Numismatic Object Identification Using Fusion of Shape and Local Descriptors*

R. Huber-Mörk[1], M. Zaharieva[2], and H. Czedik-Eysenberg[1]

[1] Austrian Research Centers GmbH - ARC, smart systems Division, Business unit High Performance Image Processing, A-2444 Seibersdorf, Austria
reinhold.huber@arcs.ac.at
[2] Vienna Univ. of Technology, Institute for Computer Aided Automation, Pattern Recognition & Image Processing Group, A-1040 Vienna, Austria
maia@prip.tuwien.ac.at

Abstract. Reliable object identification is an essential task in the process of recognition and traceability of stolen cultural heritage. Existing approaches for object recognition focus mainly on object classification. However, they are not sufficient to identify a given object among hundreds of objects of the same class. In this paper, we investigate the feasibility of computer aided identification of ancient coins. Since the shape of a coin is a very unique feature, we first apply a shape descriptor to capture its characteristics. In the next step, local features are used to describe the die information. We present experiments on a data set of 2400 images of ancient coins. The evaluation results show that our approach is competitive. Moreover, it indicates some outstanding features that show great promise for reliable object identification in the area of cultural heritage.

1 Introduction

Illegal trade and theft of coins appears to be a major part of the illegal antiques market. Methods for computer aided identification of ancient coins become significant for the areas of cultural heritage and law enforcement. Traditionally, coin identification is performed by manual search in auctions catalogues. Nowadays, there are large digital coin collections available, e.g. the collection of coins and medals at the Fitzwilliam museum, Cambridge, UK (http://www.fitzmuseum.cam.ac.uk/dept/coins/) contains more than 40 000 digital coin descriptions. 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 sings, shape, scratches, wearing, etc.

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

G. Bebis et al. (Eds.): ISVC 2008, Part II, LNCS 5359, pp. 368–379, 2008.

Fig. 1. Examples for ancient coin images acquired by scanner and camera

Several authors address the image based recognition of modern coins, e.g. using artificial neural networks [1], edge features [2,3], gradient directions [4], eigenspaces [5], and color, shape and wavelet features [6]. Recently, [7,8] focus on approaches for the classification of ancient coins. 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. In this paper, we focus on the identification of ancient coins. Our approach relies on the combination of shape and local descriptors to capture the unique characteristics of the coin shape and die information, respectively.

Figure 1 demonstrates the challenges of the identification process. Each row shows the same coin acquired by different devices at varying conditions and different orientations. At very local level, all coins bear the same characteristics. However, the coins shown in the different rows are produced by different dies. This makes the data set special and ideal to thoroughly test identification methods. All the images are coins issued in the time of, or at least in the name of, Alexander the Great. Some of the coins are from much later and were minted in places around the Black Sea, in Egypt, in modern-day Turkey, Iran and so on. All coins follow the same basic standard: On the obverse side, there is a head of Heracles in a lion-skin. On the reverse is 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 remainder of the paper is organized as follows: Section 2 presents recent approaches for object recognition based on shape and local descriptors. The identification workflow is described in Section 3. Section 4 presents the performed evaluation. The achieved results are discussed in Section 5. Finally, in Section 6 conclusions and outlook for further research are drawn.

2 Related Work

In this section we present research approaches for object recognition based on shape and local descriptors.

2.1 Shape Descriptors

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 the coin given by the set of pixel positions sampled along it. The comparison of objects characterized by their shape representations is termed shape matching [9].

In order to evaluate shape matching techniques the MPEG-7 core experiment CE-Shape-1 database part B [10] containing images of shapes with single closed contours is frequently used in the literature. A comparison of shape descriptors based on curvature-scale space, wavelets, visual parts, Zernike moments multilayer eigenvectors and directed acyclic graphs is given in [11]. 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 [12], using the shape context approach [13] to compare shapes based on the earth movers distance [14], and using hierarchical deformable shapes, the so-called shape tree [15].

