

# Exposing image resampling forgery by using linear parametric model

Tong Qiao<sup>1,3</sup>  · Aichun Zhu<sup>2,3</sup> · Florent Reiraint<sup>3</sup>

Received: 25 May 2016 / Revised: 6 December 2016 / Accepted: 26 December 2016 /  
Published online: 14 January 2017  
© Springer Science+Business Media New York 2017

**Abstract** Resampling forgery generally refers to as the technique that utilizes interpolation algorithm to maliciously geometrically transform a digital image or a portion of an image. This paper investigates the problem of image resampling detection based on the linear parametric model. First, we expose the periodic artifact of one-dimensional 1-D) resampled signal. After dealing with the nuisance parameters, together with Bayes' rule, the detector is designed based on the probability of residual noise extracted from resampled signal using linear parametric model. Subsequently, we mainly study the characteristic of a resampled image. Meanwhile, it is proposed to estimate the probability of pixels' noise and establish a practical Likelihood Ratio Test (LRT). Comparison with the state-of-the-art tests, numerical experiments show the relevance of our proposed algorithm with detecting uncompressed/compressed resampled images.

**Keywords** Image resampling forensics · Linear parametric model · Bayes' rule · Hypothesis testing

---

✉ Tong Qiao  
tong.qiao@hdu.edu.cn

Aichun Zhu  
aichun.zhu@utt.fr

Florent Reiraint  
florent.reiraint@utt.fr

<sup>1</sup> School of Cyberspace, Hangzhou Dianzi University, Hangzhou, China

<sup>2</sup> School of Computer Science and Technology, Nanjing Tech University, Nanjing, China

<sup>3</sup> LM2S, University of Technology of Troyes, Troyes, France

# 1 Introduction and contributions

With the development of digital image industry, people can utilize benefits of digital images which help them modify a photo easily and feasibly. However, some image forgers maliciously falsify digital images. Thus, it is urgent to propose a digital forensic tool against malicious image falsification and restore the trustworthiness to digital images. Digital image forensics can be classified into two categories. One is defined as active forensics, referring to as digital watermarking. Since the embedding mechanism has to be available and the credibility of information embedded in the image remains questionable, the widespread utilization of digital watermarking has been substantially limited. The other is defined as passive forensics. Different from digital image watermarking, widely-adopted passive forensics authenticates digital images without any prior information on which forensic investigators mainly focus (see [8, 27, 31, 32]).

Many methodologies have been proposed for detecting image forgery. Prior to discussion of image forensics, it should be noted that in this context, we do not categorize steganalysis including LSB replacement and matching (see [35]) into the field of image forgery detection. Using sensor pattern noise as camera signature, it was proposed to identify the origin of the acquired image (see [3, 18]). Based on the physical noise model proposed in [33], the designed detector first identified source camera model by using RAW format images. Moreover, inspired by this pioneer work, the authors of [23] proposed the improved noise model for identifying individual source camera device. By investigating the artifacts of the Color Filter Array (CFA) pattern, authors of [12, 22] distinguished a photographic image (PIM) from a computer generated (CG) one. In the practical operation of image forgery, a manipulated image is probably compressed twice. Thus, JPEG double-compressed detector is also an effective auxiliary tool to authenticate tampered images (see [7]). To distort the truth, a real object in the digital photo is possibly concealed by forgers. Thus, it is necessary to propose a forensic tool to detect copy-move or splicing manipulation (see [1, 4, 11, 17]). Besides, image hashing technique can also help forensic investigators authenticate manipulated images (see [28]).

## 1.1 State of the art

To create convincing forged images, a forger needs to use some geometric operations which require a resampling of pixels. Resampling is a post-camera processing technique. In general, it is proposed not to classify resampling among the image tempering technique. However, resampling of pixels is the basis of many forgeries because it is involved in almost all geometric modification (rotation, resizing, etc. ...). While the evidence of resampling operation within an image does not necessarily imply that the image has been tempered it is a very important clue. Thus, resampling detectors play a very important role as an auxiliary forensic tool. Additionally, resampling falsification often happens together with other forgery techniques, for instance, copy-move or splicing forgery. For instance, in order to enhance the visual realism, splicing and copy-move forgeries often require scaling and/or rotation operations which requires resampling operation.

