curves serves as global measure of shape similarity.

The mean absolute or squared distance between the magnitude values of the Fourier coefficients is used as local measure of dissimilarity, i.e.,

$$D_L = \sum_{i=v,\ldots,l-u} \frac{\|\mathrm{sd_A}(i) - \mathrm{sd_B}(i)\|_p}{l-u-v+1}, \tag{4}$$

where $\|\cdot\|_p$, $p \in \{1,2\}$ is the L_p norm. The lower $v \geq 1$ and upper offsets $u \geq 0$ for the Fourier descriptors are small constants and used to limit errors stemming from imprecise circle fitting and quantization noise.

The global shape matching is obtained from a measure of dissimilarity or similarity, e.g., from the mean squared error (MSE) or the normalized cross-correlation (NCC) coefficient $\mathrm{ncc}(u)$ for a shift of u samples

$$\mathrm{ncc}(u) = \frac{\sum_{i=1,\ldots,l} d_A(i) \cdot d_B(i+u)}{\sqrt{\sum_{i=1,\ldots,l} d_A(i)^2 \cdot \sum_{i=1,\ldots,l} d_B(i)^2}}, \tag{5}$$

where $i+u$ might exceed l and modulo addition is applied. The maximum $D_G = (1 - \max_{i=1\ldots,l} \mathrm{ncc}(i))/2$ is used as a measure of global shape match. Similarly, the MSE is given by

$$\mathrm{mse}(u) = \frac{1}{l} \sum_{i=1,\ldots,l} (d_A(i) - d_B(i+u))^2 . \tag{6}$$

In the case of MSE, the maximum $D_G = \max_{i=1\ldots,l} \mathrm{mse}(i)$ is used as a measure of global dissimilarity. The position of the minimum of D_G is related to the rotation angle between the compared coins. While the MSE requires l shifts of the signal and l evaluations of Eq. 6, the NCC is computed in a more efficient way [26]. However, the computation of the MSE becomes efficient for large databases of shape descriptors using the lower bound and early abandon criteria suggested by Keogh et al. [21].

The overall measure of shape dissimilarity becomes

$$\mathrm{D}_{AB} = \alpha D_L + (1-\alpha) D_G, \tag{7}$$

where the weighting factor $\alpha \in [0, 1]$ controls the influence of local and global dissimilarity terms.

In order to be invariant with respect to mirroring, the D_G is replaced by the minimum of global dissimilarity obtained from matching the signal and the reversed signal. Mirror invariance enables the matching of coins irrespective of which side is shown on the image.

3.3 Local coin feature description and matching

Local image features provide a mathematical description of the image pattern in a window surrounding interest points and offer thereby a set of distinctive features for an image [47]. By matching corresponding features among coin image pairs, similarities can be detected and used for identification. An important advantage of local features is that they may be used to recognize an object despite significant clutter and occlusion. This is an essential requirement for the comparison of coin images since the metallic surface and relief structure of ancient coins make their appearance highly dependent on the illumination conditions. Although local features suffer from this high local variability (see Sect. 6.3) they are less vulnerable than global features due to their capabilities in partial matching.

As a consequence, we use the Scale Invariant Feature Transform (SIFT) [28] for the comparison of coin images. SIFT features are widely used and show an outstanding performance compared to other local features [32]. The SIFT descriptor is based on gradient distribution in salient regions. At each feature location, an orientation is selected by determining the maximum of the histogram of local image gradient orientations. Subpixel image location, scale and orientation are associated with each SIFT feature vector (4×4 location grid $\times$ 8 gradient orientations). Interest points are identified at peaks (local maxima and minima) of the Difference of Gaussians (DoG) at multiple scales. All key points with low contrast or key points that are localized at edges are eliminated using a Laplacian function.

Local features can be matched by identifying the first two nearest neighbors in Euclidean space as suggested by Lowe [28]. The idea is to identify the two nearest neighbors of a descriptor and measure the distance ratio of these two neighbors. Thus, a matching is only accepted if the distance to the nearest neighbor is considerably lower than the distance to the second nearest neighbor, i.e., the matching has a certain degree of unambiguity. An essential characteristic of this approach is that a descriptor can have several matches when different descriptors from the test image are matched against the same descriptor from the training image. To overcome this problem, one can either ignore all ambiguous matches (e.g. [43]) or keep the one with the lowest distance. In spite of the loss of potentially correct matches, our experiments show that the matching performance increases significantly when all ambiguous matches are discarded. Eventually, the number of correct matches is used to measure the similarity of two coins by local features.