Recently, two algorithms have been applied to shape matching of ancient coins: the shape context description and a robust correlation algorithm [8]. Based on the basic ideas of the latter approach, i.e. the description of the shape border as the deviation from a circular shape, we will present a method called deviation from circular shape matching (DCSM) 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 performed significantly better on ancient coins.

2.2 Local Features

The application of local features in the computer vision is manifold ranging from object [16] and texture recognition [17] to robot localization [18], symmetry detection [19], wide baseline stereo matching [20], and object class recognition [21]. In spite of their success and generality, these approaches are limited by the distinctiveness of the features and the difficulty of appropriate matching [16]. For through 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 [22].

In [23] local features was successfully applied for the identification of footprint patterns. The authors use Maximally Stable Extremal Region (MSER) [24] detector to locate interest points. Following, their appearance is captured by Scale Invariant Feature Transform (SIFT) [25] and Gradient Location and

Orientation Histogram (GLOH) [22] descriptors. On a dataset of 368 different footwear patterns the authors report 85% correct classification rate. Recently, SIFT features were applied for the recognition of ancient coins[8]. The authors present evaluation on a small set of 350 coin images of three different coin types and achieved 84.24% classification and 76.41% identification rate. In contrast to existing approaches, we aim at extended evaluation of the identification efficiency of top performing local features, SIFT [25] and SURF [26], on a data set of 2400 images of the same coin type.

3 Identification Workflow

In this section we describe the main steps of the object identification workflow shown in Fig. 2.

3.1 Object Segmentation

Prior to identification or classification the location of the coin contained in the image is required. 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. Adaptive thresholding overcomes the shortcomings of single level thresholding to some extend, e.g. in [8] a method based on adaptive thresholding using a potential surface [27] was employed for the segmentation of coins.

In situations, where explicit knowledge on the properties of objects is available, this knowledge can be used to steer segmentation parameters. For example, in [28] the compactness measure is used to find an intensity threshold in

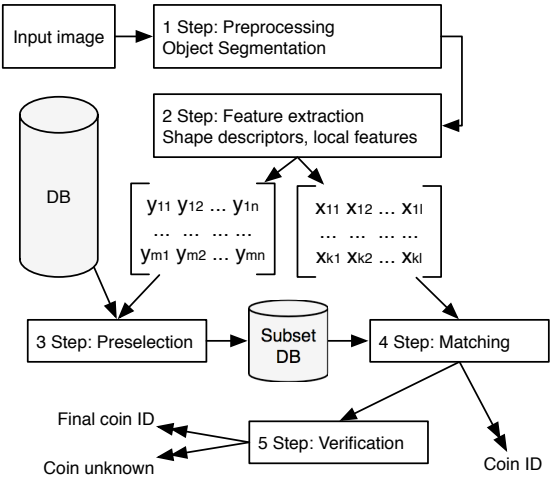

Fig. 2. Object identification workflow

images showing circular spot welds. Similarly, in [29] ancient coins are localized by thresholding the local entropy and range. We apply a similar segmentation procedure based on the use the local range feature.

Typically, the shape of modern coins is close to a circle, 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 greylevel threshold t using $c_t = 4\pi A_t / P_t^2$, where A_t is the area of the region covered by a coin and P_t is the perimeter of the coin, respectively. The perimeter P_t is derived from tracing the outer boundary of the binary region using 8-connectivity. The best threshold t is the one minimizing c_t.

3.2 Shape Descriptor

Based on the segmentation of the object, in our case a coin, we perform a border tracing on the binary image B and obtain a list of border pixels $P = (p_1, \ldots, p_k)$. In order to ensure scale invariance, the border pixels are resampled to l samples $S = (s_1, \ldots, s_l)$ using equidistantly spaced intervals with respect to the arc length. A one-dimensional descriptor D, i.e. a curve describing the border, is obtained from fitting 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 $s_i = (x_i, y_i)$. The 1D representation is given by $D = (d_1, \ldots, d_l)$, where $d_i = (\|s_i - s_c\| - r)/r, i = 1, \ldots, l$. The division by r makes the representation invariant with respect to scale.