To distinguish a resampled and non-resampled image is the subfield of digital image forensics on which this paper focuses. Forensic methods for exploiting the artifacts of resampled images can generally be divided into two following fundamental categories:

1. Methods in the first category focus on identifying the linear correlation of resampled images in the spatial domain. Authors of [21] first proposed a linear predictor to

expose the correlation existing among neighboring pixels of the resampled images by using Expectation Maximization (EM) algorithm (see [6]). Although EM algorithm was capable of estimating the linear interpolated parameters, iterative computation costs a plenty of time. Driven by this pioneer work, it was proposed to analyze the relevance between the EM detector and the second-derivative one, and design an equivalent accelerated and simplified detector in [14]. However, in place of estimating the interpolation parameters, the fixed parametric predictor indeed improves the efficiency of detection, yet the detection accuracy can not be guaranteed. The problem of how to strike the balance between the computation efficiency and estimation accuracy remains open. Instead of using a single predictor, authors of [15] proposed to use a global predictor to detect a resized image. Its limitation is that the proposed algorithm can only be used for detecting resized images excluding other geometric manipulation such rotation or shearing. In addition, these methods can detect the linear transformation, yet can not estimate the interpolation period.

2. Methods in the second category investigate the second or high-order derivative signal of resampled images in the frequency domain. By exploiting the second-derivative algorithm, it was proposed to uncover the periodicity of an interpolated image in [13]. This method can estimate the interpolation period, yet only applied in the resized images. By analyzing the variance of the  $n$ -th order derivative, it was proposed to expose the trace of a resampled image in [19]. However, the linear transformation matrix can not be approximated. Inspired by the algorithm of [13], it was proposed to analyze the relationship between the rotation angle and the frequency in the spectrum of the image's edge map (see [34]). The limitation of rotation forgery detection is that this method could not be comprehensively used. By using supervised learning method (such as Support Vector Machine, SVM), it was proposed to detect whether the inspected image was resampled (see [9, 10]). Although those methods could achieve high detection accuracy, supervised statistical learning was time-consuming. In addition, several problems such as the robustness to training and testing set mismatch remain open. For approximating the linear transformation of an interpolated image, authors of [29] proposed to analyze the second-derivative signals of the image in both the row and column directions. Although most of these methods could estimate the linear transformation matrix and interpolation period, they did not establish the detection performance statistically.

## 1.2 Contributions of the paper

First, the proposed approach detects a one-dimensional (1-D) resampled signal and estimates the interpolated period. Second, by uncovering the linear correlation of pixels existing in an image in the spatial domain, the proposed algorithm authenticates a two-dimensional (2-D) resampled image. Third, by designing a practical Likelihood Ratio Test (LRT), the proposed algorithm detects a batch of resampled images, and warrants a prescribed False Alarm Rate (FAR), and achieves good detection performance. Hence, the main contributions are the followings:

- To deal with 1-D resampled signal, it is first proposed to extract the residual noise of the inspected signal using a linear parametric model. Meanwhile, we design the detector based on the probability of residual noise (or prediction errors).
- By transforming the probability of the residual noise into Fourier spectrum, it is proposed to establish an estimator for estimating the resampling factor.

- Based on another linear parametric model, the probability of residual noise extracted from the inspected image is estimated. Then, we expose the periodic artifact of the resampled image which can be described as symmetrical bright spots scattering in Fourier domain.
- To deal with a large scale of resampled images, based on the estimated probability of residual noise, we propose to design a practical LRT and detect resampled compressed or uncompressed images.

In this context, it should be noted that we can provide a new approach of detecting resampled images based on our proposed statistical models. However, we are incapable of designing an most optimal power test with known statistical properties, which is the limitation of the proposed algorithm.

### 1.3 Organization of the paper

This paper is organized as follows. Section 2 first recalls the principle of 1-D signal resampling. Then, the linear parametric model is proposed for dealing with the nuisance parameters. Meanwhile, the 1-D resampling detector is introduced. In Section 3, the linear correlation of a 2-D resampled image is first illustrated. At the same time, it is proposed to design a test of detecting a single resampled image. Next, Section 4 presents a practical LRT for detecting a batch of resampled images. Finally, Section 5 presents numerical results of the proposed detector on real images. Section 6 concludes this paper.

