

Estimation of Lens Distortion Correction from Single Images

Miroslav Goljan and Jessica Fridrich

Department of ECE, SUNY Binghamton, NY, USA, {mgoljan,fridrich}@binghamton.edu

ABSTRACT

In this paper, we propose a method for estimation of camera lens distortion correction from a single image. Without relying on image EXIF, the method estimates the parameters of the correction by searching for a maximum energy of the so-called linear pattern introduced into the image during image acquisition prior to lens distortion correction. Potential applications of this technology include camera identification using sensor fingerprint, narrowing down the camera model, estimating the distance between the photographer and the subject, forgery detection, and improving the reliability of image steganalysis (detection of hidden data).

1. MOTIVATION

Thanks to the increasing power of processors currently used in consumer digital cameras, manufacturers began implementing various computational imaging technology inside the camera to allow the photographer to take better images using cheaper hardware and optics. Most cameras today compensate for the imperfections of the optical system by correcting for the geometrical lens distortion (LD) and chromatic aberration inside the camera before saving the resulting image to the memory card. Such processing leaves detectable traces that can be exploited for a variety of forensic purposes.

In this paper, we introduce a technique that can estimate the parameters of the lens distortion correction from a single image. The method makes use of the so-called linear pattern (LP)^{1,2} commonly present in digital images. It manifests through non-zero sums of rows and columns in the image noise residual. The LP is caused by several processes that are applied at different times in the image acquisition pipeline. The first component of the LP, L_0 , is due to sensor signal readout; this part is present in the raw sensor output. Another component, L_1 , is added during color interpolation (demosaicking) and JPEG compression. The part of the LP that is introduced *before* the LD correction is applied becomes deformed in the image corrected for the LD and thus can be used as a template to estimate the parameters of the LD. The method proposed here estimates the parameters of the inverse LD correction transformation using the energy of the LP, particularly of the component L_0 .

We work under the assumption that a forensic analyst has one image under investigation (the test image) and, optionally, knows the model of the camera that took the image. In particular, we assume no knowledge of the image EXIF header as this meta data can be replaced or lost when editing the image. Additionally, we assume that the image was not off-center cropped, rotated, or corrected for perspective distortion.* The analyst desires to estimate the amount of LD correction that was applied to the test image. There are numerous potential applications of such an estimation method:

1. It can be used as a front-end for camera identification using sensor fingerprints.² Sensor fingerprints must be resynchronized with the noise residual of the test image prior to applying a correlation-based detector.⁴ The proposed technique can be used to resynchronize the residual with the fingerprint, thus saving on processing time when a large number of fingerprints need to be checked.
2. By detecting the presence or absence of camera LD correction, one could potentially narrow down the model of the camera that took the test image, e.g., eliminate cameras whose optical system cannot not introduce distortion of a certain strength.

*The method could be potentially extended to these cases at the expense of increased computational complexity and lowered accuracy.

3. When the camera model is known, the estimated parameters could be used to determine the focal length (zoom) at which the image was taken. This could be further used to estimate the proximity of the photographer to the object.
4. The image resampling due to LD correction introduces a specific global structure of pixel correlations into the image[†] that might be locally disturbed when an object in the image (region of interest) is replaced, deleted, or manipulated. This may be useful for forgery detection.⁷
5. The symmetry of the deformed LP can also be used to determine the optical center of the image and thus detect off-center image cropping. This would require an appropriate extension of the LD model used in this paper.
6. It has been established that the reliability of steganalysis (detection of the presence of data hidden in an image) can be highly sensitive to local correlations within pixels or other resampling artifacts.⁸ Since LD correction involves resampling, knowing that an image, which is suspected to contain secretly embedded data, has been corrected for LD is important to the steganalyst. For example, the steganalyst can train a steganography detector on an appropriate cover source (images corrected for LD) to avoid the negative impact of the cover source mismatch^{3,5,9,10} and improve the reliability of steganalysis.

In the next section, we describe the lens distortion model that will be used in this paper and we also contrast this contribution with previous art. In Section 3, we define the concept of the linear pattern, describe the method for estimating the parameters of the lens distortion, and investigate further possible adjustments and their effect on parameter estimation accuracy. We also describe the method using which we determine the ground truth to allow for a quantitative assessment of the proposed method. All experimental results appear in Section 4. We experiment with a total of six cameras and report the percentage of successful determination of the lens distortion parameters including the cases when no distortion correction was applied. The paper is summarized in Section 5.

1.1 Notation and preliminaries

Everywhere in this paper, boldface font will denote vectors (or matrices) of length specified in the text. For \mathbf{X} and \mathbf{Y} two $m \times n$ matrices (or $m \times n$ dimensional vectors), the Euclidean dot product is denoted as $\mathbf{X} \cdot \mathbf{Y}$ with $\|\mathbf{X}\| = \sqrt{\mathbf{X} \cdot \mathbf{X}}$ being the L_2 (Euclidean) norm of \mathbf{X} . Denoting the sample mean with a bar, the normalized correlation and the Peak Correlation to Energy (PCE) ratio are defined as

$$\rho(\mathbf{X}, \mathbf{Y}) = \frac{(\mathbf{X} - \bar{\mathbf{X}}) \cdot (\mathbf{Y} - \bar{\mathbf{Y}})}{\|\mathbf{X} - \bar{\mathbf{X}}\| \|\mathbf{Y} - \bar{\mathbf{Y}}\|}, \quad \text{PCE}(\mathbf{X}, \mathbf{Y}) = \frac{\rho^2}{\frac{1}{mn - |\mathcal{N}|} \sum_{\mathbf{s} \in \mathcal{I} - \mathcal{N}} \rho^2(\mathbf{s})}, \quad (1)$$

where $\mathcal{I} = \{1, \dots, m\} \times \{1, \dots, n\}$ and $\mathcal{N} = \{\mathbf{s} = (s_1, s_2); (s_1 \leq 5 \vee s_1 \geq m - 5) \wedge (s_2 \leq 5 \vee s_2 \geq n - 5)\}$ is a small neighborhood of $\mathbf{s} = (0, 0)$. For brevity, we use $\rho(\mathbf{s}) = \rho(x_{i+s_1, j+s_2}, y_{i, j})$ with the remark that the indices wrap up cyclically whenever they get out of their original ranges.

2. LENS DISTORTION MODEL

A model of geometrical LD expresses how the lens distortion changes the original distance of each pixel to the image center, r' :

$$r' = T_{\mathbf{a}}(r) = r(1 + a_2 r^2 + a_4 r^4), \quad \mathbf{a} = (a_2, a_4). \quad (2)$$

This model is commonly used for modeling barrel and pincushion optical distortion.¹⁵ The parameters a_2 and a_4 control the amount and type of the distortion. Our goal is to estimate one or both of these parameters from a single image. Notice that this model reflects our assumption that the image under investigation has not been cropped off-center as it assumes the optical center to be located exactly in the image geometrical center.

[†]This structure can be captured using e.g., the so-called p-map.¹³

We also point out that our goal is not to implement the best LD correction. The goal is to estimate LD correction that *the manufacturer implemented in the camera*. A previously published related study⁴ justifies this particular form of the model (2) for applications in digital image forensics.