Two shape descriptions $D_A = (d_A(1), \ldots, d_A(l))$ and $D_B = (d_B(1), \ldots, d_B(l))$, corresponding to coins A and B, are compared by a linear combination of a matching based on a global and a local shape description. The local description is derived from the difference of Fourier shape descriptors, whereas the correlation coefficient between curves D_A, D_B serves as global descriptor. Let $F = (f(i), \ldots, f(m))$ be the complex-valued Fourier frequency domain coefficients of the 1D signal D obtained by $F = \mathcal{FFT}(D)$. Invariance with respect to rotation is obtained by taking the magnitude of each $f(i)$. Therefore, a local shape descriptor for coin A is defined by $\mathrm{SD_A} = (\mathrm{sd_A}(1), \ldots, \mathrm{sd_A}(m))$, where $\mathrm{sd_A}(i) = |f_A(i)|, i = 1, \ldots, m$.

The global shape description is obtained from the normalized cross correlation (NCC) coefficient $ncc(u)$ for a shift of u samples. The maximum $NCC = \max_{i=1\ldots,l} \mathrm{ncc}(i)$ is used as a measure of global shape match.

The dissimilarity quantifies the shape matching cost of two coins A and B

$$\mathrm{DISS_{AB}} = \alpha \sum_{i=v,\ldots,m-u} \frac{\|\mathrm{sd_A}(i) - \mathrm{sd_B}(i)\|_p}{m-u-v+1} + \beta \frac{1-NCC}{2},$$

where $\beta = 1 - \alpha, \alpha \in [0,1]$ and $\|\cdot\|_p, p \in \{1,2\}$ is the L_p norm. The weighting parameters α and β are experimentally found. The weighting factors α and β control the influence of global and local dissimilarity. 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.

3.3 Local Descriptors

The Scale Invariant Feature Transform (SIFT) [25] descriptor is based on gradient distribution in salient regions. At each feature location, an orientation is selected by determining the peak 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 Gaussian function applied in scale space. All keypoints with low contrast or keypoints that are localized at edges are eliminated using a Laplacian function.

Speeded Up Robust Features (SURF) [26] are fast scale and rotation invariant features. The descriptor captures distributions of Haar-wavelet responses within the neighborhood of an interest point. Each feature descriptor has only 64 dimensions which results in fast computation and comparison (4×4 location grid $\times$ 4 wavelet responses in horizontal and vertical direction).

Local descriptors are matched by identifying the first two nearest neighbors in Euclidean space. A descriptor is accepted only if the distance ratio to the second nearest neighbor is below a given threshold [26,25]. 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. The overcome this problem, one can either ignore all ambiguous matches (e.g. [30]) or keep the one with 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, a coin is identified as its nearest neighbor by mean of most detected matches.

4 Experiments

In this section we shortly present the image databases used in the performed evaluations. Following, we present the results achieved using shape and local descriptors as single and combined approaches.

4.1 MPEG7 CE-Shape-1 Database

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. 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

4.2 Ancient Coins Database

The data set used in the experiments is an image database provided by the Fitzwilliam Museum, Cambridge, UK, which consists of 2400 images of 240 different ancient coins. Both sides of each coin were acquired 5 times using different

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

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

techniques. Three scans in different rotations with an offset of 120° and two photos under varied illumination conditions of each coin side are available. The scans provide an image resolution of up to 700×700 pixels. The images acquired by a camera are of lower quality, i.e. maximum of 350×350 pixels, and some of them are very blurred. Figure 1 showed examples for coins contained in the data set, each row shows different images and sides of the same coin.

