POSTER: **Who was Behind the Camera?**

— Towards Some New Forensics

Jeff Yan
Linköping University, Sweden
jeff.yan@liu.se

Aurélien Bourquard
Massachusetts Institute of Technology, USA
aurelien@mit.edu

ABSTRACT

We motivate a new line of image forensics, and propose a novel approach to photographer identification, a rarely explored authorship attribution problem. A preliminary proof-of-concept study shows the feasibility of our method. Our contribution is a forensic method for photographer de-anonymisation, and the method also imposes a novel privacy threat.

KEYWORDS

Forensics, privacy, photographer identification, photographer de-anonymisation, inverse problems

1 INTRODUCTION

We consider such a research problem: given a *single* photo, how to determine who was the cameraman? This is in general a hard problem, except for selfies and except if the photographer's shadow became visible in the photo or her image was captured by a reflective object in the photo, such as a subject's eyes.

This problem is interesting to intelligence agencies. For example, a photo of a secret military facility in Russia can be valuable to the Central Intelligence Agency of USA. However, when the photo gets leaked by a mole inside the CIA, Russia's anti-spy operatives would be keen to work out who took the photo in the first place.

The problem is interesting to law enforcement agencies, too. For example, when the Scotland Yard are tipped off by a photo from an anonymous source that offers clues to a criminal investigation, it is likely to gain further information to accelerate their investigation by identifying the person behind the camera.

Moreover, the problem is also interesting to privacy researchers. The answer to the research question will likely provide novel methods of privacy intrusion by de-anonymising a photographer of any concerned photo on the Internet, and motivate novel research for protecting photographers' anonymity.

Not all photographers care that it is public knowledge that some photos are taken by them. But in some circumstances, some photographers would care if some photos are linked to them as the people behind the camera.

From the forensic perspective, a technique that does not identify the photographer 100% of the time can still be practically useful, since it will narrow down suspects to a small number. Complemented with other means such as surveillance, it is highly likely for intelligence agencies or law enforcements to pin down the concerned photographer accurately.

We first review related work, and show that existing approaches do not resolve the research question we are asking. Then, we propose a new approach, and demonstrate its feasibility by a proof-of-concept but realistic simulation study. Our method is applicable to both digital and film photography, in theory.

2 RELATED WORK

Visual stylometry. Artists like Claude Monet and Vincent van Gogh demonstrate distinctive styles in their paintings. In the past hundreds of years, people relied on stylistic analysis to tell apart genuine fine art from fakes. It became an emerging research area in recent years to apply signal processing and machine learning methods to analyse painting images for artist identification [1, 2].

Similarly, some photographers display peculiar styles in the photos they produce. For example, widely regarded as one of the best portrait photographers of all time, Yousuf Karsh is known for distinctive features in his portraits due to lighting, composition and posture. Ernst Haas showed a distinctive personal style in his impressionist colour photography, too. Therefore, it is a natural extension to develop photographer identification methods from painting artist identification.

However, a training set of photos a priori, usually of a large size, is needed for each concerned photographer to make machine learning methods to work. This approach will hardly work if the given photo is the only available one taken by a suspect photographer, since it is impossible to collect a training set of photos for the photographer. On the other hand, if a photographer's style is not sufficiently sophisticated, it is easy for somebody else to emulate. This can be exploited to fool machine learning algorithms, and to frame a photographer.

Camera fingerprint [3,4]. CCD or CMOS imaging sensors are a digital camera's heart. Due to sensor design and imperfections of the sensor manufacturing process, systematic artefacts (usually known as sensor pattern noises) form an equivalent of a digital fingerprint that can identify a camera. Such fingerprints are intrinsically embedded in each image and video clip created by a digital camera. Forensic applications of camera fingerprints include 1) source camera identification (which camera was used to produce this image?), and 2) device linking (were two images produced by the same camera?).

Camera fingerprint, in theory, can link a photo to a specific camera, if a reference fingerprint can be established for the camera,

This work is partially supported by the Knut and Alice Wallenberg Foundation.

CCS '17, October 30-November 3, 2017, Dallas, TX, USA

ACM ISBN 978-1-4503-4946-8/17/10.
https://doi.org/10.1145/3133956.3138848

e.g. when the camera is physically accessible, or a set of photos taken by the same camera is otherwise available. However, camera fingerprint does not link a photo to a specific user of the camera. This is an issue when the same camera has been used by many people. Moreover, camera fingerprint can be easily removed from each photo, entirely disabling its forensic applications. The camera fingerprint technique has been developed for digital cameras, and it does not work for traditional film photography.

Image metadata has a limited forensic application. For example, it can link a digital image to a camera model at most, not to a specific camera, let alone a photographer. On the other hand, film photography does not produce any such metadata.

All the methods discussed above do not really provide a solution to our research question.

3 A NOVEL METHOD

When a scene is photographed, a photographer's body often deflects light (via reflection and refraction) into the scene, leaving an impact on a photo created thereafter. Our hypothesis is that light rays deflected by a photographer into the photo that she is taking will give away some physical characteristics of herself.

