Towards automatic coin classification

Laurens J.P. van der Maaten* and Eric O. Postma*

Abstract. *Automatic image classification algorithms can support coin experts in their analysis and study of coins. These algorithms take digital images of coins as input and generate a class as output. Automatic classification proceeds in two stages. In the feature-extraction stage, the image is transformed into a compact representation that contains information on the presence of features. In the classification stage, the feature representations are mapped onto a class. This paper focuses on the first stage by presenting and evaluating two feature types for automatic classification of modern coins: contour features and texture features. For the second stage, a standard (nearest-neighbour or naive Bayes) classifier is used. We evaluate the classification performance obtained with both feature types on an image collection of modern coins. The classification results are promising. In addition, we test the performance on a collection of medieval coins. We show that the effectiveness of the features does not generalize to medieval coins, probably due to erroneous labelling of the images. The paper concludes by stating that automatic image classification algorithms may support coin experts in their analysis of modern coins. Future work is directed towards finding appropriate features for ancient coins.*

1. Introduction

The development of a system for automatic coin classification can serve two goals. First, it allows for sorting the large amounts of old European coins that were collected after the introduction of the euro. Second, systems for automatic coin classification can be of help for institutions working with historical coins, such as the Dutch Money Museum. Cultural heritage institutions own large collections of historical coins, which are traditionally stored in safes that are not accessible to the public. Since recently, historical coin collections are made available to the public by means of the internet. A good example is the NUMIS[1] project that allows users to view the coin heritage via their browser. Unfortunately, systems such as NUMIS are not capable of aiding users in the classification of coins. Therefore, a system such as NUMIS would benefit from a system for automatic classification of coins. Our goal is to develop a system that is capable of performing image-based automatic classification of coins. Automatic coin classification consists of two stages: feature extraction and classification. In the feature-extraction stage, visual properties of the coin image are transformed into feature representations. In the classification stage, the feature representations are mapped onto a class. Our study focuses on the feature-extraction stage by presenting and evaluating feature-extraction techniques on an image collection of modern coins. In addition, we assess to what extent these techniques can be successfully applied to a collection of medieval coins.

Until now, only a few studies have focussed on automatic coin classification [1, 2, 3, 11]. Generally, the systems investigated are limited in performance. Moreover, the systems often require specific de-

*MICC-IKAT, Maastricht University, The Netherlands, email: {`l.vandermaaten,postma`}`@cs.unimaas.nl`

[1]See `http://83.149.77.24:8080/numis/` for more information.

vices such as a proximity sensors. An image-based system with high sorting performances on modern European coins is presented by Nölle *et al.* [9]. However, this system highly relies on additional sensor data on coin size and thickness. In contrast, the objective of our coin classification approach is to obtain a high classification performance using visual data only. A second promising approach is presented by Huber *et al.* [4]. Huber *et al.* present a system for the recognition of modern coins that is based on rotation-invariant eigenfeatures.
The outline of the remainder of the paper is as follows. Section 2. gives a brief description of the image datasets we used to evaluate the classification performance of our features. In section 3., we present our two types of feature representations for automatic classification of coins in detail. Section 4. describes the general setup of our experiments. In section 5., the results of the experiments are presented. The results of our experiments are discussed in section 6.. Finally, section 7. concludes the paper.

2. Datasets

In our experiments on automatic coin classification, we used two datasets. The main dataset is the MUSCLE CIS dataset (subsection 2.1.), which we use for evaluating the effectiveness of the feature types. An additional dataset, the Merovingen coin dataset (subsection 2.2.), is employed to evaluate to what extent our feature types are appropriate for ancient coin classification.

2.1. MUSCLE CIS dataset

The MUSCLE CIS dataset[2] contains images of modern European coins that were collected after the introduction of the euro. Each image in the dataset is labelled with one of 109 different coin classes. Taken together, the coin classes have 389 different coin faces. The photographed coins were sampled from a collection of approximately 300 tons of unsorted coins, mainly collected by charity organizations.
The dataset has a fixed training set of 4,575 coins and a fixed test set of 1,100 coins. The training set contains approximately 24 images of each coin face. The training set mainly contains selected, non-degraded coins. The test set contains coins with various levels of degradation. In addition, the test set also contains a number of non-European coins, which cannot be correctly classified. We measured a percentage of unknown coins in the test set of 5.64%. This percentage determines the upper limit of the classification performance, since we do not address coin verification in this study.