4.3 Shape-Based Identification

Several publications on shape matching methods applied to the MPEG7 CE-Shape-1 database appeared recently. The method suggested in [14] 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, which coins neither are, 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 [11]. 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. 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 $\alpha = 0.025$, lower and upper offsets are $v = 3, u = 3$. The shape context parameters are the same as for them 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 3. Identification rate derived from leave-one-out estimation applied to SIFT and SURF

	Scan	Camera	Average		Scan	Camera	Average
SIFT NND-1	42.79%	86.36%	59.71%	SURF NND-1	13.23%	57.08%	30.54%
SIFT NND-2	60.58%	89.33%	71.77%	SURF NND-2	39.69%	81.60%	56.24%

Leave-N-out accuracy estimators for ancient coins are given in Tab. 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 classes, refers to training and test sets of equal sizes. The leave-$9k$-out result corresponds to training sets containing single images per class. It is clearly visible, that even in the case of a single example per class in the training set, the accuracy stays above 90 percent. The result of 99 percent for leave-one-out accuracy estimator means that only 24 images are wrongly classified. A detailed look at the wrongly 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, provided α is kept low, 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.

4.4 Local Features-Based Identification

SIFT features are 128 dimensional vectors which describe the neighborhood of interest points. The high dimensionality of the SIFT descriptor and the large amount of interest points detected is a drawback of SIFT at the matching step. On the contrary, SURF features are efficiently computed using integral images and they use 64 dimensional vector. For the data set of ancient coins the amount of points detected with SURF is reduced by 70% on the average. However, the reduction of information captured leads to loss of valuable information. We compared both descriptors in the context of two matching strategies and for different training sets.

Table 3 summarizes the results for the leave-one-out accuracy estimator. Matching of local features which eliminates all ambiguous matches (NND-2) outperforms traditional nearest neighbor distance ratio matching (NND-1) by approximately 20%. For the given coin data set best results are achieved using distance ratio of 0.8.

Noteworthy is the difference in the accuracy depending on the type of test image, see Tab. 3. It should be expected that high resolution scan images achieve higher accuracy rate than camera images. However, a detailed look at misclassified coins reveals a displacement of the boundaries of local features against the given image data. In spite of the scan as a high quality acquisition device, the

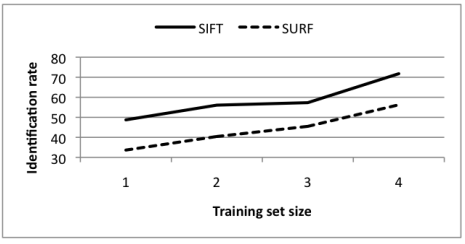

Fig. 3. Comparison of averaged identification rates for different training sets

physical measurements of the coin itself and deep die relief contribute to the significant differences in the acquired images.

Figure 3 shows the identification rate of both descriptors for different sizes of training sets. In every iteration of the experiment N random images per coin are used as training set. Since each coin side is pictured five times, $N = 1, \ldots, 4$. The presented results are averaged rates over 10 iterations for each N. The number of iterations was determined experimentally. Further iterations did not have any significant impact on the results. The results show that SIFT clearly outperforms SURF independently of size of training set and matching strategy. However, the quality and size of training set bear an essential influence on the identification performance.

4.5 Combined Approach

In our final 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 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. Figure 4 summarizes the results derived from leave-one-out estimation for different preselection sets based on the whole available coin set.

On the contrary to the previously presented results, the identification rate of both SIFT and SURF features lie close to each other. This is due to the high accuracy of the shape descriptor. Preselection of up to 10 image candidates discards a major part of the mismatches of local features. At this level, SURF becomes a competitive approach again, since it has low computational time and significantly high accuracy rate.

Table 4 shows the identification rates for the single descriptors and their combination. The shape-based preselection of size 5 was performed accordingly to the given side of the test coin image. Note, the given number for DCSM is lower than in Tab. 2. As DCSM is designed to match coins irrespective from the presented side, the accuracy obviously drops in the considered setting. The combination of both descriptors clearly holds the top performing accuracy rate of 95.16% correctly identified coins.