To summarize, basically two steps are performed to measure the local feature based similarity of two coin images. First, interest points are detected in both images by finding peaks in the DoG scale space. Second, the SIFT descriptors of the interest points are matched between the two images

using the matching strategy described above. An empirically determined threshold is used to determine whether a match is regarded as valid and unique (see also Sect. 6.3). The resulting number of valid matches accounts for the measure of similarity provided by local features.

4 Proposed method

An identification system for coins should be reliable from the point of recognition quality and efficient from the point of computing resources. Computational efficiency is especially important when coin identification is done at border control or large collections of images, e.g., internet auction sites, are processed. Depending on the scenario, either the response time for the comparison of a single coin to a data set of moderate size or the total computation time when comparing large collections of coins should be kept low. From the point of reliability, the trade-off between false acceptance rate (FAR) and false rejection rate (FRR) is an important performance measure [11]. From the point of efficiency, multi-stage classifiers, where computational less demanding classifiers are used in early stages, have been used in various applications, e.g. [10]. In particular, in a two-stage classification system, the first-stage classifier limits the overall reliability of the whole system through inclusion of false positives passed to the second level classifier and false rejections of true positive examples. In the design of multi-stage classifiers the risk should be evaluated with respect to metrics such as false alarm and rejection rates, and precision and recall [40]. The trade-off between precision and recall determines the size of the preselection set, i.e., the data passed from the first stage to the second stage classifier. Naive Bayesian fusion is used for combination of individual classifier stages. Related strategies of multi-feature information integration are explored in [39].

4.1 Preselection

It takes 0.006 seconds to compare two coins based on their shape description on a Intel Core 2 CPU with 2.5 GHz. Therefore, shape matching is suited as a preselection step to the less efficient matching based on local features which typically takes two orders of magnitude longer [1]. The size of the preselection set is determined experimentally from Precision-Recall curves. Recall measures the ratio given by true positives divided by the sum of true positives and false negatives, i.e., rec = TP/(TP + FN) and precision is given by prec = TP/(TP + FP), where FP is the number of false positives. Figure 5 shows plots of precision versus recall for a test set of 2400 images of coins. The test set, which is described in more detail in Sect. 5, contains 240 different coins with 10 images of each coin. Different settings of the shape matching

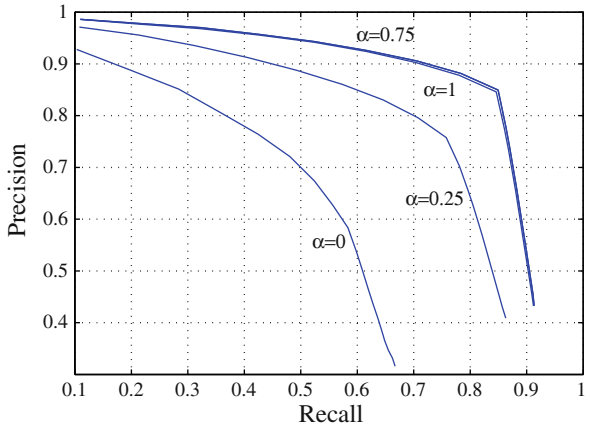


Fig. 5 Precision versus recall for shape matching

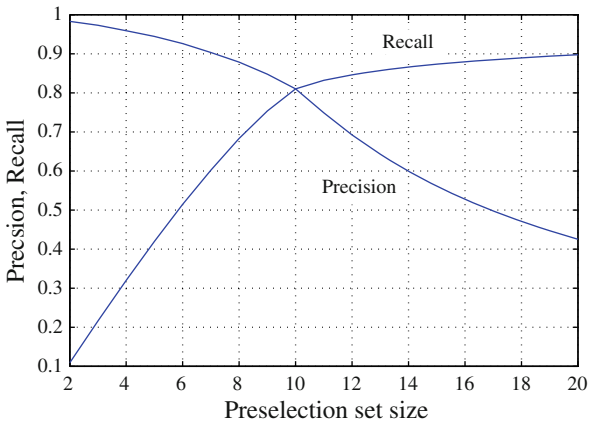


Fig. 6 Precision and recall versus preselection set size for shape matching

weight parameter α show that a relatively large value of α, which directs the matching dissimilarity towards more local influence, performs best. In order to obtain a preselection set of moderate size and high quality, i.e., the coin in question should likely be contained, a high recall is aspired. This is obtained by selecting the set size corresponding to the sudden decrease in Fig. 5. Figure 6 shows that this sudden decrease in precision versus recall corresponds to a preselection set size of 9 to 10 images. None to all coins of a specific class might be present in the preselection set, depending on the size of the preselection set and the the number of images of a coin in the data set.

