Source Camera Identification Using Footprints from Lens Aberration

Kai San Choi, Edmund Y. Lam, Kenneth K.Y. Wong

Department of Electrical and Electronic Engineering,
The University of Hong Kong, Pokfulam Road, Hong Kong.

ABSTRACT

Source camera identification is the process of discerning which camera has been used to capture a particular image. In this paper, we consider the more fundamental problem of trying to classify images captured by a limited number of camera models. Inspired by the previous work that uses sensor imperfection, we propose to use the intrinsic lens aberration as features in the classification. In particular, we focus on lens radial distortion as the primary distinctive feature. For each image under investigation, parameters from pixel intensities and aberration measurements are obtained. We then employ a classifier to identify the source camera of an image. Simulation is carried out to evaluate the success rate of our method. The results show that this is a viable procedure in source camera identification with a high probability of accuracy. Comparing with the procedures using only image intensities, our approach improves the accuracy from 87% to 91%.

Keywords: Source camera identification, forensic science, statistical classification, image processing, lens aberration.

1. INTRODUCTION

In many court cases, images taken by film cameras are used as evidence in the trials. As digital cameras are increasingly popular, one would expect that there will be an increase in the use of digital images as evidence. However, due to the availability of powerful softwares, manipulation of digital images is easy and most of the time, it leaves no obvious traces. Therefore, the credibility of an image hinders the usefulness of a digital image in many court cases. As a consequence, in image forensics, one would like to be able to ascertain that a digital image has not been tampered, and furthermore, that it is captured by a particular camera. In this paper, we focus on the scenario to distinguish between images captured by a limited number of camera models.

The existing approaches for camera identification can be grouped into three categories.

The first category is to examine the image files' headers or watermarks. Most cameras attach an Exchangeable Image File Format (EXIF) header in their images. The header contains information such as digital camera type, exposure, date and time about an image. However, this information may be maliciously altered and they may not be available if the image is resaved or recompressed. Some specially designed cameras can embed a watermark in images. The watermark can carry information about the biometric data of the photographer, camera, time and date.[1] However, majority of the digital cameras available in the market do not contain this function. Therefore, it is imperative to find another way to solve this problem.

The second category is to make use of the difference in image processing methods among camera models. Demosaicing, gamma correction, color processing, white point correction and compression are standard processes in digital cameras.[2] However, the exact processing details or algorithms used may vary from one manufacturer to another. As a result, it is believed that the output image may exhibit some traits and patterns regardless of the original image content. Kharrazi et al.[3] tried to capture those traits by proposing a vector of thirty-four features obtained from pixel intensities to represent an image. The features include average pixel value, RGB

Further author information: (Send correspondence to K.S. Choi)
K.S. Choi: E-mail: kaisan@eee.hku.hk, Telephone: (852) 2859 2696
Edmund Y. Lam: E-mail: elam@eee.hku.hk, Telephone: (852) 2241 5942
Kenneth K.Y. Wong: E-mail: kywong@eee.hku.hk, Telephone: (852) 2857 8483

Digital Photography II, edited by Nitin Sampat, Jeffrey M. DiCarlo, Russel A. Martin,
Proc. of SPIE-IS&T Electronic Imaging, SPIE Vol. 6069, 60690J, © 2006 SPIE-IS&T · 0277-786X/06/$15

pairs correlation, center of mass of neighboring distribution, RGB pairs energy ratio, wavelet domain statistics and a set of Image Quality Metrics (IQM)[4,5] which was used previously for steganalysis problems. Then a classifier is used to identify the source camera of an image.

The third category is to use the noise pattern in digital cameras. In the Ref. 6 and 7, camera identification methods by CCD noise pattern are proposed. The noise pattern is caused by several factors, e.g. pixel non-uniformity, dust on lens, dark currents resulted by defective manufacturing processes. These noise patterns are unique for each camera and can be used as camera identification purpose. Lukas et al.[7] proposed to use a Gaussian denoising filter[8] to extract the pattern noise. The reference noise pattern of a camera is obtained by averaging a number of images. Then the source camera of an image is determined by correlating the noise pattern in the image and the reference noise pattern.

