Feature-Based Forensic Camera Model Identification

Thomas Gloe

Technische Universität Dresden,
Institute of Systems Architecture,
01062 Dresden, Germany
Thomas.Gloe@tu-dresden.de

Abstract. State-of-the-art digital forensic techniques for camera model identification draw attention on different sets of features to assign an image to the employed source model. This paper complements existing work, by a comprehensive evaluation of known feature sets employing a large set of 26 camera models with altogether 74 devices. We achieved the highest accuracies using the extended colour feature set and present several detail experiments to validate the ability of the features to separate between camera models and not between devices. Analysing more than 16,000 images, we present a comprehensive evaluation on 1) the number of required images and devices for training, 2) the influence of the number of camera models and camera settings on the detection results and 3) possibilities to handle unknown camera models when not all models coming into question are available or are even known. All experiments in this paper suggest: feature-based forensic camera model identification works in practice and provides reliable results even if only one device for each camera model under investigation is available to the forensic investigator.

1 Introduction

The abundance of digital photographs in everyday life and the trust placed into them as pieces of evidence raises a need for reliable forensic techniques to test the authenticity of digital images. The increasing number of available techniques can be broadly divided into image source identification and manipulation detection [1]. This paper concentrates on the former, and, more specifically, on camera *model* identification. Camera model identification is relevant in practice if the forensic investigator ...

- knows that a given image has been taken with a digital camera (otherwise she would use a scheme targeted to distinguish between *types* of acquisition devices, e.g., scanner, digital camera, or computer-generated [2,3]),
- but does not have a set of independent images unequivocally shot with exactly the same *device* as the suspect image (otherwise she would use a method for camera *device* identification based on the intrinsic fingerprint embedded in each sensors' pattern noise [4,5,6,7,8]).

Y.Q. Shi (Ed.): Transactions on DHMS VIII, LNCS 7228, pp. 42–62, 2012.

Forensic investigators typically have to find answers to the following questions related to camera model identification: "which camera model was (most likely) used to shoot this given image?", for the identification scenario; or, if prior beliefs prevail, "has this image been shot with a Canon Ixus 70 camera?", for the validation scenario.

Techniques trying to answer these questions require characteristics covering differences between camera models, while occurring very similar between devices of one model. Common ingredients are depicted in Fig. 1 in the simplified acquisition pipeline of a digital camera. For example, aberrations introduced by the lens, are specific to models sharing the same optical system, e.g. [9,10]. Estimating the configuration of the colour filter array (CFA) or determining the employed method for colour interpolation allows a basic discrimination between groups of camera models, e.g. [11,12]. Furthermore, compression parameters [13], meta data [14] and the employed file structure can help to isolate questionable models.

This paper belongs to feature-based camera model identification based on characteristics covering coarse noise properties, colour reproduction and image quality. This approach has been first proposed by Kharrazi, Sencar and Memon [15] for the identification of digital camera models and was further investigated by Çeliktutan, Avcibas and Sankur for low resolution mobile-phone cameras [16,17]. Using small sets of digital cameras or mobile-phone cameras, reliable results for both camera and mobile-phone model identification were reported.

In our previous work [18], we started to explore the influence of the cardinality and between-model variance of the training set on the overall classification performance employing a subset of 12 camera models of the 'Dresden Image Database' [19]. The reported results with average success rates above 90% are promising and indicate the possibility to apply this scheme in practice. However, the scalability in scenarios with even more camera models, the influence of the different sets of features [15,17,18] as well as the handling of unknown camera models require further attention.

Based on all natural images in the 'Dresden Image Database' including 26 camera models with altogether 74 devices, we compare different sets of features and investigate the influence of feature selection under practical relevant conditions. We extend our previous analysis in [18] on the determined best feature set and evaluate the ability to separate between different models and not between different devices. A discussion on the influence of the number of images, devices and camera models on the detection results gives valuable clues for creating appropriate training data sets. Finally, we draw attention to handle unknown cameras in open sets of camera models and deepen our investigations presented at DAGM 2011 [20]. We believe that this paper contains valuable information for forensic investigators with typically limited resource and limited budgets to correctly train a feature-based forensic camera model identification scheme.

The remainder of the paper is organised as follows: Section 2 introduces the general scheme for camera model identification and discusses the different feature sets. Section 3 describes the employed data set and basic test settings used in our

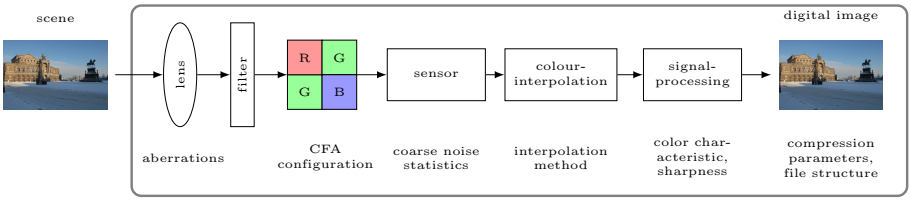


Fig. 1. Simplified image acquisition pipeline of a digital camera and possible sources of model-dependent characteristics

investigations. An detailed analysis of feature-based camera model identification employing closed model sets, where all questioned camera models are considered known, is presented in Sec. 4. Section 5 considers the problem of unknown camera models and possible false accusation in so called open sets. Finally, the paper concludes with a discussion in Sec. 6.

2 Camera Model Identification

Feature-based forensic camera model identification is motivated by differences in the internal image acquisition pipeline of digital camera models [15]. To create visually pleasant images, the manufacturers specifically fine-tune the components and algorithms for each digital camera model. Details on this fine-tuning are usually considered trade secrets. Nevertheless we can capture some variation with specific features. These features are required to be stable for all devices of one model to capture characteristics inherent in each image acquired with the same model. For instance, we can try to characterise the camera model-dependent combination of colour filter array and colour interpolation algorithm, or the algorithms implemented in the internal signal processing pipeline including, for example, the white-point correction.

The first step in feature-based forensic camera model identification is the estimation of a set of features in images taken with all camera models coming into question. Employing the estimated features, second, a machine learning algorithm can be trained to, third, determine or to validate the probable source of an image. In literature, a support vector machine (SVM) is used for machine learning and the major difference between previously proposed schemes relies in the set of employed features. The features are broadly classified into four main groups: Colour characteristics ($\mathcal{F}_{\text{col}}$) [15] describing the colour reproduction of a camera model, wavelet statistics ($\mathcal{F}_{\text{wav}}$) [21] coarsely quantifying noise, image quality metrics ($\mathcal{F}_{\text{iqm}}$) [22] measuring noise and sharpness (quality of scene reproduction by the optical system), and binary similarity measures ($\mathcal{F}_{\text{bsm}}$) [23] characterising relations between different bit planes of one colour channel as well as between different colour channels.

Kharrazi, Sencar, and Memon introduced feature-based camera model identification employing a set of 34 features $\mathcal{F}_{\text{Khar}}$ including $\mathcal{F}_{\text{col}}$, $\mathcal{F}_{\text{iqm}}$ and $\mathcal{F}_{\text{wav}}$ [15].

Another set of features $\mathcal{F}_{\text{Çeli}}$ was proposed by Çeliktutan et al. [17] containing a similar group of wavelet statistics $\mathcal{F}_{\text{wav}}$ and image quality metrics $\mathcal{F}_{\text{iqm}}$, complemented with binary similarity measures $\mathcal{F}_{\text{bsm}}$. Contrary to Kharrazi et al., $\mathcal{F}_{\text{iqm}}$ are analysed for each colour channel separately and an extended set of $\mathcal{F}_{\text{wav}}$ is calculated for up to 4 wavelet sub bands. Altogether, the feature set consists of 592 characteristics.

In our earlier work [18], we used an *extended* version $\mathcal{F}_{\text{ext}}$ of the original Kharrazi et al. feature set, consisting of 46 features. We added some new colour features and, motivated by the work of Farid and Lyu [21] as well as Çeliktutan et al., calculated an extended set of wavelet features. Furthermore, we investigated the combination of the 46 features with characteristics of lateral chromatic aberration [10,20] with little success. Lateral chromatic aberration as well as aberrations in general are dependent on focal length and focus settings. The resulting variance makes it very difficult to generate a comprehensive measurement of all possible settings necessary for reliable feature-based camera model identification.

Additionally to the three mentioned feature sets, within this paper, we investigate an *extended colour* feature set $\mathcal{F}_{\text{extc}}$, including the features of our extended set for each colour channel separately (82 features), and a *complete* feature set $\mathcal{F}_{\text{comp}}$ joining all features of the four basic feature groups: $\mathcal{F}_{\text{col}}$, $\mathcal{F}_{\text{wav}}$, $\mathcal{F}_{\text{iqm}}$ and $\mathcal{F}_{\text{bsm}}$.