We refine our research problem as follows. Photo P_1 was taken of a scene by a photographer at will, i.e. its acquisition is non-controlled; our task is to work out who took the photo. We have access to the same physical scene, and we take a photo P_2 similar to P_1, while all acquisition parameters are reproduced in a controlled manner to be the same as used for producing P_1, except that the photographer is absent. Our research questions are: 1). What differences in P_1 and P_2 can be exploited to deduce the photographer's physical characteristics? 2). Under what conditions the above measurement will work for the purpose?

We choose to answer these questions via a realistic simulation, rather than an empirical lab study, since the latter involves with experiments that are more expensive and sophisticated to set up. Specifically, we use photon mapping, a well-established ray tracing technique, to conduct a proof-of-concept feasibility study. Photon mapping realistically simulates the interaction of light with different objects. In this approach, light rays from a light source and rays from a camera are traced independently until some termination criterion is met. Then, they are connected in a second step to produce a radiance value.

3.1 Experimental Design

We use the popular POVRay software[1] for scene definition and rendering, as well as photon mapping.

3.1.1 Scene Definition. Fig. 1 illustrates the scene that we use for this study, as viewed from the camera. The ground consists of a brown surface of unit-normalized RGB colour (0.80, 0.55, 0.35). The camera capturing the scene is placed 1.5m above the ground, and 2m away from a dark wall that is modelled as a non-reflective rectangular object of 1m width and 1.9m height. This wall casts a shadow on the floor, because of a light source 3m high and 1m behind the wall. The camera's angle of view is 90°, which many lenses can achieve in photography, and the camera is oriented towards the corner formed by the floor and the wall.

The ground and wall surfaces are flat and modelled with ambient-light and diffuse-light coefficients of 0.1 and 0.9, respectively. The resolution of the rendered scene is 1600x900 pixels. The bit depth is 16 bits per colour channel, in RGB format; this allows minimizing numerical errors.

The photos P_1 and P_2 are acquisitions of the underlying 3D scene described above, from the point of view of the camera, taken with and without the presence of the photographer, respectively. Accordingly, every picture is a rendered version of the scene computed through the POVRay software.

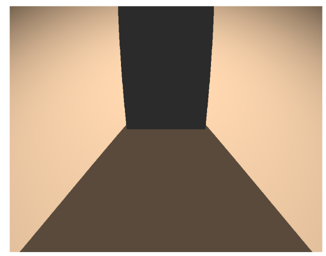

Fig. 1: Scene as viewed from the camera, without the presence of the photographer, as defined in Sect. 3.1.1. The wall (black) and its shadow on the floor (brown) are visible in the picture.

3.1.2 Photographer. When present, the photographer faces the wall in the scene and stands together with the camera. The photographer's jacket is modelled as a reflective rectangular object whose width and height are free parameters. The jacket exhibits surface irregularities in form of bumps, whose characteristic widths parallel to the surface is 30cm, and whose depth normal to the surface is left as a free parameter, as for the case of the jacket colour. Accordingly, the jacket material reflects light from the light source onto the floor of the scene.

The jacket-surface bumps are modelled by the so-called bump-mapping technique with a smooth-random-noise function. This approach allows simulating accurate surface properties without increasing the complexity of the underlying surface geometry. The light reflections from the body surface onto the floor and the wall are simulated using photon mapping with a count of 20×10^6 photons, a figure that is empirically determined to be sufficient to converge to a high-quality scene rendering. Photon mapping is essential to model the effect of reflected light from the complex jacket surface onto the rest of the scene (especially the floor) by simulating trajectories of individual photons emitted from the light source and infer their distribution accordingly.

Based on the pair of photographs P_1 and P_2, the parameters that are estimated with our method are the (a) height, (b) width, and (c) colour of the photographer's jacket, where the jacket dimensions are assumed to match the photographer's dimensions, and (d) the presence of bumps of distinct depths on the jacket surface. A non-flat surface typically exhibits bumps, whose size and depth normal to the surface may vary; a zero depth is equivalent to a flat surface. In conjunction with colour, the presence and dimensions of bumps characterize the type or class of material used. Indeed, fibres

[1] http://www.povray.org

composed of different fabrics are expected to modulate light-reflection properties differently through their surface irregularities. The observed light-brightness distribution on the photo P_1 used for estimation is thus expected to vary accordingly, which can be used for estimation.

Each of the parameters (a)-(d) is varied within a certain range and compared to corresponding estimates, using 8 data points in total, and using preassigned default values for the other parameters. The jacket width ranges from 0.5m to 1.5m. The jacket height ranges from 1m to 2m. The depth of bumps normal to the jacket surface ranges from 0 to 20cm. Following the RGB convention, the jacket colour ranges from 0 to 1 for the G channel, the other colour channels being left to their default values.