Inspired by the work on CCD noise pattern, we use the unique imprints left behind by the lens on images for camera identification. We propose to use the lens aberrations to help in classifying images originating from a number of cameras. As all lens elements inevitably produce some aberrations, they also leave unique imprints on the images being captured. Using the straight-line camera calibration algorithm described in the Ref. 9, we estimate the radial distortion parameters of the lens from an image. The radial distortion parameters, together with those features proposed by Kharrazi et al.,[3] forms the feature vector in the classification. We then use a support vector machine (SVM)[10] classifier to evaluate the success rate of the classification. In this paper, we show that lens radial distortion is an effective feature in improving the success rate of the source camera identification problem.

This paper is organized as follows. In Sect. 2, we first give a background introduction to lens radial distortion and discuss the radial distortion measurement method used in our experiment. In Sect. 3, we propose an approach to incorporate our lens radial distortion measurement into Kharrazi's feature-based method,[3] in order to increase the classification accurarcy. Experimental results for the three camera case are provided in Sect. 4. The future work and conclusion are presented in Sect. 5.

2. LENS RADIAL DISTORTION

2.1. Background of Lens Radial Distortion

Due to the design and manufacturing process, lens produces aberrations in images. There are different kinds of lens aberrations such as spherical aberration, coma, astigmatism, field curvature, lens radial distortion and chromatic distortion. However, lens radial distortion is the most severe part among the aberrations, especially when inexpensive wide-angle lenses are used.[11] In this paper, we will focus on lens radial distortion.

The radial distortion causes straight lines in the object space rendered as curved lines on the film or camera sensor. It originates from the transverse magnification M_T.[12] The transverse magnification of an image is the ratio of the image distance to the object distance. Radial distortion arises when M_T is a function of the off-axis image distance, r, rather than a constant. In other words, the lens has various focal lengths and magnifications in different areas. The radial distortion deforms the whole image even though every point is in focus. When M_T increases with r, the optical system suffers from barrel distortion. Similarly, when M_T decreases with r, the optical system suffers from pincushion distortion. Examples of barrel distortion and pincushion distortion are shown in Fig. 1 and Fig. 2.

For the manufacturing cost, most of the digital cameras are limited to lenses having spherical surfaces.[12] These spherical lenses have inherent radial distortion and must be corrected by manipulating the system variables (indices, shapes, spacing, stops, etc). However, the degree and order of compensation are varied from one manufacturer to the others or even in different camera models by the same manufacturer. Therefore, lens from different camera model may have different degree of radial distortion. Apart from the lens design, the degree of lens distortion depends on the focal length of lenses.[13] Usually, lenses with short focal length have a larger degree of barrel distortion, while lenses with long focal length suffer more the pincushion distortion. As a result, lenses from different camera leave unique imprints on the pictures being captured.

In the next section, we will introduce how to estimate the lens radial distortion from an image, and the feasibility of using it for camera identification problem.

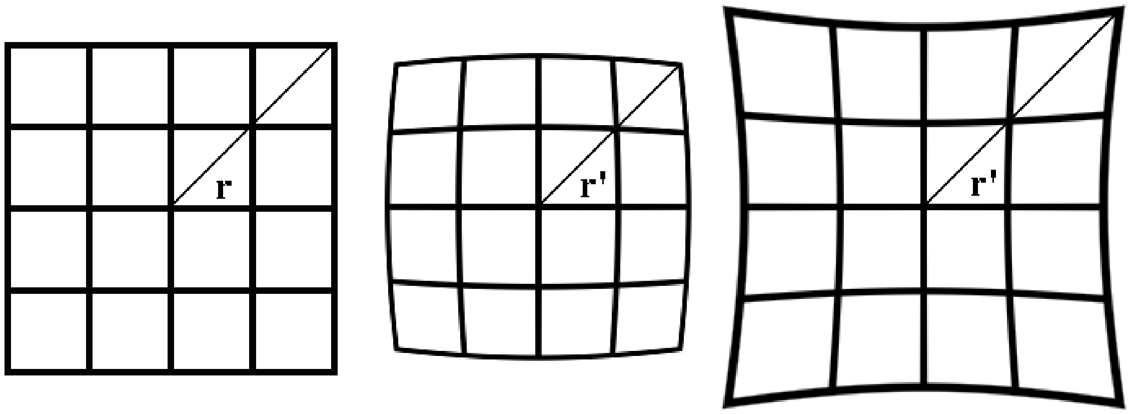

Figure 1. Distortion of a rectangular grid. Left: Undistorted grid. Middle: Grid with barrel distortion. Right: Grid with pincushion distortion.