2.1 Relationship to prior art

We would like to make a clear distinction between the method proposed here and the prior art,⁴ which focused on extending the camera identification algorithm based on sensor fingerprint to images corrected for LD. The method employed in Ref. [4] used the sensor fingerprint to estimate the distortion parameters. In this paper, we cannot use the sensor fingerprint because all the analyst has is a single image. Instead, we make use of the part of the LP that was inserted into the test image prior to the LD correction.

We also contrast the proposed work with the content of Section 6.2 from Ref. [4], where the parameter a_2 was estimated using the energy of the LP. There, a high quality fingerprint was needed for the method to work reliably. Moreover, this cited work did not consider the two-parameter model (2), which is essential for the proposed method to work reliably for single images.

3. ESTIMATING LD PARAMETERS

Our goal is to estimate the LD parameters a_2 and a_4 in (2) from a single image, when it is an unmodified JPEG output from a digital camera.[‡] We do so by repeatedly applying the inverse of the LD correction (2) with varying parameters to the image noise residual and detecting the largest local peak in its LP energy. The assumption we adopt is that the energy of the original LP is reduced by LD correction. Generally, the stronger the correction is, the lower the LP energy becomes.

The proposed parameter estimation method starts with extracting the image noise component (also called the noise residual) from each color channel. Given an $m \times n$ 8-bit color channel, e.g., the red channel $\mathbf{R} \in \{0, \dots, 255\}^{m \times n}$, its noise residual is obtained by subtracting from it its denoised version, $\mathbf{W}(\mathbf{R}) = \mathbf{R} - F(\mathbf{R})$, where F is a denoising filter. We use the filter described in Ref. [12], which employs Wiener filtering of wavelet coefficients in a Daubechies 8-tap wavelet transform (also see Appendix A of Ref. [11]). The noise residuals of all three \mathbf{R} , \mathbf{G} , \mathbf{B} channels are merged to one $m \times n$ matrix \mathbf{W} using the linear combination used for conversion from RGB to grayscale:[§]

$$\mathbf{W} = 0.2126 \times \mathbf{W}(\mathbf{R}) + 0.7152 \times \mathbf{W}(\mathbf{G}) + 0.0722 \times \mathbf{W}(\mathbf{B}). \quad (3)$$

The LD parameter estimation method is based on the assumption that the energy of the linear pattern of \mathbf{W} is maximized when the inverse of the LD correction transform is applied to it, $T_{\mathbf{a}}^{-1}(\mathbf{W})$. Before describing the individual steps of the algorithm in detail, we provide a formal definition of the linear pattern and its energy.

3.1 Linear pattern

For a zero-mean matrix \mathbf{X} , $\mathbf{X} = [x_{i,j}]_{i,j=1}^{m,n}$, $\sum_{i,j} x_{i,j} = 0$, we define the *linear pattern* of \mathbf{X} as an ordered set $\mathcal{L} = \mathcal{L}(\mathbf{X}) = \{\mathbf{c}, \mathbf{r}\}$ of column averages $\mathbf{c} = (c_1, c_2, \dots, c_n)$ and row averages $\mathbf{r} = (r_1, r_2, \dots, r_m)$ of \mathbf{X} :

$$c_j = \frac{1}{m} \sum_{i=1}^m x_{i,j}, \quad j = 1, \dots, n, \quad r_i = \frac{1}{n} \sum_{j=1}^n x_{i,j}, \quad i = 1, \dots, m. \quad (4)$$

For the purpose of displaying \mathcal{L} as an image, we introduce the matrix $\mathbf{L} = \mathbf{L}(\mathbf{X}) = [l_{i,j}]_{i,j=1}^{m,n}$, where $l_{i,j} = r_i + c_j$. Figure 1 shows an example of \mathbf{L} after scaling its elements to the range $[0, 255]$ to obtain a grayscale image. Note that the linear pattern of \mathbf{L} is again the same set $\{\mathbf{c}, \mathbf{r}\}$, $\mathcal{L}(\mathbf{L}) = \mathcal{L}(\mathbf{X})$, which justifies why we do not always differentiate between the two forms, \mathcal{L} and \mathbf{L} , when speaking about the same linear pattern in this paper.

[‡]The method would still work if the image was resized and/or centrally cropped.

[§]Although not introduced in this paper, the proposed method can be applied to individual color channels to make it work when the camera corrects for chromatic aberration.

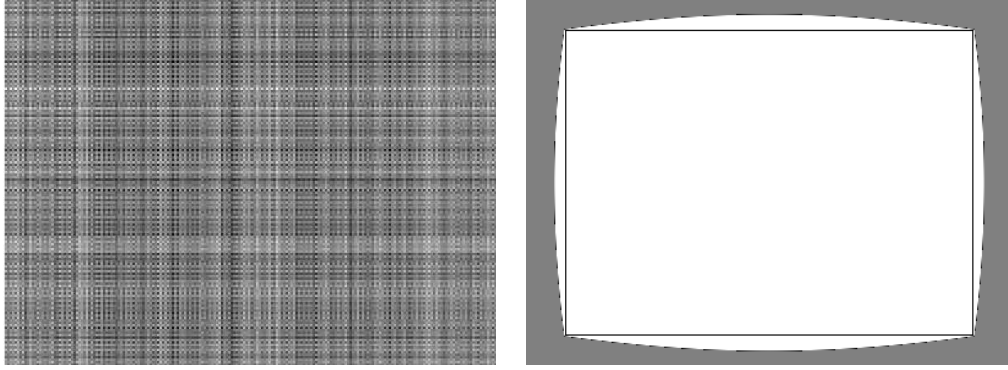


Figure 1. Left: Close up of a linear pattern \mathbf{L} obtained from one test image from Nikon S9100. Right: Padded border area (gray) appears after the barrel distortion. The LP is computed only from the inner rectangle.

The *energy of the linear pattern* $\mathcal{L} = \{\mathbf{c}, \mathbf{r}\}$ is defined as the sum of squares of L_2 norms of \mathbf{c} and \mathbf{r} :

$$E(\mathcal{L}) = \|\mathbf{c}\|^2 + \|\mathbf{r}\|^2 = \sum_{j=1}^n c_j^2 + \sum_{i=1}^m r_i^2. \quad (5)$$

3.2 Parameters estimation as an optimization problem

Typically, cameras correct for barrel distortion at the wide-angle end of the zoom (short focal length). The camera applies the inverse of the barrel transform – the pin-cushion transform. In order to undo this correction, we have to apply the original barrel distortion or, more precisely, the inverse transform to the pin-cushion correction. Notice that the image dimensions and shape change as the content may shrink or expand during the transform (see Figure 1 right). This affects the computed LP and its energy (5) because the mean values calculated from the columns and rows of \mathbf{W} get smaller with increasing barrel distortion when padding to the original dimensions with zeros. Working with all valid samples (the white area in Figure 1 right), on the other hand, creates problems due to the large fluctuations of mean values of short columns and rows, which often leads to false peaks in the LP energy. In order to avoid these unwanted border effects, we restrict all LP computations to the largest rectangle of dimensions $m' \times n'$ inscribed into the transformed image (Figure 1 right). The parameters $\mathbf{a} = (a_2, a_4)$ of the LD correction transform $T_{\mathbf{a}}$ (2) are estimated by applying an inverse mapping, $T_{\mathbf{a}}^{-1}$, to \mathbf{W} and searching for the maximum of the energy of its linear pattern (5) as a function of \mathbf{a} .

