

Estimation of Image Rotation Angle Using Interpolation-Related Spectral Signatures With Application to Blind Detection of Image Forgery

Weimin Wei, Shuozhong Wang, Xinpeng Zhang, and Zhenjun Tang

Abstract—Motivated by the image rescaling estimation method proposed by Gallagher (2nd Canadian Conf. Computer & Robot Vision, 2005: 65–72), we develop an image rotation angle estimator based on the relations between the rotation angle and the frequencies at which peaks due to interpolation occur in the spectrum of the image's edge map. We then use rescaling/rotation detection and parameter estimation to detect fake objects inserted into images. When a forged image contains areas from different sources, or from another part of the same image, rescaling and/or rotation are often involved. In these geometric operations, interpolation is a necessary step. By dividing the image into blocks, detecting traces of rescaling and rotation in each block, and estimating the parameters, we can effectively reveal the forged areas in an image that have been rescaled and/or rotated. If multiple geometrical operations are involved, different processing sequences, i.e., repeated zooming, repeated rotation, rotation-zooming, or zooming-rotation, may be determined from different behaviors of the peaks due to rescaling and rotation. This may also provide a useful clue to image authentication.

Index Terms—Image forgery, interpolation, passive authentication, rescaling, rotation.

I. INTRODUCTION

WITH the availability of powerful image editing tools, numerous image retouching techniques have become practical, which can be used to create great artistic works. However, malicious modification of image content forms a serious threat to the secure and legal usage of digital images. By skillful manipulation, forgery may be very difficult to

recognize by the naked eye. Therefore, automatic detection of image forgery has attracted much research interest. In recent years, many image forgery detection techniques have been proposed, especially passive approaches which do not require any additional information besides the image itself [1], [2]. Some published methods make use of lighting abnormality [3], blur moment invariants [4], and similarity/dissimilarity of color and structural characteristics [5].

Several techniques for image authentication based on detection of image rescaling/rotation have been reported. Popescu and Farid [6] presented their early method to find rescaling traces hidden in any portion of an image without resorting to a reference image by using expectation maximization (EM) [7]. Mahdian and Saic [8] used periodicity due to interpolation to perform blind image authentication. They introduced Radon transform on the basis of second derivative to detect rotation without estimation of the rotation angle. In another work, Mahdian *et al.* [9] used noise inconsistency to aid the detection of image rescaling for image splicing detection. Kirchner [10] constructed a rescaling detector based on periodic artifacts in the residue of a local linear predictor. He analytically derived the expected position of characteristic rescaling peaks and formed a detector that was considerably faster than [6] with comparable results. Prasad *et al.* [11] localized the tampered areas by labeling a pixel as 1 where the second derivative changes sign, otherwise labeling it as 0. A spectral analysis can reveal the resampling induced periodicity. This method can detect the presence of resampling operation, and is effective for JPEG compression and subsequent rescaling of tampered images.

In this work, we propose a method for rotation angle estimation motivated by some rescaling detection methods, in particular that proposed by Gallagher [12], develop a unified way to determine parameters of rescaling and rotation by exploring interpolation-induced spectral signatures, and apply it to image forensics.

Gallagher's method exploits periodicity in the interpolated image, which will be briefly described in Section II. In a more recent work, Suwendi *et al.* [13] presented a method to estimate rational enlargement factors in both vertical and horizontal directions.

Existing methods for image rotation angle estimation are mostly nonblind. They require prior knowledge about the original image, a reference template, or features extracted from the original. In [14], the rotation angle between the input and the reference images was obtained from the peak of angle

Manuscript received August 31, 2009; revised March 05, 2010; accepted April 08, 2010. Date of publication June 01, 2010; date of current version August 13, 2010. This work was supported by the Natural Science Foundation of China (60773079, 60872116, and 60832010), and by the High-Tech Research and Development Program of China (2007AA01Z477). The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Min Wu.

W. Wei was with the School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China. He is now with Shanghai University of Electric Power, Shanghai 200090, China (e-mail: weiwemin@hotmail.com).

S. Wang is with the School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China (e-mail: shuowang@shu.edu.cn).

X. Zhang is with School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China (e-mail: xzhang@shu.edu.cn).

Z. Tang was with the School of Communication and Information Engineering, Shanghai University, Shanghai 200072, China. He is now with Guangxi Normal University, Guilin 541004, China (e-mail: tangzj230@163.com).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TIFS.2010.2051254

histogram generated through a voting procedure. Onishi *et al.* [15] applied a modified version of Hough transform to the reference and input images, and uniquely computed the angle of rotation. Rotation angle estimation can be carried out from texture features. In [16], a steerable oriented pyramid was used to extract features for the input textures, followed by a supervised classification. Ulas *et al.* [17] studied rotation angle estimation of textures aiming at a real-time implementation. A rotation invariant template matching method based on the combination of a projection method and Zernike moments was proposed to estimate rotation angle in [18].

Nonblind methods have limited values in applications since the original image is unavailable in most practical cases. Blind detection of image rotation and blind estimation of the rotation angle are more useful. As will be shown in the following sections, blind detection of rescaling/rotation and estimation of the parameters provide a means to reveal image areas that are foreign to the rest of the image. It will also be shown that, if multiple geometrical operations are involved, the history of processing may be determined from different behaviors of the spectral peaks due to rescaling and rotation, respectively. This also provides a useful clue for image forensics.

In Section II, an overview of Gallagher's rescaling estimation method is briefly introduced. In Sections III and IV, an image rotation angle estimation technique is proposed, and experimental results presented. Section V describes a method for exploring the history of successive scaling and rotation. Section VI presents an application of the rescaling/rotation estimator to image forgery detection. Conclusions and discussions are given in Section VII.

II. OVERVIEW OF GALLAGHER'S METHOD AND ITS EXTENSION

According to Gallagher [12], a signal $s(x)$ is sampled with a step size $\Delta \in R^+$ to produce a discrete data sequence $s_m = s(m\Delta)$. The signal can be reconstructed from its samples

$$s^{(\varphi)}(x) = \sum_{m=-\infty}^{\infty} s_m \varphi\left(\frac{x}{\Delta} - m\right) \quad (1)$$

where $\varphi(\cdot)$ is the interpolation function. For linear interpolation $\varphi(x) = 1 - |x|$ where $|x| \leq 1$. Gallagher took the second derivative of the interpolated signal, and showed that variance of the second derivative is periodic with a period Δ . In [8], Mahdian *et al.* has generalized the method to the k th order derivative

$$D^{(k)}s^{(\varphi)}(x) = \sum_{m=-\infty}^{\infty} s_m D^{(k)}\varphi\left(\frac{x}{\Delta} - m\right) \quad (2)$$

where $D^{(k)}$ is an operator of the k th order derivative, and $D^{(0)}s^{(\varphi)}(x)$ is $s^{(\varphi)}(x)$ itself when $k = 0$. Mahdian *et al.* has shown that

$$V[D^{(k)}s^{(\varphi)}(x + \mu\Delta)] = V[D^{(k)}s^{(\varphi)}(x)] \quad (3)$$

where μ is an integer, and $V[\cdot]$ means variance. In a word, interpolated signals and their derivatives have inherent periodicity. We will confine to the case of $k = 2$.