3 Test Setup

A key factor in designing and evaluating techniques for camera model identification is in the composition of a suitable benchmark database. To assure that arbitrary features capture characteristics of the model rather than of the device or image contents, a database of images with comparable contents shot with multiple models and *multiple devices per model* is required. Only recently, we compiled the 'Dresden Image Database' for this very purpose [19].

Table 1 lists the 26 employed camera models, the number of corresponding devices and images, as well as basic camera specifications. The database includes both, typical consumer digital camera models and semi-professional digital SLR cameras. In some investigations the computational complexity was quite demanding and a reduced set of camera models similar to our previous work was employed [18]. We abbreviate the complete and reduced sets of camera models with symbols $\mathcal{M}_{\text{all}}$ and $\mathcal{M}_{\text{red}}$.

To provide images with similar content, all images in the database are created employing a specific image acquisition procedure. A tripod was fixed for each motif and using each device three or more scenes have been taken with different focal length settings. For logistical reasons all camera models were split into set A and B, both covering different motifs (cf. [19] for a more detailed description). Identifiers of the employed model, device id, camera settings and the acquired motif are stored together with each image and enable to analyse features under selected constraints.

Altogether, 16,958 original images stored in the JPEG format at full resolution were analysed. Compared to previous work, where a small set of digital cameras (5 camera models) [15] or a small set of low resolution mobile-phone cameras (13 mobile-phone models with altogether 16 devices) [17] was investigated, the employed data set allows the evaluation of feature-based camera model identification under comprehensive settings. Another difference is the variation in focal length and flash settings.

Analysing feature-based camera model identification requires a set $\mathcal{I}_{\text{train}} \subset \mathcal{I}$ of images to *train* and a set $\mathcal{I}_{\text{test}}$ of images to optimise the detection performance of the employed machine learning algorithm. An independent *validation* set $\mathcal{I}_{\text{val}}$ is required to report practical relevant detection results. In real scenarios, we expect different devices employed for acquiring the image under investigation and the images available for machine learning. Consequently, we partition the set of available devices $\mathcal{D}^{(m)}$ of each camera model m in a set of devices for machine learning $\mathcal{D}^{(m)}_{\text{train,test}}$ and a disjoint set of devices for validating the detection results $\mathcal{D}^{(m)}_{\text{val}} = \mathcal{D}^{(m)} / \mathcal{D}^{(m)}_{\text{train,test}}$. The expected difference of image content between images available during machine learning and images provided for classification makes a reasonable separation of the set of available motifs $\mathcal{P}$ necessary. To report detection results independently of the image content, requires images of the same motifs for each device. Ideally, a large number of different motifs for training, testing and validating feature-based camera model identification exists. However, due to the limited number of images in our database, we decided to separate $\mathcal{P}$ in a set of training $\mathcal{P}_{\text{train}}$ and a disjoint set of test and validation $\mathcal{P}_{\text{test,val}}$ motifs. The stored device (d) and motif (p) identifiers enable the assignment of each image $i \in \mathcal{I}$ to its corresponding set:

$$\mathcal{I}_{\text{train}} = \{i | d_i \in \mathcal{D}_{\text{train,test}} \wedge p_i \in \mathcal{P}_{\text{train}}\} \tag{1}$$

$$\mathcal{I}_{\text{test}} = \{i | d_i \in \mathcal{D}_{\text{train,test}} \wedge p_i \in \mathcal{P}_{\text{test,val}}\} \tag{2}$$

$$\mathcal{I}_{\text{val}} = \{i | d_i \in \mathcal{D}_{\text{val}} \wedge p_i \in \mathcal{P}_{\text{test,val}}\} \tag{3}$$

Experimental results in this paper are based on cross-validation using a fixed set of 100 different partitionings of $\mathcal{I}$ unless otherwise stated. We assigned devices and motifs to the corresponding training, test and validation sets randomly with preset cardinalities ($|\mathcal{D}^{(m)}_{\text{train,test}}| = 1$, $|\mathcal{P}_{\text{train}}| = 26$). All available camera models are included to make the identification task challenging and realistic. For some camera models only 1 device is available (e.g., models of make Agfa) and detection results are only computed for the test images. The fixed partitionings make the experiments repeatable and we present average results balancing between best and worst results.

Besides the public available set of images in the 'Dresden Image Database', we employed an additional set of snapshot images created only for some selected camera models. These images form another validation set with scene content and camera settings independent to the structured design of the official database. Due to the lack of a tripod during image acquisition, the quality is not always convincing and also blurred images are included.

All investigations in this paper make use of a support vector machine implementation developed by Chang and Lin with a radial based kernel function [24] provided through the interface of the R package e1071. More precisely, we employed the multi-class classification scheme of LIBSVM to separate between different camera models. The implementation of the library solves the multi-class problem by creating single binary SVMs for all pairs of different classes (one-versus-one) and a voting scheme determines the most probable class.

Table 1. List of digital camera models included in this study; number of devices per model, basic camera specifications and number of available images. (*) indicates models included in the reduced set $\mathcal{M}_{\text{red}}$ necessary to investigate the large feature sets ($\mathcal{F}_{\text{Çeli}}, \mathcal{F}_{\text{comp}}$).

make	model	no. devices	resolution [pixel]	sensor size [inch or mm]	focal length [mm]	no. images (flash released) set A	set B	$\sum$	snapshots
Agfa	DC-504	1	4032×3024	-	7.1	78 (14)	91 (10)	169 (24)	
Agfa	DC-733s	1	3072×2304	-	6.2–18.6	150 (21)	128 (3)	278 (24)	
Agfa	DC-830i	1	3264×2448	-	6.2–18.6	176 (50)	187 (31)	363 (81)	
Agfa	Sensor505-X	1	2592×1944	-	7.5	87 (11)	85 (9)	172 (20)	
Agfa	Sensor530s	1	4032×3024	-	6.1–18.3	195 (50)	177 (35)	372 (85)	
Canon	Ixus55	1	2592×1944	1/2.5"	5.8–17.4	224 (52)		224 (52)	
Canon	Ixus70 (*)	3	3072×2304	1/2.5"	5.8–17.4	567 (119)		567 (119)	
Canon	PowerShot A640	1	3648×2736	1/1.8"	7.3–29.2		188 (23)	188 (23)	
Casio	EX-Z150 (*)	5	3264×2448	1/2.5"	4.65–18.6	924 (178)		924 (178)	
FujiFilm	FinePix J50	3	3264×2448	1/2.5"	6.2–31.0		630 (99)	630 (99)	503 (210)
Kodak	M1063 (*)	5	2748×3664	1/2.33"	5.7–17.1	1070 (330)	1321 (289)	2391 (619)	
Nikon	CoolPix S710 (*)	5	4352×3264	1/1.72"	6.0–21.6	925 (173)		925 (173)	
Nikon	D200 (*)	2	3872×2592	23.6×15.8 mm	18 – 135/17 – 55	752 (79)		752 (79)	
Nikon	D70/D70s	2/2	3008×2000	23.7×15.6 mm	18–200	736 (78)		736 (78)	
Olympus	μ1050SW (*)	5	3648×2736	1/2.33"	6.7–20.1	1040 (342)		1040 (342)	
Panasonic	DMC-FZ50	3	3648×2736	1/1.8"	7.4–88.8		931 (115)	931 (115)	832 (426)
Pentax	Optio A40	4	4000×3000	1/1.7"	7.9–23.7		638 (90)	638 (90)	574 (445)
Pentax	Optio W60	1	3648×2736	1/2.3"	5.0–25.0		192 (23)	192 (23)	308 (290)
Praktica	DCZ5.9 (*)	5	2560×1920	1/2.5"	5.4–16.2	1019 (273)		1019 (273)	
Ricoh	GX100	5	3648×2736	1/1.75"	5.1–15.3		854 (112)	854 (112)	1794 (1106)
Rollei	RCP-7325XS (*)	3	3072×2304	1/2.5"	5.8–17.4	589 (148)		589 (148)	
Samsung	L74wide (*)	3	3072×2304	1/2.5"	4.7–16.7	686 (144)		686 (144)	
Samsung	NV15 (*)	3	3648×2736	1/1.8"	7.3–21.9	645 (110)		645 (110)	
Sony	DSC-H50	2	3456×2592	1/2.3"	5.2–78.0		541 (57)	541 (57)	375 (45)
Sony	DSC-T77	4	3648×2736	1/2.3"	6.18–24.7		725 (88)	725 (88)	863 (346)
Sony	DSC-W170	2	3648×2736	1/2.3"	5.0–25.0		405 (52)	405 (52)	244 (38)
$\sum$		74				9863 (2172)	7093 (1036)	16956 (3208)	5493 (2906)

4 Camera-Model Identification in Closed Sets

In the first part of our investigations, we assume that all camera models coming into question are known. We call sets including only known camera models *closed sets*. In practice, it might be difficult to consider all possibly employed camera models during image acquisition and unknown camera models might cause false accusations. This problem is an inherent property of multi-class SVMs, which always assign a sample to one of the trained classes. We discuss this problem in Sec. 5 and present first solutions to cover unknown models in *open sets*.