The inverse LD correction transform, $T_{\mathbf{a}}^{-1}$, was implemented in the following manner. The value of pixel x at distance r from the image center is calculated by interpolating four pixels in the image with the nearest integer coordinates of pixel x' at distance r' given by Eq. (2) and in the same direction from the center. Bilinear interpolation method was chosen for its speed as it is utilized repeatedly during the search. To avoid transforming the image a number of times and recalculating its noise residual, the inverse transforms $T_{\mathbf{a}}^{-1}$ were applied directly to the noise residual \mathbf{W} .

It is worth pointing out that the main driving parameter is a_2 . The parameter a_4 is responsible for “fine tuning” the transform and starts playing its role for pixels farther away from the image center, typically more than 1/4 of the image diagonal. This allows for a two-stage search – first running a grid search for a_2 , which is followed by a two-parameter search to determine both a_2 and a_4 . A suitable optimization method for this problem is the Nelder-Mead method¹⁴ because it is robust to noise in our objective function (see Figure 2).

During the first stage, we can either fix the value of a_4 or adopt some functional dependence of a_4 on a_2 in order to have a single-parameter search. (See Section 3.3 for more details.) When the distortion is small, $a_2 \approx 0$, the LP energy $\mathcal{L}(T_{\mathbf{a}}^{-1}(\mathbf{W}))$ mostly depends on L_1 . Hence, we cannot reliably estimate a_2 when it is small, which is why we restrict the search interval to $a_2 \in [a_{\min}, a_{\max}]$, $a_{\min} > 0$. In the first stage, we fit a second-degree polynomial to $E(\mathcal{L}(T_{(a_2, 0)}^{-1}(\mathbf{W})))$ as a function of a_2 and compute the difference $\Delta E(a_2)$ between the LP energy and the polynomial fit. Then, we locate the maximum of this difference, say, found at $a_2^{(1)}$ (see

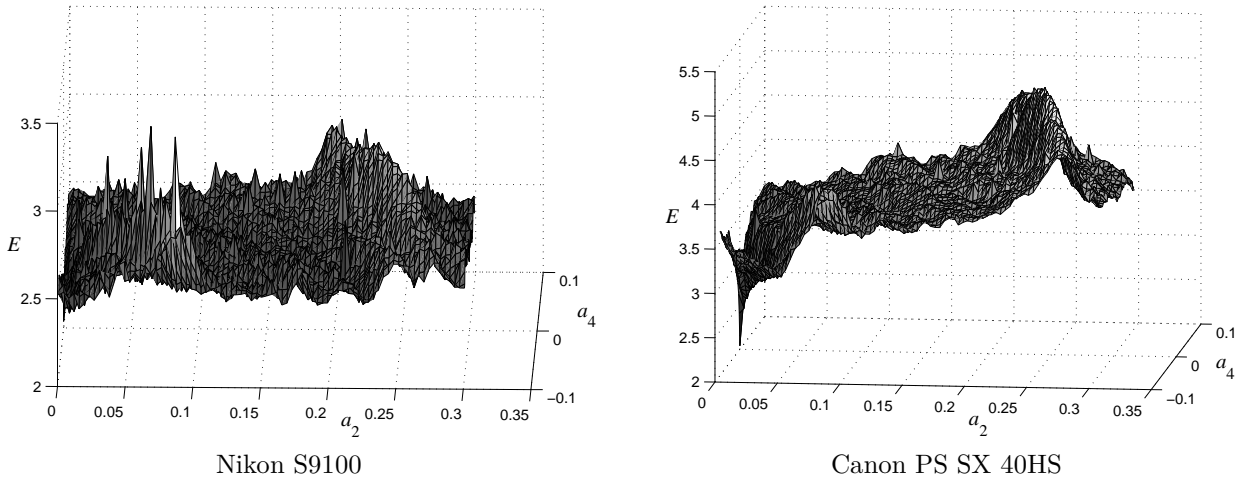


Figure 2. Linear pattern energy vs. radial transform parameters for two cameras.

an example in Figure 4). After obtaining the rough estimate $a_2^{(1)}$, we use it for a three-point initialization of the Nelder-Mead optimization by setting $\mathbf{a}^{(1)} = (x, x^2)$, $\mathbf{a}^{(2)} = (x, -x^2)$, and $\mathbf{a}^{(3)} = (0.8x, 0)$, where $x = a_2^{(1)}$. During the optimization, both parameters (\hat{a}_2, \hat{a}_4) are determined by the maximum of the objective function:

$$f(a_2, a_4) = E \left(\mathcal{L} \left(T_{(a_2, a_4)}^{-1}(\mathbf{W}) \right) \right). \quad (6)$$

The reason for subtracting the second-order polynomial fit is the following. The objective function (6) as a function of a_2 (under $a_4 = 0$) is non-monotone and non-linear. Our goal is to find its local peak corresponding to the true LD correction parameters even if this peak does not attain the global maximum value of f . Fitting higher-order polynomials would lead to over fitting.

Due to the noise in the objective function, the above estimation procedure always finds some maximum, which will occur at $a_2 \neq 0$, even when no LD correction was applied. Thus, the estimation needs to be augmented with a routine that would decide when the distortion correction has not occurred, i.e., when $\mathbf{a} = 0$. To this end, we evaluate the so-called “significance of the peak” at $\hat{\mathbf{a}} = (\hat{a}_2, \hat{a}_4)$. First, the objective function (6) is evaluated at four colinear points in the neighborhood of $\hat{\mathbf{a}}$: (a_{2i}, a_{4i}) , $a_{2i} = \hat{a}_2 + 0.01i$, $a_{4i} = \hat{a}_4 \frac{a_{2i}}{\hat{a}_2}$, $i = -2, -1, 1, 2$. After including the fifth point $(a_{2i}, a_{4i}) = (\hat{a}_2, \hat{a}_4)$, $i = 0$, where the objective function (6) was already evaluated, a second-degree polynomial $A \|a\|^2 + B \|a\| + C$ is fit to these five points using the least square method (Figure 3). The peak significance s is defined as $s = -A$. The peak is called “significant” if $s > t \triangleq 180(\alpha + 1/\alpha)^2$, where $\alpha = m/n$ is the image aspect ratio. The form of this threshold is dictated by the expected energy of a residual that does not contain any LP. Modeling such a residual as a zero-mean matrix whose elements are i.i.d. realizations of a random variable with finite variance σ^2 , the expected value of its LP energy is $(\alpha + 1/\alpha)\sigma^2$ (Appendix 6), which only depends on the image aspect ratio but not the image dimensions. The value of the multiplicative threshold was chosen by analyzing 555 images from three camera models to obtain the rate of $\approx 3\%$ of not correctly recognizing the case when no LD correction was applied.