For an image, a second-order difference vector of the pixel values are calculated along each line, and all the obtained vectors are then averaged. Due to the interpolation-induced periodicity, discrete Fourier transform (DFT) of this average vector contains peaks directly related to the scaling factor. Gallagher obtained the factor of image enlargement as $R = 1/f_{\text{int}}$ or $R = 1/(1 - f_{\text{int}})$, where $f_{\text{int}} \in (0, 1)$ is the normalized frequency at which the spectral peak occurs. Because of the symmetry of Fourier transform, the frequencies always occur in pairs and are symmetrical about $1/2$. We only consider $f_{\text{int}} \leq 1/2$ in this paper. Gallagher's result may be reorganized as follows in which f_{int} is expressed as a function of the scaling factor R :

$$f_{\text{int}} = \begin{cases} 1 - 1/R, & 1 < R \leq 2 \\ 1/R, & R > 2. \end{cases} \quad (4)$$

In fact, the method can easily be extended to image size reduction

$$f_{\text{int}} = 1/R - 1, \quad R < 1. \quad (5)$$

It is observed from (4) and (5) that each interpolation induced peak corresponds to three R values in the regions $(0, 1)$, $(1, 2]$ and $(2, \infty)$, respectively. For example, a peak at $f_{\text{int}} = 1/3$ can be caused by size-reduction with $R = 3/4$, or enlargement with $R = 3/2$ or 3 . Ambiguities exist if judged solely by f_{int} . Thus, use of *a priori* knowledge would be necessary to resolve the problem as suggested in [12].

For size-reduction, f_{int} will cross the symmetry center, $1/2$, when R is less than $2/3$ so that $1 - f_{\text{int}}$ may be used instead. Nonetheless, we limit our discussion to $R > 2/3$ (corresponding to $f_{\text{int}} < 1/2$) because the peak becomes very weak for small R . Further, we will show in the following that the case of $R < 1$ does not occur in image rotation angle estimation.

III. PROPOSED METHOD FOR IMAGE ROTATION ANGLE ESTIMATION

A. Estimation of Rotation Angle

To develop an image rotation angle estimator, consider the interpolation involved in rotation. Assume that an image is rotated about its center

$$\begin{cases} x' = x \cos \theta - y \sin \theta \\ y' = x \sin \theta + y \cos \theta \end{cases} \quad (6)$$

where (x, y) and (x', y') are the pixel coordinates before and after rotation, respectively, θ the rotation angle with positive values indicating counterclockwise rotation. In most cases, the rotated image is cropped and the same aspect ratio is kept.

As pixels are located on a rectangular grid, interpolation is needed after any geometric transformation. Consider a fixed row in the rotated image, $y' = h$; the following relation can be obtained by eliminating y :

$$x' = \frac{1}{\cos \theta}(x - h \sin \theta). \quad (7)$$

Equation (7) gives a relation between the row coordinate after rotation x' , projected onto the initial horizontal, and the initial row coordinate x . Thus we see that rotation introduces rescaling with a factor of $1/\cos \theta$, calculated from the rows before and

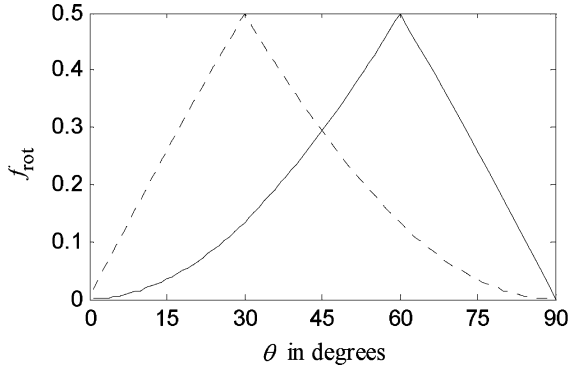


Fig. 1. Peak frequency as functions of rotation angle. The solid and dotted lines correspond to (8) and (10), giving f_{rot1} and f_{rot2} , respectively.

after rotation. We obtain a pattern of 2-D second-order difference, i.e., the edge map, of the rotated image by convolving the image with a Laplacian operator. Calculate 1-D DFT of each line of the edge map and, from (4), obtain the frequency at which the rotation-induced peak occurs (for brevity, referred to as *peak frequency* in the following) in terms of the rotation angle $\theta \in (0^\circ, 90^\circ)$

$$f_{\text{rot1}} = \begin{cases} 1 - \cos \theta, & 0^\circ < \theta \leq 60^\circ \\ \cos \theta, & 60^\circ < \theta < 90^\circ. \end{cases} \quad (8)$$

No interpolation is involved when θ is 0° or 90° . Similarly, by eliminating x from (6)

$$x' = \frac{1}{\sin \theta} (h \cos \theta - y). \quad (9)$$

This is another relation between the row coordinate after rotation x' , projected onto the initial horizontal, and the initial column coordinate y , indicating a rotation-introduced scaling factor $1/\sin \theta$. Thus we can obtain another peak frequency

$$f_{\text{rot2}} = \begin{cases} \sin \theta, & 0^\circ < \theta \leq 30^\circ \\ 1 - \sin \theta, & 30^\circ < \theta < 90^\circ. \end{cases} \quad (10)$$

Clearly, the scaling factors $1/\cos \theta$ and $1/\sin \theta$ are greater than unity. Fig. 1 shows the peak frequency as functions of the rotation angle. The two curves, corresponding to (8) and (10), respectively, are symmetrical about $\theta = 45^\circ$.

Suppose the image size is $M \times N$. We only consider its luminance component \mathbf{Y} . The following are the proposed steps for rotation angle estimation:

- 1) *Edge map generation.* Generate an edge map \mathbf{E} by convolving \mathbf{Y} with a 3×3 Laplacian operator.
- 2) *Counter initialization.* Reset the peak counter $c(n)$, i.e., let $c(n) = 0, n \in [1, N]$.
- 3) *Peak counting.* For the m th row of \mathbf{E} , denoted $\mathbf{v}^{(m)}$, $m = 1, 2, \dots, M$, calculate its discrete Fourier transform $\mathbf{V}^{(m)}$ using fast Fourier transform (FFT). If $[\mathbf{V}^{(m)}(n)]$ is the maximum within $[n - \delta, n + \delta]$, $n = 1, 2, \dots, N$, then update the counter: $c(n) \leftarrow c(n) + 1$. Here δ is chosen empirically. Experiments show that $\delta = 5$ is appropriate.
- 4) *Peak detection.* Normalize $c(n)$ to get a frequency-peak count diagram. Only half of the diagram is considered be-

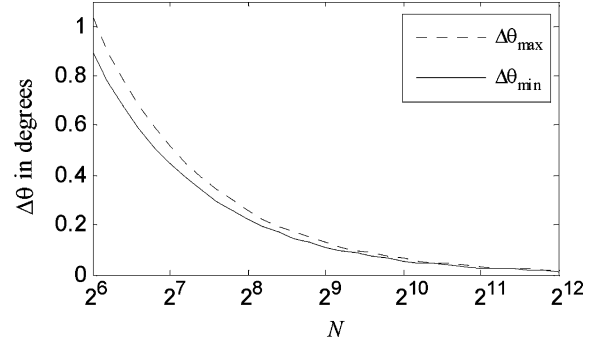


Fig. 2. Angular resolution of the detector as a function of N .

cause of the spectral symmetry. If a $c(n)$ value is greater than the median in a window W by a threshold T , a peak candidate is recorded as $\{c_p(l), f_p(l)\}$. If no peak is detected, the image is considered “not rotated.” In our experiments, $W = 5$ and $T = 2$.

- 5) *Rotation angle calculation.* Take the largest two peaks found in Step 4. Calculate the angle from (8) and (10).

B. Resolution of the Estimator

According to (8) and (10), the angular resolution of the detector sets the resolution of angle estimation. Considering $\theta = \sin^{-1} f$ as in (10), the angular resolution can be obtained

$$\Delta \theta = \frac{d\theta}{df} \Delta f = \frac{1}{N \sqrt{1 - \sin^2 \theta}} \quad (11)$$

which ranges from $\Delta \theta_{\min} = 1/N$ when θ approaches 0° to $\Delta \theta_{\max} = 2/(\sqrt{3}N)$ at $\theta = 30^\circ$. The angular resolution as a function of N is shown in Fig. 2, where the solid and dotted lines are the lower and upper bounds. Clearly, the larger the image size, the better the angular resolution, conforming to intuition. If $N = 64$, for example, the upper bound of the estimation error is in the range of $(0.89^\circ, 1.05^\circ)$, depending on the rotation angle, while for $N = 512$, the error is below 0.15° . Experimental results agree with the analysis as will be shown in Section IV-A.

