

# Contrast Enhancement-Based Forensics in Digital Images

Gang Cao, Yao Zhao, *Senior Member, IEEE*, Rongrong Ni, *Member, IEEE*, and Xuelong Li, *Fellow, IEEE*

**Abstract**—As a retouching manipulation, contrast enhancement is typically used to adjust the global brightness and contrast of digital images. Malicious users may also perform contrast enhancement locally for creating a realistic composite image. As such it is significant to detect contrast enhancement blindly for verifying the originality and authenticity of the digital images. In this paper, we propose two novel algorithms to detect the contrast enhancement involved manipulations in digital images. First, we focus on the detection of global contrast enhancement applied to the previously JPEG-compressed images, which are widespread in real applications. The histogram peak/gap artifacts incurred by the JPEG compression and pixel value mappings are analyzed theoretically, and distinguished by identifying the zero-height gap fingerprints. Second, we propose to identify the composite image created by enforcing contrast adjustment on either one or both source regions. The positions of detected blockwise peak/gap bins are clustered for recognizing the contrast enhancement mappings applied to different source regions. The consistency between regional artifacts is checked for discovering the image forgeries and locating the composition boundary. Extensive experiments have verified the effectiveness and efficacy of the proposed techniques.

**Index Terms**—Computer vision, digital forensics, image forgery, contrast enhancement, composite image.

## I. INTRODUCTION

WITH THE rapid development of digital media editing techniques, digital image manipulation becomes rather convenient and easy. While it benefits to legal image processing, malicious users might use such innocent manipulations to tamper digital photograph images. Currently, image forgeries

Manuscript received June 5, 2013; revised September 14, 2013 and December 26, 2013; accepted January 2, 2014. Date of publication January 16, 2014; date of current version February 14, 2014. This work was supported in part by the 973 Program under Grant 2011CB302204, in part by the National Natural Science Foundation of China under Grants 61025013, 61125106, 61210006, 61272355, and 61332012, and in part by PCSIRT under Grant IRT 201206. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Chiou-Ting Hsu.

G. Cao was with the Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China. He is now with the School of Computer Science, Communication University of China, Beijing 100024, China (e-mail: gangcao@cuc.edu.cn).

Y. Zhao is with the Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China, and also with the State Key Laboratory of Rail Traffic Control and Safety, Beijing 100044, China yzhao@bjtu.edu.cn).

R. Ni is with the Institute of Information Science, Beijing Jiaotong University, Beijing 100044, China, and also with the Beijing Key Laboratory of Advanced Information Science and Network Technology, Beijing 100044, China (e-mail: rni@bjtu.edu.cn).

X. Li is with the Center for OPTical Imagery Analysis and Learning (OPTIMAL), State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China (e-mail: xuelong\_li@opt.ac.cn).

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.2014.2300937

are widespread on the Internet and other security-related applications such as surveillance and recognition that utilize images are therefore impacted. The event and scene information delivered in images might become no longer believable. In the applications such as law enforcement and news recording, it is also necessary to verify the originality and authenticity of digital images, and make clear the image manipulation history to get more information. To circumvent such a problem, digital forensic techniques have been proposed to blindly verify the integrity and authenticity of digital images [1], [2].

A set of previous works deal with image manipulation detection by classifier-based approaches. In [3], three categories of statistical features including binary similarity, image quality and wavelet statistics were developed. Swaminathan *et al.* proposed to estimate both in-camera and post-camera operation fingerprints for verifying the integrity of photographs [4]. Cao *et al.* designed a new ensemble manipulation detector to simultaneously detect a wide range of manipulation types on local image patches [5]. Fan *et al.* proposed to correlate statistical noise features with exchangeable image file format header features for manipulation detection [6]. Although such techniques [3]–[6] could detect if manipulation occurred, they fail to determine which specific type of manipulation was enforced.

There also exist another category of forensic techniques which focus on detecting specific image manipulations. Since each manipulation typically leaves behind unique fingerprints on images, it is feasible to design individual tests to identify each type of enforced manipulation. The manipulation-specific detection techniques can help recover the image processing history. The prior works focus on detecting different types of alterations, which can be broadly divided into two categories: 1) non-content-changing operations including resampling [7], compression [8], sharpening filtering [9], contrast enhancement [10]–[13] and median filtering [14], [15]; 2) content-changing operations, i.e., splicing and composition [16]–[18].

Note that the prior contrast enhancement forensic algorithms work well under the assumption that the gray level histogram of an unaltered image exhibits a smooth contour. However, digital images are often stored in the JPEG format and even heavily compressed with a middle/low quality factor (Q) in real applications, such as the Internet and mobile phones. It is well-known that the low quality JPEG compression usually generates blocking artifacts, which might cause unsmoothness and even locally dense peak bins in the gray level histogram as shown in Fig. 1. In such a scenario, the existing approaches [10]–[12] fail to detect contrast enhancement in the previously low quality JPEG-compressed images, since the assumption of

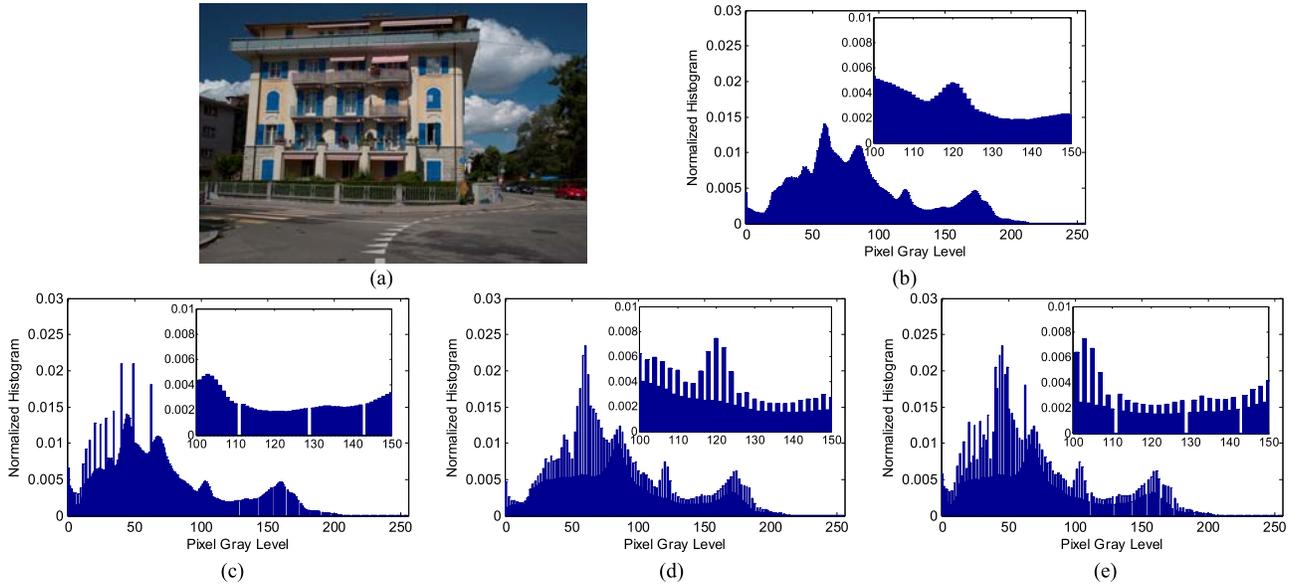


Fig. 1. Gray level histogram of contrast-enhanced images. (a) Raw image; histogram of (b) the raw image and its (c) enhanced ( $r = 1.2$ ), (d) JPEG ( $Q = 50$ ) and (e) JPEG ( $Q = 50$ ) followed by enhancement ( $r = 1.2$ ) versions. Histogram of the green channel and its locally magnified version are shown here.  $r$  denotes gamma value.

smoothness becomes dissatisfied. To solve such a problem, we propose to detect the global contrast enhancement not only in uncompressed or high quality JPEG-compressed images, but also in low quality ones. The main strategy relies on the blind identification of zero-height gap bins.

Besides global contrast enhancement, the detection of local contrast enhancement is also significant. A valuable application is to identify the cut-and-paste type of forgery images, in which the contrast of one source object region is enhanced to match the rest. In [10], although the composite image created by enhancing *single* source region could be identified, those enhanced in *both* source regions may not. In this paper, a new method is proposed to identify both single-source-enhanced and both-source-enhanced composite images. Peak/gap pattern of the pixel value mapping applied to each source region is self-learned from the detected blockwise peak/gap positions. Then composition boundary is located by detecting the inconsistency between the position vectors in different regions.