This way, peaks that are not significant are deemed as random fluctuations rather than due to the presence of a LP, and our routine decides that no LD correction was applied, $\mathbf{a} = 0$.

3.3 List of improvements

The proposed parameter estimation method solely relies on the strength of the linear pattern, which is a very weak signal. In the process, the LP is transformed twice, once with $T_{\mathbf{a}}$ and once with $T_{\hat{\mathbf{a}}}^{-1}$. In practice, these two mapping are only approximately inverse to each other even when $\hat{\mathbf{a}} = \mathbf{a}$ because each transform has to apply some kind of interpolation when evaluating the pixel values on the regular lattice. Inverting this interpolation is

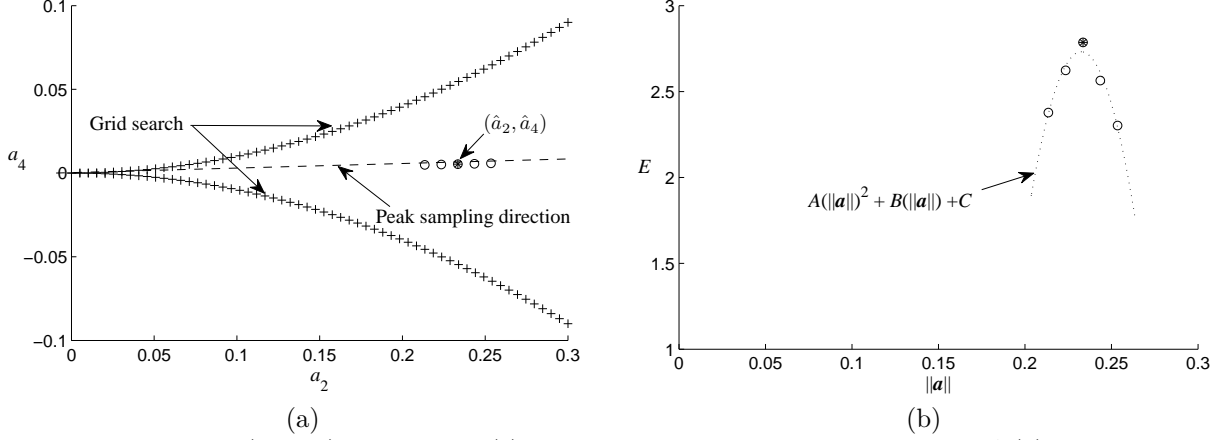


Figure 3. Sampled points (a_{2i}, a_{4i}) , $i = 0, \dots, 4$ (a), the peak significance is defined as $s = -A$ (b). See text for more details.

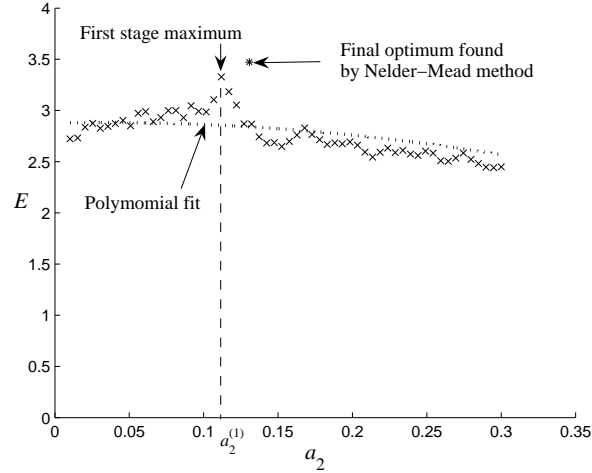


Figure 4. Example of the grid search in the first stage and the optimal (peak) value from the second stage for a test image from Canon PS SX 40HS taken at focal length = 5.649 mm, $\hat{a}_2 = 0.130$, $\hat{a}_4 = -0.0145$.

not perfect either because the interpolation method (its kernel) is not publicized by camera manufacturers and may vary across camera models. The image content is another factor that influences the results of the proposed method.

To alleviate the above issues, we investigated several potential improvements, which we divide into subsections based on their effectiveness.

3.3.1 Realized improvements

In this subsection we expose the details of the method implementation as it was used to carry our experiments in Section 4.

First of all, the image noise residual \mathbf{W} as well as its every transformed form $T_{\mathbf{a}}^{-1}(\mathbf{W})$ is normalized to a unit sample variance before the LP energy is computed. This step makes the optimized objective function “better behaved.”

In order to avoid the strong influence of the component L_1 , which was introduced *after* the LD correction, on the measured LP energy, we subtract it before applying the inverse transform, and we measure only the residual energy of L_0 . A drawback of this approach is that we cannot completely separate L_1 from L_0 . Subtracting the LP from \mathbf{W} not only eliminates L_1 but also suppresses L_0 in the central part in the image, where the distortion

is the smallest. If no LD correction is applied, both components of the LP are subtracted fully, which means that a small LD correction is harder to estimate than a large one.

A partial remedy to this problem is a variation of this approach. Instead of subtracting the entire LP, $\overline{\mathbf{W}} = \mathbf{W} - \mathbf{L}(\mathbf{W})$, we really wish to subtract only a certain “scaled amount” of $\mathbf{L}(\mathbf{W})$ to obtain a residual whose LP energy would correspond to an undistorted residual with no LP. The expected LP energy of a random matrix whose elements are i.i.d. realizations of a random variable with variance σ^2 is $(\alpha + 1/\alpha)\sigma^2$, where $\alpha = m/n$ is the aspect ratio of \mathbf{W} (Appendix 6). Estimating σ^2 as the sample variance of \mathbf{W} , it is shown in Appendix that by transforming \mathbf{W} in the following manner:

$$\widetilde{\mathbf{W}} = [\tilde{w}_{i,j}]_{i,j=1}^{m,n}, \quad \tilde{w}_{i,j} = w_{i,j} - c_j - r_i + c_j \frac{1}{\|\mathbf{c}\|} \sqrt{\frac{n\sigma^2}{m}} + r_i \frac{1}{\|\mathbf{r}\|} \sqrt{\frac{m\sigma^2}{n}}, \quad (7)$$

we give the resulting residual $\widetilde{\mathbf{W}}$ the required energy $(\alpha + 1/\alpha)\sigma^2$. In (7), c_j, r_i are as defined by (4). The advantage of this transform of \mathbf{W} manifests itself primarily when the test image has not been corrected for LD as in this case near $a_2 = 0$ the energy $E(\mathcal{L}(\mathbf{W}))$ exhibits a false peak, $E(\mathcal{L}(\overline{\mathbf{W}})) = 0$, while $E(\mathcal{L}(\widetilde{\mathbf{W}})) = (\alpha + 1/\alpha)\sigma^2$.