4.1 Benchmarking Feature Sets

Before we will take a closer look at the ability of the features to separate between different camera models and not between different devices, we will first analyse the performance of the five feature sets described in Sec. 2.

Therefore, we optimised the basic parameters of a SVM with a radial based kernel function using grid-search with γ in the range of $2^{3,2,\ldots,-15}$ and C in the range of $2^{-5,-4,\ldots,15}$. The optimisation was done separately for each of the fixed 100 partitionings to report average and standard deviation of the accuracies (i.e., the average detection rate over all camera models) under optimal parameter settings in Tab. 2. The required computational time increased exceptionally when we tried to optimise a SVM for $\mathcal{M}_{\text{all}}$ in combination with the large feature sets ($\mathcal{F}_{\text{Çeli}}, \mathcal{F}_{\text{comp}}$) and we decided to calculate the results only for $\mathcal{M}_{\text{red}}$.

Table 2. Average accuracies for five different feature sets over 100 fixed partitionings (standard deviation is enclosed in brackets). Best accuracies are in bold for $\mathcal{M}_{\text{all}}, \mathcal{M}_{\text{red}}$ in combination with $\mathcal{I}_{\text{test}}, \mathcal{I}_{\text{val}}$.

		feature set				
	model set	$\mathcal{F}_{\text{Khar}}$	$\mathcal{F}_{\text{ext}}$	$\mathcal{F}_{\text{extc}}$	$\mathcal{F}_{\text{Çeli}}$	$\mathcal{F}_{\text{comp}}$
test data $\mathcal{I}_{\text{test}}$	$\mathcal{M}_{\text{red}}$ $(w_{\text{Khar,grey}})$	95.62% (1.60)	97.45% (1.27)	97.94% (1.10)	94.07% (1.61)	94.00% (1.61)
	$\mathcal{M}_{\text{red}}$ $(w_{\text{Khar,col}})$	-	-	**98.06**% (1.02)	95.32% (1.55)	95.30% (1.49)
	$\mathcal{M}_{\text{red}}$ $(w_{\text{DB8,grey}})$	95.57% (1.60)	97.13% (1.34)	97.73% (1.18)	94.29% (1.45)	94.29% (1.42)
	$\mathcal{M}_{\text{red}}$ $(w_{\text{DB8,col}})$	-	-	97.83% (1.03)	95.23% (1.27)	95.24% (1.26)
	$\mathcal{M}_{\text{all}}$ $(w_{\text{Khar,grey}})$	89.29% (2.46)	91.69% (2.33)	92.71% (2.25)	-	-
	$\mathcal{M}_{\text{all}}$ $(w_{\text{Khar,col}})$	-	-	93.08% (1.95)	-	-
	$\mathcal{M}_{\text{all}}$ $(w_{\text{DB8,grey}})$	89.26% (2.51)	91.51% (2.32)	92.85% (2.20)	-	-
	$\mathcal{M}_{\text{all}}$ $(w_{\text{DB8,col}})$	-	-	**93.12**% (1.94)	-	-
validation data $\mathcal{I}_{\text{val}}$	$\mathcal{M}_{\text{red}}$ $(w_{\text{Khar,grey}})$	93.14% (1.93)	95.27% (1.70)	96.12% (1.49)	91.50% (1.79)	91.47% (1.76)
	$\mathcal{M}_{\text{red}}$ $(w_{\text{Khar,col}})$	-	-	**96.36**% (1.32)	92.80% (1.66)	92.70% (1.67)
	$\mathcal{M}_{\text{red}}$ $(w_{\text{DB8,grey}})$	93.18% (1.94)	94.96% (1.80)	96.01% (1.56)	92.66% (1.73)	92.65% (1.74)
	$\mathcal{M}_{\text{red}}$ $(w_{\text{DB8,col}})$	-	-	96.18% (1.40)	93.75% (1.58)	93.69% (1.63)
	$\mathcal{M}_{\text{all}}$ $(w_{\text{Khar,grey}})$	85.62% (2.94)	88.60% (2.87)	89.98% (2.63)	-	-
	$\mathcal{M}_{\text{all}}$ $(w_{\text{Khar,col}})$	-	-	90.67% (2.42)	-	-
	$\mathcal{M}_{\text{all}}$ $(w_{\text{DB8,grey}})$	85.79% (2.86)	88.70% (2.85)	90.41% (2.59)	-	-
	$\mathcal{M}_{\text{all}}$ $(w_{\text{DB8,col}})$	-	-	**90.93**% (2.48)	-	-

We also experimented with the influence of the selected wavelet filter to calculate the wavelet statistics. Namely, we tested a wavelet filter originally employed by Kharrazi et al. (w_{Khar}) and a Daubechies 8 filter (w_{DB8}) commonly used for the estimation of sensor noise [4]. For the feature sets $\mathcal{F}_{\text{extc}}$, $\mathcal{F}_{\text{Çeli}}$ and $\mathcal{F}_{\text{comp}}$, we included also tests with and without averaging the wavelet statistics of each colour channel (w_{grey}, w_{col}).

It can be expected, that increasing the number of known camera models increases also the chance, that different models share similar characteristics. The decrease of accuracy in case of $\mathcal{M}_{\text{all}}$ compared to $\mathcal{M}_{\text{red}}$ demonstrates this effect and might result in less accurate decisions when using sets with considerably more than 26 models. Furthermore, the results employing $\mathcal{I}_{\text{test}}$ (images acquired with the same device as the training data) are always better than the results for $\mathcal{I}_{\text{val}}$ (images acquired with other devices). This indicates small variations in the analysed characteristics between devices of one model.

We obtained the best results in this scenario employing $\mathcal{F}_{\text{extc}}$ for both sets of camera models $\mathcal{M}_{\text{all}}$ and $\mathcal{M}_{\text{red}}$. Differences in the results related to the two

employed wavelet filters are negligible and depend on the employed feature set. In contrast, using the wavelet statistics for each colour channel separately, increases the accuracy slightly in all cases. Contrary to our expectations, adding binary similarity measures $\mathcal{F}_{\text{bsm}}$ did not increase the identification performance for camera models available in the 'Dresden Image Database'. In fact, the larger number of features seem to complicate the training of the SVM. Çeliktutan et al. therefore applied feature selection to reduce the number of features and to increase the identification performance.

4.2 Feature Selection

Based on the previous tests, we decided to try feature selection to investigate the possibility to improve the accuracy. We employed sequential forward floating search (SFFS) [25] on $\mathcal{F}_{\text{extc}}$ and $\mathcal{F}_{\text{comp}}$. Features in $\mathcal{F}_{\text{extc}}$ were analysed using $\mathcal{M}_{\text{red}}$ and $\mathcal{M}_{\text{all}}$ and, again due to the computational constraints, $\mathcal{F}_{\text{comp}}$ was tested only with $\mathcal{M}_{\text{red}}$ resulting in altogether three test configurations. Furthermore, we restricted all following experiments to wavelet statistics based on $w_{\text{DB8,col}}$. Knowing that the selection of motifs and devices employed for training might influence feature selection, we applied SFFS to the first 10 of the 100 fixed partitionings of $\mathcal{I}$.

Figure 2 depicts the relation between accuracy and number of selected features for two partitionings with the best and worst achieved maximum accuracy for $\mathcal{I}_{\text{test}}$. In this example, we used $\mathcal{F}_{\text{extc}}$ together with $\mathcal{M}_{\text{red}}$ and it was possible to obtain an accuracy above 98% for $\mathcal{I}_{\text{test}}$ for all 10 partitionings of $\mathcal{I}$. In contrast, the results for $\mathcal{I}_{\text{val}}$ are worse and are only above 98% in the best case.

SFFS tries to select features according to their importance on the overall classification accuracy and we would expect a similar order of selected features

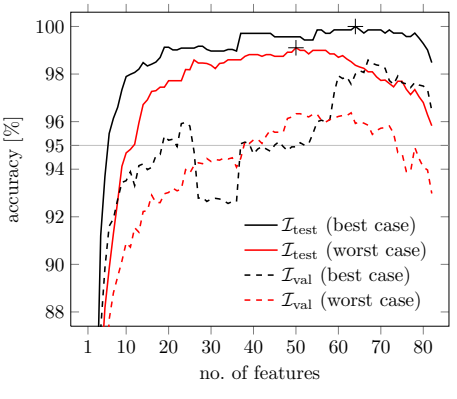


Fig. 2. Number of selected features of $\mathcal{F}_{\text{extc}}$ in relation to the achieved accuracy using $\mathcal{M}_{\text{red}}$. 2 partitionings with the best and worst accuracy for $\mathcal{I}_{\text{test}}$ after feature selection are depicted (both above 98%). Symbols + mark the highest accuracy obtained in both cases.

for all 10 partitionings depending on the employed feature set. While there are indeed some features always in the range of the first 30 most important features, the concrete partitioning of devices and motifs employed for training made the ordering of most features more variable than expected. Considering the worst case in Fig. 2, the increase in accuracy is not always continuous and makes the selection of an universal set of features even more difficult.