2.2. Merovingen coin dataset

The Merovingen coin dataset consists of 4,659 early-medieval coins from the Merovingen dynasty, photographed on both sides. An example of a coin from the dataset is shown in Figure 1(a).
The coins in the dataset can be classified into four types of classes: (1) city, (2) mint master, (3) currency, and (4) nation. Table 1 lists the number of classes for the four types, as well as the mean class size and its standard deviation. The number of coins per classification type differs because not all coins can be classified accordingly. For instance, for some coins the currency is known, whereas the mint master is not. The high standard deviations in Table 1 reveal that the class distributions in the dataset are severely skewed.

[2]See `http://muscle.prip.tuwien.ac.at/` for more information.

Class type	*No. of classes*	*Mean class size*	*St. dev. of class size*
City	18	53	125
Mint master	19	69	121
Currency	4	859	1,469
Nation	12	199	438

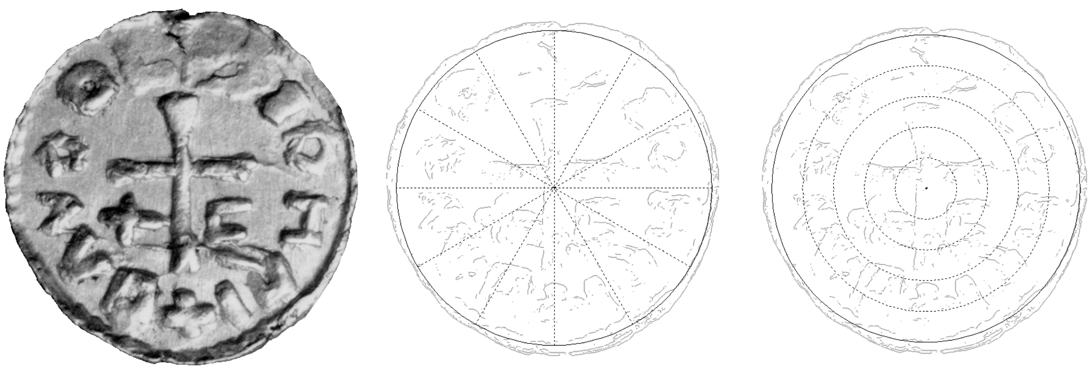
Table 1. Number and size of Merovingen coin classes.

3. Feature extraction

In coin classification, classification performances depend largely on the visual properties extracted from the coin images. Three main properties are: (1) the face or picture on the coin, (2) the texture of the coin, and (3) the text on the coin. We focus on the first two properties, leaving the third (that requires special optical character recognition techniques) to future work. In subsection 3.1., we present two contour features that provide a representation of the face or picture on the coin. In subsection 3.2., we present two texture features to represent the texture of the coin.

3.1. Contour features

Contour features provide a way of representing the contours (i.e. the edges) of an image. The extraction of contour features consists of two stages: (1) the extraction of the contour image and (2) the representation of this contour image in statistical features.
The extraction of the contour image is performed by convolving a coin image with two orthogonal Sobel kernels and performing a threshold operation. The outer border edges of the coin are removed



(a) Original coin image. (b) Edge angle histogram. (c) Edge distance histogram.

Figure 1. Edge histogram examples.

from the edge image, since they correspond to the outer border of the coin, which does not contribute to the classification process. From the resulting contour image, edge-based statistical features are computed.
Edge-based statistical features represent statistical distributions of edge pixels in a contour image. We tested two edge-based statistical features for coin classification: multi-scale edge angle histograms and multi-scale edge distance histograms.
Edge angle histograms represent the distribution of edge pixels in a coarsely discretized polar space.