4.2 Fusion of shape and local descriptors

Both classes of descriptors deliver distance measures between the coin in question and all other images in the database. In this case, a two-stage rank based strategy is possible, i.e., a small subset is preselected based on shape comparison and further processed using local features based matching. This strategy was explored in a previous paper [18]. A more

elaborate strategy combines probabilities which are derived from distance measures [17] and are combined by a rule of combination, e.g., the product rule [22]. The product rule of combination is equivalent to naive Bayes fusion of classifiers. Naive Bayes fusion of classifiers in turn coincides with Bayes classification over composite descriptors if the individual features are conditionally independent [42]. Conditional independence between shape and local features, as well as between coin sides, can be assumed.

From ranking the shape dissimilarity D_{AB} for shapes given in Eq. 7 for shape B matched to shape A results in a preselection set $\mathcal{P}$. From an observed shape description A we derive a conditional probability for a coin side label $\mathcal{L}$ assigned to B. The conditional probability for a $P_{\text{shape}}(\mathcal{L}|A)$ is estimated to be inversely proportional to the dissimilarity given in Eq. 7 between coin A and coin B labelled $\mathcal{L}$:

$$P_{\text{shape}}(\mathcal{L}|A) = \frac{1}{\mathrm{D}_{AB} \sum_{C \in \mathcal{P}} 1/\mathrm{D}_{AC}}, \tag{8}$$

where the summation term in the denominator accounts for normalization.

A similar argument is applied to derive a conditional probability for observed local descriptors X matched to local descriptors Y labelled $\mathcal{L}$ and corresponding to an image contained in the preselection set $\mathcal{P}$:

$$P_{\text{local}}(\mathcal{L}|X) = \frac{M_{XY}}{\sum_{Z \in \mathcal{P}} M_{XZ}} \tag{9}$$

where M_{XY} denotes the number of matches between the query image with local descriptors X and the coin side image with local descriptors Y and the denominator accounts for normalization.

As local and shape features describe different properties of a coin, it is reasonable to assume statistical independence between shape and local feature measurements. Thus, the combination is performed by the product rule [22]:

$$\begin{aligned} P(\mathcal{L}|A, X) &= P(\mathcal{L}_{\text{shape}} = \mathcal{L}_{\text{local}}|A, X) \\ &= P_{\text{shape}}(\mathcal{L}|A) \cdot P_{\text{local}}(\mathcal{L}|X), \end{aligned} \tag{10}$$

where $\mathcal{L}_{\text{shape}}$ and $\mathcal{L}_{\text{label}}$ are labels derived from shape and local descriptions.

4.3 Fusion of coin sides

The idea of fusion of different descriptor outputs is extended to a fusion of more than one image of a coin. Typically, a coin is presented by images of the obverse and reverse side. Additionally, for modern coins the designs on reverse and obverse sides are oriented either into the same direction or are turned upside down. In modern coin sorting machines this property is used to increase classification accuracy [17].

The arguments presented in Sect. 4.2 are extended to a fusion of coin sides in a straightforward fashion. Equation 10 is extended to the following four terms

$$\begin{aligned} P(\mathcal{L}|A_i, X_i) = {} & P_{\text{shape}}(\mathcal{L}|A_1) \cdot P_{\text{local}}(\mathcal{L}|X_1) \\ & \cdot P_{\text{shape}}(\mathcal{L}|A_2) \cdot P_{\text{local}}(\mathcal{L}|X_2), \end{aligned} \tag{11}$$

where A_i and X_i corresponds to shape and local feature descriptions of the ith coin side.

5 Data set and evaluation strategy

In order to evaluate shape matching techniques, the MPEG-7 core experiment CE-Shape-1 database part B [19] containing images of shapes with single closed contours is frequently used in the literature. Part B of the MPEG-7 CE-Shape-1 database contains a total number of 1400 binary images with 70 different classes, each of which contains 20 images. Figure 7 shows some of the images in the database. The recognition rate of an algorithm in this data set is commonly measured by the so called bull's-eye test. For every image in the database, the occurrences of images belonging to the same class among the 40 most similar images are counted for the query image. The final score of the test is the ratio of the overall number of correct hits divided by the maximum number of possible hits. In our case, the number of possible hits is $1400 \times 20 = 28{,}000$.

To evaluate our approach on coin data, we use an image database provided by the Fitzwilliam Museum, Cambridge, UK, which consists of 2400 images of 240 different ancient coins of the same class. Figure 8 shows four of the coins contained in the data set. Each row shows the same coin acquired by different devices at varying conditions and different orientations. In particular, each side of each individual coin was acquired at three different angles of rotation using a scanner device and two acquisitions of each side were made using a digital camera and varying illumination. At first sight, all coins bear the same characteristics. However, the coins shown in the different rows are produced by different dies. What makes this data set special and ideal to thoroughly test identification methods, is that all the coins are very similar. All the images are issued in the time of, or at least in the name of, Alexander the Great who came to power in Macedonia in 336 BCE and died as emperor in 323 BCE. Some of the coins are from much later and were minted in places around the Black Sea, in Egypt, in modern-day Turkey, Iran, etc. All coins follow the same basic standard: on the obverse side there is the head of Heracles in a lion skin. The reverse side shows the god Zeus, seated left on a throne. Nevertheless, there is a huge range of detail in the minor variations that experts use to deduce the mint and date of the coin.