Therefore, we used all 30 sets of the best selected features (3 test configurations $\times$ 10 partitionings of $\mathcal{I}$) and calculated accuracies in combination with all 100 fixed partitionings. For each of the 3 test configurations we determined the set of selected features maximising the average accuracies of $\mathcal{I}_{\text{test}}$ over all partitionings and report the results in Table 3. Note that in case of $\mathcal{F}_{\text{comp}}$ the feature selection was computationally demanding and we aborted the calculation after 48h for each of the 10 partitionings. At that time the first 40 best features were selected, explaining the slightly lower accuracies. Even if we directly compare the accuracies between the best 40 features of $\mathcal{F}_{\text{extc}}$ and $\mathcal{F}_{\text{comp}}$, the results are very similar. Furthermore, a comparison of the accuracies in Tab. 3 with the results in Tab. 2 shows only small differences and negligible changes in the overall performance.

Table 3. Best average accuracies obtained after feature selection in $\mathcal{F}_{\text{extc}}$ and $\mathcal{F}_{\text{comp}}$. The employed set of camera models during feature selection is indicated in brackets after each feature set and the depicted accuracies are calculated in combination with both sets of camera models.

		SFFS on feature set		
model set	$\mathcal{I}$	$\mathcal{F}_{\text{extc}}$ ($\mathcal{M}_{\text{red}}$)	$\mathcal{F}_{\text{extc}}$ ($\mathcal{M}_{\text{all}}$)	$\mathcal{F}_{\text{comp}}$ ($\mathcal{M}_{\text{red}}$)
$\mathcal{M}_{\text{red}}$ ($w_{\text{DB8,col}}$)	$\mathcal{I}_{\text{test}}$	97.72% (1.27)	**97.79**% (0.95)	97.05% (1.17)
$\mathcal{M}_{\text{all}}$ ($w_{\text{DB8,col}}$)	$\mathcal{I}_{\text{test}}$	92.89% (1.93)	**93.59**% (1.96)	91.49% (2.13)
$\mathcal{M}_{\text{red}}$ ($w_{\text{DB8,col}}$)	$\mathcal{I}_{\text{val}}$	96.08% (1.60)	95.97% (1.11)	**96.16**% (1.20)
$\mathcal{M}_{\text{all}}$ ($w_{\text{DB8,col}}$)	$\mathcal{I}_{\text{val}}$	90.77% (2.50)	**91.68**% (2.31)	90.47% (2.08)

Feature selection might help to increase the overall performance in case of a specific partitioning of $\mathcal{I}$, but it is difficult to find a universally valid reduced set of features. For the following tests only results employing $\mathcal{F}_{\text{extc}}$ are reported. Table 4 gives an example of the average detection rate for each camera model employing $\mathcal{I}_{\text{val}}$ ($\mathcal{I}_{\text{test}}$ is used in cases where only one camera model is available). The concrete detection rate depends on the camera model and, for example, images acquired with a Nikon S710 or a Ricoh GX 100 can be reliably identified with low false acceptance rate. For cameras like the Sony T77 the accuracy drops to 84.9% due to a higher similarity to camera models of the same manufacturer W170 and H50 as well as as a higher similarity to the Panasonic DMC-FZ50.

The next Section presents a detailed analysis of the similarity between different camera models and between devices of one camera model. Using $\mathcal{F}_{\text{extc}}$ we determined $\gamma = 2^{-9}$ and $C = 2^8$ as appropriate parameters for all 100 partitionings to decrease computational requirements.

4.3 Intra- and Inter-camera Model Similarity

The ability to separate between different camera models and not between different devices is a basic requirement for all camera model identification schemes. Therefore, characteristics should be chosen in a way that the intra-camera model similarity is high, i.e., feature values of devices of the same model are similar. In contrast, the inter-camera model similarity between different camera models should be minimised.

To give a visualisation of intra- and inter-camera model similarity between all employed 74 devices for $\mathcal{F}_{\text{extc}}$, we calculated for each device the centroid over all corresponding feature values. Applying multi-dimensional scaling the centroids mapped to 2D are depicted in Fig. 3. The visualisation clearly shows a spatial grouping of devices of the same camera model and supports the assumption, that the employed features are able to separate between camera models.

In a more detailed plot in Fig. 4, we illustrate the dissimilarity between feature values of single images of different camera models together with the similarity between devices of the same model. In Figure 4a the separation between the two camera models Nikon S710 and Sony T77 is clearly possible, while it is not possible to separate between different devices. In contrast, the dissimilarity of images made with camera models of the same manufacturer is sometimes lower and makes a separation more difficult. Figure 4b depicts an example for this case.

To investigate intra- and inter-camera model similarity in more detail, we trained the feature-based camera model identification scheme for each device of one camera model – contrary to its original purpose. We calculated average results iterating over the fixed set of 100 partitionings. Table 5 shows the results for separating between all devices of Nikon S710, Sony T77 and Sony W170. Similar to the observations in Fig. 4, a clear separation between Nikon and Sony camera models is possible. Furthermore, discriminating between the camera models T77 and W170 both manufactured by Sony are slightly worse and indicate possible manufacturer-specific dependencies. In all cases it is not possible to separate between devices of the same camera model with acceptable accuracy.

The two devices of the SLR camera model Nikon D200 are a noteworthy exception. In case of these two devices, device identification using feature-based camera model identification is indeed possible with high accuracy (see Tab. 6). In contrast, the four devices of the SLR model D70/D70s are indistinguishable from each other. Reconsidering the detailed results in Tab. 4, the low intra-camera model similarity between the D200 devices might explain the low detection rate of 80% in detecting the correct camera model.

4.4 Influence of the Number of Images, Devices and Models

The available resources during a forensic investigation are limited in terms of time and money. Consequently, it is important to know how many images and devices have to be considered to train a feature-based camera model scheme for a specific set of models reliably. Furthermore, it is important to consider relations between the number of questioned camera models and the accuracy in detecting the correct camera model.

Table 4. Average detection rates for $\mathcal{I}_{\text{val}}$ and, if only one device is available, for $\mathcal{I}_{\text{test}}$ employing $\mathcal{F}_{\text{extc}}$ and all 100 fixed partitionings of $\mathcal{I}$. Models with only one device available are indicated by (*). The average accuracy is 90.93% for $\mathcal{I}_{\text{val}}$, 93.12% for $\mathcal{I}_{\text{test}}$ and 91.18% for the depicted combination of $\mathcal{I}_{\text{val}}$ and $\mathcal{I}_{\text{test}}$.

