

Fast Forgery Detection with the Intrinsic Resampling Properties

Cheng-Chang Lien, Cheng-Lun Shih, Chih-Hsun Chou

Department of Computer Science and Information Engineering, Chung Hua University, Hsinchu, Taiwan, China E-mail: cclien@chu.edu.tw Received June 21, 2010; revised July 20, 2010; accepted July 25, 2010

Abstract

With the rapid progress of the image processing software, the image forgery can leave no visual clues on the tampered regions and make us unable to authenticate the image. In general, the image forgery technologies often utilizes the scaling, rotation or skewing operations to tamper some regions in the image, in which the resampling and interpolation processes are often demanded. By observing the detectable periodic distribution properties generated from the resampling and interpolation processes, we propose a novel method based on the intrinsic properties of resampling scheme to detect the tampered regions. The proposed method applies the pre-calculated resampling weighting table to detect the periodic properties of prediction error distribution. The experimental results show that the proposed method outperforms the conventional methods in terms of efficiency and accuracy.

Keywords: Image Forgery, Resampling, Forgery Detection, Intrinsic Properties of Resampling

1. Introduction

In recent years, with the rapid progress of image processing software, it becomes a great challenge to verify whether the digital image is tampered or not because the image processing software can create a sophisticated digital forgery and leave no visual clues on the tampered regions. For example, the *Liberty Times* newspaper in January 2008 (newspaper in Taiwan) published a photograph shown in **Figure 1(b)** in which the picture "Miss Wang" had been removed intentionally.

In general, the digital forgery detection methods can be roughly categorized into the active [1-4] and passive methods [5-16]. In the active methods [1-4], the digital watermarking or signatures are hid in the image for the purpose of authentication [1-4]. In addition, the embedded watermarks need to be robust enough to resist the various kinds of image attacks. On the contrary, the passive approaches [5-17] do not need any prior information for the forgery detection and can be further categorized into the methods of detecting copy-pasted regions, defocus blur edges, resampling, sensor noise pattern, different lighting conditions and block artifact inconsistency.

In [5], the author provided a method to identify the digital forgery regions that are copied and pasted from

the same image by applying the method of block matching. However, the matching process can fail if the tampered region is cropped from different images. Zhou et al. [6] proposed a method to identify the digital forgeries by using the edge preserving smoothing filter in which the manual blur edge is discriminated from the defocus blur edge and the erosion operation is applied for detecting the manual blur edge. Another typical method developed by Popescu [7] detected the digital forgeries by tracing the characteristic of the resampled signals. The major concept of this method is to apply the EM (expectation/ maximization) algorithm to acquire the resampling coefficients and then calculate the resampling probability map. Based on the spectral analysis of the probability map, the magnitude peak can be used to identify the forgery patterns. Moreover, Popescu [8] utilized the specific interpolation coefficients of color filter array for each brand of digital camera to identify the digital forgery. Kirchner [9] proposed a more efficient method by directly applying the converged resampling coefficients to detect the tempered regions. As same as tracing the periodic characteristic of the resampled signals, Prasad [10] and Mahdian [11,12] proposed their method to extract the periodical property of the resampled signals based on analyzing the periodic characteristic of the covariance of the second order derivatives. In [13,14], Lukáš et al.



(a)

(b)

Figure 1. (a) The original image; (b) The tampered image.

proposed a method that utilize the imaging sensor noise as a unique stochastic characteristic to detect the forgeries. Johnson *et al.* [15] discovered that the light condition of the tampered area will be inconsistent to the original image. For the compressed image, Ye *et al.* [16] proposed a method based on the different block artifacts caused by different quantization tables.

Generally, each kind of digital forgery detection method can solve only one kind of forgery pattern. In this study, we only address on the detection of resampling forgery. Two related researches addressed on the detection of resampling forgery are the methods proposed by Popescu [7] and Mahdian [11]. However, there exist two major drawbacks in the above-mentioned algorithms. For the Popescu's method [7], high computation cost in the iterative computing procedure is required. It takes almost 5 minutes to generate the probability map for the image with resolution 512×512 pixels. For the method proposed by Mahdian [11], we found that the derivative kernel used in [11] will destroy the periodicity of the correlation function at the high texture regions. Hence, in this study, we try to investigate and analyze the intrinsic properties of resampling scheme and develop a new more efficient algorithm based on the intrinsic properties of resampling.