The described data set of ancient coin images is available on the internet from the download section of the COINS project web page http://www.coins-project.eu.

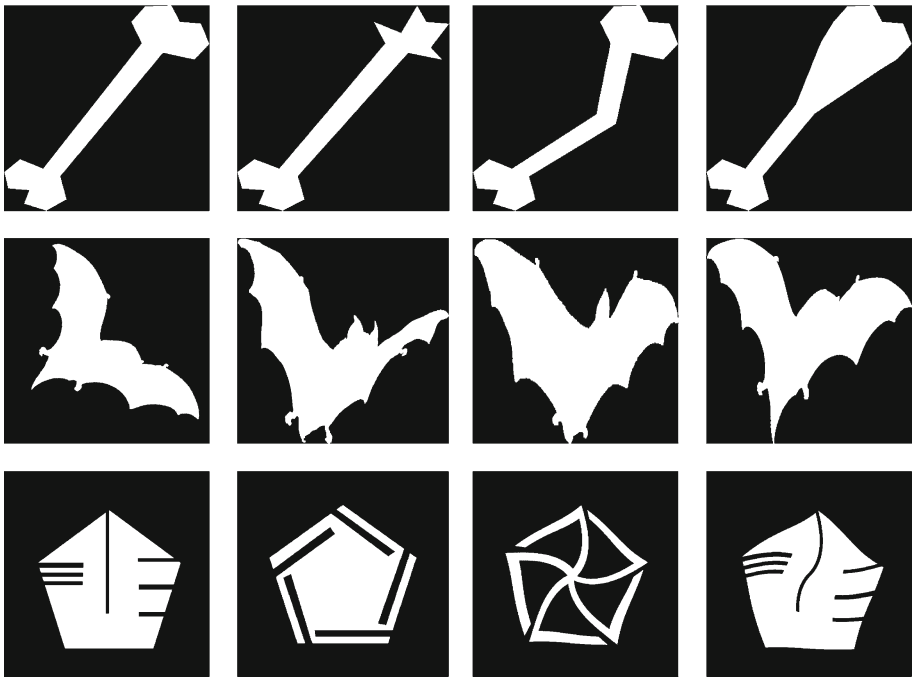


Fig. 7 Examples for binary images contained in MPEG-7 CE-Shape-1 data set

Fig. 8 Examples for ancient coin images acquired by scanner and camera

6 Results

We present a comparison of the DCSM method to a state-of-the-art shape matching method and results for combination of DCSM and SIFT features matching as well as the combination of shape matching results for reverse and obverse sides of coins.

6.1 Results of shape matching on MPEG7 shape data

Several publications on shape matching methods applied to the MPEG7 CE-Shape-1 database appeared recently. The method suggested in [27] is based on shape context (SC) description and earth movers distance (EMD) matching is reported to perform best. The suggested DCSM method is also compared to the SC description. For baseline comparison, we used the L_2 distance instead of the EMD. As EMD becomes especially powerful when it comes to articulated non-rigid shapes. As coins are rigid objects, we decided not to use EMD in order to keep our experiments comparable to the ancient coins data sets. Bull's-eye test results for SC and DCSM on the MPEG7 data set are reported in the first line of Table 1.

We used 100 sample points, 12 angle bins and 5 logarithmically scaled distance bins in SC matching. Clearly, DCSM is not well suited for the given task as it makes too much assumptions about the object class and, thus, performs worse. However, even though our method is not designed for such data, it achieves an accuracy better than some of the methods reported in [24]. According to the published numbers, it is slightly better than multilayer eigenvectors and Zernike moments, and even outperforms methods based on wavelets and directed acyclic graphs.

6.2 Results of shape matching on coin data

Results on the coins data set are reported using results from the bull's eye test as well as leave-N-out accuracy estimators. The best performance of DCSM was obtained using 256 sample points, a weighting parameter $\alpha = 0.975$, and lower and upper offsets of $v = 3, u = 3$. The measure for global shape dissimilarity calculation was the mean squared error, see Eq. 6. The shape context parameters are the same as for the MPEG7 shape data. The second line of Table 1 gives bull's eye test results for SC and DCSM applied to coins. In contrast to the MPEG7 shapes data set, the DCSM method is clearly superior for coins.