	identified as																								
model	504	733	830	505	530	I55	I70	A64	Z15	J50	M10	S71	D20	D70	μ	FZ5	A40	W60	DCZ	GX	732	L74	NV	H50	T77	W17
DC-504*	**88.0**	-	-	-	6.8	0.6	-	-	-	-	-	0.2	1.2	1.2	-	-	0.1	-	1.8	-	-	-	-	-	-	-
DC-733s*	-	**92.6**	0.2	-	0.1	-	1.5	-	1.4	-	0.3	-	-	0.6	-	-	-	0.2	0.1	0.6	-	2.2	0.1	-	-	-
DC-830i*	-	0.5	**97.4**	-	0.2	-	0.1	-	0.8	0.6	-	-	0.2	-	0.1	0.1	-	-	-	-	-	0.1	-	-	-	-
505-X*	-	0.7	-	**87.6**	0.7	1.5	-	-	-	-	2.2	-	-	1.2	-	-	0.1	-	5.9	-	-	0.1	-	-	-	-
530s*	3.8	-	-	0.1	**90.2**	3.1	-	-	-	-	-	-	1.1	0.5	-	0.1	0.5	-	0.5	-	-	-	-	-	-	-
Ixus 55*	-	0.1	-	0.3	1.3	**94.9**	-	-	-	-	0.1	-	-	0.9	-	-	-	-	2.4	-	-	-	-	-	-	-
Ixus 70	-	6.0	0.6	0.1	0.1	0.1	**87.3**	-	0.2	0.1	-	-	-	1.9	-	-	-	-	-	-	-	3.5	-	-	-	-
A640*	-	-	-	-	1.5	-	-	**92.3**	-	-	0.2	-	0.9	-	3.3	-	0.3	0.6	-	-	-	-	0.9	-	0.1	-
Z150	-	2.5	2.4	-	0.1	-	-	-	**93.5**	-	0.3	0.1	0.4	0.2	-	0.2	-	-	0.1	-	-	-	0.2	-	0.1	-
J50	-	-	8.8	-	-	-	-	-	-	**90.3**	-	-	-	-	-	0.1	-	-	-	-	0.3	-	-	0.3	0.1	-
M1063	0.1	-	-	-	0.1	-	-	0.1	-	-	**98.9**	-	0.4	0.1	-	-	-	-	-	-	-	-	0.2	-	-	-
S710	-	-	-	-	-	-	-	0.2	-	-	0.1	**98.8**	-	-	-	0.4	0.3	-	-	-	-	-	0.2	-	-	-
D200	6.1	-	0.5	-	1.2	-	-	0.8	0.6	-	0.8	-	**80.2**	-	0.9	5.6	0.5	0.6	-	0.1	-	-	1.0	0.2	0.2	0.6
D70(s)	-	0.7	0.7	0.3	0.3	0.1	1.2	-	0.2	-	0.9	0.1	0.1	**92.5**	-	-	-	-	0.9	-	-	1.6	0.2	-	-	-
μ1050	-	0.7	0.8	-	0.1	-	-	3.4	0.1	-	0.1	-	0.2	-	**85.5**	1.2	0.1	4.5	0.1	0.6	-	-	2.6	-	-	-
FZ50	0.2	-	0.4	-	-	-	-	0.1	-	0.3	1.4	-	0.6	-	-	**90.2**	-	-	-	-	-	-	0.1	3.1	1.4	2.1
A40	1.5	-	-	-	1.1	-	-	0.5	0.2	-	1.1	1.8	0.2	-	0.4	0.2	**89.6**	0.1	-	1.8	-	-	1.2	-	0.1	0.1
W60*	0.1	-	-	-	-	-	-	0.2	0.2	0.6	0.1	-	0.4	-	6.7	0.4	-	**91.0**	-	-	-	-	0.2	-	-	-
DCZ5.9	-	-	0.1	4.6	0.1	2.3	-	-	0.1	-	0.1	-	-	1.2	-	-	-	-	**91.4**	-	-	-	-	-	-	-
GX100	0.1	0.2	-	-	-	-	-	0.2	-	-	-	-	0.1	-	0.8	-	-	-	-	**97.2**	-	-	0.1	0.2	0.1	0.9
7325XS	-	-	0.1	0.1	0.5	-	-	-	-	1.1	0.2	-	-	-	-	-	-	-	0.2	-	**97.2**	0.2	0.2	0.1	-	-
L74	-	4.3	0.6	0.6	-	-	1.7	-	-	-	-	-	-	1.6	-	-	-	-	-	-	-	**91.0**	-	-	-	-
NV15	0.1	-	-	-	0.5	-	0.4	2.6	0.1	0.2	0.3	-	1.4	-	4.1	2.5	0.1	0.2	-	1.2	-	-	**85.9**	-	0.2	-
H50	-	-	0.7	-	-	-	0.1	-	-	-	-	-	0.2	-	0.1	3.0	-	-	-	-	-	-	-	**92.7**	1.4	1.7
T77	-	-	0.1	-	-	-	-	0.1	0.3	0.2	-	-	0.6	-	0.1	6.5	-	-	-	0.5	-	-	0.2	2.9	**84.9**	3.6
W170	0.1	-	-	-	-	-	-	0.1	-	-	-	-	0.1	-	0.1	2.8	-	-	-	0.1	-	-	-	3.5	3.6	**89.5**

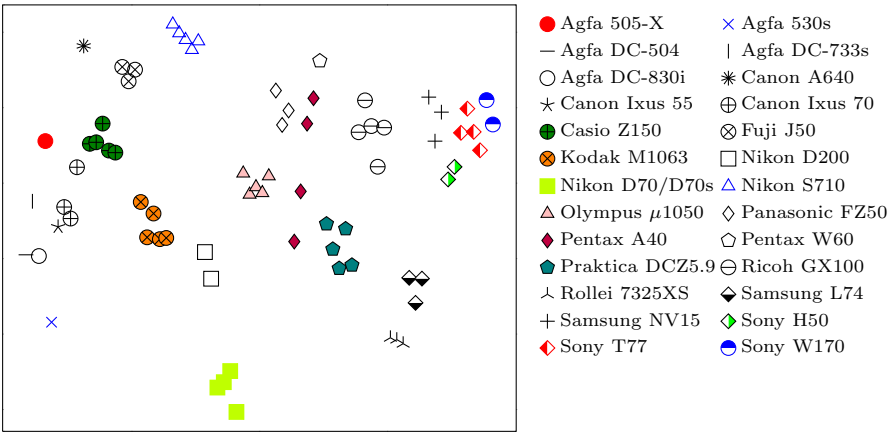
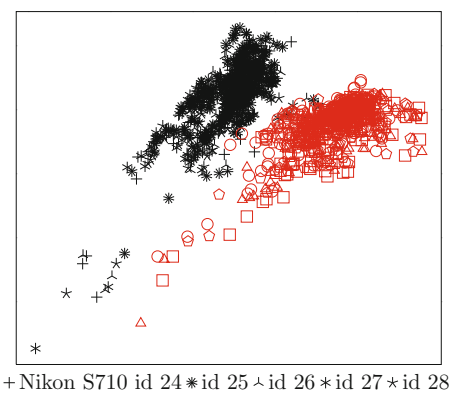
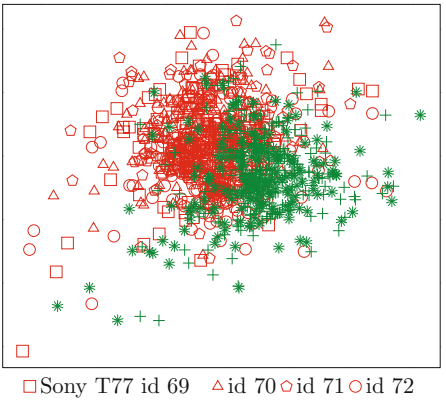


Fig. 3. Visualisation of intra- and inter-camera model similarity. The centroid of all feature values of all images of each device is mapped to 2D using multi-dimensional scaling and different devices of one camera model are depicted by the same symbol. Devices of the same camera model are closer to each other whereas devices of different models are farther apart.



(a) Nikon S710 (5×) and Sony T77 (4×)



(b) Sony T77 (4×) and W170 (2×)

Fig. 4. Visualisation of the similarity of feature values between all images of pairs of camera models using the two most distinctive principal components. Different symbols of the same colour indicate different devices of one camera model. Camera models are visually separable, but a differentiation between devices of the same model is not possible.

Table 5. Intra- and inter-camera similarity between devices of three different camera models averaged over all 100 fixed partitionings (overall accuracy 43.76%, accuracy S710 32.23%, accuracy T77 49.92%, accuracy W170 60.27%).

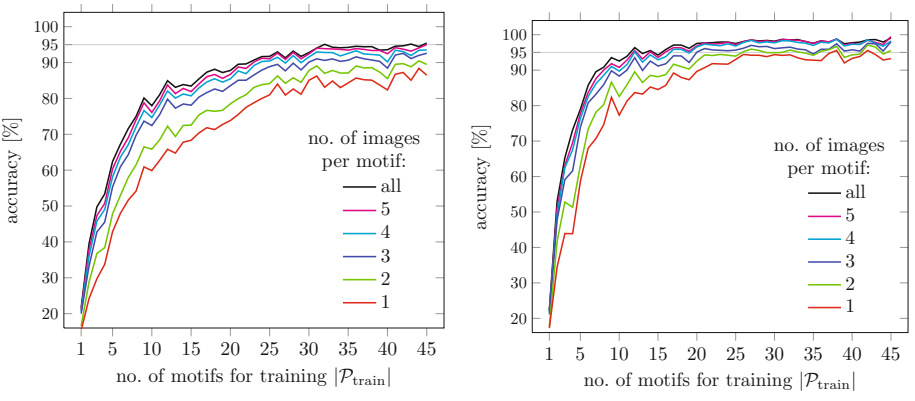
		identified as										
device		id 24	id 25	id 26	id 27	id 28	id 69	id 70	id 71	id 72	id 73	id 74
S710	id 24	**42.8**	15.0	12.1	20.2	9.8	-	-	-	-	-	-
S710	id 25	**24.1**	21.5	17.9	20.7	15.9	-	-	-	-	-	-
S710	id 26	19.8	20.6	**26.4**	14.6	18.5	-	-	-	-	-	-
S710	id 27	26.7	17.3	12.6	**33.1**	10.3	-	-	-	-	-	-
S710	id 28	15.3	16.0	16.6	14.7	**37.4**	-	-	-	-	-	-
T77	id 69	0.3	0.2	0.1	-	0.4	**55.1**	17.6	12.8	10.3	2.2	0.8
T77	id 70	0.4	-	0.3	-	0.3	23.6	**50.9**	14.9	7.7	1.0	0.8
T77	id 71	0.1	-	0.1	-	0.1	21.4	24.8	**39.3**	11.8	1.8	0.6
T77	id 72	0.2	-	0.1	-	0.4	16.4	10.5	11.6	**54.4**	3.8	2.6
W170	id 73	-	-	-	-	-	0.8	0.2	0.2	1.0	**55.6**	42.2
W170	id 74	-	-	-	-	-	1.0	0.1	0.3	1.4	32.2	**64.9**

