IDENTIFICATION OF IMAGE ACQUISITION CHAINS USING A DICTIONARY OF EDGE PROFILES

Thirapiroon Thongkamwitoon, Hani Muammar, and Pier Luigi Dragotti

Communications and Signal Processing Group, Imperial College, London, UK

ABSTRACT

This paper presents a novel framework for the identification of an image acquisition chain using dictionaries of edge profiles. We investigate how edges, one of the most common features present in images, are transformed during the image acquisition chain. A dictionary of edge profiles is constructed corresponding to edges obtained from known devices and at different stages in the chain. The processing chain of a query image is then identified by feature matching using the maximum inner product criteria. Experiments have shown that the proposed method is able to identify the sources and acquisition stages of query images. It also has good performance in both recapture detection and chain identification applications for natural scene images.

Index Terms— Image Forensics, Recapture Detection, Dictionary of Edge Profiles, Image Acquisition Chain, Line Spread Function, Source Identification

1. INTRODUCTION

In recent years the focus of attention by the research community has turned to the authenticity and integrity verification of digital images. Despite the fact that image forensics is a new area of research, several forensic techniques aimed at verifying the fidelity of digital images have been proposed in the literature. This is due to the high demand for tools in this area. The majority of research conducted in image forensics centers on image source identification and forgery detection. A common approach for source identification is to search for traces of unique *footprints* left in an image by different stages of the acquisition process which are correlated to the intrinsic properties of known image sources. Footprints are also residual artefacts generated during a particular imaging process and are often used as evidence of image tampering. Image forgery can be detected by the presence of processing footprints or the inconsistency of source footprints across images. This may indicate the fusion of images from different sources. A comprehensive overview of image forensic techniques based on specific footprints can be found in [1].

Today, digital images can be repeatedly captured, manipulated, and shared online by multiple users and can, therefore, undergo several stages of processing over their lifetime. However, an individual footprint determined from a specific image source or a clue derived from local processing alone may be insufficient to provide a detailed description of the processing history of an image. Furthermore, the individual footprints may, in turn, be altered by the processing chain.

The presence of footprints introduced by the individual stages of a processing chain was first explored in [2]. The method models the processing block of a single camera with a series of in-camera operations and is able to distinguish original images from their post-acquisition tampered versions. In [3] a method is proposed for the detection of JPEG recompression using the statistical features of DCT coefficients. Forgery detection based on resampling in a digital domain is studied in [4] and [5] with the assumption that tampering is often associated with scaling and rotation.

Problems in image reacquisition have also been studied in terms of recapture detection. In [6] a recapture detection algorithm that uses a combination of colour and resolution features and a Support Vector Machine (SVM) classifier is presented. The method in [7] employs multiple physical features including surface gradient, specularity distribution, contrast, histogram, and blurriness in order to detect images recaptured from printed material. The blurriness feature discussed in this research is based on off-focus capture, mismatched resolution, and depth of field limits. In [8], a source camera can be identified from a printed image using the Photo Response Non-Uniformity (PRNU) pattern that is unique to each image sensor.

Most approaches for recapture detection rely on classifiers that use statistical parameters that are extracted from an image and a combination of features derived from training sets. Furthermore, existing methods detect recapture using evidence obtained from a single event. In our scheme, we apply the knowledge that the shape of an edge profile in an image is determined by the different stages of the image acquisition chain. Using a *structured* dictionary built in a controlled way from edges found in an image our proposed method is capable not only of detecting recapture but also of retrieving the

This work is partially supported by the REWIND Project, funded by the Future and Emerging Technologies (FET) programme within the 7th Framework Programme for Research of the European Commission, FET Open grant number:268478. Thirapiroon Thongkamwitoon is supported by The Office of NBTC and the Royal Thai Scholarship, THAILAND.

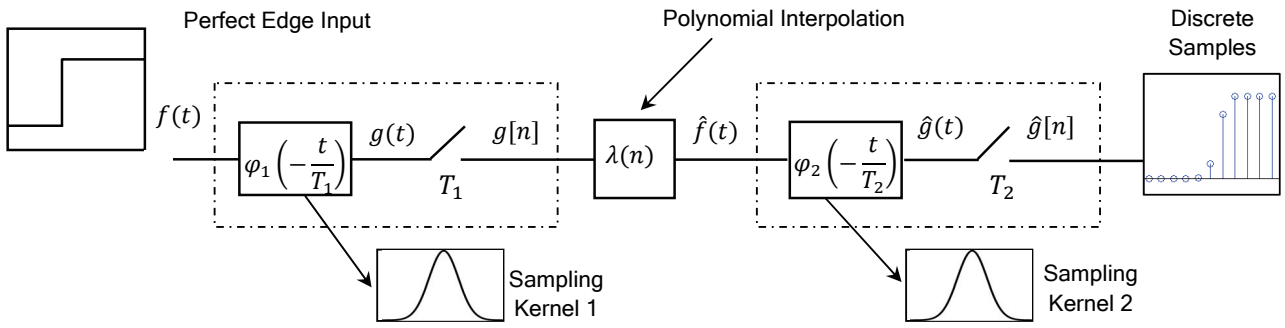


Fig. 1. Setup of signal diagram used in the research problem statement.