Based on the periodical property that the original values can be selected from the resampling process, some of the reconstructed values would exactly overlap the original values in resampled signal and then the error between the predicted value and the resampled value would be very small. By analyzing the prediction error distribution generated by the weighting tables from different resampling rates, we can detect the digital forgeries. To enhance the periodical property, the projection operation is used for creating one-dimensional prominent periodical patterns. In addition, both of the vertical and horizontal predicting error variations are considered simultaneously.

The rest of this paper is organized as follows. In Section 2, two typical forgery detection methods are described. In Section 3, a new forgery detection method based on the intrinsic properties of resampling is proposed, which can detect the tampered regions more efficiently. In Section 4, we present the efficiency and accuracy analyses among the proposed method and other approaches. Finally, we summarize the contributions and future works in Section 5.

2. Related Works

In this section, two typical forgery detection methods for the resampling forgery techniques are introduced. These methods detect the forgery by tracing the interpolation clues of resampled signal

2.1. The Popescu's Method

A well known forgery detection method proposed by Popescu [7] assume that the interpolated samples are the linear combination of their neighboring pixels and try to train a set of resampling coefficients to estimate the probability map. In this method, a digital sample can be categorized into two models: M_1 and M_2 . M_1 denotes the model that the sample is correlated to their neighbors; while M_2 denotes that the sample isn't correlated to its neighbors. The resampling coefficients can be acquired by the EM algorithm. In the E-step, the probability for M_1 model for every sample is calculated. In the M-step, the specific correlation coefficients are estimated and updated continuously. The detailed description of the forgery detection algorithm is described in the sequel.

2.1.1. E-Step

The conditional probability for sample y[i] belonging to M_1 model is calculated by the following formula.

$$\Pr\left\{y[i]|y[i] \in M_{1}\right\}$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \exp\left[\frac{-\left(y[i] - \sum_{k=-N}^{N} \alpha_{k} y[i+k]\right)^{2}}{2\sigma^{2}}\right]$$
(1)

2.1.2. M-Step

Minimize the quadratic error function defined in Equation (2) by updating the correlation coefficients α iteratively.

$$E(\vec{\alpha}) = \sum_{i} \omega(i) \left(y[i] - \sum_{k=-N}^{N} \alpha_{k} y[i+k] \right)^{2}$$
(2)

where $\omega(i) \equiv \Pr\{y[i] \in M_1 | y[i]\}.$

After applying the Popescu's method to the image, we can obtain a probability map. The peak ratio of frequency response of the probability map can be used to identify the digital forgery. Figure 2 illustrates that the peaks of frequency response exist in the tampered image. On the contrary, no peaks exist in the original image shown in Figure 2(a).

2.2. The Mahdian's Method

Another method proposed by Mahdian and Saic [11] demonstrates that the interpolation operation can exhibit periodicity in their derivative distributions. To emphasize the periodical property, they employ the radon transformation to project the derivatives along a certain orientation. The radon transformation is defined as:



Figure 2. The frequency response of the probability maps generated from Popescu's method for the original image, resampled images with up-sampling rate 10% and 20% respectively.

$$\rho D^{2} \left\{ b \right\} \left(x, y \right) = \int_{L} \left| D^{2} \left\{ b \left(x, y \right) \right\} \right| dl$$
(3)

where, b(x, y) denotes the pixel in the block with size of $R \times R$ and D^2 {*} denotes the derivative kernel of order 2. The radon transform along angle θ (0 ~ 179°) is defined in Equation (4).

$$\rho_{\theta}(x') = \int_{-\infty}^{\infty} \left| D^{2} \{ b(x, y) \} \right| \cdot (x' \cos \theta) - y' \sin \theta, x' \sin \theta + y' \cos \theta dy'$$
(4)