The rest of this paper is organized as follows. In Section II, we revisit the previous works on contrast enhancement forensics in digital images. In Section III and IV, two algorithms are proposed to detect global contrast enhancement and identify the source-enhanced composite image, respectively. Experimental results and discussions are presented in Section V. The conclusion is drawn in Section VI.

## II. PREVIOUS WORKS

### A. Global Contrast Enhancement Detection

Generally, contrast enhancement is inherently a pixel value mapping,  $y = m(x)$ , where  $m(\cdot)$  is the mapping function,  $x, y = 0, 1, 2, \dots, 255$  are the pixel gray level before and after mapping, respectively. Without special statement, 8-bit grayscale images are considered in this paper. It is shown in [10]–[12] that the histogram of unaltered images typically

conforms to a smooth envelope, while that of enhanced images presents peak/gap artifacts as indicated in Fig. 1(b) and (c). Based on such observation, Stamm and Liu [10] proposed a contrast enhancement detection algorithm as follows.

- 1) Get the test image's gray level histogram  $h(x)$ .
- 2) Compute  $g(x) = p(x)h(x)$ , where  $p(x)$  is a pinch-off function for combating the false peak artifacts in saturated images.
- 3) Calculate the high frequency energy metric  $F$  as

$$F = \frac{1}{N} \sum_w |b(w)G(w)|. \quad (1)$$

Here,  $G(w)$  is the discrete Fourier transform of  $g(x)$ .  $b(w) = 1$ , if  $|w| \geq c$ ;  $b(w) = 0$ , otherwise, where  $c$  is a cutoff frequency.  $N$  is the number of pixels in the image.

- 4) Determine if contrast enhancement has been performed by applying thresholding test. If  $F$  is greater than the decision threshold, contrast enhancement is detected.

In [11] and [12], such peak/gap artifacts are further exploited to estimate the pixel valued mapping involved in contrast enhancement. The estimation is valid only when the contrast enhancement operation has been detected before. In [13], a method is designed to identify an operation chain, which is composed by double JPEG compression interleaved by a linear contrast enhancement. The special type of contrast enhancement, i.e., linear pixel value stretching, can be detected on condition that the operation chain occurs. In this paper, we would address the detection of general pixel value mappings, which are typically nonlinear and widely used in latest techniques [21]–[23].

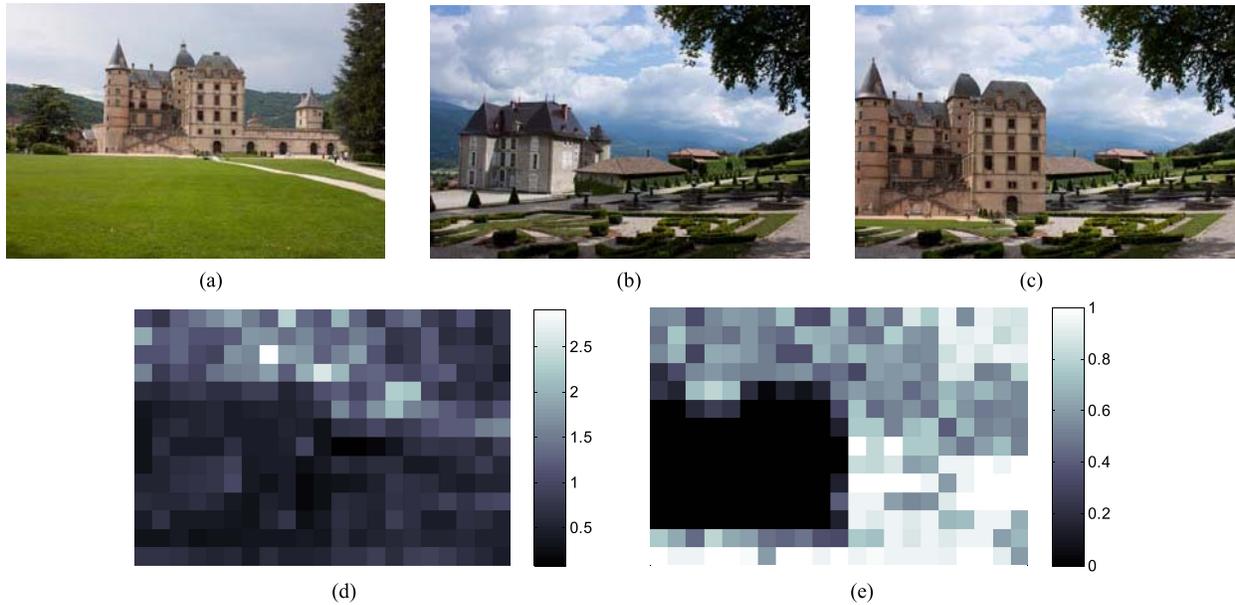


Fig. 2. Detection results on a composite image. (a) and (b) Two source images; (c) both-source-enhanced composite image; blockwise metric map of (d) Stamm and Liu method [10] and (e) our proposed method. Blocksize = 200.

### B. Local Contrast Enhancement Detection

In [10], the global contrast enhancement detection method is applied to image blocks for detecting local enhancement. In practice, the two source images used for creating a composite image may have different color temperature or luminance contrast. In order to make the composite image more realistic, contrast enhancement is typically performed in one or both source regions. In such scenarios, the composite image created by applying contrast enhancement to a single source region can be detected by the prior method [10]. However, if contrast enhancement is enforced in both regions, such a method may fail since all blocks are detected as enhanced ones. As shown in Fig. 2, a composite image (c) is created by copying a region of the source image (a) and pasting onto another source image (b). For matching the regional contrast perception, the two source regions are enhanced by different mappings using the ‘Curve’ tool in Adobe Photoshop. Fig. 2(d) shows the detection result of Stamm and Liu method [10], where blockwise  $F$  values across the whole image are so consistent that the tampered region could not be distinguished correctly. Such a result illustrates that the prior method fails to detect both-source-enhanced composite images.

## III. GLOBAL CONTRAST ENHANCEMENT DETECTION

In this section, we propose an efficient algorithm which can detect the global contrast enhancement in both uncompressed and previously JPEG-compressed images.

### A. Peak/Gap Artifacts Incurred by JPEG Compression

The prior methods [10]–[12] fail to detect contrast enhancement in the previously middle/low quality JPEG-compressed images. To investigate the reasons behind such ineffectiveness, the impact of JPEG compression on histograms is analyzed.

JPEG compression of an 8-bit gray-scale image includes the steps: pixel value shift from the range  $[0, 255]$  to  $[-128, 127]$ ,  $8 \times 8$  blocks division, block discrete cosine transform (DCT), quantization and entropy encoding [24]. In natural images, more or less flat regions exist. Once applying DCT, only DC coefficient is large while the others are small in flat blocks. After quantization, it is reasonable to assume that there exist the blocks in which merely the quantized DC coefficient is non-zero. After inverse DCT, the decompressed flat block denoted by  $I_b$  can be computed as

$$I_b(i, j) = \text{round}(qn_d/8 + 128). \quad (2)$$

Here,  $i, j = 0, 1, 2, \dots, 7$  are local pixel coordinates.  $\text{round}(\cdot)$  is the rounding operator.  $q$  is DC quantization step.  $n_d = L, L + 1, L + 2, \dots, U$  denotes the quantized DC coefficient, where  $L = \text{round}(-128 \times 8/q)$  and  $U = \text{round}(127 \times 8/q)$  are deduced from the definitions in JPEG compression standard.