The resolution is independent of the number of rows M , provided it is not too small, because it does not affect the position of the peak but only the strength with respect to the off-peak spectral magnitudes. When M is very small, say, less than 32, the peak becomes faint and may be indiscernible. For portraits, it is better to use columns since a larger N provides higher angular resolution according to (11). We will only consider rows in this paper.

C. Distinguishing Rescaling and Rotation

Both rescaling and rotation use interpolation, leading to detectable peaks in the DFT of the edge map along the row or column direction. These peaks can be used to estimate the interpolation parameters. However, the two operations behave differently in some aspects so that they can be distinguished.

In Step 3 of Section III-A, 1-D DFT of each row is first calculated, and the average is taken over all rows to get the horizontal spectrum. An alternative method would be to average the rows first, and then take 1-D DFT. Assume that entries of the edge

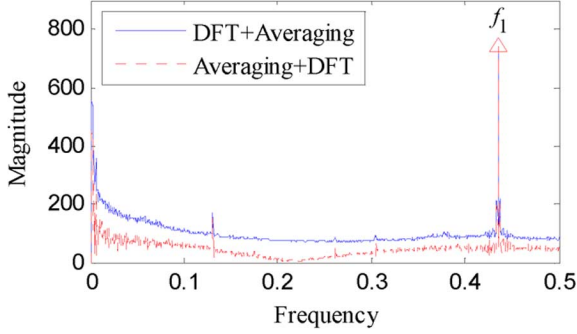


Fig. 3. Averaged horizontal spectra of the rescaled Pepper's edge map. Both peaks at f_1 have similar heights.

map are $E(m, n)$, $m \in [1, M]$, $n \in [1, N]$. Using the first procedure, i.e., DFT + Averaging (DA), we have

$$E_{DA}(\omega) = \frac{1}{M} \sum_{m=1}^M |F[E(m, n)]| \quad (12)$$

where the operator F means discrete Fourier transform. The second procedure [Averaging + DFT (AD)] can be expressed as

$$\begin{aligned} E_{AD}(\omega) &= \left| F \left[\frac{1}{M} \sum_{m=1}^M E(m, n) \right] \right| \\ &= \frac{1}{M} \left| \sum_{m=1}^M F[E(m, n)] \right|. \end{aligned} \quad (13)$$

Curves of $E_{DA}(\omega)$ and $E_{AD}(\omega)$ have similar shapes, with the former larger than the latter in general cases. But the difference in peaks heavily depends on whether the operation is rescaling or rotation.

Consider rescaling first. For example, enlarging the 512×512 Pepper by $R = 2.3$, the average spectrum of its edge map is shown in Fig. 3. In the following, large DC components in all similar plots are removed for display clarity. In this figure, a sharp peak appears at $f_1 = 1/R = 0.435$ in both curves, with a similar magnitude. We then plot the magnitudes and phase angles of the spectra at this peak frequency of all $512 \times 2.3 = 1177$ rows as shown in Fig. 4. It is observed that signs of the phase angles are the same for all rows, leading to effective enhancement of the peak-to-background ratio of the spectral curves when averaging the magnitudes, regardless of which averaging method, (12) or (13), is used.

Now for image rotation, the rescaling factors are $1/\cos \theta$ and $1/\sin \theta$. The average spectra obtained from (12) and (13) greatly deviate from each other due to differences in the spectral phases. For example, rotating Pepper by 23° , the average spectra of the magnitudes are shown in Fig. 5, in which the DA method produces a curve with two clear peaks at $f_1 = 0.082$ and $f_2 = 0.393$, whereas the AD method produces a rather smooth curve without discernible peaks. This can be explained by observing the spectral phases of all rows at these two frequencies. In Fig. 6, the phase angles fluctuate around zero; therefore, the magnitudes cancel out when averaged using (12).

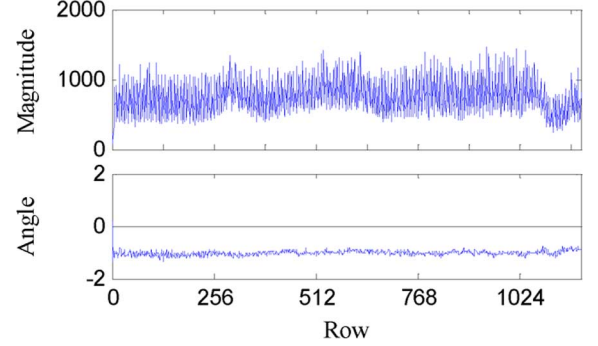


Fig. 4. Spectral magnitudes and phase angles (in radians) of all rows of the rescaled image at f_1 . All phase angles have the same sign, leading to adding-up of the magnitude when averaged using the AD method.

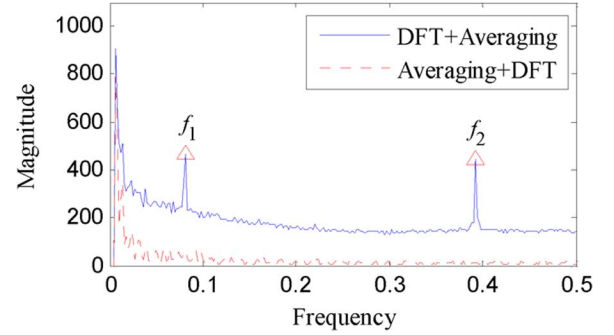


Fig. 5. Average row spectra derived from rotated Pepper.

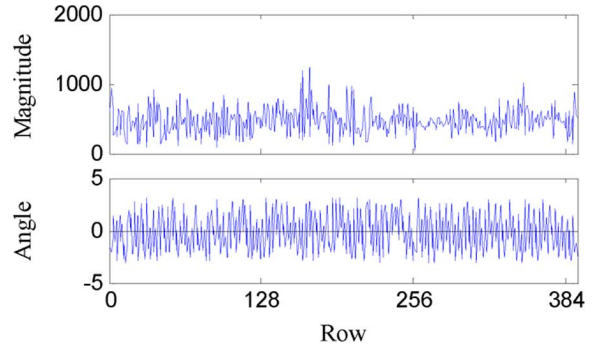


Fig. 6. Spectral magnitudes and phase angles (in radians) of all rows of the rotated image at $f_1 = 0.082$. The phase angles fluctuate around zero, leading to cancellation of magnitudes when averaged using the AD method.

In summary, when both (12) and (13) produce striking peaks, the image is rescaled. When only (12) produces peaks but (13) does not, the image is rotated.

IV. EXPERIMENTS ON ROTATION ANGLE ESTIMATION

A. Rotation Angle Estimation and Robustness Against Attacks

In the experiment, we took 200 uncompressed color and gray-level images of different types from the USC-SIPI database [20], sized 256×256 , 512×512 , and 1024×1024 . These images were rotated using the watermark attacking tool StirMark 4.0 [19] by angles from 1° through 5° with a 1° interval, and from 5° through 45° with a 5° interval, and cropped to remove the useless borders while keeping the aspect ratios. Fig. 7 shows such an example. The original image was 512×512 , rotated

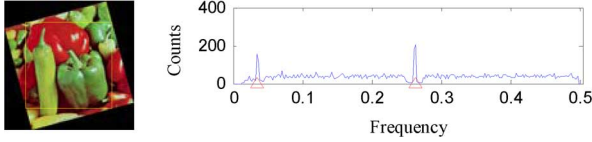


Fig. 7. Image rotated by 15° and cropped, and corresponding spectrum derived from the edge map.

by 15° resulting in a smaller image of 419×419 . The averaged horizontal spectrum of the edge map of the rotated-cropped image was obtained and presented in the figure. Peaks appeared at $f_{\text{rot}1} = 0.035$ and $f_{\text{rot}2} = 0.262$, giving an estimated rotation angle 15.2° , very close to the actual value.