parameters in the image chain, thereby allowing individual devices in the chain to be identified.

The paper is organised as follows. In Section 2, we formulate a problem statement in order to model a series of processes in a chain of acquisition events with controlled conditions. A model for image acquisition and its unique properties is also described in this section. In section 3 we present our proposed framework with a dictionary setup and feature matching technique. The experiments and performance evaluation are discussed in Section 4 and conclusions are presented in Section 5.

2. PROBLEM SETUP

2.1. Problem Statement

In this section, a simplified model of an image acquisition chain is defined in order to enable us to study the degradations introduced into an edge by the image acquisition and reproduction chain. The only prior in this model is an input which is described by an ideal edge $f(t)$. The model incorporates the fundamental components of an imaging chain that comprise the processes of image acquisition, image reproduction, and image re-acquisition as shown in Figure 1.

The acquisition blocks are modeled by an A/D conversion step (See detail in section 2.2). Each of the blocks has two main parts: a sampling kernel and a sampling operation for a 1-D signal. The first and the second sampling kernels represent the impulse responses of the first and second acquisition devices respectively. The images may be sampled at different sampling rates T_1, and T_2.

After the first acquisition stage, a digitised version, $g[n]$, of the perfect edge, $f(t)$, is obtained. The discrete samples are then transformed back to the continuous domain by interpolation before being reacquired to the digitised form, $\hat{g}[n]$, by the second acquisition device. As a result of the reacquisition process, footprints that are present in the image, $\hat{g}[n]$, will have inherited characteristics from the original digitised signal, $g[n]$.

We now consider an unknown query digital image $q[n]$ given only the prior knowledge that it is the result of the acquisition of a perfect edge $f(t)$. We would like to answer the following questions:

1) What stages in the chain are the samples $q[n]$ from? (we are referring to the single captured $g[n]$ and recaptured $\hat{g}[n]$ images)

2) How can we identify the acquisition elements $\varphi_1(t)$ and $\varphi_2(t)$ (or $\varphi_1(t)$ in a single capture case) with access to $q[n]$ alone?

2.2. Image Acquisition Model

Image acquisition is a process that converts the spectral energy of a real scene to colour signals in the form of digital numeric values corresponding to the location of pixels in the image. The standard components of an acquisition device include the optical units, image sensor, and post processing units. Figure 2(a) describes the process of image acquisition which can be parametrically modeled as shown in Figure 2(b). The light-field image $f(x, y)$, that is captured by

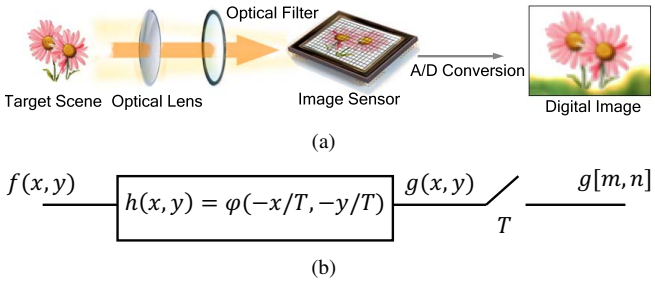


Fig. 2. Simplified image acquisition models represented by (a) acquisition components and (b) signal diagram.

an acquisition device, becomes blurred by the imperfection of the optical units and sensor. These imperfections can be modeled by the filter $h(x, y)$. The blurred image $g(x, y) = f(x, y) * h(x, y) = f(x, y) * \varphi(x/T - m, y/T - n)$ is then uniformly sampled with sampling period T, where $\varphi(x, y)$ is a sampling kernel which is a time reverse version of the transfer function $h(x, y)$. The discrete samples $g[m, n]$, representing a digital image, can be derived from sampling a blurred

image, $g(x, y)$, at the location $[m, n]$ as:

$$\begin{aligned} g[m,n] &= g(m,n) \\ &= \iint f(x,y)\varphi(x/T-m, y/T-n)dxdy \\ &= \langle f(x,y), \varphi(x/T-m, y/T-n)\rangle, \end{aligned} \quad (1)$$

where $x, y \in \mathbb{R}$ and $m, n \in \mathbb{Z}$. The parameters x, y represent the coordinate of the image plane while m, n are the row and column pixel indices of the image respectively.