Let  $c = q/8 = T_1/T_2$ , such that  $T_1, T_2 \in \mathbb{Z}$  are relatively prime.  $I_b(i, j)$  via  $n_d$  is inherently a *scale round mapping*, which has been proved to definitely incur a certain periodicity [10]. Specifically, let  $\Psi_c(b) = \{n_d | b = \text{round}(cn_d)\}$  be the set of  $n_d$  mapped to the same value  $b$ . The cardinality  $|\Psi_c(b)|$  is periodic in  $b$  with period  $T_1$ . Note that  $|\Psi_c(b)|$  is nonconstant if  $q \neq 1, 2, 4, 8$ . Because of such periodic and nonconstant properties, candidate peaks may appear in the gray level histogram bin at  $k$ , where

$$k \in \left\{ b' \mid b' = \arg \max_{b \in \{0, 1, \dots, 255\}} (|\Psi_c(b)|) \right\}. \quad (3)$$

In fact, two factors affect the de facto presence of histogram peak bins in a JPEG image: 1) the flatness; 2) JPEG quality factor. The larger flat regions and larger DC quantization step would cause more apparent peak bins. Even for the block with sparse non-zero quantized AC coefficients, after

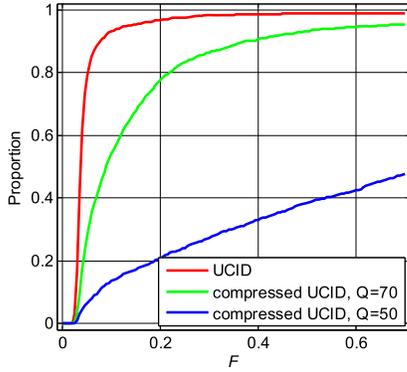


Fig. 3. Cumulative distribution of  $F$  values computed from UCID images and their JPEG versions.

decompression the number of pixel gray levels still decreases to some extent. So the global histogram is still prone to discontinuity.

Strength of the histogram peak/gap artifacts incurred by JPEG compression can also be measured by the metric  $F$  defined in (1). Fig. 3 shows the cumulative distribution of  $F$  computed from UCID images and their compressed versions [25]. It demonstrates that  $F$  of more than 85% uncompressed images lie below 0.06, which can be taken as the approximate threshold for determining peak occurrence. After compression,  $F$  values of the images increase and most of them become larger than 0.06. Moreover, the smaller the  $Q$  is, the larger the  $F$  becomes. Such results validate that peak/gap artifacts really exist in most middle/low quality JPEG images.

### B. Peak/Gap Artifacts Incurred by Contrast Enhancement

To investigate the impact of contrast enhancement on histograms, the histogram of an enhanced image,  $h_Y$ , is written as

$$h_Y(y) = \sum_x h_X(x) \ell(m(x) = y), \quad (4)$$

where  $\ell(\cdot)$  denotes the indicator function defined as  $\ell(u) = 1$ , if  $u = 1$ ;  $\ell(u) = 0$ , otherwise.  $h_X$  is the histogram of the original image. As stated in [11], this equation shows that every value of  $h_Y$  must equal a single value of  $h_X$ , a sum of distinct  $h_X$  values, or zero, depending on the mapping curve's tangent slope. As a result, impulsive peaks will occur in  $h_Y$  at  $y$  values to which multiple  $x$  values were mapped, such as  $y = 254$  in Fig. 4. Similarly, gaps will appear in  $h_Y$  at  $y$ , for example,  $y = 4$  in such an illustration, to which no  $x$  values were mapped.

However, there exists notable difference between the peak/gap artifacts from contrast enhancement and those from JPEG compression. As shown in Fig. 1(c) and (e), the gap bins with zero height, where no primary pixel values are mapped to, always appear in enhanced images. On the contrary, the zero-height gap bins are absent in compressed images since there is lack of a distinct pixel value mapping applied to all pixels. A regular pixel value mapping relationship exists in flat regions, but not in other regions. Therefore, the zero-height gap feature can be used to detect global contrast enhancement in both uncompressed and compressed images.

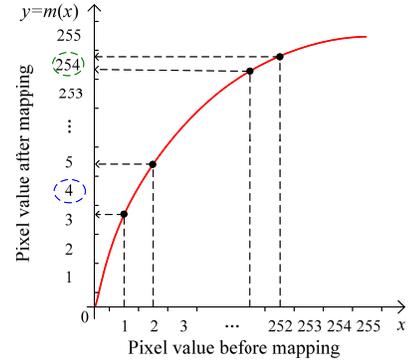


Fig. 4. Peak/gap incurred by pixel value mapping. Here, a peak/gap bin occurs at  $y = 254$  and  $4$ , respectively.

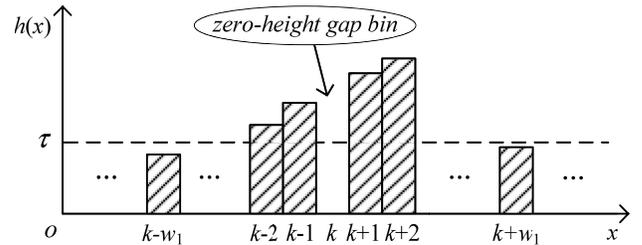


Fig. 5. Definition of a zero-height gap bin at  $k$ .

### C. Proposed Contrast Enhancement Detection Algorithm

Based on the above analyses, we propose the contrast enhancement detection algorithm as follows.

- 1) Get the image's normalized gray level histogram  $h(x)$ .
- 2) Detect the bin at  $k$  as a zero-height gap bin if it satisfies:

$$\begin{cases} h(k) = 0 \\ \min\{h(k-1), h(k+1)\} > \tau \\ \frac{1}{2w_1+1} \sum_{x=k-w_1}^{k+w_1} h(x) > \tau. \end{cases} \quad (5)$$

Here, the first sub-equation assures that the current bin is null. To define a gap bin, the second sub-equation keeps two neighboring bins larger than the threshold  $\tau$ , as shown in Fig. 5. To exclude the zero-height gap bins which may be incorrectly detected in histogram trail-ends, the average of neighboring  $(2w_1 + 1)$  bins should be larger than  $\tau$ , as constrained by the third sub-equation. Experiments show that  $w_1 = 3$  and  $\tau = 0.001$  are appropriate. Note that we focus on the detection of isolated zero-height gap bins but not connected bins, which are rarely present in the middle of histograms.

- 3) Count the number of detected zero-height gap bins, denoted by  $N_g$ . If it is larger than the decision threshold, contrast enhancement is detected, else not.

## IV. IDENTIFY SOURCE-ENHANCED COMPOSITE IMAGES

In this section, a novel algorithm is proposed to identify the source-enhanced composite image created by enforcing contrast adjustment on either single or both source regions. The outline of our technique is shown in Fig. 6. Since positional

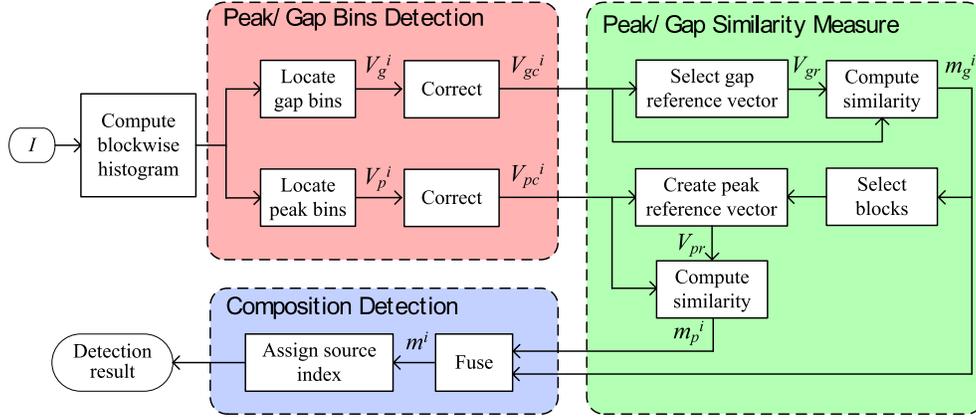


Fig. 6. Flowchart of the proposed composite image detection technique.

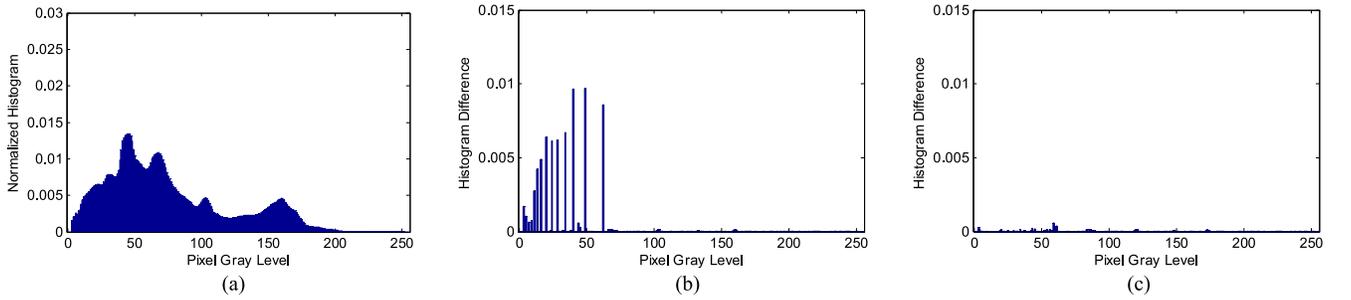


Fig. 7. Peak bins location. (a) Filtered gap-filled version of the histogram shown in Fig. 1(c); histogram difference for the (b) enhanced and (c) unenhanced image.

distribution of the peak/gap bins incurred by contrast enhancement is unique to the involved pixel value mapping, such positional information could serve as fingerprinting feature for identifying different contrast enhancement manipulations [11]. Consistency between the peak/gap artifacts detected in different regions is checked for discovering composite images.