Nearest neighbor, bilinear, and bicubic interpolations were used in the rotation. We recorded the peak frequencies and calculated the estimated angles. If the error is less than 0.5° , we call it a *correct estimate*. Fig. 8(a) presents the percentages of correct estimates C with respect to the rotation angle θ , based on an experiment on the 200 test images.

To evaluate robustness of the angle estimator against post-processing of the image after rotation, three kinds of manipulations were tested: contamination of zero mean Gaussian noise with standard deviation 0.03 (pixel values were normalized to $[0, 1]$), median filtering using a 3×3 mask, and arbitrary cropping with 50% of the area cut off. Fig. 8(b)–(d) shows the results. The rate of correct estimation dropped in all these cases. The bilinear and bicubic methods produced similar results with the correct estimation rate slightly higher in the bilinear case because it introduced more striking spectral peaks. Although a bilinear-interpolated function is continuous in its values, it is discontinuous in the derivatives of the first and higher orders. In contrast, a bicubic-interpolated function is continuous in its values and the first- and second-order derivatives, and discontinuous in the derivatives of the third and higher orders, producing weaker spectral peaks as compared to bilinear interpolation. More sophisticated methods of interpolation keep even higher order derivatives continuous, and therefore, are hard to detect.

The nearest neighbor method produced poorer results. The reason is that ruggedness of the interpolated signal introduces spurious peaks in the spectrum, which, thus, greatly complicates the detection. Nonetheless, the nearest neighbor method is not widely used in practice and, if used, it is easy to detect by revealing the unusual spectral peaks.

From Fig. 8, the rate of correct estimation is low at small angles, similar to the results given in [6] and [8], as in this case the interpolation-induced peak is near the DC component, therefore, hard to detect. Another observation is a dip at 30° , corresponding to $f_{\text{rot}2} = 0.5$, which is one end of the frequency range in which the peak search is performed. The single-side search produces inferior results. If judged by human eyes, the result may be improved.

B. Influence of JPEG Compression and Experiments

The 8×8 blocks in the JPEG coding can cause spectral peaks at $f = i/8$, $i = 1, 2$, and 3. When the quality factor is low, the peaks become quite large. Fig. 9 shows the spectrum of Baboon

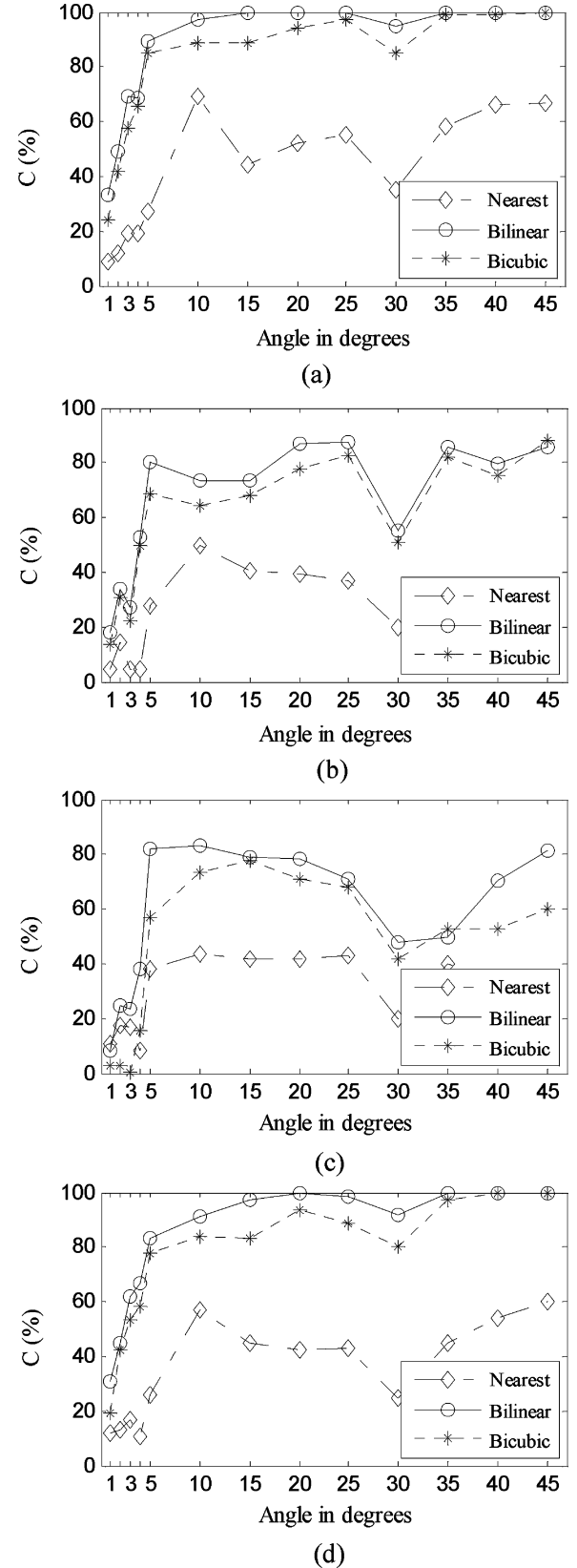


Fig. 8. Rate of correct estimation versus rotation angle. (a) Without attack. (b) Contaminated by additive Gaussian noise. (c) After median filtering. (d) Cropped.

(originally 512×512), rotated by 13° and cropped, and then JPEG coded with a quality factor $Q = 85$. Peaks at $i/8$, $i = 1, 2$,

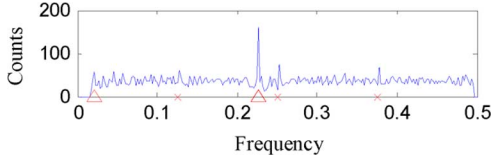


Fig. 9. Spectrum of the edge map of Baboon, rotated by 13° , cropped, and JPEG-compressed.

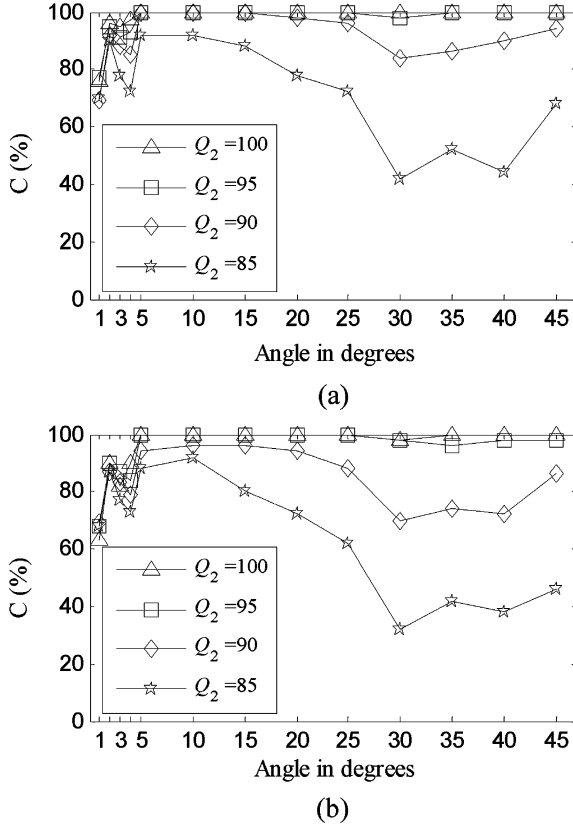


Fig. 10. Rate of correct estimation versus rotation angle based on 100 test images. (a) Tested images uncompressed. (b) Tested images JPEG-compressed with $Q_1 = 85$.

and 3 due to block-DCT are marked with crosses, and the rotation-induced peaks by triangles. Since the JPEG-caused peaks are small, we can still pick the largest peak at $f_{\text{rot}} = 0.227$ to give an accurate estimate of 13.09° .