This is illustrated in Figure 1(b). The height of the histogram bins correspond to the number of edge points in the polar segments. By definition, the edge angle histogram is not rotation invariant, but the modulus of its Fourier transform is [13]. Multi-scale edge angle histograms are edge angle histograms for a number of scales, ranging from coarse to fine. Edge angle histograms for various number of bins are measured, and combined into one feature vector.
The distances of edge pixels from the centre of the coin are collected in a histogram called the edge distance histogram. The histogram bins represent the relative frequency of points lying within a concentric ring (see Figure 1(c)). Edge distance histograms are rotation invariant by definition. Multi-scale edge distance histograms are multiple edge distance histograms, one for each scale ranging from coarse to fine. The histogram bin heights are all combined into one feature vector.

3.2. Texture features

Depending on the scale of analysis and the erosion of the surface, highly detailed artwork and markings on a coin may be regarded as texture. We selected two types of wavelet features that are capable of representing the texture of coins: (1) Gabor wavelet features (subsubsection 3.2.1.) and (2) Daubechies wavelet features (subsubsection 3.2.2.).

3.2.1. Gabor wavelet features

Gabor wavelet features are features based on convolutions of coin images with Gabor filters [8]. Gabor filters are biologically inspired filters exhibiting a response to visual input that is similar to that of neurons in the human primary visual cortex [5]. The Gabor filter is the product of a 2-dimensional Gaussian function and a complex sinusoid. It is described by the equation

$$G(x,y) = \frac{1}{2\pi} e^{-\frac{1}{2}(x'^2+y'^2)} e^{i\kappa x'} - e^{-\frac{\kappa^2}{2}} \tag{1}$$

In this equation the variable κ is given by

$$\kappa = \sqrt{2\ln(2)} \frac{2^\phi + 1}{2^\phi - 1} \tag{2}$$

In this equation, ϕ is the bandwidth in octaves. The value of ϕ is typically $0.5 < \phi < 1.5$. The variables x' and y' define the orientation of the sinusoid, and thereby of the function response. They are defined by the equations

$$x' = x\cos\theta + y\sin\theta \tag{3}$$

$$y' = -x\sin\theta + y\cos\theta \tag{4}$$

In these equations, θ is the orientation of the filter in radians.
In the literature, various Gabor-based features are proposed, e.g., sparse object representations [6] and simple Gabor features [7]. In our experiments, we used Gabor histograms, that are commonly used in texture classification tasks [12]. The extraction of Gabor histograms consists of two stages: (1) Gabor filtering and (2) histogram extraction. In the first stage, the coin image is convolved with Gabor filters at various scales and orientations. In the second stage, image histograms are computed for all resulting convolution images. The combination of the image histograms of the resulting images forms the Gabor histogram of the image.
On the Gabor histograms, we apply PCA to reduce the dimensionality of the feature vectors. We selected the SPCA procedure [10] over the normal PCA procedure because of computational and

memory constraints. The resulting feature vectors are 200-dimensional. The percentage of variance in the original Gabor histograms described by the resulting feature vectors is unknown, since the SPCA procedure does not provide insight in this percentage.

3.2.2. Daubechies wavelet features

Daubechies wavelet features are based on the computation (i.e. expansion) of wavelet coefficients for coin images. Their use is widespread in image analysis applications [14]. In our experiments, we used the Daubechies D4 wavelet transform. The discrete Daubechies D4 wavelet transform is specified by the coefficients $h_0 = \frac{1+\sqrt{3}}{2\sqrt{2}}$, $h_1 = \frac{3+\sqrt{3}}{2\sqrt{2}}$, $h_2 = \frac{3-\sqrt{3}}{2\sqrt{2}}$, and $h_3 = \frac{1-\sqrt{3}}{2\sqrt{2}}$. The wavelet function is specified by the equation

$$c_i = h_3 s_{2i} - h_2 s_{2i+1} + h_1 s_{2i+2} - h_0 s_{2i+3} \tag{5}$$

The scaling functions that is used to create dilated versions of the wavelet is specified by the equation

$$a_i = h_0 s_{2i} + h_1 s_{2i+1} + h_2 s_{2i+2} + h_3 s_{2i+3} \tag{6}$$

In these equations s_i indicates the value of the signal s at time step i.
In order to construct wavelet features, we compute 2-level, 3-level, and 4-level wavelet coefficients of the coin images. In order to make our classifiers more robust due to changes in rotation, we perform the wavelet feature computation for 16 rotated versions of the coin images.
On the obtained wavelet coefficients, we apply the SPCA procedure [10] to reduce the dimensionality of the feature vectors to 200.