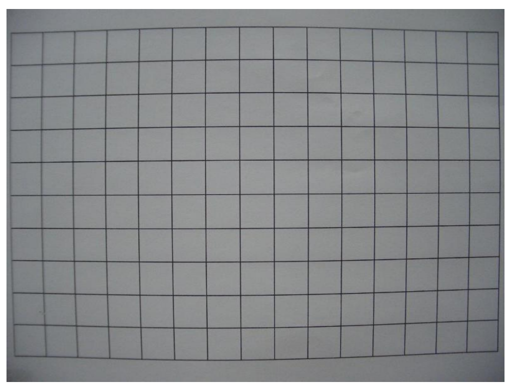

Figure 2. A rectangular grid taken by Camera B. The grid has barrel distortion.

2.2. Measuring Lens Radial Distortion

The lens radial distortion model can be written as an infinite series. Usually, the first order radial symmetric distortion parameter k_1 can achieve enough accuracy.[9] In order to achieve a higher accuracy, we use the first order and sencond order distortion parameters as an estimation of the degree of distortion in an image. The lens radial distortion can be written as[9]:

$$r_u = r_d + k_1 r_d^3 + k_2 r_d^5 \tag{1}$$

where r_u and r_d are the undistorted radius and distorted radius respectively, and k_1 and k_2 are the first order and second order distortion parameters respectively. The radius is the radial distance $\sqrt{x^2 + y^2}$ of a point (x, y) from the center of distortion. In this paper, we simply take the center of an image as the center of distortion.

Devernay[9] introduced a straight line method for computing radial distortion. He exploited the fundamental property that if a lens does not have radial distortion, every straight line in space should be projected as a straight line onto an image. By using this property, an iterative process is exployed to estimate the distortion parameter k_1. The process is to first extract the distorted line segments by sub-pixel edge dectection and polygonal approximation. The distortion error is measured between the distorted line segments and their corresponding straight lines. Then the distortion parameters, k_1 and k_2, are optimized to minimize distortion error. The optimization process is repeated until the relative change of distortion error is less than a threshold. An implementation of Devernay's algorithm in Matlab is available on the Internet.[14] We modify that program to estimate radial distortion parameters, k_1 and k_2, for every image under consideration.

3. USING LENS RADIAL DISTORTION AS FEATURES

According to the Ref. 3, Kharrazi et al. proposed to use a number of features to capture the photometric effects left behind by the camera on the images. Apart from the color processing algorithms, we believe that an output image is also affected by the intrinsic lens aberrations on the cameras. One kind of the aberrations is called radial distortion. As explained in the previous section, lens radial distortion causes straight lines in an object appeared as curves in the output image. In order to capture the radial distortion characteristics of different cameras, we propose to incorporate lens radial distortion parameters, k_1 and k_2, as two new features to aid the classification. We believe that our new features can capture the geometric footprints left behind by the camera (lens) on the images and can serve as complementary features to the features proposed by Kharrazi et al.

For each image under consideration, a vector of thirty-six features is extracted from the image. The feature vector consists of thirty-four features proposed by Kharrazi et al. and our lens radial distortion parameters, k_1 and k_2. Assuming that a collection of images is available for each possible camera, they are then used to train a classifier for distinguishing between images originating from a specific camera.

4. EXPERIMENTAL RESULTS

Two set of experiments were performed. The first set of experiment is a feasibility test of using lens radial distortion in image classification. The second set of experiment shows that our approach has a statistically significant improvement in accuracy over the procedures using only image intensities.

4.1. Cameras and test images

In order to observe the effectiveness of lens radial distortion in classifying images originating from a digital camera, we conducted two set of experiments. In our experiments, we used three different cameras which are recent models from three famous manufacturers. The Camera A and Camera B were used to produce 1600×1200 images, while Camera C was used to take 2560×1920 images. The images were taken with no flash, auto-focus, no manual zooming, best JPEG compression quality, and other settings set to the default values. The configurations of the cameras are shown in Tab. 1. A picture data set was prepared by taking 100 images with each camera randomly around the university campus. Figure 3 presents some of the samples from our image data set. After collecting the data set, the poposed measures were calculated for each image.