In order to reduce the possibility of missing the peak in the LP energy, we implemented the following measure. The first search over a_2 is run twice: for $a_4 = a_2^2$ and $a_4 = -a_2^2$. This dependence $a_4(a_2)$ was chosen empirically from experiments based on the observed relationship between both parameters. The optimal value of a_4 mostly happens to lie between these two values, still including the possible (and likely) case when $a_4 = 0$. Thus, we obtain two intermediate estimates, from which we choose the one having the larger ΔE and denote it as $a_2^{(2)}$. This estimate replaces $a_2^{(1)}$, while the choice of the initial points for the Nelder-Mead optimization remains as described in Section 3.2.

3.3.2 Possible improvements

Here, we list a few ideas that did not consistently improve our test results but may be worth investigating.

1. Replacing $E(\mathcal{L})$ (5) with $E_4(\mathcal{L})$ defined as follows. Start with breaking down the noise residual $\mathbf{W} = [w_{i,j}]_{i,j=1}^{m,n}$ into four sub-matrices corresponding to four pixel types determined by the Bayer Color Filter Array (CFA),⁶ normalizing them to a unit sample variance, and evaluating the total energy as the sum of energies of all four sub-matrices. The total energy derived from \mathbf{W} (assuming n, m even) is

$$E_4(\mathcal{L}) = \sum_{k=1}^4 E(\mathcal{L}(\mathbf{W}_k/\sigma_k)), \quad (8)$$

$$\mathbf{W}_1 = [w_{i,j}], \quad i = 1, 3, \dots, m-1; \quad j = 1, 3, \dots, n-1,$$

$$\mathbf{W}_2 = [w_{i,j}], \quad i = 2, 4, \dots, m; \quad j = 1, 3, \dots, n-1,$$

$$\mathbf{W}_3 = [w_{i,j}], \quad i = 1, 3, \dots, m-1; \quad j = 2, 4, \dots, n,$$

$$\mathbf{W}_4 = [w_{i,j}], \quad i = 2, 4, \dots, m; \quad j = 2, 4, \dots, n,$$

where

$$\sigma_k^2 = \text{Var}(\mathbf{W}_k), \quad k = 1, \dots, 4, \quad (9)$$

are the sample variances computed from all elements of the matrices \mathbf{W}_k . The normalization by the standard deviation equalizes the contribution of each pixel type to the total energy E .

2. Normalizing E (5) by $\frac{1}{n'+m'}$, where n' and m' are the dimensions of the largest rectangle inscribed into $T_{\mathbf{a}}^{-1}(\mathbf{W})$. This would make sense as E may be intuitively expected to decrease with decreasing n' and m' . However this is not generally true and it may depend on the properties of the interpolation method. We observed that E was not decreasing when $T_{\mathbf{a}}^{-1}(\mathbf{W})$ was evaluated using bi-linear interpolation.

3. It remains unclear whether a more advanced interpolation may bring any improvement unless it perfectly inverts the original interpolation applied in the camera. However, camera manufactures do not reveal this level of detail about their in-camera processing. Moreover, perfect inversion will never be possible as some information will always be lost due to additional in-camera image processing, which includes the JPEG compression.

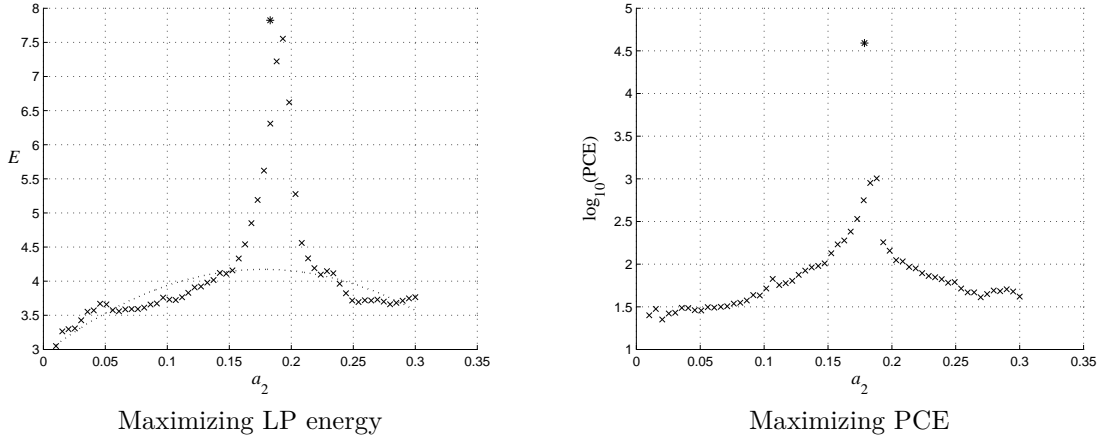


Figure 5. Example of an independent verification of the estimated parameter a_2 for an image from Canon PS SX 230HS when maximizing the LP energy and the PCE.

We note that we considered skipping the first-stage grid search, which would significantly improve computation time. However, in this case, we were not able to initialize the Nelder-Mead optimization to consistently find the desired maximum and avoid being trapped in an incorrect local maximum.

3.4 Ground truth

In order to evaluate the accuracy of the proposed estimation procedure, we need to know the ground truth. To this end, we employed an independent method of LD estimation based on camera fingerprints matching.⁴ First, we estimate the camera fingerprint \mathbf{F} from images that did not undergo LD, e.g., images with focal length at 20 mm or larger for our tested cameras. The PCE (1) between \mathbf{F} and $T_{\mathbf{a}}^{-1}(\mathbf{W})$, $\text{PCE}(\mathbf{F}, T_{\mathbf{a}}^{-1}(\mathbf{W}))$, is maximized at the true value of \mathbf{a} .

For the three cameras that we had access to, i.e., Canon PS SX 230HS, Nikon S9100, and Panasonic DMC ZS7, we estimated the camera fingerprint from ten flat field images at focal length near 20 mm using the maximum likelihood estimator.² The PCE (1) was collected along with the LP energy during each search for the parameter \mathbf{a} . We observed that the points in the parameter space that maximized the LP energy and the PCE were always very close to each other.

To provide an additional insight and evidence, we executed the search algorithm as described in Sections 3.2–3.3.1 with the objective function (the LP energy) replaced with $\log_{10}(\text{PCE}(\mathbf{F}, T_{\mathbf{a}}^{-1}(\mathbf{W})))$. Figure 5 shows an example of the PCE on the log scale evaluated at different values of \mathbf{a} during the search. The output of this algorithm is an alternative estimate of \mathbf{a} , which we expect to be of higher precision thanks to the side information in the form of the camera fingerprint. This is confirmed by contrasting the statistical spread of the estimates obtained using PCE as the objective function in Figure 6 with Figure 7 showing the results for the proposed method.

The problem with uncertain ground truth can be eliminated altogether if our goal is estimating the focal length instead of the parameters of the LD correction. This is possible if the relationship between the focal length and the parameter a_2 is known. Estimating focal length will have one obvious limitation – one will not be able to obtain the focal length estimate when no LD correction is applied (typically when the focal length is above 9–11 mm).