4. Experimental setup

This section describes the general setup of our experiments. In all experiments, we determine the generalization performance of classifiers that were trained using the features described in section 3..
In our experiments on the MUSCLE CIS dataset, we used the fixed training set of 4,550 coins and the fixed test set consisting of 1,100 coins (see subsection 2.1.). We improved our results on the MUSCLE CIS dataset, by applying a preselection based on a measurement of the total surface area of the depicted coin. In this preselection, we assume that the average area of a coin class is within a small range (7%) of the area of the unclassified coin.
On the Merovingen coin dataset, we performed our experiments using 10-fold cross validation. We performed the classification on this set using a naive Bayes classifier, in order to exploit the skewness in class priors in the dataset (see subsection 2.2.).
In the experiments, we use edge-distance histograms with 2, 4, 8, and 16 bins to construct the multi-scale edge-distance histogram. For the multi-scale edge-angle histograms, we combined edge angle histograms of 4, 8, 16, and 32 bins. For the Gabor histograms, we performed experiments using 5 scales and 16 rotations of the Gabor wavelet.

5. Results

The section presents the classification performances for the two feature types presented in section 3. on the MUSCLE CIS dataset. Subsequently, the performances on the Merovingen coin dataset are presented.
The classification performances on the MUSCLE CIS dataset are listed in Table 2. The results reveal that a combination of multi-scale edge distance histograms and Gabor histograms yields the best

Approach	*Class. perf.*
Area	40%
Multi-scale edge distance histogram	68%
Multi-scale edge angle histogram	17%
Gabor histograms	55%
Wavelet features	46%
Area + MSEDH	75%
Area + MSEDH + Gabor histograms	76%

Table 2. Classification performances on the MUSCLE CIS dataset.

Approach	*City*	*Mint master*	*Currency*	*Nation*
Area	16%	10%	61%	17%
Multi-scale edge distance histogram	12%	8%	50%	20%
Multi-scale edge angle histogram	8%	6%	34%	14%
Gabor histograms	5%	6%	25%	8%
Wavelet features	8%	5%	25%	6%

Table 3. Classification performances on the Merovingen coin dataset.

classification performance. We analysed the classifications of individual coins and found that, in general, a wrong classification is due to one of the two following reasons: (1) the offered coin type was not in the training set (see subsection 2.1.) or (2) the coin is very dark. An example of a very dark coin is shown in Figure 2. For very dark coins, not all edge pixels in the coin are found, thereby decreasing the quality of the contour features. In addition, very dark coins have a lack of contrast, which decreases the quality of the texture features. We performed an additional test on a set of 100 manually selected high-contrast coins. On this set, we measured a classification performance of the combination Area + EMDH of 89%.

The classification performances for the Merovingen coin dataset are presented in Table 3. The

Figure 2. Example of a dark coin.

results reveal a low identification performance compared to those obtained for the modern coins (see Table 2), especially when the skewness of the class distributions (see subsection 2.2.) is taken into account. Therefore, the results indicate that features, which achieve high classification performances on modern coins, do not yield successful performance on the Merovingen coin dataset.

6. Discussion

From the results presented in section 5., we make two observations.
First, the classification performances on modern coins are promising. Obviously, future work is necessary to further improve the results. This work should focus on the enhancement of edges and contrast in dark coins, since a lack of clear edges is the main cause of incorrect classifications. Second, we observe that our feature-extraction schemes do not generalize to the Merovingen coin dataset. We surmise this is due to the differences in the nature between modern coins and medieval coins. We observe four differences between the modern and medieval coins. First, medieval coins are not fabricated in a factorial process, whereas modern coins are. As a consequence, the positions of the stamps on medieval coins vary per coin. In addition, medieval coin dies deteriorate quickly. Hereby, the coin die produces a different stamp on each coin. Second, medieval coins are often strongly eroded due to frequent use and by being buried in the soil. Although modern coins are often also damaged by abrasion or dirt, their stamps are usually better observable than those of medieval coins. Third, the variety in the stamps of medieval coins is smaller than the variety in stamps of modern coins. Fourth, expert classifications of medieval coins suffer from inconsistencies and errors. Figure 3 shows an example of inconsistent labelling of the Merovingen dataset. It is likely that the low classification performances on the Merovingen coin dataset are partly due to the inconsistencies in the labelling.
The observations above lead to the question how a system for automatic classification of medieval coins should work. The text on medieval coins is highly discriminating between coin classes. However, it is unlikely that current state-of-the-art in character recognition is capable of reading the texts on medieval coins. Therefore, an inscription-based approach is not promising. Experts indicate that classification of medieval coins is performed by looking at differences in style and by applying a large number of non-documented rules. Therefore, we foresee the development of a semi-automatic adaptive system that allows the expert to indicate relevant parts of classified coins. In combination with our contour and texture features, such a ”human-in-the-loop” approach allows the expert knowledge to be gradually incorporated into the system.