Table 6. Intra- and inter-camera similarity between devices of camera models D70/D70s and D200 averaged over all 100 fixed partitionings (overall accuracy 60.35%, accuracy D200 95.82%, accuracy D70/D70s 42.61%).

		identified as					
device		id 29	id 30	id 31	id 32	id 33	id 34
D200	id 29	**95.4**	4.5	-	-	-	-
D200	id 30	3.8	**96.2**	-	-	-	-
D70/D70s	id 31	-	0.2	**34.2**	25.3	22.0	18.3
D70/D70s	id 32	-	0.1	23.0	**44.5**	16.8	15.6
D70/D70s	id 33	-	0.1	20.1	16.6	**47.6**	15.6
D70/D70s	id 34	0.1	0.2	17.1	20.7	17.8	**44.1**

No. of Images and Devices for Training. Pictures taken from different motifs differ significantly in image content compared to images capturing different scenes of the same motif with different focal length settings. Based on the stored identifiers in the database, we investigate the number of images for training each camera model under consideration of the number of motifs $|\mathcal{P}_{\text{train}}|$ and the number of images $|\mathcal{I}_{\text{train}}^{(\text{d,p})}|$ per motif p and device d. More precisely, we created a specific set of 225 fixed partitionings of $\mathcal{I}$ similar to the procedure in Sec. 3: First, we selected randomly one device per model for training ($|\mathcal{D}_{\text{train,test}}^{(m)}| = 1$) and, second, we varied the number of motifs for training $|\mathcal{P}_{\text{train}}|$ in the range from 1 to 45. We repeated both steps 5 times resulting in 225 partitionings. To investigate the influence of the number of images per motif p and device d, we varied $|\mathcal{I}_{\text{train}}^{(\text{d,p})}|$ in the range from 1 to the maximum number available. Note that there are 47 motifs for camera models in set A available, whereas set B only holds 30 or 36 motifs, depending on the camera model. To keep distinct motifs for training and testing, we left always a minimum of 3 testing motifs for each camera model ($|\mathcal{P}_{\text{test,val}}| \geq 3$).

Figure 5 illustrates the relation between the accuracy and the number of images for training for $\mathcal{M}_{\text{all}}$ and $\mathcal{M}_{\text{red}}$. The depicted accuracy is averaged over all



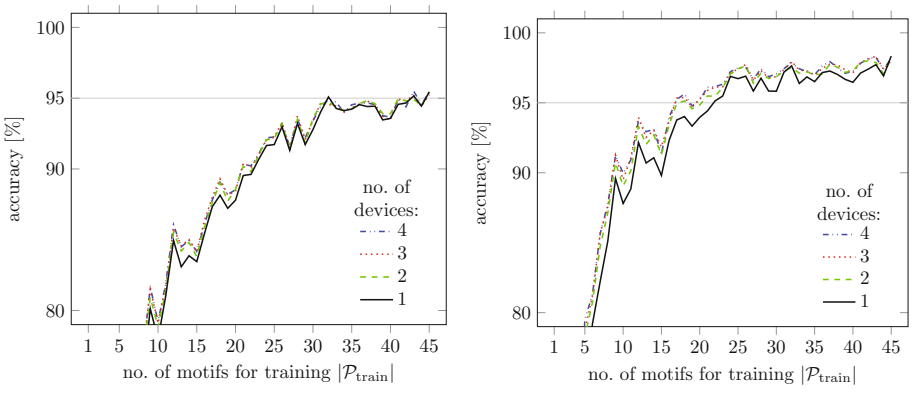
(a) model set $\mathcal{M}_{\text{all}}$ **(b)** model set $\mathcal{M}_{\text{red}}$

Fig. 5. Relation between the number of motifs available for training $|\mathcal{P}_{\text{train}}|$ and the accuracy for $\mathcal{I}_{\text{val}}$. Additionally, the influence of the number of images per motif p (and per device d) $|\mathcal{I}_{\text{train}}^{(\text{d,p})}|$ is depicted.

5 fixed sets of training devices $\mathcal{D}_{\text{train,test}}$. Employing the smaller set of camera models $\mathcal{M}_{\text{red}}$ results in a faster increase of accuracy compared to $\mathcal{M}_{\text{all}}$. The influence of the number of different motifs is higher than the influence of the number of images per motif. Nevertheless, adding images with different camera settings is still important to achieve the maximum possible accuracy. After adding 30 motifs in case of $\mathcal{M}_{\text{all}}$ and 20 motifs in case of $\mathcal{M}_{\text{red}}$, we observe only a minor increase in accuracy. Nonetheless, it is important to employ a notable number of images capturing different motifs to get best detection rates. Depending on the number of camera models considered during an investigation, we suggest $|\mathcal{I}_{\text{train}}^{(\text{d,p})}| \geq 3$ acquired with different camera settings for a minimum of $|\mathcal{P}_{\text{train}}| \geq 30$ motifs to get reasonable accuracies. Whenever possible, the number of motifs should be increased.

In another experiment, we investigated the relation between the number of devices $|\mathcal{D}_{\text{train,test}}^{(m)}|$ and motifs $|\mathcal{P}_{\text{train}}|$ for training. We used the previously introduced set of 225 fixed partitionings and selected randomly 1 up to 4 devices for training, where available. In case of camera models with more than one device in the database, we always left a minimum of one distinct device for validation ($|\mathcal{D}_{\text{val}}^{(m)}| \geq 1$).

The average results are depicted in Fig. 6 for $\mathcal{M}_{\text{all}}$ and a set of 15 camera models including 3 or more devices per model. Different to the previous results on the influence of the number of motifs and images per motif, increasing the number of devices has a negligible influence on the accuracy. Consequently, practical investigation should focus available resources on the acquisition of enough images per camera model covering different motifs and camera settings. Borrowing or purchasing more than one device per camera model is less important regarding the average case. Reconsidering the results on intra- and inter-camera



(a) model set $\mathcal{M}_{\text{all}}$ **(b)** only camera models with $|\mathcal{D}^{(m)}| \geq 3$

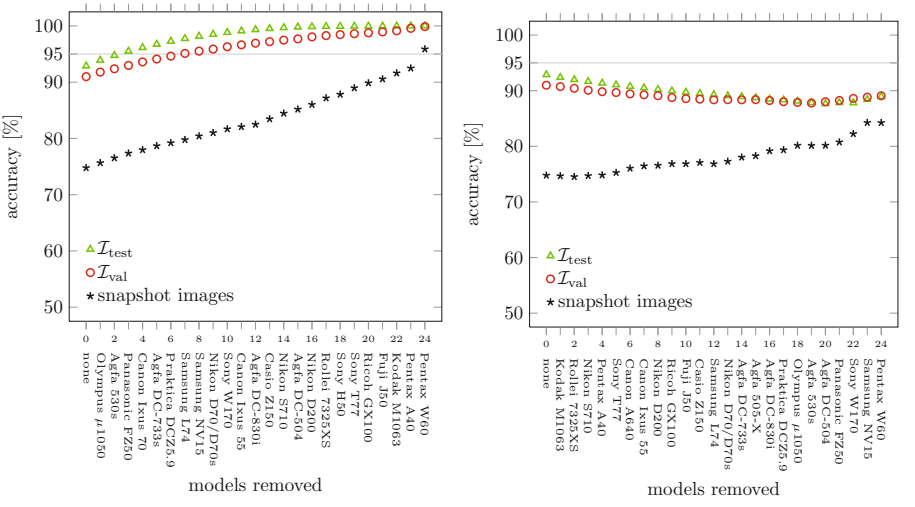
Fig. 6. Relation between number of motifs $|\mathcal{P}_{\text{train}}|$, devices per model $|\mathcal{D}^{(m)}_{\text{train,test}}|$ and accuracy for $\mathcal{I}_{\text{val}}$. Increasing the number of devices per camera model has a negligible influence on the accuracy.

model similarity of the Nikon D200 in Sec. 4.3, it might be still necessary to employ more than one device in some cases.

No. of Camera Models. Comparing the results depicted in Fig. 5a and b clearly indicates a decrease in accuracy when a larger set of camera models is considered. To investigate the influence of the number of camera models in more detail, we conducted two experiments employing the 100 fixed partitionings introduced in Sec. 3. Starting with $\mathcal{M}_{\text{all}}$, in the first experiment we removed models one by one in order to maximise the accuracy on the reduced set. In the second experiment we did the opposite and removed models in order to minimise the accuracy. With both experiments we try to gauge the range of correct identification results in relation to the number of employed camera models in a good and bad scenario.