Table 1 Bull's-eye test results for shape-matching using SC and DCSM applied to MPEG7 CE Shape-1 and ancient coins

	SC (%)	DCSM (%)
MPEG7 CE Shape-1	76.79	71.75
Ancient coins	50.64	93.75

Table 2 Identification accuracy derived from Leave-N-out estimation applied to DCSM shape-matching of coins

	$N = 1k$	$N = 5k$	$N = 9k$
Accuracy (%)	99.00	98.18	90.29

Leave-N-out accuracy estimators for ancient coins are given in Table 2. The leave-one-out estimator is the average nearest-neighbor classification score obtained from matching each of the 2400 images to all other images. The leave-5k-out, where k is the number of coins, refers to training and test sets of equal sizes. The leave-9k-out result corresponds to training sets containing single images per coin. Note, that it is practically not possible to evaluate all possible combinations of partitions of images into training and test sets. Therefore, the leave-N-out accuracies are estimated from a smaller number of different partitions. It was shown, that such a procedure, also known as n-fold cross-validation, delivers accurate estimates for real-world data sets [23]. It is clearly visible, that even in the case of a single example per coin in the training set, the accuracy stays above 90%. The result of 99% for leave-1k-out accuracy estimator means that only 24 images are incorrectly classified. A detailed look at the incorrectly classified images revealed that they are characterized by a rough border. The rough border is caused by the discretization of the imaging process and disturbs the Fourier space description.

Another interesting feature of our approach is that, provided α is kept large, DCSM is largely invariant with respect to mirroring. Therefore, when using the obverse side of a coin as reference, the images showing the reverse side of the same coin are also found to be highly matching.

6.3 Results of local features matching on coin data

In contrast to a previous paper [18], local feature matching is based on SIFT features only and no further experiments with SURF [1] were conducted. As a result from the experiments, SIFT provides a higher matching rate and its drawback compared with SURF, the longer computational time, is of low impact as preselection from DCSM is performed.

When using only local features, coin identification can be achieved by determining the maximum number of matches between the query image and the images of the data set. On our data set matching of local features with the elimination of all ambiguous matches achieves an identification rate of 71.77%, whereas traditional nearest neighbor distance ratio matching achieves only 59.71%. Thus, for all further experiments ambiguous matches were eliminated. For the given coin data set best results are achieved using a distance ratio

Fig. 9 Example scan images

Table 3 Identification rates derived from leave-one-out accuracy estimation

Descriptor	DCSM	SIFT	DCSM + SIFT
Accuracy (%)	97.04	71.77	98.54

of 0.8, i.e., a matching is valid only if the ratio of the distances to the first and second nearest neighbor does not exceed 0.8.

A detailed look at misclassified coins demonstrates the boundaries of local features for the given image data. Figure 9 shows three scan pictures of the same coin. In spite of the scan as a high quality acquisition device, the physical measurements of the coin itself and deep die reliefs contribute to the significant differences in the acquired images. In all three images different regions of the coin are highlighted and produce different local features. Hence, matching across such images leads to low scores and possible mismatches.

6.4 Results of combined shape and local features matching

In this experiment we combine shape and local descriptors to increase the identification rate. Preselection based on shape matching allows for the restriction of required comparisons for local features matching. As a result we achieve speed up of the identification process and higher accuracy rate. Since our shape descriptor is mirroring invariant, preselection can be performed either on the whole available coin data, i.e., the preselected set can contain images of the second coin side, or preselection can be performed on the relevant coin side directly.

As a conclusion from the experiments reported in Sect. 4.2, the preselection size was set to 10. Therefore, for the experiments presented here, $P_{\text{shape}}(\mathcal{L}|A)$ is computed for the 10 images with lowest dissimilarity and $P_{\text{local}}(\mathcal{L}|X)$ for the same 10 images. The final decision is made according to the product rule given in Eq. 10.

Table 3 shows the identification rates for the single descriptors and their combination with a leave-one-out evaluation scheme. The shape-based preselection of size 10 was performed accordingly to the given side of the test coin image. The DCSM alone gives an identification rate of 97.04% on the whole data set of 2400 images. For a preselection size of 10, there are only 13 cases (0.54 %) where the correct coin is not contained in the preselected set. Consequently, local feature matching on the preselected set and fusion with the label probabilities from DCSM lead to an identification rate of 98.54 %.