To study the JPEG effects, we took 100 uncompressed images sized 512×512 and 1024×1024 from [20]. The images were rotated by the same set of angles as in the experiments of the previous subsection, using bicubic interpolation, and JPEG-compressed with different Q values. Rotation angles were estimated. The results are shown in Fig. 10(a), giving the rates of correct estimation (when estimation error is less than 0.5°) for different angles and different Q values. We observe that the performance is worsened with Q decreasing. It is more so when the rotation angles are large.

The second part of the experiment is conducted on images that have been JPEG compressed. In this case, the block DCT-related peaks originally at frequencies of $i/8$ will be slightly shifted after rotation. When the rotated image is stored in the

JPEG format, there will be three groups of peaks: the rotation-induced peaks, peaks near $i/8$ due to the first JPEG and rotation, and those at $i/8$ due to the second JPEG. In this case, performance of the rotation angle estimator is inevitably worsened as compared to the previous case where images before rotation were uncompressed.

We JPEG-compressed the same 100 images with $Q_1 = 85$, and then went through the same rotation-compression-estimation procedure. The results are presented in Fig. 10(b). We can see that JPEG before rotation has moderate effects on the performance when the rotation angle is less than 20° . The influence becomes more serious as the angle increases. Performance at small angles appeared better than Fig. 8 because the smallest images (256×256) were not used here.

V. SUCCESSIVE SCALING AND ROTATION

In image forgery detection, it is useful to learn the processing chain and the parameters used in each step [21], [22]. As rotation and rescaling behaves differently, we now show that, in certain cases, different combinations of these two operations can be distinguished and the parameters estimated. The four possibilities are double zooming (DZ), rotation-zooming (RZ), zooming-rotation (ZR), and double rotation (DR).

A. Double Zooming

Assume the first scaling factor is R_1 . If the second operation were not performed, the interpolation-caused peaks in the DFT of the edge map would have been at frequencies $1/R_1$, $1 - 1/R_1$ or $1/R_1 - 1$. These peaks do not appear because of the second operation. The second zooming with a factor of R_2 produces spectral peaks at $1/R_2$, $1 - 1/R_2$ or $1/R_2 - 1$, and, as a result of successive operations, at several composite frequencies that are multiplications of single zooming frequencies such as

$$f_{\text{DZ}} = 1/(R_1 R_2)$$

and

$$f_{\text{DZ}} = (1/R_1)(1 - 1/R_2). \quad (14)$$

These frequencies are generated by the total rescaling factor $R_1 R_2$. If f_{DZ} is greater than $1/2$, $1 - f_{\text{DZ}}$ is used instead. The same applies to the subsequent subsections, and will not be restated.

To give an explanation to the occurrence of peaks at composite frequencies, taking double zooming-in as an example, let us consider a row in the original image and rescale it by R_1 using bilinear/bicubic interpolation. Stretch the sequence of interpolated samples to fit the pixel grid so that the interval between samples becomes one. This gives a row in the zoomed-in image, denoted $I_1(m)$. Resample the sequence $I_1(m)$ at an interval $\Delta_2 = 1/R_2$, to produce a new sequence. In the same way, stretch the new sequence to fit the grid, giving a row in the double-zoomed image, denoted $I_2(m)$. This image has a zooming factor R_2 with respect to $I_1(m)$, and $R_1 R_2$ with respect to the original. According to Gallagher, variance of the second derivative of $I_2(m)$ has a period, i.e., the normalized peak frequency, $1/R_1 R_2$ in addition to $1/R_2$.

For example, we enlarge Pepper by a factor of 2.3, and enlarge it again by 1.5. The average magnitudes of the spectra are

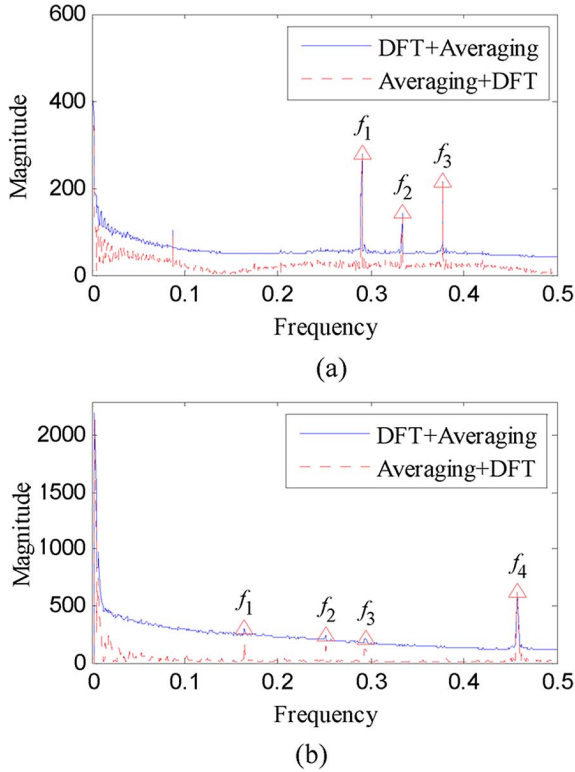


Fig. 11. Average spectra of the successively rescaled Pepper. (a) Zoomed-in by 2.3 and then zoomed-in by 1.5. (b) Zoomed-in by 2.3 and then zoomed-out by 0.8.

shown in Fig. 11(a). Three peaks are detected at $f_1 = 0.290$, $f_2 = 0.333$, and $f_3 = 0.377$. The second zooming of 1.5 corresponds to a peak at f_2 . The other two are composite frequencies. The zooming factors can, therefore, be calculated from these frequencies. The same applies to zooming-out. Fig. 11(b) shows an example in which the Pepper is first zoomed-in by a factor 2.3, and then zoomed-out by 0.8. The peaks at $f_2 = 0.251$ correspond to the zooming-out, and $f_1 = 0.164$, $f_3 = 0.293$, and $f_4 = 0.457$ are composite frequencies.

B. Rotation-Zooming

In image forgery using a copy-move method, the inserted object may be rotated and rescaled to merge into the surroundings. Assume that the rotation angle is θ , corresponding to the peak frequencies at $\sin \theta$, $\cos \theta$, $1 - \sin \theta$, and $1 - \cos \theta$ that would have appeared without further geometric transformation. After rotation, the image is rescaled with a factor R , leading to peaks at $1/R$, $1 - 1/R$, and $1/R - 1$, and a number of composite frequencies such as

$$f_{RZ} = (1/R) \sin \theta$$

and

$$f_{RZ} = (1/R)(1 - \cos \theta). \quad (15)$$

Suppose Pepper is rotated by 20° followed by rescaling with a factor 1.3. The average spectra of the rows in the edge map are

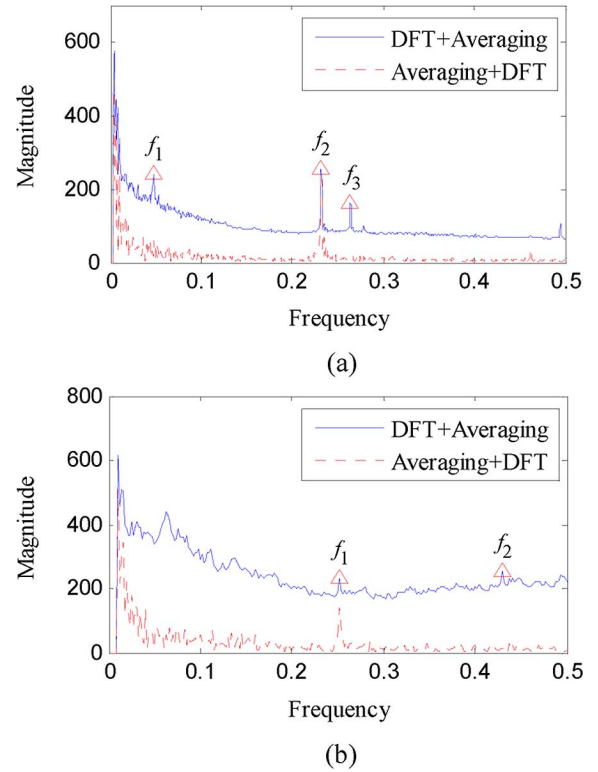


Fig. 12. Average spectra of rotated-zoomed Pepper. (a) Rotated by 20° and zoomed-in by 1.3. (b) Rotated by 20° and zoomed-out by 0.8.

presented in Fig. 12(a). Peaks at $f_1 = 0.048$, $f_2 = 0.231$, and $f_3 = 0.264$ are detected, where f_1 and f_3 correspond to rotation-zooming, and f_2 corresponds to the 1.3 rescaling. All three peaks appear in the DA curve and only the one at f_2 in both curves, indicating a sequence of rotation-zooming. Fig. 12(b) shows another example in which the image is rotated by 20° and then zoomed-out with a factor 0.8. The peaks appear at $f_1 = 0.252$ and $f_2 = 0.430$, where the latter is due to rotation-zooming.