Figure 7 depicts the average accuracy for $\mathcal{I}_{\text{test}}$ and $\mathcal{I}_{\text{val}}$ in relation to the number of removed camera models for the two experiments over 100 fixed partitionings. Furthermore, we determined the number of occurrences each camera model was removed first, second and so on, and specified the camera models with the highest count on the x-axis of the two plots. Reducing the number of camera models considered during an investigation does not necessarily increase the accuracy in detecting the correct camera model. In fact, an inconvenient combination of camera models can even for small sets result in low detection rates. It might be contradictory at first glance, that removing a camera model results in a lower accuracy, but by doing this only worse separable camera models remain and the average accuracy decreases.

We plotted also the accuracy of correctly identifying the set of snapshot images mentioned in Sec. 3. These images are photographed freehand employing many different focal length settings, and were acquired in several cases in dark



(a) models removed to maximise accuracy **(b)** models removed to minimise accuracy

Fig. 7. Relation between number of camera models considered during an investigation and accuracy. Depending on the combination of camera models in question, the accuracy can be also low for small sets of models.

environments with active flash. The depicted accuracy is considerably less compared to $\mathcal{I}_{\text{test}}$ and $\mathcal{I}_{\text{val}}$. To study the cause of this bad result, we trained the feature-based camera model identification scheme for all 8 camera models where snapshot images are available. We either employed the standard motifs or the snapshot images of one device of each camera model for training. Similar to Fig. 7 the average accuracy for correctly identifying the employed model to acquire a snapshot image is quite low (78.4%) when we employ our standard motifs for training. However, if we flip training and validation data and use all snapshot images of one device of each camera model for training, we get a high accuracy (93.1%) for correctly assigning standard motifs to the corresponding camera model. The results emphasise the importance of using different camera settings during the preparation of images for training feature-based camera model identification.

5 Camera-Model Identification in Open Sets

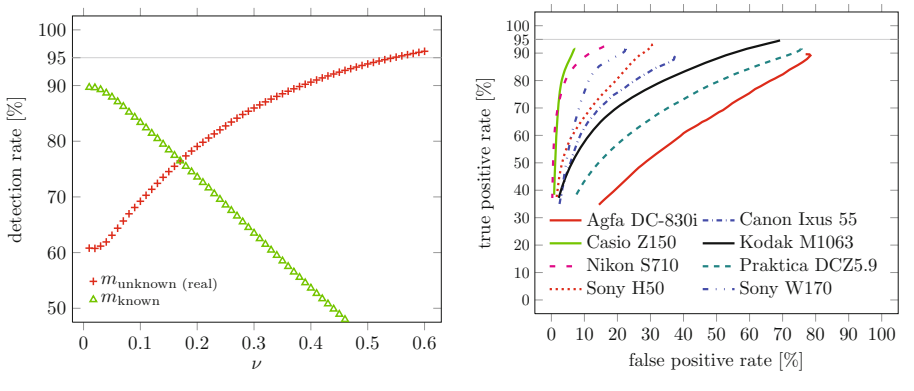
Although forensic investigators may put considerable effort into the creation of a large reference database for training, it is very unlikely that this database will ever be comprehensive. Therefore a residual risk remains that the camera model of a given image is not in the database. For critical forensic applications (i.e., in criminal court), this case should be detected with almost certainty to avoid potential false accusations under all circumstances.

We started to investigate two different approaches in [20] based on one-class SVMs and binary SVMs to handle this task:

One-Class SVM. The idea of one-class classification is to transform the feature vectors of one single class via a kernel to find a small region that captures most vectors. The concept was proposed in the seminal work of Schölkopf, Platt, Shawe-Taylor, Smola and Williamson [26]. In order to apply one-class SVMs to camera model identification, we train one SVM for each known camera model m_{known} using $\mathcal{I}_{\text{train}}$ of the fixed 100 partitionings (cf., Sec. 3).

The absence of negative information (i.e., feature vectors of unknown camera models) comes with a price, and one should not expect as good results as when they are available. The average detection rates using $\mathcal{I}_{\text{val}}$ (and $\mathcal{I}_{\text{test}}$ for models m with $|\mathcal{D}^{(m)}| = 1$) for unknown $m_{\text{unknown (real)}}$ and known models m_{known} are depicted in Fig. 8a in relation to the one-class SVM parameter ν. We iterated ν in the range $0.01, 0.02, \ldots, 0.6$. Depending on the parameter setting we can either obtain higher detection rates for unknown or for known camera models and reach a trade-off between both with $\nu = 0.17$ and 76% average accuracy.

An alternative approach is to optimise ν for each camera model separately. Figure 8b visualises the performance for selected models using ROC curves. Here, true and false positive rates indicate the percentage of images assigned to a known model correctly or incorrectly. The results for models Casio EX-Z150 and Agfa DC-830i show the best and worst case for our employed set of camera models ($\mathcal{M}_{\text{all}}$) and ROC curves of all other camera models are in between. Ideally, we want very low false positive rates (high probability to identify $m_{\text{unknown (real)}}$) and high true positive rates (high probability to identify m_{known}). However, the figure indicates that not all observed results are convincing with respect to practical scenarios in this regard.



(a) detection rates in relation to ν averaged over all models $\mathcal{M}_{\text{all}}$

(b) ROC curves for selected models (curves are incomplete due to the chosen range of ν)

Fig. 8. Averaged results using one-class SVMs and 100 fixed partitionings of $\mathcal{I}$

Binary SVM. In addition to one-class SVMs, also binary SVMs can be exploited in the open-set problem. In Section 4 the multi-class problem is solved by creating single binary SVMs for all pairs of different classes (one-versus-one) and a voting scheme determines the most probable class (cf. Sec. 3). The general lack of training data for unknown camera models makes the application of the standard multi-class classification scheme difficult. One solution to find an approximation of unknown models could be to use known models as samples of unknown models.

To experimentally investigate this approach, we iterate over all combinations of each one known and one unknown camera model, m_{known} and $m_{\text{unknown (real)}}$, respectively. Depending on the combination, we sample training data for unknown models from all remaining models $m_{\text{unknown (train)}} \in \mathcal{M}_{\text{all}}/\{m_{\text{known}}, m_{\text{unknown (real)}}\}$ and employ the fixed set of 100 partitionings of $\mathcal{I}$. Figure 9 summarises average detection rates for correctly identifying our two trained classes m_{known}, $m_{\text{unknown (train)}}$ and the 'real' unknown models $m_{\text{unknown (real)}}$ in relation to the selected known camera model m_{known}. Depending on the known camera model, detecting trained models m_{known}, $m_{\text{unknown (train)}}$ works well. Compared to the results in Sec. 4, the detection rates for known camera models are lower because of the differences in the implementation. On average also a reliable detection of $m_{\text{unknown (real)}}$ is possible, but in practice especially worst cases need attention. Depending on the combination of m_{known} and $m_{\text{unknown (real)}}$ the correct identification of unknown models is sometimes much more difficult and the detection results are not always convincing. For example, detecting $m_{\text{unknown (real)}}$ = Nikon S710 results in the lowest detection rate (7.33%) in our test scenario when we trained a binary SVM with m_{known} = Pentax A40. In the opposite combination the detection rate for the unknown

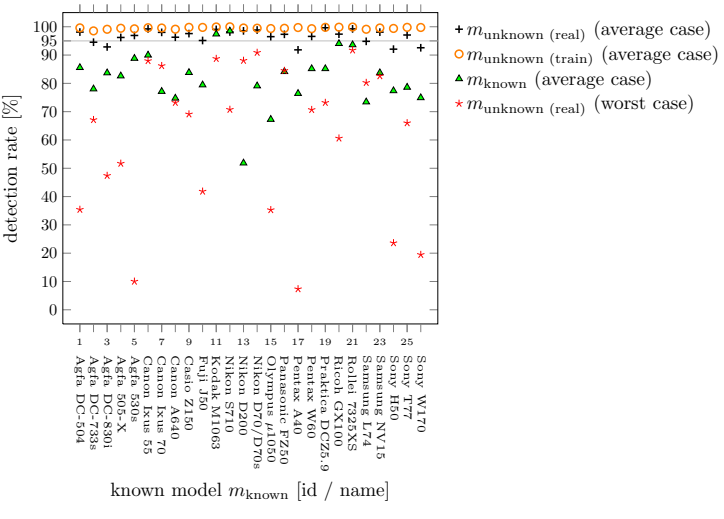


Fig. 9. Averaged results using binary SVMs and 100 fixed partitionings of $\mathcal{I}$

model is considerably better (70.74%, $m_{\text{unknown (real)}}$ = Pentax A40, m_{known} = Nikon S710). Note, detection rates for known models remain stable even in the worst case scenario and do not decrease by more than 3 percentage points.