After projecting all the derivatives to one direction and forming 1-D projection vectors, the autocovariance function can be used to emphasize the periodicity and defined as:

$$R_{\rho_{\theta}}(k) = \sum_{i} \left(\rho_{\theta}(i+k) - \overline{\rho_{\theta}} \right) \left(\rho_{\theta}(i) - \overline{\rho_{\theta}} \right)$$
(5)

Then, the Fourier transformation of $R_{\rho_{\theta}}$ are also computed to identify the periodic peaks which can indicate the existing of digital forgery. The simulation results are shown in **Figure 3**. It shows that the resampled image can have strong peaks in the frequency response of the derivative covariance.

3. Forgery Detection Using the Resampling Intrinsic Properties

There exist two major drawbacks in the above-mentioned algorithms. For the Popescu's method [7], high computation cost in the iterative computing procedure is required. It takes almost 5 minutes to generate the probability map for an image with resolution 512×512 pixels. For the method proposed by Mahdian [11], we found that the derivative kernel used in [11] can reduce the periodicity of the correlation function at the high texture region. Hence, in this study we try to investigate and analyze the intrinsic properties of resampling process and develop a new more efficient algorithm. The system flowchart is shown in **Figure 4** and the detailed function for each block will be described in the following subsections.

3.1. Intrinsic Properties of Resampled Signal

In this section, we firstly introduce the procedures of general resampling process. The up-sampling process is illustrated in **Figure 5(a)** and the original values are denoted as red bars. **Figure 5(b)** shows that interpolation operation fills the empty points with the linear combination of the adjacent signals' values which are denoted as yellow bars. Finally, the samples selected for decimation process which are denoted as blue bars are shown in **Figure 5(c)**. Through the observation of the resampling process, it gives us an important clue to design a new



3.2. Periodicity of the Prediction Error

Every resampled value denoted as blue bar in Figure 5 can be approximated by the linear combinations of the adjacent original values denoted as red bar with different weights according to their positions, *i.e.*, the weighting in the linear interpolation algorithm is propositional to the distance to their neighbors. Here, we pre-calculate the weighing table (shown in Table 1) for each resampling rates. If the resampling rate is known, then the original values can be approximated by the linear combination of the interpolated values. Based on the periodical property of the original values selected from resampling, some of the approximated values would exactly overlap the original values in resampled signal (see the green bar in Figure 6). Ideally, the error between the predicted value and the resampled value would be very small at the position where the original value is resampled (the green bar in Figure 6). Moreover, the variation of the prediction error will distribute periodically. The weighting table WT [i], i = 1, 2, ..., N, should be calculated in advance for





Copyright © 2010 SciRes.



Figure 5. An example for illustrating the intrinsic property of resampled signal. The scaling factor used here is 6/5. (a) The up-sampling for the original values (red bars); (b) Linear interpolation denoted as yellow bars; (c) Down sampling of signal in (b). The resampled signal is denoted as blue bars. The blue bars labeled the white node denote that the original values are chosen.



Figure 6. The values (red bar) could be predicted by the resampled values (blue bar). After a certain periodical time interval, the predicted value will overlap the original value denoted as green bar.

Table 1.	Weighting	table for	resampling	rate 6/5

	$WT_{\rm L}[i]$	$WT_{\rm R}[i]$
1	1/6	5/6
2	2/6	4/6
3	3/6	3/6
4	4/6	2/6
5	5/6	1/6

each resampling rate. The prediction process is described in **Figure 6**.