4. EXPERIMENTS

To test the proposed estimation procedure, we prepared a database of images from three cameras made by Canon, Nikon, and Panasonic, as well as images from three other Canon and Nikon models downloaded from the Flickr image sharing portal (see Table 1). For the first three cameras, we also prepared their fingerprints of comparable quality obtained from ten flat field exposures to independently verify the results. All test images

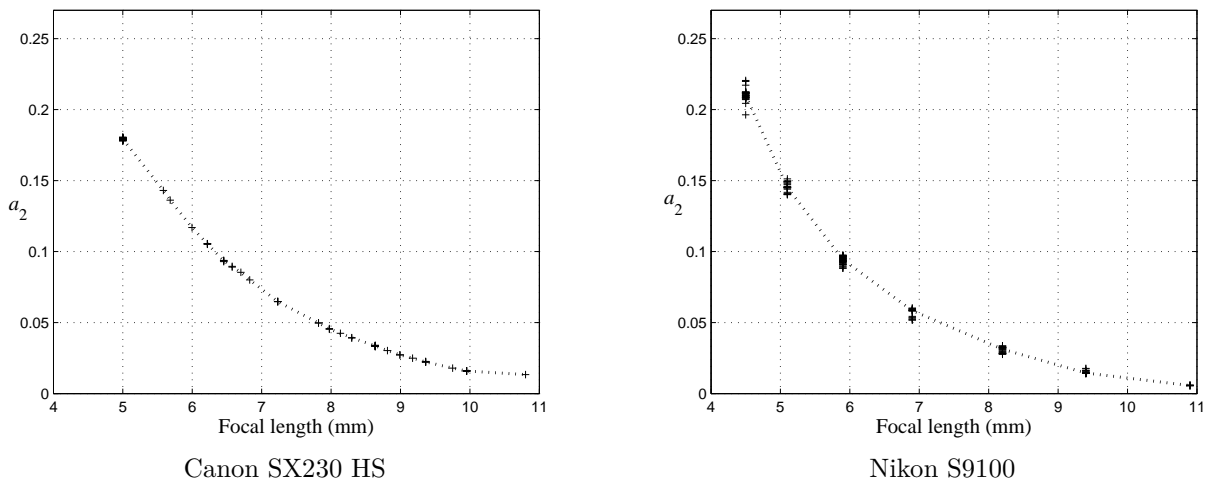


Figure 6. Parameter \hat{a}_2 as a function of focal length estimated by maximizing $\text{PCE}(\mathbf{F}, T_a^{-1}(\mathbf{W}))$ for two cameras. The crosses correspond to individual images.

Camera Model	Zoom (mm)	Sensor	Test Images (#)	f_c	Image Dimension
Canon PS SX 230HS	5.0 – 70	12.1 Mp CMOS	100 + 89	11.2	3000 × 4000
Nikon S9100	4.5 – 81	12.1 Mp CMOS	150 + 39	11.2	3000 × 4000
Panasonic DMC ZS7	4.1 – 49.2	12.1 Mp CCD	91 + 86	8.0	3000 × 4000
Canon PS SX 210IS	5.0 – 70	14.1 Mp CCD	80 + 44 (F)	11.2	3240 × 4320
Canon PS SX 40HS	4.3 – 150.5	12.1 Mp BSI-CMOS	62 + 96 (F)	9.0	3000 × 4000
Nikon P510	4.3 – 180	16.1 Mp CMOS	82 + 18 (F)	9.0	3456 × 4608

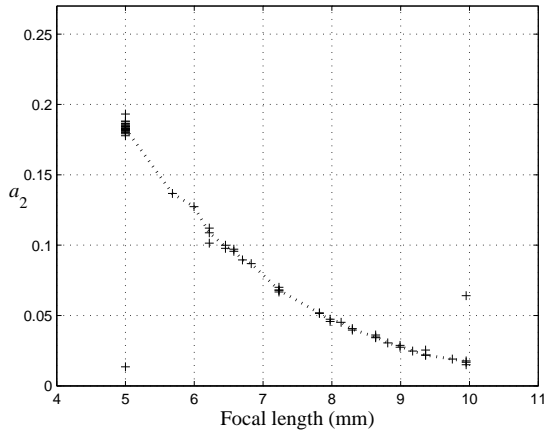
Table 1. Camera models and test images. Images downloaded from Flickr are marked by (F).

were acquired at the full resolution as indicated in Table 1, at varying focal lengths, and were chosen randomly from a larger number of images available. The number of test images indicated in this table is split to images with focal length that is below/above a cut-off focal length f_c . The cut-off is an approximate value above which the LD correction is not present.[¶]

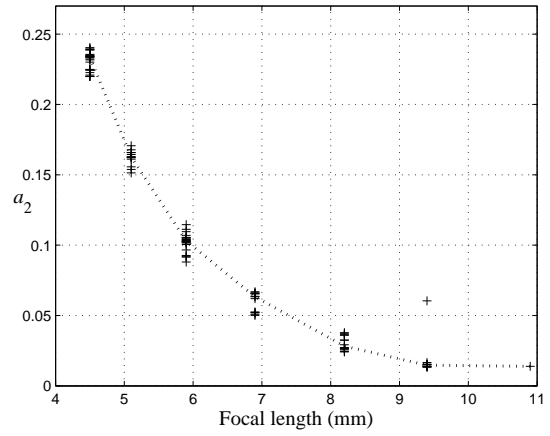
The grid search for parameter a_2 was run for 60 equidistant samples on the interval $[0.0, 0.3]$. We set $a_{\min} = 0.01$ and $a_{\max} = 0.30$. Figure 7 shows the estimated leading parameter a_2 vs. the image focal length retrieved from the image EXIF header. The plotted data points are only those obtained from significant peaks. The focal length information helps us determine the ground truth needed for evaluating the results. Although we do not know the exact relationship between a_2 and the focal length, we can assume that this relationship does not include any other hidden parameters. Therefore, the parameter a_2 should fall on one curve when estimated from images coming from the same camera (or more cameras of the same model). For illustration, the dependence of a_2 on the focal length is rendered by a dotted line connecting the median estimate values for each focal length. These figures only include images with focal length below f_c where the LD correction was apparently detected.

The results from images with focal length above f_c are not shown because most of these images were correctly recognized as having an undistorted LP, i.e., $a_2 = 0$ (and by default $a_4 = 0$). A small number of incorrectly estimated parameters were encountered for images that were corrected for LD. These errors can be seen as outliers in Figure 7 and their numbers are shown in the rightmost column of Table 2.

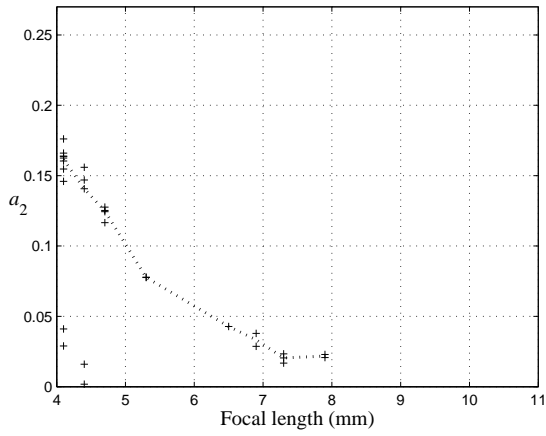
[¶]The knowledge of the cut-off focal length is not used in the process of parameters estimation.