Table 1. Cameras used in experiments and their properties

Camera brand	Resolution	Image Format
Camera A	1600×1200	JPEG
Camera B	1600×1200	JPEG
Camera C	2560×1920	JPEG

4.2. Image classification by lens radial distortion only

The first experiment is to show the feasibility of using lens radial distortion in classifying images originating from a three-camera model. Firstly, we obtain the lens distortion parameters, k_1 and k_2, for all images. Then the steps for training and testing the SVM classifier[15] are as follows:

1. 40 images were randomly selected to train a classifier.
2. The rest of the images were used to test the classifier.

Figure 3. Sample images obtained using the Camera A.

3. Step 1 and Step 2 were repeated 50 times and the average classification accuracy is obtained.

The success rate of the classification for Camera A, Camera B and Camera C are 97.8%, 92% and 84.8% respectively. The average accuracy obtained is 91.53% and the corresponding confusion matrix is in Tab. 2.

Table 2. The confusion matrix for camera identification by lens radial distortion only.

		Predicted (%)		
		Camera A	Camera B	Camera C
Actual(%)	Camera A	**97.8**	1.1	1.1
	Camera B	5.6	**92**	2.4
	Camera C	3.13	12.07	**84.8**
Average Accuracy (%)	**91.53**			

The scatter plot of the lens radial distortion paramters, k_1 and k_2, is presented in Fig. 4. Camera A, Camera B and Camera C are represented by '$\times$', 'o' and '+' respectively. It can be seen that the lens radial distortion parameters can be clearly separable into three groups. Those outliers in the plot are images with too short straight lines or too few straight lines. The classification results in Tab. 2 and the scatter plot in Fig. 4 show that it is feasible to identifying the source camera of a digital image by lens radial distortion.

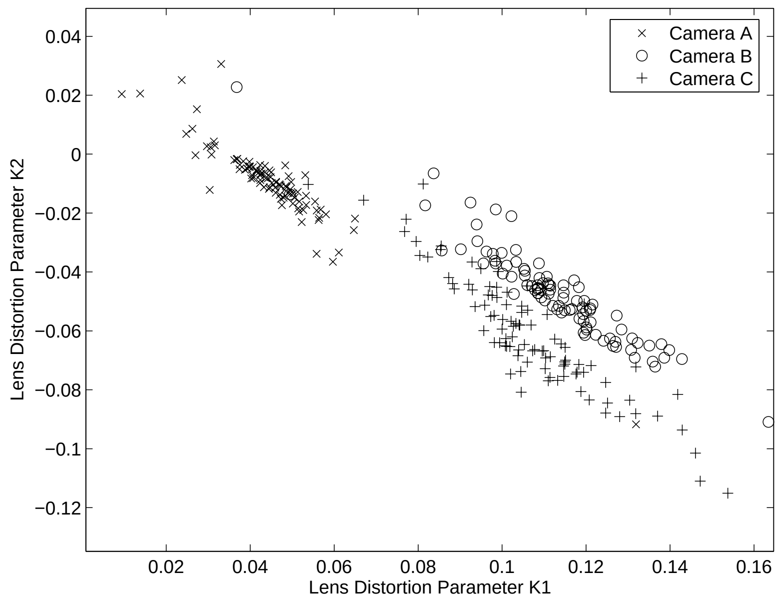

Figure 4. The scatter plot of lens radial distortion parameters, k_1 and k_2, in Camera A ($\times$), Camera B (o) and Camera C (+) camera in our experiment.

4.3. Image classification by lens radial distortion and Kharrazi's proposed features

In this section, we investigate the improvement in accuracy by adding radial distortion to Kharrazi's proposed statistics. Based on Kharrazi's features, we evaluate the accuracies of the classification with and without radial distortion. In these two experiments, we use the SVM classifier available in the LibSvm package.[15] The procedures for training and testing the classifier are the same as the Sect. 4.2. The average accuracy of the system with and without lens radial distortion parameter is 87.38% and 91.39% respectively. There is a 4.01% improvement in accuracy with our proposed lens radial distortion feature. The corresponding confusion matrices of these two experiments are in Tab. 3 and Tab. 4 respectively.

Table 3. The confusion matrix for camera identification by features proposed by Kharrazi et al. only

		Predicted (%)		
		Camera A	Camera B	Camera C
Actual(%)	Camera A	**83.7**	14.87	1.43
	Camera B	15.9	**83.33**	0.77
	Camera C	2.43	2.47	**95.1**
Average Accuracy (%)	**87.38**			

Table 4. The confusion matrix for camera identification by lens radial distortion and features proposed by Kharrazi et al.