C. Zooming-Rotation

Let the zooming factor be R , and the subsequent rotation angle θ . The rotation introduces peaks at frequencies $\sin \theta$, $\cos \theta$, $1 - \sin \theta$, and $1 - \cos \theta$, and several composite frequencies such as

$$f_{ZR} = (1/R) \sin \theta, \quad f_{ZR} = (1 - 1/R) \sin \theta$$

and

$$f_{ZR} = (1/R) \cos \theta. \quad (16)$$

Assuming Pepper is enlarged by 2.3, and then rotated by 20° , six peaks are found at $f_1 = 0.062$, $f_2 = 0.149$, $f_3 = 0.194$, $f_4 = 0.343$, $f_5 = 0.409$, and $f_6 = 0.470$ as shown in Fig. 13(a). The rotation of 20° corresponds to f_1 and f_4 , and the others are composite frequencies, e.g., $f_6 = 1 - (1 - 1/2.3) \cos 20^\circ = 0.470$. Since the last operation is rotation, no peak is found in the average spectrum calculated from (13). In Fig. 13(b), the average spectra for the image are shown with

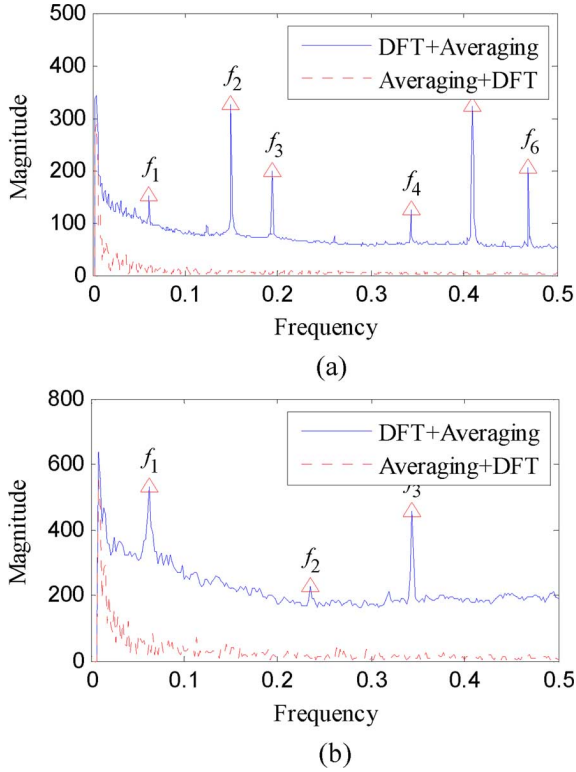


Fig. 13. Average spectra of zoomed-rotated Pepper. (a) Zoomed-in by 2.3 and then rotated by 20° . (b) Zoomed-out by 0.8 and then rotated by 20° .

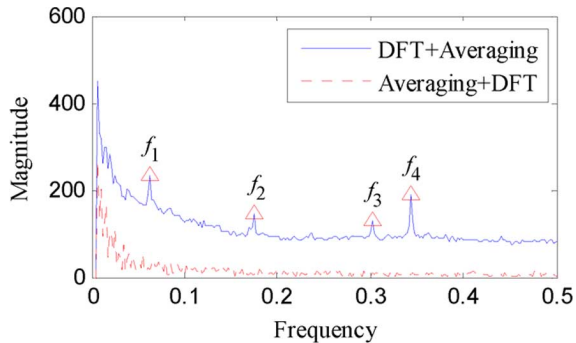


Fig. 14. Average spectra of Pepper rotated by 20° and 20° in succession.

peaks at $f_1 = 0.062$, $f_2 = 0.236$, and $f_3 = 0.343$, where f_2 is due to zooming by 0.8 and rotation by 20° .

D. Double Rotation

Suppose the first rotation angle is θ_1 corresponding to peaks at frequencies $\sin \theta_1$, $1 - \sin \theta_1$, $\cos \theta_1$, and $1 - \cos \theta_1$ that would have appeared if no further operations were performed. The second rotation causes a number of peaks, e.g., $\sin \theta_2$ and $1 - \sin \theta_2$, as well as composite frequencies such as

$$f_{DR} = \sin \theta_1 \cos \theta_2$$

and

$$f_{DR} = \cos \theta_1 \sin \theta_2. \quad (17)$$

Assume the Pepper is rotated twice by 20° and 20° successively. The spectra are shown in Fig. 14, with four peaks at frequencies $f_1 = 0.063$, $f_2 = 0.176$, $f_3 = 0.303$, and $f_4 = 0.344$.

The second rotation corresponds to the peaks at f_1 and f_4 . The others are composite frequencies. As in the previous case, no peak exists in the average spectrum based on (13) because the final operation is rotation. Note that double rotation is different from a single rotation by $(\theta_1 + \theta_2)$.

E. Discussion and Experiments

Based on the above study, we can easily identify the nature of the most recent operation, rescaling or rotation, and estimate its parameters. Confining to double operation, if the second operation is zooming, the nature of the first operation can also be determined by observing the difference between the AD and DA curves. In this case, if the same peaks appear in both curves as in Fig. 11, the first operation should be zooming. If, on the other hand, the DA curve has more peaks than the AD curve as in Fig. 12, the first operation is rotation. However, if the second operation is rotation, there is no simple way to tell whether the previous operation is zooming or rotation.

Taking the same set of 100 images used in Section IV-B, we performed an experiment to check the double operation-caused peak frequencies. Bilinear interpolation was used. Zooming ratios 0.8, 1.3, 1.8, and 2.3, and rotation angles 5° , 10° , 15° , and 20° were tested. With the combinations DZ, RZ, ZR, and DR, a total of 64 computations were done for each image. Rotation was followed by cropping as before. Parameters of the first operations were derived from the detected spectral peaks in automatic detection. If the estimation error of the first rotation angle was less than 0.5° , or that of the first scaling factor less than 0.2, we say the estimate was *correct*. Percentages of correct estimates are listed in Table I. Reasonably high correct rates were observed in cases of DZ, RZ, and ZR except for zooming-out. Double rotation showed lower correct rates, and double zooming-out was hard to detect. In fact, for many cases in which automatic detection was difficult, peaks due to double operation were still visible. Examples can be seen in Figs. 12(b), 13(b) and 14, in which some discernible peaks fall into the three categories marked with asterisks in Table I. Thus the correct rates would be higher if detected by human inspection.

VI. APPLICATIONS IN IMAGE FORGERY DETECTION

The technique described in the previous sections can be used to combat image forgery involving interpolation. Many techniques of image authentication have been introduced such as copy-move detection [23] and exposing forgery based on the color filter array (CFA) interpolation [24], [25]. In this section, we describe an application of the proposed method of rescaling/rotation detection and parameter estimation to image authentication, and present experimental results. It is hoped that, as said in [26], the method will add a new tool to the arsenal of forensic analysts.