#### A. Blockwise Peak/Gap Bins Location

To locate composition, the test image  $I$  is first divided into nonoverlapping blocks. For the  $i$ -th block, peak/gap bins in its gray level histogram are located as follows. Here and below,  $i = 1, 2, \dots, N_b$ , where  $N_b$  is the number of divided blocks.

1) *Gap Bins Location*: The zero-height gap bins are detected as in Section III-C. The position of detected gap bins is labelled as  $V_g^i = [V_g^i(0), V_g^i(1), \dots, V_g^i(k), \dots, V_g^i(255)]$ , where  $V_g^i(k) = 1$ , if the bin at  $k$  is a gap;  $V_g^i(k) = 0$ , otherwise.

2) *Peak Bins Location*: Peak bins which behave as impulse noise can be located by median filtering. Specifically, the gap bins are first filled with the average of neighboring bins, then median filtering is applied to the gap-filled histogram. As shown in Fig. 7(a), the filtered histogram possesses a smooth contour. Lastly, peak positions are located by thresholding the difference between the gap-filled histogram and its filtered version. The histogram differences for the enhanced and primary example images are shown in Fig. 7(b) and (c), respectively. It can be seen that peak bins are not detected in the primary image. Record the detected peak positions as

$V_p^i = [V_p^i(0), V_p^i(1), \dots, V_p^i(k), \dots, V_p^i(255)]$ , where  $V_p^i(k) = 1$  refers to a peak.

In practice, the peak/gap bins which are theoretically computed from contrast enhancement mapping may not appear since the histogram is too narrow to cover such bin positions. To address such a factor, Effective Detection Range (EDR) for the  $i$ -th blockwise peak/gap position vector, denoted by  $\Omega_i$ , is defined as the set of gray levels around which the histogram bins are not all zeros. In other words, the histogram bins at the gray levels out of EDR are zeros. Because of the narrow histograms incurred by low resolution, EDR of most position vectors actually contains limited gray levels.

To further decrease detection errors, the extracted peak/gap positions are corrected by retaining the co-existing peak/gap positions in most blocks, which are deemed to be determined by the involved mappings. Specifically, we apply a simple and strict threshold-based binarization to  $C_g = \sum_{i=1}^{N_b} V_g^i / N_b$ . The detected co-existing gap positions are recorded as  $V_g = [V_g(0), V_g(1), \dots, V_g(k), \dots, V_g(255)]$ , where  $V_g(k) = 1$ , if  $C_g(k)$  is larger than the threshold;  $V_g(k) = 0$ , otherwise. To eliminate the gap bins which might not be caused by contrast enhancement, the corrected gap position vector,  $V_{gc}^i$ , is generated as

$$V_{gc}^i = V_g^i \odot V_g, \quad (6)$$

where  $\odot$  denotes Hadamard product, i.e.,  $V_{gc}^i(k) = V_g^i(k) V_g(k)$ ,  $k = 0, 1, 2, \dots, 255$ . Similarly, the corrected peak position vector  $V_{pc}^i$  can also be obtained.

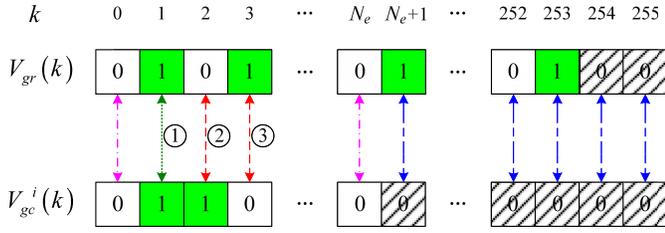


Fig. 8. Correspondence between  $V_{gr}(k)$  and  $V_{gc}^i(k)$ . Here, EDRs of  $V_{gr}$  and  $V_{gc}^i$  are  $\Omega_{gr} = \{0, 1, 2, \dots, 253\}$ ,  $\Omega_i = \{0, 1, 2, \dots, N_e\}$ , respectively,  $N_e \in \mathbb{Z}$ . The positions out of EDR are marked by the box with shadow. The gap-involved pair is matched in the case ① and mismatched in the cases ②③.

### B. Gap Based Similarity Measure

To discriminate two source regions, we should first set a reference position vector for either one. Then each block can be classified by the similarity between its position vector and the reference one. It is reasonable to deem that the blocks with approximate similarity come from the same source image.

The reference position vector should not be selected from splicing boundary. Fortunately, the block with the largest number of zero-height gap bins is believed to locate within one source region. In boundary blocks, the interaction between the pixels from different source regions makes the number of zero-height bins decrease. As a result, the reference gap position vector for its located source region (marked as  $S_1$ ) can be set as  $V_{gr} = V_{gc}^k$ , where

$$k = \arg \max_{i \in \{1, 2, \dots, N_b\}} (\|V_{gc}^i\|). \quad (7)$$

Here,  $\|\cdot\|_1$  denotes 1-norm of a vector. EDR of the reference gap position vector is marked as  $\Omega_{gr} = \Omega_k$ .

To measure the overall similarity between the gap position vectors  $V_{gc}^i$  and  $V_{gr}$ , each gap-involved pair  $V_{gc}^i(k)$  and  $V_{gr}(k)$  should be investigated firstly. Since the histogram at the gray levels out of EDR does not deliver any effective peak/gap information left by contrast enhancement, the element pairs in the intersection of two EDRs are used to measure the similarity. As shown in Fig. 8, there exist three possible correspondences for a gap-involved pair. They are,

- ①  $V_{gr}(k) = 1, V_{gc}^i(k) = 1;$
- ②  $V_{gr}(k) = 0, V_{gc}^i(k) = 1;$
- ③  $V_{gr}(k) = 1, V_{gc}^i(k) = 0.$

We can see that the gap-involved pair is matched in the case ① but mismatched in ②③. The overall similarity between  $V_{gc}^i$  and  $V_{gr}$  is determined by such three cases' frequency. In the intersection between two EDRs, the ratio between the number of matched pairs and that of total gap-involved pairs is defined as the similarity. The pair is completely matched in the case ①. The more pairs occur as such, the more similar the two gap position distributions are. In the case ②, the detected gap is not marked in the reference vector. In the case ③, the gap in the reference vector is absent in the unlabeled histogram. Both ② and ③ attribute to the different contrast enhancement mapping applied to the image region out of the reference block.

Based on the above analyses, the similarity between  $V_{gc}^i$  and  $V_{gr}$ , denoted by  $m_g^i$ , can be defined as the equation (8). Here,

$$m_g^i = \frac{\sum_{k \in \Omega_i \cap \Omega_{gr}} V_{gc}^i(k) \cdot V_{gr}(k)}{\sum_{k \in \Omega_i \cap \Omega_{gr}} V_{gc}^i(k) \cdot V_{gr}(k) + \overline{V_{gc}^i(k)} \cdot V_{gr}(k) + V_{gc}^i(k) \cdot \overline{V_{gr}(k)}}, \quad (8)$$

$\overline{V}$  is the element-wise complement of  $V$ , that is,  $\overline{V}(k) = 1 - V(k)$ ,  $k = 0, 1, 2, \dots, 255$ .  $\Omega_i \cap \Omega_{gr}$  denotes the intersection between EDRs of  $V_{gc}^i$  and  $V_{gr}$ . When no gap-involved pair is detected in the EDR intersection, we mark  $m_g^i = -1$ .

We can see that  $m_g^i$  becomes larger if more gaps co-exist.  $m_g^i$  approaches the maximum 1 if all gaps are matched, and decreases to 0 if all gap-involved pairs behave as ②③. Correspondingly, the possibility that the two blocks undergo the same pixel value mapping ranges from the largest to the smallest.