Overall, the experiments considering open sets reveal the difficulties to handle unknown camera models in practice. While average detection rates of binary SVMs for known and unknown camera models are promising, detailed investigations illustrate that the approximation of unknown models with samples of known models is prone to be incomplete. Hence, worst case detection rates might be considerably lower. Similarly, the absence of training data for unknown models in case of one-class SVMs results in unsatisfactory identification results. In summary, we see both approaches as very first attempts to handle unknown camera models. Further research is necessary to improve the prevention of false accusations.

6 Concluding Remarks

This paper complements existing work on feature-based forensic camera model identification by an experimental in-depth analysis on a large set of 26 camera models and altogether 74 devices. We investigated different sets of features known from the literature and achieved the best results using the 82 features of the extended colour feature set $\mathcal{F}_{\text{extc}}$. While the improvement by adding binary similarity measures BSM was negligible in our test scenario, the computational requirements increased considerably for the set of all 26 camera models. In practice, the optimal feature set depends on the employed set of models and maybe also on the type of cameras (digital still camera or mobile phone camera). Furthermore, our investigations on feature selection show, that finding a set of optimal features is difficult due to dependencies on the selected devices and motifs for training. Hence, we decided to continue with all features in $\mathcal{F}_{\text{extc}}$.

Our analysis of intra- and inter-camera model similarity gives empirical evidence, that the employed features are appropriate to differentiate between camera models and not between devices. The influence of the number of images, devices and models on the correct identification was investigated and our results emphasise the importance to employ many images of different motifs together with different acquisition settings to train each considered model optimally. Employing more than one device per model is only necessary in exceptional cases like the Nikon D200, where it was indeed possible to separate between devices.

The last part of the paper deals with the practically relevant case of incomplete training data in open sets of camera models and presents results using one-class and binary SVMs. The detection rates depend clearly on the combination of known and unknown camera models. While binary SVMs enable to reliably separate known and unknown models in most cases, finding a broad enough training set for all cases is difficult and further investigations are necessary.

The presented results provide important clues for forensic investigators to select appropriate parameters for training a specific set of camera models as well as to assess detection results in a real case. A direction for future work

is a detailed analysis of the influence of image processing on the identification results. First investigations presented at DAGM 2010 [20] are worse compared to the results known from previous literature [16,17] and need further attention to identify robust subsets of features. Also different JPEG compression settings implemented in digital cameras may obscure a correct model identification and call for a careful investigation under controlled settings.

Calculating and analysing the features for all employed images required more than 15,000 hours of computation time. Therefore, supplementary material is made available via 'Dresden Image Database' website (`https://forensics.inf.tu-dresden.de/ddimgdb/publications/modelid`). We hope this not only saves ressources in the development of new forensic techniques but als provides a good starting point for introducing students to image forensics research.

Acknowledgements. The author thanks Nicolas Cebron and Rainer Böhme for fruitful discussions on machine learning and the detection of unknown camera models during the preparation of the extended abstract for DAGM 2010 [20]. We thank also Matthias Kirchner for commenting the final manuscript.

This project was partly funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG). Computational resources to calculate the experimental results were provided by the centre for high-performance computing (ZIH) at our university.

References

1. Böhme, R., Freiling, F., Gloe, T., Kirchner, M.: Multimedia Forensics Is Not Computer Forensics. In: Geradts, Z.J.M.H., Franke, K.Y., Veenman, C.J. (eds.) IWCF 2009. LNCS, vol. 5718, pp. 90–103. Springer, Heidelberg (2009)
2. Khanna, N., Mikkilineni, A.K., Chiu, G.T.C., Allebach, J.P., Delp, E.J.: Forensic classification of imaging sensor types. In: Delp, E.J., Wong, P.W. (eds.) Proceedings of SPIE: Security and Watermarking of Multimedia Content IX, vol. 6505, p. 65050U (2007)
3. Lyu, S., Farid, H.: How realistic is photorealistic? IEEE Transactions on Signal Processing 53(2), 845–850 (2005)
4. Fridrich, J.: Digital image forensics. IEEE Signal Processing Magazine 26(2), 26–37 (2009)
5. Gloe, T., Franz, E., Winkler, A.: Forensics for flatbed scanners. In: Delp, E.J., Wong, P.W. (eds.) Proceedings of SPIE: Security, Steganography, and Watermarking of Multimedia Contents IX, vol. 6505, p. 65051I (2007)
6. Gou, H., Swaminathan, A., Wu, M.: Robust scanner identification based on noise features. In: Delp, E.J., Wong, P.W. (eds.) Proceedings of SPIE: Security and Watermarking of Multimedia Content IX, vol. 6505, p. 65050S (2007)
7. Khanna, N., Mikkikineni, A.K., Chiu, G.T.C., Allebach, J.P., Delp, E.J.: Scanner identification using sensor pattern noise. In: Delp, E.J., Wong, P.W. (eds.) Proceedings of SPIE: Security and Watermarking of Multimedia Content IX, vol. 6505, p. 65051K (2007)
8. Goljan, M., Fridrich, J., Filler, T.: Large scale test of sensor fingerprint camera identification. In: Delp, E.J., Dittmann, J., Memon, N., Wong, P.W. (eds.) Proceedings of SPIE: Media Forensics and Security XI, vol. 7254, pp. 7254–18 (2009)

9. Johnson, M.K., Farid, H.: Exposing digital forgeries through chromatic aberration. In: Proceedings of the Multimedia and Security Workshop (MM&Sec 2006), pp. 48–55 (2006)
10. Gloe, T., Borowka, K., Winkler, A.: Efficient estimation and large-scale evaluation of lateral chromatic aberration for digital image forensics. In: Memon, N.D., Dittmann, J., Alattar, A.M., Delp, E.J. (eds.) Proceedings of SPIE: Media Forensics and Security II, vol. 7541, pp. 7541–7 (2010)
11. Swaminathan, A., Wu, M., Liu, K.J.R.: Nonintrusive component forensics of visual sensors using output images. IEEE Transactions on Information Forensics and Security 2(1), 91–106 (2007)
12. Cao, H., Kot, A.C.: Accurate detection of demosaicing regularity for digital image forensics. IEEE Transactions on Information Forensics and Security 4(4), 899–910 (2009)
13. Farid, H.: Digital image ballistics from JPEG quantization: A followup study. Technical Report TR2008-638, Department of Computer Science, Dartmouth College, Hanover, NH, USA (2008)
14. Kee, E., Johnson, M.K., Farid, H.: Digital image authentication from JPEG headers. IEEE Transactions on Information Forensics and Security 6(3), 1066–1075 (2011)
15. Kharrazi, M., Sencar, H.T., Memon, N.: Blind source camera identification. In: Proceedings of the 2004 IEEE International Conference on Image Processing (ICIP 2004), pp. 709–712 (2004)
16. Çeliktutan, O., Avcibas, İ., Sankur, B.: Blind identification of cellular phone cameras. In: Delp, E.J., Wong, P.W. (eds.) Proceedings of SPIE: Security and Watermarking of Multimedia Content IX, vol. 6505, p. 65051H (2007)
17. Çeliktutan, O., Sankur, B., Avcibas, İ.: Blind identification of source cell-phone model. IEEE Transactions on Information Forensics and Security 3(3), 553–566 (2008)
18. Gloe, T., Borowka, K., Winkler, A.: Feature-Based Camera Model Identification Works in Practice – Results of a Comprehensive Evaluation Study. In: Katzenbeisser, S., Sadeghi, A.-R. (eds.) IH 2009. LNCS, vol. 5806, pp. 262–276. Springer, Heidelberg (2009)
19. Gloe, T., Böhme, R.: The 'Dresden Image Database' for benchmarking digital image forensics. In: Proceedings of the 25th Symposium on Applied Computing (ACM SAC), vol. 2, pp. 1585–1591 (2010)
20. Gloe, T., Cebron, N., Böhme, R.: An ML perspective on feature-based forensic camera model identification. In: DAGM Workshop Pattern Recognition for IT Security, Darmstadt, Germany, September 21 (2010)
21. Farid, H., Lyu, S.: Higher-order wavelet statistics and their application to digital forensics. In: IEEE Workshop on Statistical Analysis in Computer Vision (2003)
22. Avcıbaş, İ., Sankur, B., Sayood, K.: Statistical evaluation of image quality measures. Journal of Electronic Imaging 11(2), 206–223 (2002)
23. Avcibas, İ., Kharrazi, M., Memon, N.D., Sankur, B.: Image steganalysis with binary similarity measures. EURASIP Journal on Applied Signal Processing 2005(17), 2749–2757 (2005)
24. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machines. Software (2001), `http://www.csie.ntu.edu.tw/~cjlin/libsvm`
25. Pudil, P., Novovičová, J., Kittler, J.: Floating search methods in feature selection. Pattern Recognition Letters 15(11), 1119–1125 (1994)
26. Schölkopf, B., Platt, J.C., Shawe-Taylor, J., Smola, A.J., Williamson, R.C.: Estimating the support of a high-dimensional distribution. Neural Computation 13(7), 1443–1471 (2001)