## 2 One-dimensional signal resampling and its artifact

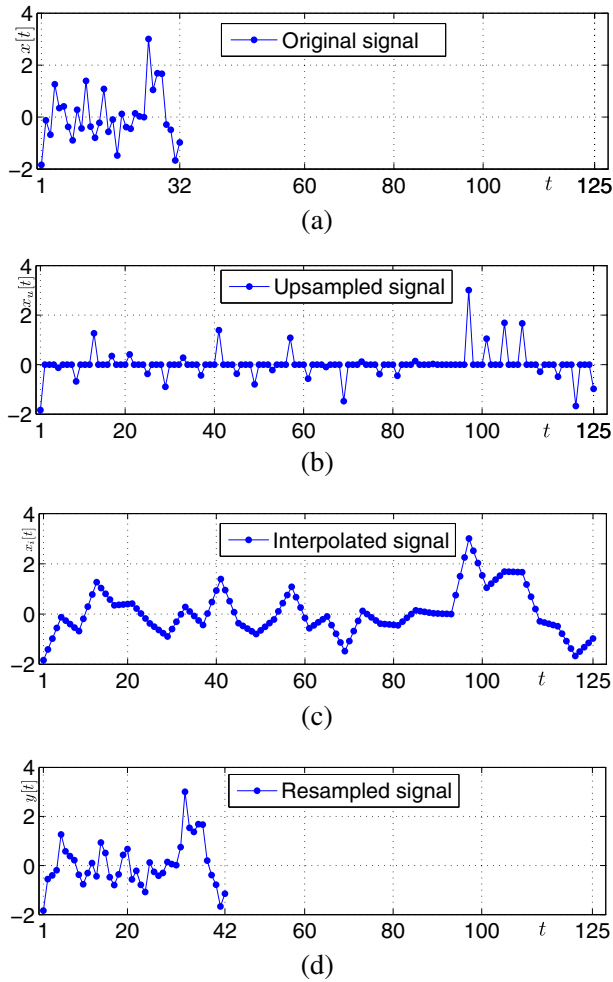
### 2.1 Problem statement of one-dimensional signal resampling

In general, resampling can be proceeded by three steps: up-sampling; interpolation; down-sampling (see details in [20]). Since linear interpolation is widely-adopted, for simplicity and clarity, it is assumed that the interpolation algorithm studied in the present paper is linear. In the following paragraph, a simple illustration of the resampling process is displayed (see Fig. 1).

Without loss of generality, a 1-D discrete signal is created (see Fig. 1a), denoted the vector  $\mathbf{x} = \{x[t]\} \ t \in \{1, \dots, T\}$ , where  $T$  denotes the length of the vector. First, by up-sampling with the factor  $p$ , a new up-sampled vector  $\mathbf{x}_u = \{x_u[t]\}$  is generated with the length  $p(T - 1) + 1$  (see Fig. 1b). In the case of  $t = 1, 2, \dots, T$ ,  $x_u[p(t - 1) + 1] = x[t]$ ; otherwise,  $x_u[t] = 0$ . Second, by convolving the vector  $\mathbf{x}_u$  with a linear low-pass filter  $h[t]$ , the interpolated vector  $\mathbf{x}_i = \{x_i[t]\}$ ,  $t \in \{1, \dots, p(T - 1) + 1\}$  is obtained, where  $x_i[t] = x_u[t] * h[t]$  (see Fig. 1c). Third, by down-sampling with the factor  $q$ , a new down-sampled vector  $\mathbf{x}_d = \{x_d[t]\}$  with the length  $\lfloor \frac{p}{q}(T - 1) + 1 \rfloor$ , where  $x_d[t] = x_i[q(t - 1) + 1]$ ,  $t = 1, 2, \dots, \lfloor \frac{p}{q}(T - 1) + 1 \rfloor$  (see Fig. 1d). Let us denote  $\mathbf{y} = \{y[t]\}$  resampled signal where  $\mathbf{y} \equiv \mathbf{x}_d$ . It should be noted that  $p, q \in \mathbb{N}^+$ . Then, it is immediately proposed that the resampled signal  $\mathbf{y}$  is described by the following linear equation:

$$\mathbf{y} = \mathbf{B}_{\frac{p}{q}} \mathbf{x} \quad (1)$$

where  $\mathbf{y} = \{y_m\}$ ,  $m \in \{1, \dots, M\}$ , denoted a resampled signal,  $\mathbf{x} = \{x_n\}$ ,  $n \in \{1, \dots, N\}$ , denoted an original signal. A  $M \times N$  linear resampling matrix  $\mathbf{B}_{\frac{p}{q}}$  controls the



**Fig. 1** Illustration of resampling 1-D signal by a factor  $\frac{p}{q} = \frac{4}{3}$ : (a) 1-D original signal, (b) up-sampled signal by a factor  $p = 4$ , (c) interpolated signal, (d) resampled signal after down-sampling by using a factor  $q = 3$

resampling procedure, including up-sampling, interpolation, and down-sampling. For instance, the linear resampling matrix  $\mathbf{B}_{\frac{4}{3}}$  (see Fig. 1) takes the form:

$$\mathbf{B}_{\frac{4}{3}} = \begin{pmatrix} 1 & 0 & 0 & 0 & \dots \\ 0.25 & 0.75 & 0 & 0 & \dots \\ 0 & 0.5 & 0.5 & 0 & \dots \\ 0 & 0 & 0.75 & 0.25 & \dots \\ 0 & 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

where  $\mathbf{B}_{\frac{4}{3}}$  has the period 4. Since the linear resampling matrix  $\mathbf{B}_{\frac{p}{q}}$  introduces the periodicity, the resampled signal  $\mathbf{y}$  owns the period  $p$ . Moreover, the expectation of the periodic signal  $y_m$  should be linear combination from the neighbouring samples, which is defined by:

$$\hat{y}_m = \sum_{k=-L}^L \alpha_k y_{m+k} \tag{2}$$

where  $\alpha_k$  denotes the weighted factor with  $\alpha_0 = 1$  and  $2L$  ( $L \in \mathbb{N}^+$ ) the number of used neighbouring samples. In the present paper the moments of (2) is not in detail, the interested reader is referred to [21]. It should be noted that the linear combination (2) can not hold true in each sample along the signal  $y_m$ , but only be defined in the interval of  $p$ . In the practical resampling, if an up-sampling factor  $p$  set as 1 while a down-sampling factor  $q > 1$ , the resampled signal does not expose the periodicity, and then its expectation can not be formulated by linear combination (2).

It is proposed to use a linear parametric model to deal with a 1-D resampled signal  $\mathbf{y}$ . To this end, the vector  $\mathbf{y} = \{y_1, \dots, y_M\}^T$  is defined by:

$$\mathbf{y} \sim \mathcal{N}(\boldsymbol{\mu}, \sigma^2 \mathbf{I}_M) = \boldsymbol{\mu} + \boldsymbol{\epsilon} \tag{3}$$

where  $\boldsymbol{\mu} = (\mu_1, \dots, \mu_M)^T$  denotes the expectation of  $\mathbf{y}$  and  $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_M)^T$  denotes residual noise following the Gaussian model with the variance  $\sigma^2 \mathbf{I}_M$  and  $\mathbf{I}_M$  the identity matrix of size  $M \times M$ . Furthermore,  $\boldsymbol{\mu}$  can be described with the following linear parametric model:

$$\boldsymbol{\mu} = \mathbf{A}\boldsymbol{\omega} \tag{4}$$

where  $\mathbf{A}$  is a known full rank matrix of size  $M \times N$ , with  $M > N$ , and nuisance parameters  $\boldsymbol{\omega}$  is a  $N \times 1$  vector of describing the expectation of  $\mathbf{y}$ . In the present paper, the matrix  $\mathbf{A}$  can be described as:

$$\mathbf{A} = \begin{pmatrix} a_1 & \dots & a_L & a_{L+2} & \dots & a_{2L+1} \\ a_2 & \dots & a_{L+1} & a_{L+3} & \dots & a_{2L+2} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_m & \dots & a_{L+m-1} & a_{L+m+1} & \dots & a_{2L+m} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ a_{M-2L} & \dots & a_{M-L-1} & a_{M-L+1} & \dots & a_M \end{pmatrix},$$

where  $m$  denotes an index of each column vector and  $2L$  a number of used neighbouring samples, with  $M > 3L$ . It should be noted that the expectations of the first and last  $L$  samples are not considered in our proposed model. In practice,  $y$  denotes an observation of  $a$ .