		Predicted (%)		
		Camera A	Camera B	Camera C
Actual(%)	Camera A	**90.67**	8.07	1.27
	Camera B	10	**88.37**	1.63
	Camera C	2.33	2.53	**95.13**
Average Accuracy (%)	**91.39**			

5. CONCLUSION

In this paper, we examine the lens footprints left on the images for identifying the source camera of a digital image. We propose to use the lens radial distortion on the images for this problem. A classifier based on lens radial distortion is built and used to evaluate the effectiveness of this feature. We show that it is feasible to use the lens radial distortion to classify images originating from a three-camera model. We also propose to incorporate our lens radial distortion with the statistics obtained from image intensities for image classification. We demonstrate that comparing with the procedures using only statistics from image intensities, our approach shows a statistically improvement in accuracy.

Lukas et al.[7] have pointed out that reliable camera identification should be approached from multiple directions, combining with evidences obtained from various methods. Our proposed lens radial distortion is another evidence for solving source camera identification problem. Our initial results are encouraging. They can be improved by using a more sophisticated method to estimate the lens distortion from an image. There is one major limitation of our proposed feature. When images from a specific camera are taken by zoom lens with different manual zooming, the measured lens distortion parameter will span over a large range. This may decrease the success rate of the classification. However, we believe that this technique will still prove useful in a number of different situations — for example images are taken by SLR cameras with prime lens (fixed focal length lens) or we have knowledge that the images from zoom lens are taken without manual zooming.

REFERENCES

1. P. Blythe and J. Fridrich, "Secure digital camera," in *Digital Forensic Research Workshop*, 2004.
2. J. Adams, K. Parulski, and K. Spaulding, "Color processing in digital cameras," *IEEE Micro* **18**(6), pp. 20–30, 1998.
3. M. Kharrazi, H. T. Sencar, and N. Memon, "Blind source camera identification," in *International Conference on Image Processing*, pp. 709–712, 2004.
4. I. Avcibas, N. Memon, and B. Sankur, "Steganalysis using image quality metrics," *IEEE Transactions on Image Processing* **12**(2), pp. 221–229, 2003.
5. I. Avcibas, B. Sankur, and K. Sayood, "Statistical evaluation of image quality metrics," *Journal of Electronic Imaging* **11**(2), pp. 206–223, 2002.
6. N. Saitoh, K. Kurosawa, K. Kuroki, N. Akiba, Z. Geradts, and J. Bijhold, "CCD fingerprint method for digital still cameras," in *Investigative Image Processing II*, *Proc. SPIE* **4709**, pp. 37–48, 2002.
7. J. Lukáš, J. Fridrich, and M. Goljan, "Determining digital image origin using sensor imperfections," in *Image and Video Communications and Processing*, *Proc. SPIE* **5685**, pp. 16–20, 2005.
8. M. K. Mıhçak, I. Kozintsev, and K. Ramchandran, "Spatially adaptive statistical modeling of wavelet image coefficients and its application to denoising," in *International Conference on Acoustics, Speech, and Signal Processing*, **6**, pp. 3253–3256, 1999.
9. F. Devernay and O. Faugeras, "Automatic calibration and removal of distortion from scenes of structured environments," in *Investigative and Trial Image Processing*, *Proc. SPIE* **2567**, pp. 62–67, 1995.
10. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Wiley, New York, 2001.
11. J. Perš and S. Kovačič, "Nonparametric, model-based radial lens distortion correction using tilted camera assumption," in *Proceedings of the Computer Vision Winter Workshop 2002*, pp. 286–295, 2002.
12. E. Hecht, *Optics*, Addison Wesly, San Francisco, California, 2002.
13. B. Tordoff and D. W. Murray, "Violating rotating camera geometry: the effect of radial distortion on self-calibration," in *Proc. 15th International Conference on Pattern Recognition*, **1**, pp. 423–427, 2000.
14. P. D. Kovesi, *Matlab and Octave Functions for Computer Vision and Image Processing*, Software available at http://www.csse.uwa.edu.au/~pk/research/matlabfns/.
15. C.-C. Chang and C.-J. Lin, *LIBSVM: a Library for Support Vector Machines*, Software available at http://www.csie.ntu.edu.tw/~ccjlin/libsvm, 2001.