In **Figure 6**, the interpolated values can be computed as:

$$B_{i} = R_{i-1} \times WT_{L}[i-1] + R_{i} \times WT_{R}[i-1]$$
(6)

Then, the predicted resampling values can be computed as:

$$pre_{1} = R_{2} = \frac{B_{2} - R_{1}^{*}WT_{L}[i]}{WT_{R}[i]}$$

$$pre_{2} = R_{3} = \frac{B_{3} - pre_{1}^{*}WT_{L}[i]}{WT_{R}[i]}$$
(7)

$$pre_{m} = R_{N} = \frac{B_{N} - pre_{m-1}^{*}WT_{L}[i]}{WT_{R}[i]} = B_{N+1}$$

Finally, the prediction error within the certain sliding window can be computed as:

Prediction error =
$$|B_{N+1} - pre_m|$$
 (8)

For the case of resampling rate 120%, the difference between pre_5 and B_7 will be very small. When the sliding window for calculating the sample prediction is moving (shown in **Figure 7**), the prediction error will increase and then decrease to the minimum value until the sliding window moves to the next periodical position (B_{14} , B_{21} ...). Such a periodical property makes the sequence of prediction error distribute periodically shown in **Figure 8**. In order to enhance this property, the projection operation is also performed for every row and column (two directions are considered separately) before we utilize the frequency analysis to detect the forgery patterns (peaks in frequency response). If the test samples are not resampled or the wrong weighting table is selected, the distribution of prediction error would be irregular.



Figure 7. The sliding window for calculating the sample prediction using the pre-calculated weighting table.







(c)

1357912162024283236404448525660646872768084889296101107113





Figure 8. (a) The original image; (b) Resampled image with up-sampled rate 10%; (c) The magnitudes of row-based prediction error variation projection of (a); (d) The magnitudes of row-based prediction error variation projection of (b); (e) The frequency response of (c); (f) The frequency response (d).

To develop an automatic forgery detection method, there are two main criteria should be considered. The first one is the position where the peak occurs and the second one is the peak ratio. According to the different weighting tables (different resampling rate) for the forgery detection and the specific periodical property for each resampling rate, the expected position where the peak occurs could be forecasted. Then, we can match the peak position to the forecasted position where the specific resampling rate generates for identifying the existence of digital forgery. If the ratio is larger than a specified threshold, we can identify that existence of digital forgery. Finally, the flowchart of the proposed system is shown in Figure 9. To detect the tampered region, the image is scanned from left-top to right-bottom with different block sizes. In each block, the proposed method is applied to detect the tampered regions.

4. Experimental Results

In this section, the efficiency and accuracy for Popescu'd method [7], Mahdian's method [11], and the proposed method are analyzed. The experimental database is constructed with 160 gray level images with resolution 512 \times 512 and each image is partial tampered in BMP format. The image tampering is based on the resampling process with the different bi-linear sampling rates: 105%, 110%, 120% and 125%. The forgery detections are performed by scanning the image with the block size of 128 \times 128 pixels.

Before analyzing the accuracy of forgery detection, we firstly describe the detection rules for the Popescu's [7], Mahdian's [11], and our methods. Here, the forgery detection of Popescu's and Mahdian's methods is deter-



Figure 9. The flowchart of the proposed method.

mined by evaluating whether the ratio of peak-to-average frequency response is larger than a predefined threshold value or not. The ratio of peak-to-average frequency response is defined as:

$$R_{Pop \, \text{sec}\, u} = R_{Mahdian} = \frac{magnitude_{\text{maximum}}}{magnitude_{\text{average}}}$$

For our method, the forgery detection is determined by evaluating whether the ratio of forecasted peak-to-average frequency response is larger than a predefined threshold value or not. The ratio of forecasted peak-to-average frequency response is defined as:

$$R_{our} = \frac{magnitude_{\text{forecasted position}}}{magnitude_{\text{average}}}$$

The resampled image with rate 120% shown in **Figure 10(a)** is used as the tampered image for analyzing the detection accuracy for the three methods. **Figure 10(b)** shows the probability map produced by the Popescu's method and **Figure 10(c)** shows the frequency response of the probability map. **Figure 11(a)** shows the radon transformation of the derivative along horizontal direc-

tion generated by Mahdian's method and Figure 11(b) shows its auto-covariance. Figure 11(c) shows the frequency response of the auto-covariance values. Based on the proposed method, the prediction error generated by the novel algorithm is shown in Figure 12(a). Figure 12(b) presents the frequency response of the prediction error. An obvious drawback of the Mahdian's method is that the weak periodical patterns occur at the high texture regions shown in Figure 11(c). The accuracy analyses of forgery detections for different resampling rates are analyzed in Table 2.