### C. Peak Based Similarity Measure

Since the image block usually owns a narrow histogram, the theoretical gap bins might be unavailable in not a few blocks. Such blocks can not be assigned to either source region merely based on the gap information. However, the narrow histogram without gap bins might carry with peak bins. As such the peak bins should also be exploited to identify mappings.

The reference peak position vector  $V_{pr}$  is created by combining the peak position vectors which are more possible from the source region of  $V_{gr}$ , namely  $S_1$ . Such creation of  $V_{pr}$  is reliable since the collected peak information is relatively accurate. Specifically, we have

$$V_{pr}(k) = \ell \left( \sum_{n \in N_R} V_{pc}^n(k) > 0 \right). \quad (9)$$

Here,  $k = 0, 1, 2, \dots, 255$ .  $\ell(\cdot)$  is the indicator function.  $N_R = \{i | m_g^i > t_g\}$ , where  $t_g$  is the threshold used to select the candidate blocks. Experimental results show that  $t_g = 0.5 \sim 0.8$  is suitable. Correspondingly, EDR of  $V_{pr}$  is computed as  $\Omega_{pr} = \cup_{n \in N_R} \Omega_n$ .

As  $m_g^i$  defined in (8), the similarity between  $V_{pc}^i$  and  $V_{pr}$ , marked as  $m_p^i$ , is defined in the same form by replacing the gap variables with the corresponding peak ones. If no peak-involved pair exists in EDR intersection, we mark  $m_p^i = -1$ .

### D. Similarity Maps Fusion for Composition Detection

Before fusion, the blockwise similarities  $m_g^i = -1$  or  $m_p^i = -1$  are updated by averaging the available neighboring effective measurements. The resulting similarity for the  $i$ -th unlabeled block, denoted by  $m^i$ , can be generated by fusing the peak/gap based similarities. When  $m_g^i \neq -1$  and  $m_p^i \neq -1$ , we compute

$$m^i = \left( m_g^i + m_p^i \right) / 2. \quad (10)$$

When  $m_g^i = -1$  or  $m_p^i = -1$ ,  $m^i = \max(m_g^i, m_p^i)$ . Note that  $m^i = -1$  occurs scarcely in both-source-enhanced composite images.

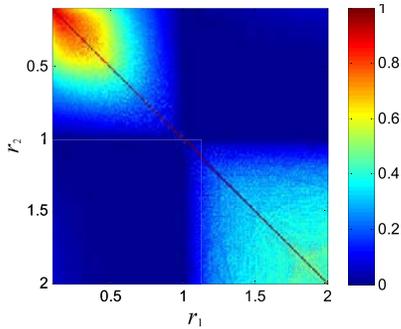


Fig. 9. Distribution of  $m^i$  values computed under two gamma mappings with the parameter  $r_1$  and  $r_2$ .

If all the blocks in an unsaturated region own  $m^i = -1$ , while the blocks out of the region have  $m^i \geq t$ , the test image can be identified as a single-source-enhanced composite image. Otherwise, each block is classified as: 1) if  $m^i \geq t$ , contrast enhancement mapping applied to  $S_1$  is detected; 2) if  $m^i < t$ , a different contrast enhancement mapping applied to the other source region is detected. Hence, the both-source-enhanced composite image is detected if two different mappings are detected in two complementary regions, respectively. The threshold  $t$  is experimentally set as 0.2.

Similarity between the peak/gap positions determined theoretically by two mappings can be measured by  $m^i$ . Two different mappings could be distinguished theoretically when average of the ratios between the number of shared peak/gap bins and that of total peak/gap-involved pairs keeps smaller than  $t$ . As a case study, Fig. 9 shows  $m^i$  distribution under two gamma mappings with gamma values  $r_1, r_2$ . EDR is set as  $[0, 255]$ . There exists the approximate distribution: (i)  $m^i \equiv 1$ , if  $r_1 = r_2$ ; (ii)  $m^i > 0.2$ , if  $r_1, r_2 < 0.7$  or  $r_1, r_2 > 1.2, r_1 \neq r_2$ ; (iii)  $m^i < 0.2$ , otherwise. In fact, each scenario might appear in applications, since the contrast and brightness of source images are diverse. The scenario (iii) covers the largest working range. The above analyses show that our method can work efficiently in such a scenario, but may fail in other scenarios.

## V. EXPERIMENTAL RESULTS AND DISCUSSION

To verify the efficacy of our proposed forensic methods, extensive experiments are performed in three test image sets.

### A. Test Data

The test data covers the following three image sets:

*Dataset 1*: 800 unaltered photograph images in raw format, where 700 images are randomly selected from BOSS public dataset [26], captured by 7 cameras and 100 images for each. The other 100 images are captured by Cannon 5D II. The size ranges from  $2000 \times 3008$  to  $5212 \times 3468$  pixels.

*Dataset 2*: Uncompressed Colour Image Database (UCID) [25]. It has 1338 uncompressed images on various topics such as natural scenes, man-made objects, indoors and outdoors.

*Dataset 3*: 1100 natural photographs, captured by several different cameras, stored in JPEG format and with size from  $1200 \times 900$  to  $2832 \times 2128$  pixels. All camera settings

including exposure and white balance are set automatic for resembling what probably common users would do.

As the tests in prior works [10], the green color channel from each test image is used as a set of unaltered grayscale images. The enhanced versions are generated via the widely used contrast enhancement operations including gamma correction  $m(x) = \text{round}(255(x/255)^r)$ ,  $r = 0.5, 0.7, 0.9, 1.1, 1.5, 2.0$ , and  $s$  mapping  $m(x) = \text{round}(255(\arcsin(2x/255 - 1)/\pi + 1/2))$ .

### B. Global Contrast Enhancement Detection

For testing the performance of our proposed contrast enhancement detection technique, each test image is classified by determining if it is contrast-enhanced or not using a series of decision thresholds. The probabilities of detection ( $P_d$ ) and false alarm ( $P_{fa}$ ) determined by thresholds are calculated as the percentage of the enhanced images correctly classified and that of the unenhanced images incorrectly classified, respectively. The receiver operating characteristic (ROC) curves are generated for evaluation.

Contrast enhancement detection ROC curves on *Dataset 1* are shown in Fig. 10. It indicates that our proposed algorithm achieves high detection rates even under low  $P_{fa}$ .  $P_d$  attains 100% when  $P_{fa} = 1\%$ . Note that the same high  $P_d$  is always gained when the raw images are JPEG-compressed with different Qs. Gamma correction with various  $r$  values and  $s$  mapping are detected accurately. Such results demonstrate the effectiveness of our method both in uncompressed and compressed images. Corresponding test results of Stamm and Liu method [10] are shown in Fig. 10(b)–(e), which indicate that the detection rate decreases with the reduction of Q and becomes random guess at  $Q = 50$ . In contrast, our method gains high detection rates in the case of low quality compression ( $Q = 70, 50$ ), while keeps comparative performance in the other cases.

To assess the performance on extensive datasets, the same experiments are redone on *Dataset 2* and *Dataset 3*. From the results shown in Figs. 11 and 12, we can observe that comparative detection performance is still obtained. Such results on various types of test images further validate the reliability and efficacy of our proposed technique.

The ratio of samples misclassified by our method is listed in Table I. Here, the misclassified images refer to the unenhanced ones with  $N_g > 0$  and the enhanced ones with  $N_g = 0$ . Results show that except the case of  $Q = 50$  on *Dataset 2*, the number of misclassified unenhanced images is always zero. The ratio of misclassified samples is too small to affect the overall performance. Such misclassification should attribute to the variation of histograms. For example, the histogram of an enhanced image may be so narrow that zero-height gap bins are not yielded, while that of an unenhanced image may be unsmooth.