2.3. Sampling Kernel and Line Spread Function

The response of an acquisition device to an input image can be modeled by the sampling kernel $\varphi(x, y)$ of the device. The response, which is frequently described by a Point Spread Function (PSF), represents the degree to which light from a point source is spread after it has been captured through the acquisition device. Thus the PSF determines the amount of blur that is introduced into the image. Typically, the PSF is unique to a device and its 3-dimensional shape depends on several factors such as the optics used during acquisition. The unique response of the PSF can, therefore, be used to trace back to the acquisition device that generated the image.

Edges are typical features that are present in almost every type of computer generated and natural image. We assume that the PSF is separable and therefore a one-dimensional (1D) model can be used to adequately characterise the sampling kernels. The blurring properties of 1D signals can be measured by the Line Spread Function (LSF) of the acquisition device. One way of determining the LSF, assuming that the PSF is circularly symmetric, is by integrating over cross-sections of the PSF along points that lie on a straight line.

For the purposes of simplifying our model we consider only 1D signals over the entire framework. With reference to Figure 2(b) the response, $g(t)$, to a perfect edge, $f(t)$, can be computed by $g(t) = f(t) * \varphi(-t/T)$, where $\varphi(t)$ is the LSF of the acquisition device. In our research, we computed the LSF of different acquisition devices using the SFRMAT 3.0 software [9] which is based on the slanted edge testing standard described in the international standard, ISO 12233.

3. IDENTIFICATION OF ACQUISITION CHAINS

In this section, we create a framework for the identification of the acquisition chain. A collection of dictionary elements was constructed corresponding to the discrete edge profiles that resulted from the acquisition of a perfect edge using a wide range of known capture devices at different stages of the imaging chain. Identification of a query image was done by determining the best match between an edge profile found in a query image and an edge profile from the dictionary. This enabled us to predict the device used to capture the image.

3.1. Dictionary of Key Footprints of Acquisition Devices

A dictionary $\mathcal{D}$ was built with a collection of features that represent blurred patterns introduced into a perfect edge by all individual possible devices and all possible stages in the chain. Dictionary elements, $\phi_{i,\gamma}$, for a single capture with a device, i, can be created as follows:

$$\phi_{i,\gamma} = \varphi_i * u[n-\gamma], \quad (2)$$

where φ_i is the 1D LSF of the device, i, $u[n]$ is a unit step function that represents a perfect edge, and γ are all possible shifts of edges in the window of interest.

Next, dictionary elements for recapture were built by simulating the actual image reproduction and reacquisition processes of the display and camera, respectively. Edge profiles from a single capture were linearly interpolated to increase the number of samples by 16 times before being resampled again by the second acquisition device. All possible combinations of acquisition devices were simulated at the second acquisition stage.

In this research, we captured images using three camera models, a Canon EOS 400D with Canon EF-S 18-135mm IS lens, a Canon PowerShot SX200 IS, and an Olympus E-P2 with 14-42mm lens kit. The cameras are referred to as C1, C2, and C3, respectively. To ensure that the Line Spread Functions for each device did not vary each time the camera was used, the focal length of each camera was fixed to the maximum optical zoom setting possible. Thus, for C1, C2 and C3 the focal length was set at 135mm, 35mm, and 42mm, respectively. We constructed a set of dictionary elements corresponding to the edge profiles present in images taken with the three cameras in the single capture and recapture camera combinations presented in Table 1. Some examples of the first derivative of edge profiles from cameras C1 and C2 and their recaptured version, C1-C2, is shown in Figure 3.

Table 1. DICTIONARY ELEMENTS USED IN THE EXPERIMENTS

Single Capture		**Recapture**	
Dictionary Element	Source Camera	Dictionary Element	Recapture Order
ϕ_1	C1	ϕ_4	C1 andC1
ϕ_2	C2	ϕ_5	C1 andC2
ϕ_3	C3	ϕ_6	C1 and C3
		ϕ_7	C2 and C1
		ϕ_8	C3 and C1

3.2. Edge Feature Matching

The purpose of feature matching is to find a dictionary element that provides the best representation of the observed edge from a query image. The best feature ϕ_m was chosen from a set of dictionary elements ϕ_i in the dictionary $\mathcal{D}$ by using the maximum inner product criterion.