(b)



Figure 10. (a) The tampered image; (b) The probability map generated by the Popescu's method; (c) The frequency

response of (b).



Figure 11. (a) The radon transformation output of Figure 13 by the Mahdian's method; (b) The autocovariance of (a); (c) The frequency response of (b).

The ROC curves with different up-sampling rates for Popescu's, Mahdian's and our methods are shown in **Figure 13**. In this Figure, the detection accuracy of Popescu's method is the highest one and the detection accuracy of our method is close to the Popescu's curve. However, our method is the fastest one that will be mentioned later. The detection accuracy of Mahdian's method is the lowest because the detection accuracy is affected by the high texture regions.



Figure 12. (a) The prediction error of the tampered image shown in Figure 10, which is generated by the proposed method; (b) The frequency response of (a).

Table 2. The accuracy analysis for the methods of our, Popescu's and Mahdian's with 40 resampled images for different rates.

	Popescu's method			Our method			Mahdian's method					
Up-sampling	5%	10%	20%	25%	5%	10%	20%	25%	5%	10%	20%	25%
Positive	40	40	40	40	40	40	40	40	40	40	40	40
Negative	40	40	40	40	40	40	40	40	40	40	40	40
True positive	40	39	40	40	38	39	40	40	21	22	37	37
True negative	40	40	40	40	35	37	38	38	25	33	28	30
Accuracy	100%	98.7%	100%	100%	91.2%	95%	97.5%	97.5%	57.5%	68.7%	81.2%	83.7%



Figure 13. The ROC curves of (a) Up-sampling 5%; (b) Up-sampling 10%; (c) Up-sampling 20%; (d) Up-sampling 25%.

In addition, we compare the efficiency among Popescu's [7], Mahdian's [11] and our methods with the PC of 1.8 GHz. The efficiency analysis is shown in **Figure 14**. Here, we perform the efficiency analysis from block size 64×64 to 512×512 and assume there are 21 weighting tables for 21 resampling rates used in [7]. Because the iterative EM algorithm is very time-consuming, the efficiency of Popescu's method is the lowest. On the contrary, the highest efficiency is presented in Mahdian's method because the operations in his method are very simple. It's worthy to conclude that detection accuracy and efficiency of our method can approach both of the benefits of Popescu's and Mahdian's methods.

Figures 15-16 show the detection results of the pro-





(c)



Figure 14. Efficiency analysis.





(d)



Figure 15. (a) Original image; (b) Image with up-sample rate 5%; (c) Forgery image composed from (a), (b); (d) Detection result with 64×64 block size; (e) Detection result with 128×128 block size.



Figure 16. (a) Original image; (b) Forgery image composed from up-sample (a) 10% and put the bottle near beside; (c) Detection result with 64 × 64 block size; (d) Detection result with 128 × 128 block size.

-posed method for different resampling rates with two block sizes. In **Figure 15**, the man's head in **Figure 15(b)** is cropped and replaced the head region in **Figure 15(a)** to synthesize the forgery image shown in **Figure 15(c)**. **Figure 15(d)** and **Figure 15(e)** show the detection result with 64×64 and 128×128 block sizes. **Figure 16(a)** shows an original bottle image and **Figure 16(b)** shows that a resized bottle is put on the left side of the tampered image. **Figures 16(c)** and **16(d)** show the detection results with different block sizes. Here, we observe that the detection accuracy for the smaller block size is lower than the accuracy with larger block size because more periodical patterns can be collected in larger blocks.

5. Conclusions

In this paper, we propose a novel method based on the intrinsic properties of resampling scheme to detect the forgery regions with the pre-calculated resampling weighting tables and the detecting of periodic patterns for the vertical and horizontal prediction error. In Popescu's method, high accuracy can be obtained with high computation cost. On the contrary, in Mahdian's method, the detecting accuracy can be affected on the high texture regions. The detection accuracy and efficiency of our method can approach both of the benefits of Popescu's and Mahdian's methods. The detection accuracy of our method is about 95% and the time for detecting a 512×512 image needs only 50 seconds.