### C. Source-Enhanced Composition Detection

To evaluate the proposed composition detection scheme, we first test the example forged image shown in Fig. 2(c). In Fig. 2(d),  $F$  values of the blocks outside ‘sky’ region are consistent and smaller than those within such a region. Through a simple thresholding, the ‘sky’ region would be

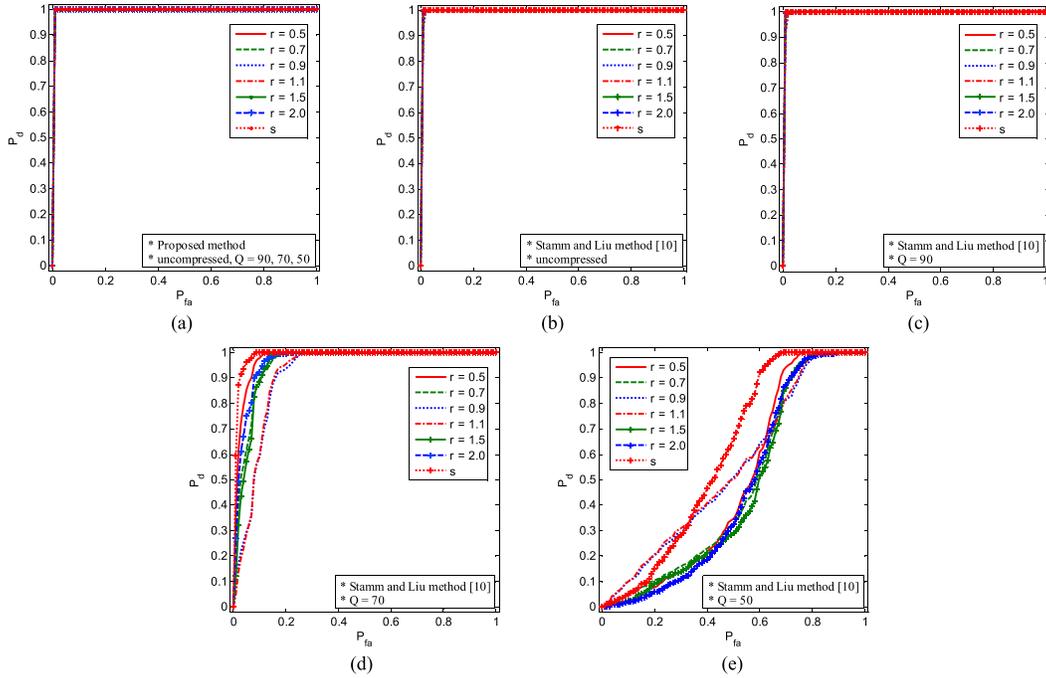


Fig. 10. Contrast enhancement detection ROC curves on *Dataset 1*. Our proposed method: (a) curves in uncompressed and compressed ( $Q = 90, 70$  and  $50$ ) scenarios are the same. Stamm and Liu method: (b) uncompressed, (c)  $Q = 90$ , (d)  $Q = 70$ , (e)  $Q = 50$ . Here, ‘ $r$ ’ denotes gamma value and ‘ $s$ ’ denotes  $s$  mapping.

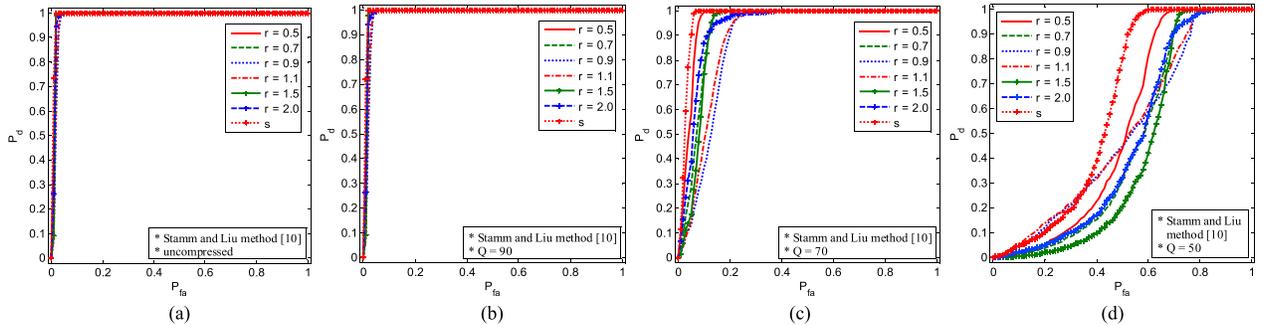


Fig. 11. Contrast enhancement detection ROC curves on *Dataset 2*. Our proposed method: curves are the same as those in Fig. 10(a). Stamm and Liu method: (a) uncompressed, (b)  $Q = 90$ , (c)  $Q = 70$ , (d)  $Q = 50$ .

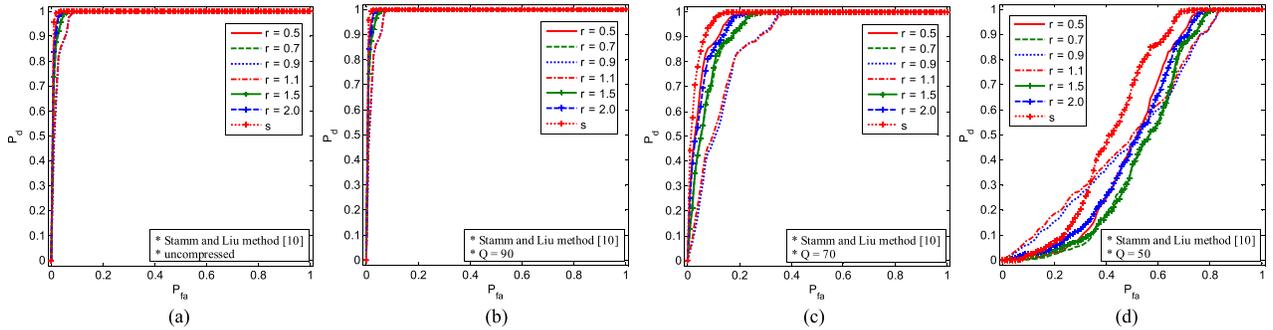


Fig. 12. Contrast enhancement detection ROC curves on *Dataset 3*. Our proposed method: curves are the same as those in Fig. 10(a). Stamm and Liu method: (a) uncompressed, (b)  $Q = 90$ , (c)  $Q = 70$ , (d)  $Q = 50$ .

wrongly detected as tampered by Stamm and Liu method [10]. In Fig. 2(e), the result of our method is shown as a blockwise  $m^i$  map. If  $m^i$  approaches 1, it means that the block is more possible to be enforced by a certain mapping. While  $m^i$  approaches 0, the block is more likely processed by a different

mapping. It reveals that our method could detect both-source-enhanced composition and locate tampered regions accurately.

We also perform quantitative tests to evaluate the proposed algorithm. For each image in *Dataset 3*, contrast enhancement  $r_1$  is performed in a rectangular region which owns half size of

TABLE I

PROPORTION OF MISCLASSIFIED SAMPLES IN EVALUATING CONTRAST ENHANCEMENT DETECTION ALGORITHM. THE NUMBERS IN PERCENTAGE PORTIONED BY ‘\’ ARE GAINED FROM Dataset 1, 2, AND 3 RESPECTIVELY. ‘Q = 100’ DENOTES UNCOMPRESSED

Q	Unenhanced	Enhanced						
		r						s
		0.5	0.7	0.9	1.1	1.5	2	
100	0\0\0	0.9\0.1\0.9	0.9\0.1\0.9	0.4\0.1\0.6	1.6\0.1\0	1.0\0.1\0	1.6\0.1\0	3.5\0.1\1.9
90	0\0\0	1.1\0.5\0.9	0.8\0.4\0.9	0.4\0.3\0.6	1.5\0\0	1.0\0\0	1.5\0\0	3.4\1.2\2.0
70	0\0\0	0.9\0.5\0.9	0.9\0.4\0.9	0.4\0.3\0.6	1.5\0\0	1.1\0\0	1.5\0\0	3.6\1.2\1.9
50	0\0.2\0	0.9\0.4\0.9	0.9\0.4\0.9	0.6\0.4\0.6	1.9\0\0	1.4\0\0	1.9\0\0	4.8\2.8\1.9