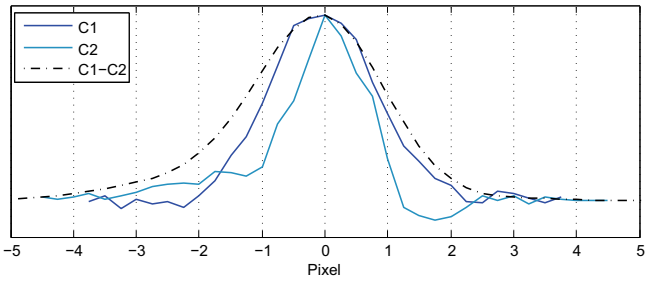


Fig. 3. Comparative plots of differentiated versions of the edge profiles from cameras C1 and C2 and their recaptured version C1-C2

$$\phi_m = \arg\max_{\phi_i \in \mathcal{D}} |\langle \mathbf{q}, \phi_i \rangle|, \tag{3}$$

where $\mathbf{q}$ is the vector of edge profiles from a query image and ϕ_m is the dictionary element with the best match. The sampling kernel and the stage in the chain that the image was from was then predicted by identifying the process that maximised the inner product between the edge feature and the dictionary element.

4. SIMULATION RESULTS AND DISCUSSIONS

4.1. Kernel and Acquisition Stage Identification

In this research we assumed that the LSFs of the capture devices were space-invariant. Inaccuracies introduced in capture device identification were due to the presence of image noise and to edge profiles that differed greatly from the profile of a ideal edge. Software was developed for selecting a region of interest (ROI) in a query image corresponding to a window of width 128 pixels and height 100 pixels. The ROIs were manually selected and centred on the in-focus areas with sharpest edges. A total of one hundred individual horizontal lines were used in dictionary matching in order to obtain a statistically significant match. Identification was based on two important results: the average inner product (%) and the confidence. The former represented the average inner product between a query edge and each dictionary element over 100 lines of a chosen area while the latter was the probability that the query image was classified into that category.

4.2. Simulation on Perfect Edges

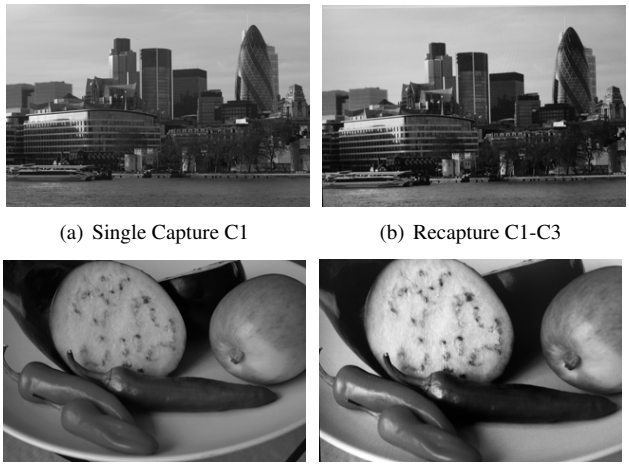
In this section, the proposed algorithm was tested with a set of query images that included both single capture and recaptured images. For the single capture case, a scene comprising a sheet of paper with an artificially generated sharp, high contrast, straight edge was captured digitally using camera C1. The image was displayed on a computer screen and recaptured using camera C2. The average inner product and confidence was computed between the recaptured edge profile and the dictionary elements and the results are summarised in Table 2.

Table 2. RESULTS FOR IMAGE CHAIN IDENTIFICATION USING A SINGLE CAPTURED IMAGE FROM C1 AND A RECAPTURED IMAGE FROM C1-C2

	Single Capture (C1)		Recapture (C1-C2)	
Dictionary Element	Average Innerproduct	Conf.	Average Innerproduct	Conf.
ϕ_1	**97.075**	**0.94**	78.650	
ϕ_2	87.818	0.01	68.030	
ϕ_3	95.903	0.05	79.623	
ϕ_4	92.799		84.998	0.10
ϕ_5	87.100		**86.996**	**0.68**
ϕ_6	87.807		85.766	0.22

4.3. Simulation on Real Images

The experiment was extended by applying the method to natural images. A collection of 160 single capture (original) and recaptured images containing natural scenery was used. For the group of images originating from a single capture, three sets of 20 images captured hand-held by cameras C1, C2 and C3 were used. For the recaptured images the three cameras were combined to create 5 groups of 20 recaptured images based on the following order of capture: C1-C1, C1-C2, C1-C3, C2-C1, and finally C3-C1. All the recaptured images were taken using a fixed tripod under controlled lighting conditions. Some examples are shown in Figure 4.