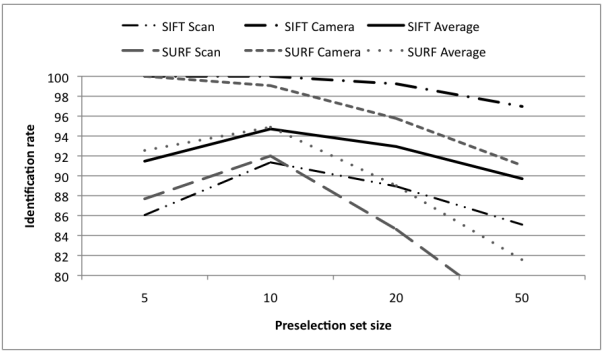

Fig. 4. Identification accuracy using shape descriptors as preselection

Table 4. Identification rates derived from leave-one-out accuracy estimation

Shape descriptor	Local descriptor		Combination	
DCSM	SIFT	SURF	DCSM + SIFT	DCSM + SURF
90.23%	71.77%	56.24%	94.11%	95.16%

5 Discussion

From numismatic point of view, restricting the range of possible classes an unknown ancient coin can be assigned to, is already of advantage. As the presented results show, an automated identification of coins is a feasible task for the computer vision. An essential impact on the accuracy rate bear the image quality of the training set and the image acquisition process. The combination of shape and local descriptors takes advantage of the very specific nature of ancient coins. Both the manufacturing process (hammer-struck from manually engraved dies) as well as the alteration process (wear, scratches, etc.) contribute to the uniqueness of an ancient coin. Since a shape descriptor focuses on the unique feature and narrows the identification process, local descriptors capture die information and assure the final decision by selecting the correct coin type. Using solely local feature matching will not solve the identification problem since local patterns appear repeatedly not only on the same coin but also across different coins. In the scenarios of coin and coin die identification the spatial relationship of local features is an essential characteristic. Further research is required to optimize spatial constrained local features to capture die specific information (e.g. mint signs).

6 Conclusions

We presented a a system for the identification of ancient coins based on the combination of shape and local descriptors. Due to their nature, the shape of an ancient coin is a well distinguishable feature for an automatic identification.

Furthermore, local features are used to capture die information and discriminate coins of different coin types. We presented results on a data set of 2400 coin images. Independently of the applied local descriptor the combination of shape and local features outperform the accuracy rate of the single features and achieved an identification rate of 95.16%.

References

1. Fukumi, M., Omatu, S., Takeda, F., Kosaka, T.: Rotation-invariant neural pat. rec. system with application to coin recognition. IEEE Trans. on Neur. Netw. 3, 272–279 (1992)
2. Nölle, M., Penz, H., Rubik, M., Mayer, K.J., Holländer, I., Granec, R.: Dagobert – a new coin recognition and sorting system. In: Proc. of Intl. Conf. on Digital Image Computing – Techniques and Applications, pp. 329–338 (2003)
3. Van Der Maaten, L., Poo, P.: COIN-O-MATIC: A fast system for reliable coin classification. In: Proc. of the Muscle CIS Coin Competition Workshop, Berlin, Germany, pp. 7–18 (2006)
4. Reisert, M., Ronneberger, O., Burkhardt, H.: An efficient gradient based registration technique for coin recognition. In: Proc. of Muscle CIS Coin Competition Workshop, Berlin, Germany, pp. 19–31 (2006)
5. Huber, R., Ramoser, H., Mayer, K., Penz, H., Rubik, M.: Classification of coins using an eigenspace approach. Pat. Rec. Let. 26, 61–75 (2005)
6. Vassilas, N., Skourlas, C.: Content-based coin retrieval using invariant features and self-organizing maps. In: Intl. Conf. on Artif. Neur. Netw., pp. 113–122 (2006)
7. Van Der Maaten, L., Postma, E.: Towards automatic coin classification. In: Proc. of Conf. on Electronic Imaging and the Visual Arts, Vienna, Austria, pp. 19–26 (2006)
8. Zaharieva, M., Huber-Mörk, R., Nölle, M., Kampel, M.: On ancient coin classification. In: Intl. Symp. on Virtual Reality, Archaeology and Cultural Heritage, pp. 55–62 (2007)
9. 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)
10. 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)
11. Latecki, L.J., Lakämper, R., Eckhardt, U.: Shape descriptors for non-rigid shapes with a single closed contour. In: Proc. Conf. on Comp. Vis. and Pat. Rec., pp. 424–429 (2000)
12. McNeill, G., Vijayakumar, S.: 2D shape classification and retrieval. In: Proc. of Intl. Joint Conf. on Artif. Intell., pp. 1483–148 (2005)
13. Belongie, S., Malik, J., Puzicha, J.: Shape matching and object recognition using shape contexts. IEEE Trans. on Pat. Anal. and Mach. Intell. 24, 509–522 (2002)
14. Ling, H., Okada, K.: An efficient earth mover´s distance algorithm for robust histogram comparison. IEEE Trans. on Pat. Anal. and Mach. Intell. 29, 840–853 (2007)
15. Felzenszwalb, P., Schwartz, J.: Hierarchical matching of deformable shapes. In: Proc. of Comp. Vis. and Pat. Rec., pp. 1–8 (2007)
16. Ferrari, V., Tuytelaars, T., Gool, L.V.: Simultaneous object recognition and segmentation from single or multiple model views. Intl. J. of Comp. Vis. 67, 159–188 (2006)