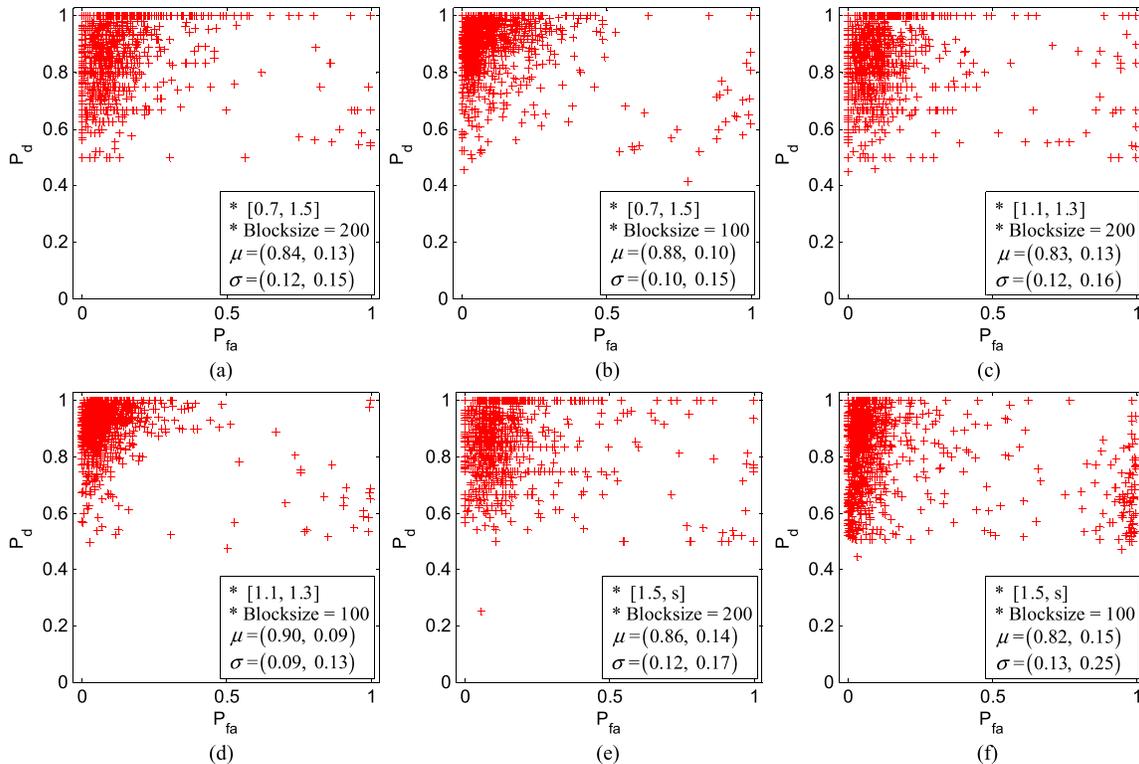


Fig. 13. Accuracy for locating tampered regions in each composite image. Different mapping pairs  $[r_1, r_2]$  and block sizes are tested. (a)  $[0.7, 1.5]$ , 200; (b)  $[0.7, 1.5]$ , 100; (c)  $[1.1, 1.3]$ , 200; (d)  $[1.1, 1.3]$ , 100; (e)  $[1.5, s]$ , 200; (f)  $[1.5, s]$ , 100.  $\mu, \sigma$  denote the mean and standard deviation of all  $(P_d, P_{fa})$  values, respectively.

TABLE III

RESULTS FOR EVALUATING COMPOSITION DETECTION ALGORITHM UNDER PRECOMPRESSION.  $Q_1, Q_2$  ARE QUALITY FACTORS OF THE JPEG COMPRESSIONS APPLIED TO BACKGROUND AND CENTRAL REGIONS, RESPECTIVELY

$(Q_1, Q_2)$	(100, 100)	(90, 90)	(70, 70)	(50, 50)	(30, 30)	(70, 50)	(70, 30)	(50, 30)	(60, 40)
$\mu (P_d)$	0.80	0.80	0.80	0.79	0.79	0.79	0.79	0.79	0.79
$\mu (P_{fa})$	0.23	0.22	0.22	0.27	0.23	0.23	0.21	0.26	0.25
$\sigma (P_d)$	0.15	0.16	0.15	0.17	0.15	0.16	0.15	0.17	0.17
$\sigma (P_{fa})$	0.29	0.29	0.28	0.32	0.29	0.29	0.28	0.30	0.31

the global image and random position. Contrast enhancement  $r_2$  is enforced in the other region to simulate the both-source-enhanced composition. To measure the accuracy for locating tampered regions, the proportion of blocks classified correctly in the selected rectangular region is defined as  $P_d$ , while the proportion of blocks classified incorrectly in the other region is measured as  $P_{fa}$ . As shown in Fig. 13, the mean of  $P_d$  values achieves above 0.83 in the cases of  $[r_1, r_2] = [0.7, 1.5]$  and

$[1.1, 1.3]$ , while the mean of  $P_{fa}$  values is smaller than 0.13. It should be mentioned that the detection for  $[1.1, 1.3]$  is rather challengeable since such two mappings share more peak/gap positions. However, the mean of  $P_d$  values in such a case still achieves 0.83 while keeping  $P_{fa} = 0.13$  for Blocksize = 200.

However, the overall performance decreases to some extent when  $s$  mapping is involved. The mean of  $P_d$  values keeps at the same level, while both mean and standard deviation of

TABLE II

DETECTION PERFORMANCE FOR DIFFERENT SIMILARITY MEASURES

Similarity Map	Gap-based	Peak-based	Fused
$\mu(P_d)$	0.79	0.77	0.88
$\mu(P_{fa})$	0.10	0.24	0.10
$\sigma(P_d)$	0.18	0.19	0.13
$\sigma(P_{fa})$	0.16	0.30	0.18

$P_{fa}$  values increase to some extent. The reason behind lies in the overlap of the peak/gap positions incurred by  $s$  mapping and gamma correction. Note that the performance is relatively better for Blocksize = 100. Such superiority should attribute to the increment of available blocks, while the histogram of unaltered blocks still keeps smooth and covers adequate gray levels.

To evaluate the performance for individual similarity measures, the same tests are repeated by exploiting the gap-based ( $m_g^i$ ), peak-based ( $m_p^i$ ) and fused similarity map ( $m^i$ ), respectively. The settings  $[r_1, r_2] = [1.1, 1.3]$ , Blocksize = 100 and the half-downsized images in *Dataset 3* are adopted illustratively. As shown in Table II, the performance for the fused similarity map is higher than those for individual maps. The gap-based result is slightly better than the peak-based result, since the zero-height gap information should be comparatively reliable.

Experiments are also conducted for evaluating the composition detection method when the source images have been compressed before enhancement. The same tests are done in the quarter-downsized images in *Dataset 1*. Such setting simulates the middle resolution widely appeared in practical applications such as the Internet.  $Q_1, Q_2$  denote the quality factors of JPEG compressions applied to the background and central tampered regions, respectively. Without loss of generality,  $[r_1, r_2] = [1.1, 1.3]$ , Blocksize = 100 are tested. As shown in Table III, the results verify that the detection performance keeps unchanged when source regions are pre-compressed at various  $Q_s$ .

#### D. Discussion on Performance Limits

It should be pointed out that the proposed forensic algorithms are not robust against post-processing which affects the accurate detection of zero-height gap bins. Tests show that the detection rates become random guess even under mild post-processing such as bicubic upscaling ( $\times 1.1$ ), unit variance Gaussian additive noise and JPEG ( $Q = 90$ ). As such the proposed methods are suitable to work in the scenario that contrast enhancement is the last step of manipulation applied to images.

Note that anti-forensic techniques [19], [20] adversely conceal abnormal histogram traces, and make the histogram of an enhanced image behave as an unaltered one. Our proposed forensic techniques could not resist against such targeted anti-forensic attacks. As expected from any other forensic technique, the proposed methods have their own limitations. However, if a large set of forensic methods are developed and enforced cooperatively, it would be difficult for a malicious

user to create a forged image which can invalidate all forensic methods [10].

Histogram modifications techniques, such as histogram equalization and the intelligent approaches [21]–[23], are the most widespread indirect contrast enhancement techniques. In such techniques, a target pixel value mapping is generally pre-computed and then applied to the global image for adjusting contrast automatically. The zero-height gap artifacts would still be yielded so that such contrast enhancement can be detected.

## VI. CONCLUSION

In this paper, we proposed two contrast enhancement based forensic algorithms via histogram peak/gap artifacts analysis. First, we extended to detect the global contrast enhancement in both uncompressed and previously JPEG-compressed images. The zero-height gap bins in gray level histograms were novelly exploited as identifying features. Large-scale experiments showed that our contrast enhancement detector achieved high performance, i.e.,  $P_d = 100\%$  at  $P_{fa} = 1\%$ . Second, we proposed a new method to discover the both-source-enhanced composite image, which invalidated the previous detection methods. The composition boundary was accurately located by detecting the inconsistency between detected blockwise peak/gap positional distributions. The tests on both a specific composite image and quantitative synthetic samples verified the efficacy of our proposed composition detector.