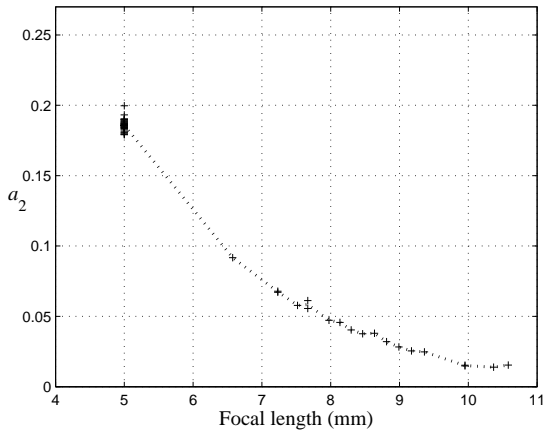
Canon SX230 HS



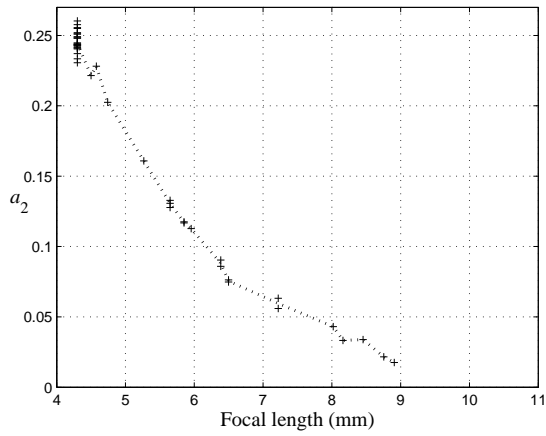
Nikon S9100



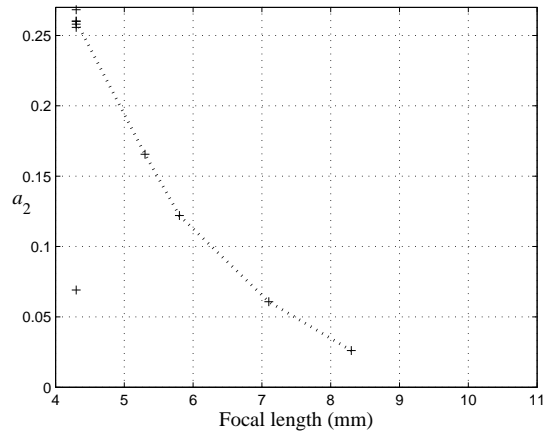
Panasonic DMC ZS7



Canon PS SX 210IS



Canon PS SX 40HS



Nikon P510

Figure 7. Estimated parameter \hat{a}_2 as a function of the focal length for six cameras.

Experiments revealed that the performance of the presented method varies significantly across camera makes and models. This is captured with the “correct estimation” rates in Table 2. This rate was between 60–80% for

Camera Model	Correct Estimation	Incorrect Estimation	Missed No-LD
Canon PS SX 230HS	62/100	2/100	2/89
Nikon S9100	102/150	1/150	7/39
Panasonic DMC ZS7	30/91	4/91	6/86
Canon PS SX 210IS	56/80	0/80	0/44
Canon PS SX 40HS	50/62	0/62	7/96
Nikon P510	11/82	1/82	0/18

Table 2. The number of correctly / incorrectly estimated a_2 for images in the LD correction range and missed cases of no LD correction.

images from four out of six camera models, while the LD correction in images from Nikon P510 and Panasonic DMC ZS7 was the hardest to detect and estimate.

This method may not be universally applicable. We ran tests for images downloaded from Flickr that originated from Sony DSC-H70, a camera equipped with a 16.1 Mp Super HAD CCD imaging sensor. The test results were mostly failures for this camera. We hypothesize that this might have been caused by a different lens distortion model used by the manufacturer.

The experiments point out an advantage of the two-parameter model over the one-parameter model (a_2 only) in its ability to output the parameter \hat{a}_2 with a smaller error (not shown in this work), rather than in its ability to estimate a_4 . The two-parameter model may produce less accurate results when the camera on-board software uses a one-parameter radial transform ($a_4 = 0$).

5. CONCLUSION

Digital images are increasingly often used as inputs for forensic analysis, data mining, and intelligence gathering. Digitally represented data can be, however, subjected to a variety of processing that might affect its semantic meaning and value, which may in turn negatively impact its interpretation and lead to erroneous decisions. Techniques that can reveal the processing history of digital media are thus of crucial importance. In this paper, we describe another tool for the forensic expert toolbox – one that can blindly estimate from a single image the parameters of a lens-distortion correction applied to it. Since the vast majority of today’s compact digital cameras employ the lens distortion correction as part of their in-camera (or off-line) processing, the proposed method can help investigators narrow down the camera that took a given image, reveal portions of the image that have been replaced or maliciously modified, estimate the focal length at which the image was taken and thus establish the distance between the scene and the photographer. Knowing that an image was corrected for lens distortion may also be useful when analyzing it for presence of steganographically embedded data.

The proposed method uses the so-called linear pattern inserted into the image during the acquisition. This pattern, which is originally a set of orthogonal lines, gets distorted when a lens distortion correction is applied. We use it as a template to estimate the lens distortion correction parameters. In particular, when the (approximately) same inverse distortion is applied, the pattern again becomes rectilinear, which manifests itself in the increased energy of pixels’ row and column means. Maximizing this energy w.r.t. parameters of the lens distortion enables us to estimate the distortion correction without the need for any other auxiliary information. The proposed method works from a single image. Tests on images coming from six cameras showed that this method can reasonably accurately estimate the applied lens distortion correction.

The method as proposed only works under some simplifying assumptions, namely that no other geometrical transformation, besides scaling or centralized cropping, was applied to the image. In principle, at the expense of computational complexity and somewhat less accurate performance, the approach could be extended to images that were additionally off-center cropped, rotated, or corrected for perspective distortion. The proposed method may not work reliably for heavily processed images, such as aggressive denoising or low-quality JPEG compression, because the template signal, the linear pattern, is relatively weak.

6. ACKNOWLEDGMENTS

The work on this paper was supported by Air Force Office of Scientific Research under the research grant number FA9950-12-1-0124. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation there on. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of AFOSR or the U.S. Government.