A. The Method

Common image forgery methods of copy-move and image splicing compose a new image with fake parts from varied sources, in many cases accompanied by rescaling and rotation of image patches. We detect foreign objects by revealing unusual spectral peaks of local second-order differences in the

TABLE I
RATES OF CORRECT ESTIMATES FOR DOUBLE OPERATIONS BASED ON
AUTOMATIC DETECTION OF SPECTRAL PEAKS. SHORT BARS INDICATE THAT
CORRECT RATES ARE LESS THAN 30%. ASTERISKS SHOW CATEGORIES
CORRESPONDING TO FIGS. 12(b), 13(b), AND 14 IN WHICH PEAKS ARE VISIBLY
RECOGNIZABLE ALTHOUGH AUTOMATIC DETECTION IS HARD

			2nd operation							
			zooming				rotation			
			0.8	1.3	1.8	2.3	5°	10°	15°	20°
1st operation	zooming	0.8	—	—	40	43	—	—	—	—(*)
		1.3	36	79	87	84	38	57	70	42
		1.8	40	78	77	75	62	42	73	76
		2.3	45	73	83	77	73	70	70	68
	rotation	5°	—	77	74	77	61	68	64	66
		10°	—	68	70	61	67	54	46	45
		15°	—	74	76	68	52	49	37	—
		20°	—(*)	72	72	53	48	40	—	—(*)

image. Assume the image to be checked has a size of $M \times N$. The detection steps are as follows:

- 1) *Differentiation*. Take the luminance components \mathbf{Y} of the image, and get the Laplacian edge map \mathbf{E} .
- 2) *Blocking*. Divide the edge map into overlapped blocks, each sized $B \times B$ with overlapping area between each pair of blocks being $L \times B (L < B)$. The block size should not be too small, say, at least 32×32 . We recommend 64×64 or larger, depending on the size of the image being checked. Larger overlapping areas lead to better accuracy but require more computations.
- 3) *Block-wise forgery detection*. For each block, generate an average spectrum derived from discrete Fourier transform of all rows in the block and, if peaks exist, record the corresponding frequencies to determine its status of rescaling and/or rotation. If no peak is found, the block is discarded.
- 4) *Frequency histogramming*. Produce a histogram of the peak frequencies using the above results. Let the number of bins equal B so that every possible frequency can be represented.
- 5) *Foreign object identification*. Scan the nonzero bins in the histogram. Blocks corresponding to populated frequencies in the histogram are labeled as “suspicious.” A foreign object can be identified by connected suspicious blocks. The nature of forgery and the processing parameters including the rescaling factor and the rotation angle are found from the record obtained in Step 3. Sporadic suspicious blocks may be ignored based on the human judgment.

B. Experimental Verification

We generated forged images using Photoshop to modify an original authentic image sized 2256×1504 as shown in Fig. 15(a). The authentic image was obtained by converting a RAW picture acquired with a digital single lens reflex camera to the JPEG format using Photoshop with the highest quality index 12. The image is cropped to an appropriate size without rescaling. We took a rectangular area in the image around a fruit fly, rotated it by 15° , and put it back to the image to cover the original insect, as in Fig. 15(b). Careful postprocessing was conducted to make the forged picture look natural. Detection was carried out with a block size 128×128 and overlapping areas of 96×128 , giving the result shown in Fig. 15(c). The



Fig. 15. Image forgery detection: a fruit fly is inserted after rotation and/or enlargement. (a) Original image. (b) Object rotated. (c) Detection result. (d) Object rotated and enlarged. (e) Detection result.

modified insect was detected. False alarms occurred sporadically. This can be overcome by omitting isolated abnormal blocks or small areas containing only a few such blocks, determined by a threshold based on human judgments.

In Fig. 15(d), the fruit fly was rotated by 25° and enlarged with a factor of 2.1. The original insect was carefully erased. The detection result is shown in Fig. 15(e). Two peaks occurred at $f_1 = 0.20$ and $f_2 = 0.48$ when (12) was used in the averaging. However, only f_2 appeared when using (13) for averaging. A rescaling factor of 2.1 was obtained from f_2 . Frequency f_1 is a result of rotation-zooming with the rotation angle $\theta = \sin^{-1}(0.20 \times 2.1) = 25^\circ$. Since the 2.1 times rescaling produced a large peak, no false alarm occurs in this case.

The experiment was performed on a computer with Pentium D CPU running at 2.8 GHz and a 512-MB memory, and using Matlab. The computations took about 210 s.

In another example, a large image sized 3092×2061 is used as shown in Fig. 16(a), with the same setting of the blocks and overlaps as in the previous experiment. The original image was generated in the same way as in Fig. 15(a). An eagle (592×395 BMP), shown in Fig. 16(b), was rotated by 25° without rescaling, and added to the sky. To make the modification natural, the Photoshop's layers and masks were carefully used to avoid visible artifacts. The composite image shown in Fig. 16(c) was saved in the JPEG format at Photoshop's quality index 12. The forged image was reloaded into the computer memory for detection. The JPEG-induced spurious peaks at frequencies of $i/8$ were identified and ignored. The result is shown with a detected peak at $f = 0.42$. The second peak was too small to be identified. Therefore, two possible angles $\theta = 25^\circ$ or 35° can be estimated from the single peak frequency (see Fig. 1). The computation took 403 s.

In Fig. 17, an inserted insect in the foreground and a fake background on the left side were detected. The insect was rotated by 25° counterclockwise before insertion. The added background was rotated by 18° clockwise, Gaussian blurred with a radius of 3 pixels, feathered around the boarder areas, and reduced to opacity of 50% before being put onto the image. The image sized 1995×1330 was in the JPEG format. It took 157 s to detect.

Fig. 18 shows an example of image splicing detection. A panoramic image sized 951×318 was obtained by combining

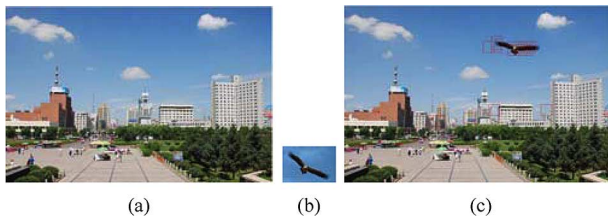


Fig. 16. Image forgery detection: a bird is added in the sky after rotation of 25° . (a) Original image. (b) Eagle. (c) Forged image and detection result.



Fig. 17. Image forgery detection: an insect is put on the flower after rotation by 25° . Background is added after rotation by 18° and Gaussian blurring with a radius of 3 pixels, with transparency of 50%. (a) Original image. (b) Forged image. (c) Detection result.

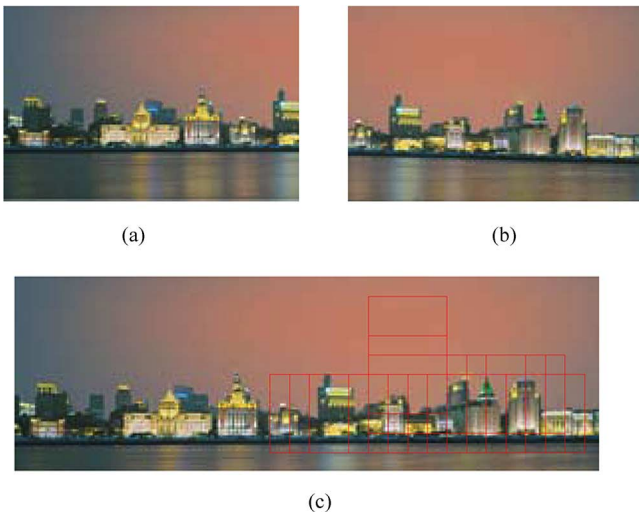


Fig. 18. Image splicing detection: the right half has been rotated by 2° . (a) Image 1. (b) Image 2. (c) Combined image and detection result.

two uncompressed pictures. The computation indicates that the right half has been rotated by 2° . The detection took 13 s. Rotation of flat areas in the sky was difficult to detect since interpolation had very little effects there. This is not important in the detection though.