(a) Coin classified as *Frankish*.

(b) Coin classified as *Frisian*.

Figure 3. Illustration of dataset inconsistency.

7. Conclusions

We have presented two effective feature types for the classification of modern coins. Our results revealed a combination of contour and texture features to yield the best performance. Furthermore, our work shows that the same features do not perform well on medieval coin data. We provided insight in the differences between modern and medieval coins, and proposed ideas for future work on the classification of medieval coins. Future work should also include the creation of a medieval coin dataset without inconsistencies, and a sufficient number of instances per class.

Acknowledgements

This work was supported by NWO/CATCH under grant 640.002.401. We thank the Dutch Money Museum for providing the Merovingen coin dataset.

References

[1] M. Adameck, M. Hossfeld, and M. Eich. Three color selective stereo gradient method for fast topography recognition of metallic surfaces. *Proceedings of the SPIE*, 5011:128–139, 2003.

[2] P. Davidsson. Coin classification using a novel technique for learning characteristic decision trees by controlling the degree of generalization. In *Proceedings of the* 9^{yh} *IEA/AIE*, pages 403–412, 1996.

[3] M. Fukumi, S. Omatu, F. Takeda, and T. Kosaka. Rotation-invariant neural pattern recognition system with application to coin recognition. *IEEE Transactions on Neural Networks*, 3(2):272–279, 1992.

[4] R. Huber, H. Ramoser, K. Mayer, H. Penz, and M. Rubik. Classification of coins using an eigenspace approach. *Pattern Recognition Letters*, 26(1), 2005.

[5] J.P. Jones and L.A. Palmer. An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *Journal of Neurophysiology*, 58(6):1233–1258, 1987.

[6] N. Krüger, G. Peters, and C. von der Malsburg. Object recognition with a sparse and autonomously learned representation based on banana wavelets. Technical Report IR-INI 96-11, Ruhr-Universität Bochum, 1996.

[7] V. Kyrki, J.-K. Kamarainen, and H. Kälviäinen. Simple Gabor feature space for invariant object recognition. *Pattern Recognition Letters*, 25(3):311–318, 2004.

[8] T.S. Lee. Image representation using 2D Gabor wavelets. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):959–971, 1996.

[9] M. Nölle, H. Penz, M. Rubik, K. Mayer, I. Holländer, and R. Granec. Dagobert - a new coin recognition and sorting system. In *Proceedings of the DICTA 2003*, pages 329–338, 2003.

[10] M. Partridge and R. Calvo. Fast dimensionality reduction and Simple PCA. *Intelligent Data Analysis*, 2(3):292–298, 1997.

[11] P.A. Passeraub, P.-A. Besse, C. de Raad, O. Dezuari, F. Quinet, and R.S. Popovic. Coin recognition using an inductive proximity sensor microsystem. In *Proceedings of TRANSDUCERS '97*, volume 1, pages 389–392, 1997.

[12] M.V. Ramakrishna and J. Lin. Gabor histogram feature for content-based image retrieval. In *Proceedings of the* 5^{th} *DICTA*, pages 149–154, 1999.

[13] T. Szoplik and H.H. Arsenault. Rotation-variant optical data processing using the 2-d nonsymmetric fourier transform. *Applied Optics*, 24(2):168–172, 1985.

[14] P.N. Topiwala. *Wavelet image and video coding*. Kluwer Academic Publishers, 1998.