In our study, each parameter of interest is varied and estimated independently while other parameters stay constant at their default values. This allows inferring preliminary yet indicative proof-of-principle results where confounding factors are minimized.

3.1.3 Noise levels. For every parameter and estimations we work on, we consider several simulated noise levels in rendered scenes, both with and without the photographer, corresponding to various signal-to-noise ratio (SNR) levels (defined as the ratio between the energies of the signal and of the noise) for the photographs in decibels (dB). Specifically, we consider cases with 25, 30, 35, 40, 45, 50, and $+\infty$ dB. The noiseless case corresponds to a perfect replication of photo acquisition conditions, including the camera being the same model. Other noise levels help to lessen our tight control, by allowing for example a camera different from the one used by the original photographer, and by accounting for the presence of sensor noise.

3.3 Results and Discussions

Due to space limits, we only present selected results but omit details of our estimation methods. Figs. 2-3 show that the photographer's width and height can be estimated based on the rendered scene (estimates in pixels, according to the resolution of the rendered scenes), even though the relationship is not linear. In particular, the increase in the estimated values is monotonic with respect to the original scene-parameter values for these geometric features, which allows for further calibration. The estimation starts to break down below a certain SNR level due to noise.

Finally, Fig. 4 shows that the estimated bump depth value (determined using normalized gradients on the rendered scenes) increases with the corresponding scene parameter. Noise also affects the estimation results because it increases the perceived surface irregularity viewed from the rendered scene, even though Gaussian filtering was used to mitigate the effect. From a forensic perspective, the bump-depth estimate could provide information on fabric materials of the clothes the photographer was wearing. While it may be hard to pinpoint exact materials, our method could potentially classify them into several categories.

REFERENCES

[1] C. R. Johnson Jr, E. Hendriks, I. J. Berezhnoy, E. Brevdo, S. M. Hughes, I. Daubechies, J. Li, E. Postma, and J. Z. Wang. Image processing for artist identification. IEEE Signal Processing Magazine, 25(4):37–48, 2008.
[2] K Alfianto Jangtjik, M-C Yeh, K-L Hua. Artist-based Classification via Deep Learning with Multiscale Weighted Pooling. ACM Multimedia 2016: 635-639
[3] J. Lukas, J. Fridrich, and M. Goljan. Digital camera identification from sensor pattern noise. IEEE Trans. Inf. Forensics & Security, 1(2):205 214, 2006.
[4] J Yan. Novel security and privacy perspectives of camera fingerprints. Twenty-fourth International Workshop on Security Protocols, Apr 2016. Springer LNCS.

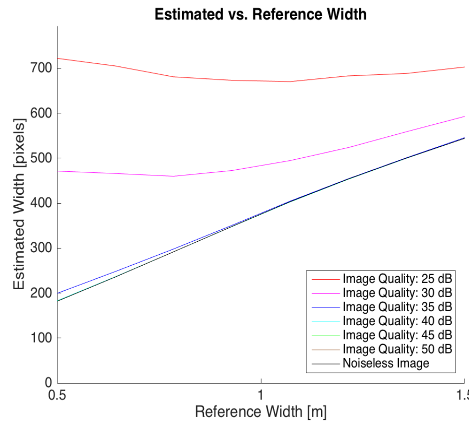

Fig. 2: Estimated vs. reference width (0.5~1.5m range)

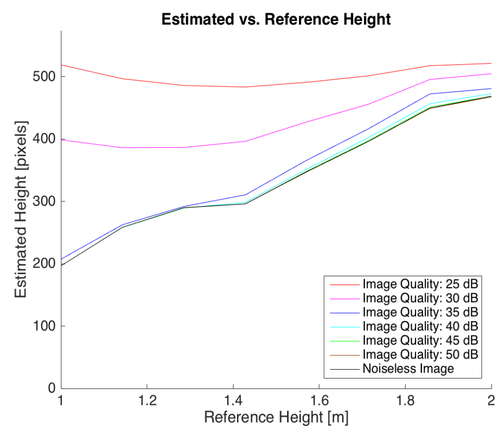

Fig. 3: Estimated vs. reference height (1~2m range)

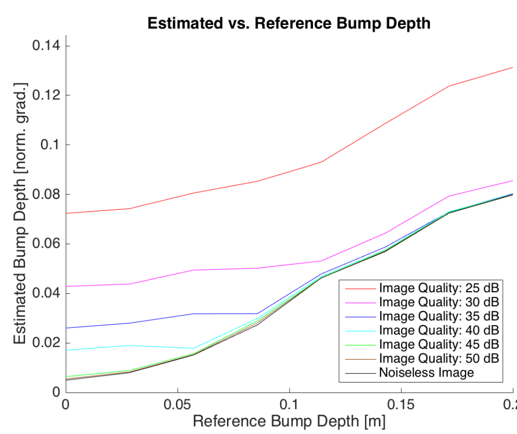

Fig. 4: Estimated vs. reference bump depth (0~20cm range)