17. Lazebnik, S., Schmid, C., Ponce, J.: A sparse texture representation using local affine regions. IEEE Trans. on Pat. Anal. and Mach. Intell. 27, 1265–1799 (2005)
18. Murillo, A.C., Guerrero, J.J., Sagüés, C.: Surf features for efficient robot localization with omnidirectional images. In: IEEE Intl. Conf. on Robotics and Automation, pp. 3901–3907 (2007)
19. Loy, G., Eklundh, J.O.: Detecting symmetry and symmetric constellations of features. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3952, pp. 508–521. Springer, Heidelberg (2006)
20. Tuytelaars, T., Gool, L.V.: Matching widely separated views based on affine invariant regions. Intl. J. of Comp. Vis. 59, 61–85 (2004)
21. Mikolajczyk, K., Leibe, B., Schiele, B.: Local features for object class recognition. In: IEEE Intl. Conf. on Comp. Vis., vol. 2, pp. 1792–1799 (2005)
22. Mikolajczyk, K., Schmid, C.: A performance evaluation of local descriptors. IEEE Trans. on Pat. Anal. and Mach. Intel. 27, 1615–1630 (2005)
23. Pavlou, M., Allinson, N.M.: Automatic extraction and classification of footwear patterns. In: Corchado, E., Yin, H., Botti, V., Fyfe, C. (eds.) IDEAL 2006. LNCS, vol. 4224, pp. 721–728. Springer, Heidelberg (2006)
24. Matas, J., Chum, O., Urban, M., Pajdla, T.: Robust wide baseline stereo from maximally stable extremal regions. In: Proc. of the Britisch Machine Vision Conf., London, vol. 1, pp. 384–393 (2002)
25. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Intl. J. of Comp. Vis. 60, 91–110 (2004)
26. Bay, H., Tuytelaars, T., Gool, L.V.: Surf: Speeded up robust features. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3951, pp. 404–417. Springer, Heidelberg (2006)
27. Yanowitz, S., Bruckstein, A.: A new method for image segmentation. Comp. Vis., Graphics and Image Proc. 46, 82–95 (1989)
28. Ruisz, J., Biber, J., Loipetsberger, M.: Quality evaluation in resistance spot welding by analysing the weld fingerprint on metal bands by computer vision. Intl. J. of Adv. Manuf. Tech. 33, 952–960 (2007)
29. Zambanini, S., Kampel, M.: Segmentation of ancient coins based on local entropy and gray value range. In: Proc. of Comp. Vis. Winter Workshop, pp. 9–16 (2008)
30. Sivic, J., Schaffalitzky, F., Zisserman, A.: Object level grouping for video shots. Intl. J. of Comp. Vis. 67, 189–210 (2006)