Appendix A

If \mathbf{X} is a random matrix with elements i.i.d. realizations of a zero-mean random variable with variance σ^2 , the expected value of its LP energy $E(\mathcal{L}(\mathbf{X}))$ is $(\alpha + 1/\alpha)\sigma^2$, where $\alpha = m/n$ is the aspect ratio of \mathbf{X} . This follows immediately from the fact that $c_j \sim N(0, \sigma^2/m)$ for all $j = 1, \dots, n$ and $r_i \sim N(0, \sigma^2/n)$ for all $i = 1, \dots, m$:

$$E \left\{ \sum_{j=1}^n c_j^2 + \sum_{i=1}^m r_i^2 \right\} = \sum_{j=1}^n E\{c_j^2\} + \sum_{i=1}^m E\{r_i^2\} = n \times \frac{\sigma^2}{m} + m \times \frac{\sigma^2}{n}. \quad (10)$$

Let $[z_{i,j}]_{i,j=1}^{m,n}$ be an arbitrary zero-mean matrix with linear pattern $\mathcal{L}(\mathbf{Z}) = \{\mathbf{r}, \mathbf{c}\}$. Let $\mathbf{G} = [g_{i,j}]_{i,j=1}^{m,n}$, $g_{i,j} = \frac{c_j}{\|\mathbf{c}\|} \sqrt{\frac{n\sigma^2}{m}} + \frac{r_i}{\|\mathbf{r}\|} \sqrt{\frac{m\sigma^2}{n}}$. Then, the energy of the linear pattern of $\mathbf{Y} = [y_{i,j}]$, $y_{i,j} = z_{i,j} - c_j - r_i + g_{i,j}$, is

$$E(\mathcal{L}(\mathbf{Y})) = \sum_{j=1}^n \left(\frac{1}{m} \sum_{i=1}^m y_{i,j} \right)^2 + \sum_{i=1}^m \left(\frac{1}{n} \sum_{j=1}^n y_{i,j} \right)^2 \quad (11)$$

$$= \sum_{j=1}^n \left(\frac{1}{m} \sum_{i=1}^m (z_{i,j} - c_j - r_i) + \frac{1}{m} \sum_{i=1}^m g_{i,j} \right)^2 + \sum_{i=1}^m \left(\frac{1}{n} \sum_{j=1}^n (z_{i,j} - c_j - r_i) + \frac{1}{n} \sum_{j=1}^n g_{i,j} \right)^2 \quad (12)$$

$$= \sum_{j=1}^n \left(\frac{1}{m} \sum_{i=1}^m g_{i,j} \right)^2 + \sum_{i=1}^m \left(\frac{1}{n} \sum_{j=1}^n g_{i,j} \right)^2 = \left(\frac{n}{m} + \frac{m}{n} \right) \sigma^2, \quad (13)$$

because

$$\sum_{i=1}^m (z_{i,j} - c_j - r_i) = \sum_{i=1}^m \left(z_{i,j} - \frac{1}{m} \sum_{i=1}^m z_{i,j} \right) - \frac{1}{n} \sum_{i,j=1}^{m,n} z_{i,j} = 0 \quad \forall j = 1, \dots, n, \quad (14)$$

and, similarly, $\sum_{j=1}^n (z_{i,j} - c_j - r_i) = 0$ for all $i = 1, \dots, m$. Moreover,

$$\sum_{j=1}^n \left(\frac{1}{m} \sum_{i=1}^m g_{i,j} \right)^2 = \sum_{j=1}^n \left(\frac{1}{m} \sum_{i=1}^m \frac{c_j}{\|\mathbf{c}\|} \sqrt{\frac{n\sigma^2}{m}} + \frac{r_i}{\|\mathbf{r}\|} \sqrt{\frac{m\sigma^2}{n}} \right)^2 = \sum_{j=1}^n \frac{n\sigma^2}{m} \frac{c_j^2}{\|\mathbf{c}\|^2} = n \frac{\sigma^2}{m}, \quad (15)$$

and similarly $\sum_{i=1}^m \left(\frac{1}{n} \sum_{j=1}^n g_{i,j} \right)^2 = m\sigma^2/n$.

REFERENCES

1. M. Chen, J. Fridrich, and M. Goljan. Digital imaging sensor identification (further study). In E.J. Delp and P.W. Wong, editors, *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, volume 6505, pages 0P–0Q, San Jose, California, January 28–February 1, 2007.
2. J. Fridrich. Digital image forensic using sensor noise. *IEEE Signal Processing Magazine*, 26(2):26–37, 2009.
3. J. Fridrich, J. Kodovský, M. Goljan, and V. Holub. Breaking HUGO – the process discovery. In T. Filler, T. Pevný, A. Ker, and S. Craver, editors, *Information Hiding, 13th International Conference*, volume 6958 of Lecture Notes in Computer Science, pages 85–101, Prague, Czech Republic, May 18–20, 2011.

4. M. Goljan and J. Fridrich. Sensor-fingerprint based identification of images corrected for lens distortion. In N. Memon, A. Alattar, and E. Delp III, editors, *Proceedings of SPIE, Electronic Imaging, Media Watermarking, Security, and Forensics 2012*, volume 8303, pages 0H 1–13, San Francisco, CA, January 23–25, 2012.
5. G. Gül and F. Kurugollu. A new methodology in steganalysis : Breaking highly undetectable steganography (HUGO). In T. Filler, T. Pevný, A. Ker, and S. Craver, editors, *Information Hiding, 13th International Conference*, volume 6958 of Lecture Notes in Computer Science, pages 71–84, Prague, Czech Republic, May 18–20, 2011.
6. J. R. Janesick. *Scientific Charge-Coupled Devices*, volume Monograph PM83. Washington, DC: SPIE Press, The International Society for Optical Engineering, January 2001.
7. M.K. Johnson and H. Farid. Exposing digital forgeries through chromatic aberration. In J. Fridrich and J. Dittmann, editors, *Proceedings of the 8th ACM Multimedia & Security Workshop*, pages 48–55. ACM Press, New York, September 26–27, 2006.
8. J. Kodovský and J. Fridrich. Steganalysis in resized images. In *Proc. of IEEE ICASSP*, pages 2857–2861, Vancouver, Canada, May 26–31, 2013.
9. J. Kodovský, V. Sedighi, and J. Fridrich. Study of cover source mismatch in steganalysis and ways to mitigate its impact. In A. Alattar, N. D. Memon, and C. Heitzenrater, editors, *Proceedings SPIE, Electronic Imaging, Media Watermarking, Security, and Forensics 2014*, volume 9028, San Francisco, CA, February 3–5, 2014.
10. I. Lubenko and A. D. Ker. Steganalysis with mismatched covers: Do simple classifiers help? In J. Dittmann, S. Katzenbeisser, and S. Craver, editors, *Proceedings of the 14th ACM Multimedia & Security Workshop*, pages 11–18, Coventry, UK, September 6–7, 2012. ACM Press, New York.
11. J. Lukáš, J. Fridrich, and M. Goljan. Digital camera identification from sensor pattern noise. *IEEE Transactions on Information Forensics and Security*, 1(2):205–214, June 2006.
12. M.K. Mihcak, I. Kozintsev, K. Ramchandran, and P. Moulin. Low-complexity image denoising based on statistical modeling of wavelet coefficients. *IEEE Signal Processing Letters*, 6(12):300–303, December 1999.
13. A.C. Popescu and H. Farid. Exposing digital forgeries by detecting traces of resampling. *IEEE Transactions on Signal Processing*, 53(2):758–767, February 2005.
14. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. Cambridge University Press, New York, NY, USA, 3 edition, 2007.
15. H. Wolfgang. Correcting lens distortions in digital photographs. 2010. http://www.imagemagick.org/Usage/lens/correcting_lens_distortions.pdf.