(a) Single Capture C1 (b) Recapture C1-C3

(c) Single Capture C3 (d) Recapture C3-C1

Fig. 4. Examples of images used in the experiments (a) Single capture from C1 (b) Recaptured image from C1 and C3 (c) Single capture from C3 (d) Recaptured image from C3- C1

In the experiment, images were classified into 8 groups based on the degree of similarity between the dictionary elements and edge profiles. According to the source of dictionary elements, images that were classified in groups 1-3 were categorised as single capture whereas recaptured images were classified in groups 4-8.

Table 3. SOURCE IDENTIFICATION RESULTS

Query Image		Dictionary Elements								Performance (%)
		Single Capture			Recapture					
		ϕ_1	ϕ_2	ϕ_3	ϕ_4	ϕ_5	ϕ_6	ϕ_7	ϕ_8	
Single capture	C1	**18**		1	1					90
	C2	1	**15**	3				1		75
	C3	1	1	**16**	1	1				80
Recapture	C1-C1	1			**17**	1			1	85
	C1-C2				1	**15**	1	3		75
	C1-C3				1	1	**16**		2	80
	C2-C1				1	1		**17**	1	85
	C3-C1	1			1		2		**16**	80

The results from the experiment are summarised in Table 3. From the 100 recaptured pictures, 98 were correctly identified as recaptured and 2 were misclassified as single capture. This resulted in a true positive rate of 98% and a false negative rate of 2% and suggests that the algorithm performed very well for the case of recapture detection. For the case of single capture detection, 56 out of the 60 single capture images were correctly classified resulting in a 93.3 % true negative rate and a 6.67% false positive rate.

As observed in Table 3 the eight sets of 20 query images were classified into 8 groups (ϕ_1 to ϕ_8 inclusive) based on the similarity between edge profiles in the images and the dictionary elements. The single capture images acquired by cameras, C1, C2 and C3 were correctly identified with an accuracy of 90%, 75% and 80% respectively resulting in an average performance of 81.67% over all three cameras. The performance of groups that contained camera C1 (Canon EOS 400D) was slightly higher than groups that did not contain C1. This is probably due to the use of manual settings on C1 but not on cameras C2 and C3 where the feature was not available. For the group of recaptured images, the average identification performance was 81%, resulting in an average overall performance for single and recaptured images of 81.25%.

5. CONCLUSIONS

A proposed framework for the identification of image acquisition chains with application to recapture detection and source camera identification has been presented. The query images were classified based on the unique properties of edge profiles found in the image resulting from the different stages of the image acquisition chain.

A dictionary of edge profiles was constructed from edges that were transformed by the line spread functions resulting from the combination of all known image capture devices and acquisition chains. The maximum inner product was used to determine the best match between an edge found in a query image and a dictionary element. The algorithm was tested with sets of ideal synthetic edges and real natural images and the results suggest that the method is capable of delivering accurate image recapture detection and image chain identification.

6. REFERENCES

[1] Hany Farid, "Image forgery detection – a survey," *IEEE Signal Processing Magazine*, vol. 2, no. 26, pp. 16–25, 2009.

[2] Ashwin Swaminathan, Min Wu, and K. J. Ray Liu, "Digital image forensics via intrinsic fingerprints," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 1, pp. 101–117, 2008.

[3] Jan Lukás and Jessica Fridrich, "Estimation of primary quantization matrix in double compressed JPEG images," in *Proc. of DFRWS*, 2003.

[4] A.C. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of resampling," *Signal Processing, IEEE Transactions on*, vol. 53, no. 2, pp. 758 – 767, Feb. 2005.

[5] Babak Mahdian and Stanislav Saic, "Blind authentication using periodic properties of interpolation," *IEEE Transactions on Information Forensics and Security*, vol. 3, no. 3, pp. 529–538, 2008.

[6] Hong Cao and Alex C. Kot, "Identification of recaptured photographs on LCD screens," in *ICASSP*, 2010, pp. 1790–1793.

[7] Xinting Gao, Tian Tsong Ng, Bo Qiu, and Shih-Fu Chang, "Single-view recaptured image detection based on physics-based features," in *IEEE International Conference on Multimedia and Expo (ICME)*, July 2010.

[8] Miroslav Goljan, Jessica Fridrich, and Jan Luk, "Camera identification from printed images," in *Symposium on Electronic Imaging Science, Security, Forensics, Steganography, and Watermarking of Multimedia Contents X*, 2008.

[9] P. Burns, *Slant edge analysis tool SFRMAT 3.0*, http://www.imagescienceassociates.com, 2010.