Fig. 10 Image of obverse and reverse side of a coin

Table 4 Identification accuracy derived from leave-one-out estimation applied to DCSM shape matching of reverse and obverse coin side images and fusion of results

Side	Obverse	Reverse	Fused
Accuracy (%)	98.00	96.08	98.83

6.5 Fusion of matching results for coin sides

In the final experiment we study the fusion of reverse and obverse side based on shape. Therefore, the data set is grouped into pairs of images where each pair of images shows obverse and reverse coin sides as shown in Fig. 10. Note the low quality of some of the images present in the considered data set. The results for each side are combined using naive Bayes fusion. Table 4 lists the identifications results for the obverse side, the reverse side, and the fusion of both sides.

A further fusion with local features from the obverse and reverse side as described in Sect. 4.3 was tested as well, resulting in the same identification rate of 98.83%. Please note that the identification rate from DCSM corresponds to only 14 false identifications out of 1200 which strongly limits the scope of improvement by local features.

7 Conclusion

In this paper we presented an approach for object identification based on the combination of shape and local descriptors and applied it to the task of ancient coins identification. Due to their nature, the shape of ancient coins is a well distinguishable feature for an automatic identification. DCSM was used for the matching of coin shapes, whereas the die of the coin was matched by means of local features. From the output of each of these two methods individual coin label probabilities were estimated and finally fused. We presented

results on a data set of 2400 coin images. The combination of shape and local features outperformed the accuracy rate of the single features and achieved an identification rate of 98.83%. Future work will include the evaluation of the presented approach on an extended data set of coin images of different coin types. Furthermore, we will extend our work towards die and mint sign identification based on spatially constrained local features.

Acknowledgment The authors would like to thank Dr. Mark Blackburn from the Fitzwilliam Museum, Cambridge, UK, for providing test images and numismatic expertise.