6. References

- R. B. Wolfgang and E. J. Delp, "A Watermark for Digital Image," *Proceedings of the International Conference on Image Processing*, Vol. 3, 1996, pp. 219-222.
- [2] R. B. Wolfgang, C. I. Podilchuk and E. J. Delp, "Perceptual Watermarks for Digital Images and Video," *Proceedings of the IEEE, Special Issue on Identification and Protection of Multimedia Information*, Vol. 87, No. 7, 1999, pp. 1108-1126.
- [3] M. Wu and B. Liu, "Watermarking for Image Authentication," *IEEE International Conference on Image Proc*essing, Vol. 2, 1998, pp. 437-441.
- [4] M. Yeung and F. Mintzer, "An Invisible Watermarking Technique for Image Verification," *Proceedings of the International Conference on Image Processing*, Vol. 1,

22

1997, pp. 680-683.

- [5] J. Fridrich, D. Soukal and J. Lukáš, "Detection of Copy-Move Forgery in Digital Images," *Proceedings of the Digital Forensic Research Workshop*, Cleveland, 2003.
- [6] L. Zhou, D. Wang, Y. Guo and J. Zhang, "Blue Detection of Digital Forgery Using Mathematical Morphology," *Technical Report, KES AMSTA*, Springer-Verlag, Berlin, Heidelberg, 2007, pp. 990-998.
- [7] A. C. Popescu and H. Farid, "Exposing Digital Forgeries by Detecting Traces of Resampling," *IEEE Transactions* on Signal Processing, Vol. 53, No. 2, 2005, pp. 758-767.
- [8] A. C. Popescu and H. Farid, "Exposing Digital Forgeries in Color Filter Array Interpolated Images," *IEEE Transactions on Signal Processing*, Vol. 53, No. 10, 2005, pp. 3948-3959.
- [9] M. Kirchner, "Fast and Reliable Resampling Detection by Spectral Analysis of Fixed Linear Predictor Residue," *MM & Sec*'08, *Proceedings of the Multimedia and Security Workshop*, 2008, pp. 11-20.
- [10] S. Prasad and K. Ramakrishnan, "On Resampling Detection and its Application to Detect Image Tampering," *Proceedings of the 2006 IEEE International Conference* on Multimedia and EXPO, 2006, pp. 1325-1328.
- [11] B. Mahdian and S. Saic, "Blind Authentication Using Periodic Properties of Interpolation," *IEEE Transactions* on Information Forensics and Security, Vol. 3, No. 3,

2008, pp. 529-538.

- [12] B. Mahdian and S. Saic, "Detection of Resampling Supplemented with Noise Inconsistencies Analysis for Image Forensics," *International Conference on Computational Sciences and its Applications*, Vol. 81, No. 4, 2008, pp. 546-556.
- [13] J. Lukáš, J. Fridrich and M. Goljan, "Detecting Digital Image Forgeries Using Sensor Pattern Noise," *Proceedings of the SPIE Conference on Security, Steganography* and Watermarking of Multimedia Contents, Vol. 6072, 2006, pp. 362-372.
- [14] J. Lukáš, J. Fridrich and M. Goljan, "Digital Camera Identification from Sensor Pattern Noise," *IEEE Transactions on Information Security and Forensics*, Vol. 1, No. 2, 2006, pp. 205-214.
- [15] M. K. Johnson and H. Farid, "Exposing Digital Forgeries in Complex Lighting Environments," *IEEE Transactions* on *Information Forensics and Security*, Vol. 2, No. 4, 2007, pp. 450-461.
- [16] S. Ye, Q. Sun and E. Chang, "Detecting Digital Image Forgeries by Measuring Inconsistencies of Blocking Artifact," *IEEE International Conference on Multimedia* and EXPO, 2007, pp. 12-15.
- [17] M. C. Stamm and K. J. R. Liu, "Forensic Detection of Image Tampering Using Intrinsic Statistical Fingerprints in Histograms," *Proceedings of the APSIPA Annual Summit and Conference*, Sapporo, 2009.