VII. CONCLUSION

Interpolation operations are performed in image rescaling and rotation, which introduce periodicity in the image. This can be used to estimate the factor of digital image rescaling. We have developed an image rotation angle estimator based on interpolation artifacts with satisfactory accuracy. It has been shown that the estimation scheme works well for small images, for example, as small as 64×64 . Combined with rescaling factor estimation, capability of locating rotated small image patches and estimating the rotation angle makes it possible to detect image forgery involving geometrical transformation.

Experiments have shown effectiveness of the technique. The method can also be used to discover the image's geometric transformation history involving rescaling and rotation, providing a useful clue in passive image authentication.

Passive image authentication using interpolation-signatures as presented in this paper can locate the areas of tampering, provide the related parameters, and reveal the order of the processing sequence in some cases. Since no time-consuming iteration is involved and the major operation performed in the detection is FFT, computation complexity of the algorithm is not high.

As curves plotted in Fig. 1 are symmetrical, the rotation angle estimator always gives two results, one below 45° and the other above, leading to ambiguity. This may be overcome by using some *a priori* knowledge. In practice, however, rotation by more than 45° is rare.

Blind detection of image forgery is a difficult task. For the copy-move type of image tampering, reliable detection of very small image areas is still a challenge. When postprocessing is done such as JPEG coding with a low quality factor, detection becomes more difficult. Moreover, to evade rescaling/rotation detection, more sophisticated interpolation methods can be used, and image manipulations may be done to make the rescaling traces undetectable [27]. All these have become motivations for the development of further improved forensic techniques.

REFERENCES

- [1] H. Farid, "Image forgery detection," *IEEE Signal Process. Mag.*, vol. 26, no. 2, pp. 16–25, Mar. 2009.
- [2] A. Swaminathan, M. Wu, and K. J. R. Liu, "Component forensics," *IEEE Signal Process. Mag.*, vol. 26, no. 2, pp. 38–48, Mar. 2009.
- [3] M. K. Johnson and H. Farid, "Exposing digital forgeries in complex lighting environments," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 2, pp. 450–461, Jun. 2007.
- [4] B. Mahdian and S. Saic, "Detection of copy-move forgery using a method based on blur moment invariants," *Forensic Sci. Int.*, vol. 171, no. 2–3, pp. 180–189, 2007.
- [5] W. Luo, J. Huang, and G. Qiu, "Robust detection of region-duplication forgery in digital image," in *Proc. 18th Int. Conf. Pattern Recognition*, Hong Kong, 2006, pp. 746–749.
- [6] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of resampling," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pt. 2, pp. 758–767, Feb. 2005.
- [7] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Royal Statist. Society, Series B: Methodol.*, vol. 39, no. 1, pp. 1–38, 1977.
- [8] B. Mahdian and S. Saic, "Blind authentication using periodic properties of interpolation," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 3, pp. 529–538, Sep. 2008.
- [9] B. Mahdian and S. Saic, "Detection of resampling supplemented with noise inconsistencies analysis for image forensics," in *Proc. Int. Conf. Computational Sciences and Its Applications*, 2008, pp. 546–556.
- [10] M. Kirchner, "Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue," in *Proc. 10th ACM Workshop Multimedia and Security*, 2008, pp. 11–20.
- [11] S. Prasad and K. Ramakrishnan, "On resampling detection and its application to detect image tampering," in *Proc. IEEE Int. Conf. Multimedia and EXPO*, 2006, pp. 1325–1328.
- [12] A. C. Gallagher, "Detection of linear and cubic interpolation in JPEG compressed images," in *Proc. 2nd Canadian Conf. Computer and Robot Vision*, Washington, DC, 2005, pp. 65–72.
- [13] A. Suwendi and J. P. Allebach, "Nearest-neighbor and bilinear resampling factor estimation to detect blockiness or blurriness of an image," *J. Electron. Imag.*, vol. 17, no. 2, pp. 023005–, 2008.
- [14] Y. Xiong and F. Quek, "Automatic aerial image registration without correspondence," in *Proc. 4th IEEE Int. Conf. Computer Vision Systems*, 2006, pp. 25–33.

- [15] H. Onishi and H. Suzuki, "Detection of rotation and parallel translation using Hough and Fourier transforms," in *Proc. IEEE Int. Conf. Image Processing*, 1996, pp. 827–830.
- [16] H. Greenspan, S. Goodman, and R. Perona, "Rotation invariant texture recognition using a steerable pyramid," in *Proc. 12th IAPR Int. Conf. Pattern Recognition*, 1994, vol. 2, pp. 162–167.
- [17] C. Ulas *et al.*, "Rotation angle estimation algorithms for textures and their real-time implementation on the FU-SmartCam," in *Proc. 5th Int. Symp. Image and Signal Processing and Analysis*, 2007, pp. 469–475.
- [18] M. Choi and W. Kim, "A novel two stage template matching method for rotation and illumination invariance," *Pattern Recognit.*, vol. 35, no. 1, pp. 119–129, 2002.
- [19] Stirmark Benchmark 4.0 [Online]. Available: <http://www.petitcolas.net/fabien/watermarking/stirmark/>
- [20] USC-SIPI Image Database [Online]. Available: <http://sipi.usc.edu/database/>
- [21] A. Swaminathan, M. Wu, and K. J. R. Liu, "Digital image forensics via intrinsic fingerprints," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 1, pp. 101–117, Mar. 2008.
- [22] W.-H. Chuang, A. Swaminathan, and M. Wu, "Tampering identification using empirical frequency response," in *Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP'09)*, Apr. 2009, pp. 1517–1520.
- [23] J. Fridrich, D. Soukal, and J. Lukáš, "Detection of copy-move forgery in digital images," in *Digital Forensic Research Workshop*, Cleveland, OH, Aug. 2003.
- [24] A. C. Popescu and H. Farid, "Exposing digital forgeries in color filter array interpolated images," *IEEE Trans. Signal Process.*, vol. 53, no. 10, pp. 3948–3959, Dec. 2005.
- [25] A. C. Gallagher and T. Chen, "Image authentication by detecting traces of demosaicing," in *IEEE Computer Society Conf. Computer Vision and Pattern Recognition Workshops*, 2008, pp. 1–8.
- [26] H. Farid, "Exposing digital forgeries from JPEG ghosts," *IEEE Trans. Inf. Forensics Security*, vol. 4, no. 1, pp. 154–160, Mar. 2009.
- [27] M. Kirchner and R. Böhme, "Hiding traces of resampling in digital images," *IEEE Trans. Inf. Forensics Security*, vol. 3, no. 4, pp. 582–592, Dec. 2008.



Weimin Wei received the M.Eng. degree from Wuhan University, China, in 2004, and the Ph.D. degree from Shanghai University, China, in 2010.

Currently, he is with School of Computer and Information Engineering, Shanghai University of Electric Power, China. His research interests include image processing, digital forensics, and data mining.



Shuozhong Wang received the B.Sc. degree from Peking University, China, in 1966 and the Ph.D. degree from the University of Birmingham, Birmingham, England, in 1982.

Currently, he is a Professor of Shanghai University, Shanghai, China. His research interests include image processing, multimedia security, and underwater acoustics.



Xinpeng Zhang received the B.Sc. degree in computation mathematics from Jilin University, China, in 1995, and the M.Eng. and Ph.D. degrees in communication and information systems from Shanghai University, Shanghai, China, in 2001 and 2004, respectively.

Currently, he is a Professor of Shanghai University. His research interests include image processing and digital forensics.



Zhenjun Tang received the M.Eng. degree from Guangxi Normal University, Guilin, China, in 2006, and the Ph.D. degree from Shanghai University, in 2010.

Currently, he is with Guangxi Normal University. His research interests include image processing and information security of digital media.