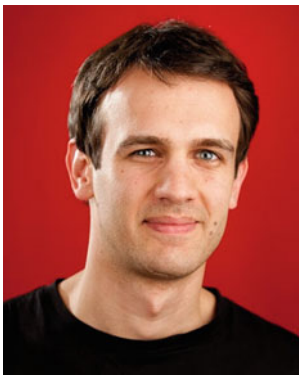
References

1. Bay, H., Tuytelaars, T., Gool, L.V.: SURF: speeded up robust features. In: Proceedings of European Conference on Computer Vision, LNCS, vol. 3951, pp. 404–417. Springer (2006)
2. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Anal. Mach. Intell. **24**(4), 509–522 (2002)
3. Cooley, J.W., Tukey, J.W.: An algorithm for the machine calculation of complex fourier series. Math. Comput. **19**, 297–301 (1965)
4. Davidsson, P.: Coin classification using a novel technique for learning characteristic decision trees by controlling the degree of generalization. In: Proc. of Int. Conf. on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems, pp. 403–412 (1996)
5. Elkins, N.T.: A survey of the material and intellectual consequences of trading in undocumented ancient coins: a case study on the north american trade. Frankfurter elektronische Rundschau zur Altertumskunde **7**, 1–13 (2008). http://s145739614.online.de/fera/ausgabe7/Elkins.pdf
6. Felzenszwalb, P., Schwartz, J.: Hierarchical matching of deformable shapes. In: Computer Vision and Pattern Recognition, pp. 1–8 (2007)
7. Ferrari, V., Tuytelaars, T., Gool, L.V.: Simultaneous object recognition and segmentation from single or multiple model views. Int. J. Comput. Vis. **67**(2), 159–188 (2006)
8. Fukumi, M., Omatu, S., Takeda, F., Kosaka, T.: Rotation-invariant neural pattern recognition system with application to coin recognition. IEEE Trans. Neural Netw. **3**, 272–279 (1992)
9. Fürst, M., Kronreif, G., Wögerer, C., Rubik, M., Holländer, I., Penz, H.: Development of a mechatronic device for high speed coin sorting. In: Proc. of IEEE Int. Conf. on Industrial Technology, vol. 1, pp. 185–189. Maribor, Slovenia (2003)
10. Giusti, N., Sperduti, A.: Theoretical and experimental analysis of a two-stage system for classification. IEEE Trans. Pattern Anal. Mach. Intell. **24**(7), 893–904 (2002)
11. Golfarelli, M., Maio, D., Maltoni, D.: On the error-reject trade-off in biometric verification systems. IEEE Trans. Pattern Anal. Mach. Intell. **19**(7), 786–796 (1997)
12. Goodman, M.: Numismatic Photography. Zyrus Press (2008)
13. Hibari, E., Arikawa, J.: Coin discriminating apparatus. European Patent EP1077434 (2001)
14. Hoberman, G.: The Art of Coins and Their Photography. Lund Humphries (1982)
15. Hoßfeld, M., Chu, W., Adameck, M., Eich, M.: Fast fast 3D-vision system to classify metallic coins by their embossed topography. Electron. Lett. Comput. Vis. Image Anal. **5**(4), 47–63 (2006)
16. Howgego, C.J.: The potential for image analysis in numismatics. In: Images and Artefacts of the Ancient World, pp. 109–113 (2005)
17. Huber, R., Ramoser, H., Mayer, K., Penz, H., Rubik, M.: Classification of coins using an eigenspace approach. Pattern Recogn. Lett. **26**(1), 61–75 (2005)
18. Huber-Mörk, R., Zaharieva, M., Czedik-Eysenberg, H.: Numismatic object identification using fusion of shape and local descriptors. In: Proceedings of International Symposium on Visual Computing, pp. 368–379, Las Vegas, NV, USA (2008)
19. Jeannin, S., Bober, M.: Description of core experiments for mpeg-7 motion/shape. Technical Report ISO/IEC JTC 1/SC 29/WG 11 MPEG99/N2690 (1999)
20. Kampel, M., Huber-Mörk, R., Zaharieva, M.: Image-based retrieval and identification of ancient coins. IEEE Intell. Syst. **24**(2), 26–34 (2009)
21. Keogh, E., Wei, L., Xi, X., Vlachos, M., Lee, S.H., Protopapas, P.: Supporting exact indexing of arbitrarily rotated shapes and periodic time series under euclidean and warping distance measures. VLDB J. **18**(3), 611–630 (2009)
22. Kittler, J., Hatef, M., Duin, R., Matas, J.: On combining classifiers. IEEE Trans. Pattern Anal. Mach. Intell. **20**(3), 226–239 (1998)
23. Kohavi, R.: A study of cross-validation and bootstrap for accuracy estimation and model selection. In: Proceedings of International Joint Conference of Artificial Intelligence, pp. 1137–1145 (1995)
24. Latecki, L.J., Lakämper, R., Eckhardt, U.: Shape descriptors for non-rigid shapes with a single closed contour. In: Proceedings of Conference on Computer Vision and Pattern Recognition, pp. 424–429 (2000)
25. Lazebnik, S., Schmid, C., Ponce, J.: A sparse texture representation using local affine regions. IEEE Trans. Pattern Anal. Mach. Intell. **27**(8), 1265–1799 (2005)
26. Lewis, J.: Fast normalized cross-correlation. In: Proceedings of Vision Interface, pp. 120–123 (1995)
27. Ling, H., Okada, K.: An efficient earth mover's distance algorithm for robust histogram comparison. IEEE Trans. Pattern Anal. Mach. Intell. **29**(5), 840–853 (2007)
28. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)
29. Loy, G., Eklundh, J.O.: Detecting symmetry and symmetric constellations of features. In: Proceedings of European Conference on Computer Vision, LNCS, vol. 3952, pp. 508–521. Springer (2006)
30. McNeill, G., Vijayakumar, S.: 2D shape classification and retrieval. In: International Joint Conference on Artificial Intelligence, pp. 1483–1488 (2005)
31. Mikolajczyk, K., Leibe, B., Schiele, B.: Local features for object class recognition. In: Proceedings of IEEE International Conference on Computer Vision, vol. 2, pp. 1792–1799 (2005)
32. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE Trans. Pattern Anal. Mach. Intell. **27**(10), 1615–1630 (2005)
33. Murillo, A.C., Guerrero, J.J., Sagüés, C.: SURF features for efficient robot localization with omnidirectional images. In: Proceedings of IEEE International Conference on Robotics and Automation, pp. 3901–3907 (2007)
34. Neubarth, S., Gerrity, D., Waechter, M., Everhart, D., Phillips, A.: Coin discrimination apparatus. Canadian Patent CA2426293 (1998)
35. Nölle, M., Penz, H., Rubik, M., Mayer, K.J., Holländer, I., Granec, R.: Dagobert—a new coin recognition and sorting system. In: Proceedings of International Conference on Digital Image Computing—Techniques and Applications, pp. 329–338 (2003)
36. Onodera, A., Sugata, M.: Coin discrimination method and device. United States Patent US2002005329 (2002)
37. Reisert, M., Ronneberger, O., Burkhardt, H.: An efficient gradient based registration technique for coin recognition. In: Proceedings of Muscle CIS Coin Competition Workshop, pp. 19–31. Berlin, Germany (2006)