The proposed contrast enhancement based forensic methods could work particularly well when contrast enhancement is performed as the last step of manipulation. In the future work, we would try to improve the robustness of such methods against postprocessing, such as JPEG compression. It is also essential to enhance the security on countering the existing and potential anti-forensic techniques.

## ACKNOWLEDGMENT

The authors would like to thank Prof. Alex C. Kot for his valuable discussions and suggestions on this work. They also thank the anonymous reviewers for their comments and suggestions that greatly improved the manuscript.

## REFERENCES

- [1] H. Farid, "Image forgery detection," *IEEE Signal Process. Mag.*, vol. 26, no. 2, pp. 16–25, Mar. 2009.
- [2] B. Mahdian and S. Saic, "A bibliography on blind methods for identifying image forgery," *Image Commun.*, vol. 25, no. 6, pp. 389–399, Jul. 2010.
- [3] S. Bayram, I. Avcubas, B. Sankur, and N. Memon, "Image manipulation detection," *J. Electron. Imag.*, vol. 15, no. 4, pp. 04110201–04110217, 2006.
- [4] 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.
- [5] H. Cao and A. C. Kot, "Manipulation detection on image patches using FusionBoost," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 3, pp. 992–1002, Jun. 2012.
- [6] J. Fan, H. Cao, and A. C. Kot, "Estimating EXIF parameters based on noise features for image manipulation detection," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 4, pp. 608–618, Apr. 2013.
- [7] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of resampling," *IEEE Trans. Signal Process.*, vol. 53, no. 2, pp. 758–767, Feb. 2005.

- [8] T. Bianchi and A. Piva, "Detection of non-aligned double JPEG compression based on integer periodicity maps," *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, pp. 842–848, Apr. 2012.
- [9] G. Cao, Y. Zhao, R. Ni, and A. C. Kot, "Unsharp masking sharpening detection via overshoot artifacts analysis," *IEEE Signal Process. Lett.*, vol. 18, no. 10, pp. 603–606, Oct. 2011.
- [10] M. C. Stamm and K. J. R. Liu, "Forensic detection of image manipulation using statistical intrinsic fingerprints," *IEEE Trans. Inf. Forensics Security*, vol. 5, no. 3, pp. 492–506, Sep. 2010.
- [11] M. C. Stamm and K. J. R. Liu, "Forensic estimation and reconstruction of a contrast enhancement mapping," in *Proc. IEEE Int. Conf. Acoust., Speech Signal*, Dallas, TX, USA, Mar. 2010, pp. 1698–1701.
- [12] G. Cao, Y. Zhao, and R. Ni, "Forensic estimation of gamma correction in digital images," in *Proc. 17th IEEE Int. Conf. Image Process.*, Hong Kong, 2010, pp. 2097–2100.
- [13] P. Ferrara, T. Bianchi, A. De Rosaz, and A. Piva, "Reverse engineering of double compressed images in the presence of contrast enhancement," in *Proc. IEEE Workshop Multimedia Signal Process.*, Pula, Croatia, Sep./Oct. 2013, pp. 141–146.
- [14] M. Kirchner and J. Fridrich, "On detection of median filtering in digital images," in *Proc. SPIE, Electronic Imaging, Media Forensics and Security II*, vol. 7541, San Jose, CA, USA, Jan. 2010, pp. 1–12.
- [15] C. Chen, J. Ni, and J. Huang, "Blind detection of median filtering in digital images: A difference domain based approach," *IEEE Trans. Image Process.*, vol. 22, no. 12, pp. 4699–4710, Dec. 2013.
- [16] D. Mahajan, R. Ramamoorthi, and B. Curless, "A theory of frequency domain invariants: Spherical harmonic identities for BRDF/lighting transfer and image consistency," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 197–213, Feb. 2008.
- [17] I. Yerushalmy and H. Hel-Or, "Digital image forgery detection based on lens and sensor aberration," *Int. J. Comput. Vis.*, vol. 92, no. 1, pp. 71–91, 2011.
- [18] J. O'Brien and H. Farid, "Exposing photo manipulation with inconsistent reflections," *ACM Trans. Graph.*, vol. 31, no. 1, pp. 1–11, 2012.
- [19] G. Cao, Y. Zhao, R. Ni, and H. Tian, "Anti-forensics of contrast enhancement in digital images," in *Proc. ACM Multimedia Security Workshop*, Rome, Italy, 2010, pp. 25–34.
- [20] M. Barni, M. Fontani, and B. Tondi, "A universal technique to hide traces of histogram-based image manipulations," in *Proc. ACM Multimedia Security Workshop*, Coventry, England, 2012, pp. 97–104.
- [21] S.-C. Huang, F.-C. Cheng, and Y.-S. Chiu, "Efficient contrast enhancement using adaptive gamma correction with weighting distribution," *IEEE Trans. Image Process.*, vol. 22, no. 3, pp. 1032–1041, Mar. 2013.
- [22] T. Arici, S. Dikbas, and Y. Altunbasak, "A histogram modification framework and its application for image contrast enhancement," *IEEE Trans. Image Process.*, vol. 18, no. 9, pp. 1921–1935, Sep. 2009.
- [23] S. Battiato, A. Bosco, A. Castorina, and G. Messina, "Automatic image enhancement by content dependent exposure correction," *EURASIP J. Appl. Signal Process.*, vol. 12, pp. 1849–1860, Jan. 2004.
- [24] W. B. Pennebaker and J. L. Mitchell, *JPEG Still Image Data Compression Standard*. New York, NY, USA: Van Nostrand, 1993.
- [25] G. Schaefer and M. Stich, "UCID—An uncompressed colour image database," in *Proc. SPIE, Storage and Retrieval Methods and Applications for Multimedia*, vol. 5307, San Jose, CA, USA, 2004, pp. 472–480.
- [26] P. Bas, T. Filler, and T. Pevny, "Break our steganographic system—The ins and outs of organizing BOSS," in *Proc. Inf. Hiding*, Prague, Czech Republic, 2011, pp. 59–70.



**Gang Cao** received the B.S. degree from Wuhan University of Technology, Hubei, China, in 2005, and the Ph.D. degree from the Institute of Information Science, Beijing Jiaotong University, Beijing, China, in 2013. He was with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, in 2010. He is currently a faculty member of the School of Computer Science, Communication University of China, Beijing, China. His current research interests include digital forensics, data hiding, and multimedia signal processing.



**Yao Zhao** (M'06–SM'12) received the B.S. degree from Fuzhou University, Fuzhou, China, in 1989, and the M.E. degree from Southeast University, Nanjing, China, in 1992, both from the Radio Engineering Department, and the Ph.D. degree from the Institute of Information Science, Beijing Jiaotong University (BJTU), Beijing, China, in 1996. He became an Associate Professor at BJTU in 1998 and became a Professor in 2001. From 2001 to 2002, he was a Senior Research Fellow with the Information and Communication Theory Group, Faculty of Information Technology and Systems, Delft University of Technology, Delft, The Netherlands. He is currently the Director of the Institute of Information Science, BJTU. His current research interests include image/video coding, digital watermarking and forensics, and video analysis and understanding. He is currently leading several national research projects from the 973 Program, 863 Program, and the National Science Foundation of China. He serves on the editorial boards of several international journals, including as an Associate Editor of the *IEEE Signal Processing Letters*, Area Editor of *Signal Processing: Image Communication* (Elsevier), and Associate Editor of *Circuits, System, and Signal Processing* (Springer). He was named a Distinguished Young Scholar by the National Science Foundation of China in 2010.



**Rongrong Ni** received the Ph.D. degree in signal and information processing from the Institute of Information Science, Beijing Jiaotong University, Beijing, China, in 2005. Since 2005, she has been the faculty of the School of Computer and Information Technology and the Institute of Information Science, where she was an Associate Professor from 2008 to 2013, and was promoted to a Professor in 2013. Her current research interests include image processing, data hiding and digital forensics, pattern recognition, and computer vision. She was selected in the Beijing Science and Technology Stars Projects in 2008, and was awarded the Jeme Tien Yow Special Prize in Science and Technology in 2009. She is currently the Principal Investigator of three projects funded by the Natural Science Foundation of China. She has participated in the 973 Program, the 863 Program, and international projects. She has published more than 80 papers in academic journals and conferences, and holds six national patents.

**Xuelong Li** (M'02–SM'07–F'12) is a Full Professor with the Center for OPTical Imagery Analysis and Learning (OPTIMAL), State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China.