38. Ruisz, J., Biber, J., Loipetsberger, M.: Quality evaluation in resistance spot welding by analysing the weld fingerprint on metal bands by computer vision. Int. J. Adv. Manuf. Tech. **33**(9-10), 952–960 (2007)
39. Saber, E., Tekalp, A.M.: Integration of color, edge, shape and texture features for automatic region-based image annotation and retrieval. Electron. Imaging **7**, 684–700 (1998)
40. Senator, T.: Multi-stage classification. In: Proceedings of International Conference on Data Mining (2005)
41. Shah, G., Pester, A., Stern, C.: Low power coin discrimination apparatus. Canadian Patent CA1336782 (1986)
42. Shi, X., Manduchi, R.: A study on Bayes feature fusion for image classification. In: Proceedings of Computer Vision and Pattern Recognition Workshop, pp. 95–103, Madison, WI, USA (2003)
43. Sivic, J., Schaffalitzky, F., Zisserman, A.: Object level grouping for video shots. Int. J. Comput. Vis. **67**(2), 189–210 (2006)
44. Sonka, M., Hlavac, V., Boyle, R.: Image Processing, Analysis, and Machine Vision, 2nd edn. PWS-an Imprint of Brooks and Cole Publishing (1998)
45. Tsuji, K., Takahashi, M.: Coin discriminating apparatus. European Patent EP0798669 (1997)
46. Tuytelaars, T., Gool, L.V.: Matching widely separated views based on affine invariant regions. Int. J. Comput. Vis. **59**(1), 61–85 (2004)
47. Tuytelaars, T., Mikolajczyk, K.: Local Invariant Feature Detectors: A Survey. Now Publishers (2008)
48. Van Der Maaten, L., Postma, E.: Towards automatic coin classification. In: Proceedings of Conference on Electronic Imaging and the Visual Arts, pp. 19–26, Vienna, Austria (2006)
49. Vassilas, N., Skourlas, C.: Content-based coin retrieval using invariant features and self-organizing maps. In: Proceedings of International Conference on Artificial Neural Network, pp. 113–122 (2006)
50. Veltkamp, R.C.: Shape matching: Similarity measures and algorithms. Technical Report UU-CS Ext. rep. 2001-03, Utrecht University: Information and Computing Sciences, Utrecht, The Netherlands (2001)
51. Zaharieva, M., Huber-Mörk, R., Nölle, M., Kampel, M.: On ancient coin classification. In: Proceedings of International Symposium on Virtual Reality, Archaeology and Cultural Heritage, pp. 55–62 (2007)
52. Zambanini, S., Kampel, M.: Robust automatic segmentation of ancient coins. In: Proceedings of International Conference on Computer Vision Theory and Applications (VISAPP'09), vol. 2, pp. 273–276 (2009)

Author Biographies

Reinhold Huber-Mörk received his Ph.D. in computer science with honours from the University of Salzburg in 1999. Before joining the Austrian Institute of Technology GmbH (AIT) in 2006 he was researcher at Aerosensing Radarsysteme GmbH, c/o German Aerospace Research Establishment (DLR) and the Advanced Computer Vision GmbH (ACV), Vienna, Austria. RHM has published 50 papers on pattern recognition and image processing with applications in remote sensing and machine vision.



Sebastian Zambanini received his B.Sc. degree in media and computer science and M.Sc. degree in computer graphics and digital image processing from the Vienna University of Technology, Austria, in 2004 and 2007, respectively. Currently, he is pursuing Ph.D. degree and working as a research assistant at the Computer Vision Lab, Institute of Computer Aided Automation, Vienna University of Technology. He has collaborated in various computer vision research projects addressing the areas of medical image processing, surveillance and cultural heritage. His research interests comprise object recognition, object tracking and human pose estimation.

Maia Zaharieva is a research assistant at the Interactive Media Systems Group at the Institute of Software Technology and Interactive Systems at the Vienna University of Technology. She received her M.Sc. degree in Business Informatics from the University of Vienna in 2003. From 2003 to 2007 she was a member of the scientific staff of the Multimedia Information Systems Group at the University of Vienna. Since 2007 she is working toward her Ph.D. degree at the Vienna University of Technology. She has been active in various national and international research projects ranging from e-learning and digital preservation to image and video analysis. Her research interests include visual media analysis, processing, and retrieval.

Martin Kampel received his B.Sc. degree in Data Technologies and Computer Science; the M.Sc. degree (Diplom Ingenieur) in Computer Science (Computer Graphics, Pattern Recognition & Image Processing) in 1999; Ph.D. degree in Computer Science in 2003; and the "venia docendi" (habilitation) in Applied Computer Science in 2009, all from the Vienna University of Technology. He is an associate professor (Privatdozent) of computer vision at the Institute for Computer Aided Automation, Vienna University of Technology, and is currently engaged in research, project leading, industry consulting and teaching. His research interests are 3D Vision and Cultural Heritage Applications, Visual Surveillance, and Image Sequence Analysis. He has edited two proceedings and is the author or coauthor of more than 80 scientific publications presented at several international conferences and workshops, and is a member of the IAPR and the IEEE.