The idea of using such a linear parametric model [5, 22, 36] is that it allows an easy elimination of nuisance parameters. To apply this theory, let us define  $C(\mathbf{A}) \subseteq \mathbb{R}^M$  the column space spanned by  $\mathbf{A}$ , with  $\dim(C(\mathbf{A})) = \text{rank}(\mathbf{A}) = N$  and  $C(\mathbf{A})^\perp \subseteq \mathbb{R}^{M-N}$  its orthogonal complement, sometimes referred to as the ‘‘parity space’’, with  $\dim(C(\mathbf{A})^\perp) = M - N$ . The projection of observation vector  $\mathbf{y}$  onto the parity space  $C(\mathbf{A})^\perp$  is obtained by  $\mathbf{n} = \mathbf{W}\mathbf{y}$  where the matrix  $\mathbf{W}$  verifies, among others, the following useful properties:

$$\mathbf{W}\mathbf{A} = \mathbf{0} \text{ and } \mathbf{W}\mathbf{W}^T = \mathbf{I}_{M-N}. \tag{5}$$

Hence, the projection of observation vector  $\mathbf{y}$  onto the parity space  $C(\mathbf{A})^\perp$  yields:  $\mathbf{n} = \mathbf{W}\mathbf{y} = \mathbf{W}\boldsymbol{\epsilon} \sim \mathcal{N}(0, \sigma^2 \mathbf{I}_{M-N})$ . Note that the use of projection matrix  $\mathbf{W}$  can be replaced

by using the Maximum Likelihood Estimation (MLE). By rejecting the nuisance parameter, let us define the estimation of  $\mathbf{n}$  as

$$\hat{\mathbf{n}} = \mathbf{y} - \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{y}. \tag{6}$$

Since the resampled samples are interpolated from their neighbourhood non-resampled samples, the residual noise obtained from the difference between the resampled samples and their estimation approximately follows the Gaussian distribution with zero expectation. To verify the effectiveness of our proposed model, the simulated experiment is illustrated. A set of random variables of 1000 samples, referring to as the original signal, is generated. Then, by using a resampling factor (RF)  $\frac{p}{q} = \frac{2}{1}$ , the resampled signal is obtained (see Fig. 2a). Based on our proposed linear parametric model, the residual noise of resampled signal is acquired, which can be fitted by using the Gaussian distribution (see Fig. 2b). Similarly, it is also proposed to give the empirical results from the residual noise of original signal and its Gaussian model fitting (see Fig. 2c). Obviously, the resampled residual noise fits the Gaussian model better, which verifies our proposed assumption that residual noise of resampled signal approximately follows the Gaussian distribution. By contrast, it hardly holds true that the residual noise of original signal follows the Gaussian distribution.<sup>1</sup>

### 2.2 Exposing one-dimensional signal resampling

Based on the linear parametric model, the extracted residual noise denoted  $\mathbf{n} = \{n_i \mid i \in \{1, \dots, I\}\}$  follows the Gaussian distribution (see Fig. 2b). For simplifying the calculation, it is assumed that the residual noise from non-resampled/original signal is uniformly distributed (see [21]). Thus, the problem of detecting between the original and resampled signal can be formulated by the following two hypotheses:

$$\begin{cases} \mathcal{H}_0 = \{n_i \sim \mathcal{U}(n_{min}, n_{max}), \forall i = (1, \dots, I)\}, \\ \mathcal{H}_1 = \{n_i \sim \mathcal{N}(0, \sigma^2), \forall i = (1, \dots, I)\}, \end{cases} \tag{7}$$

where under hypothesis  $\mathcal{H}_0$ , the residual noise of the inspected signal follows the Uniform distribution; under hypothesis  $\mathcal{H}_1$ , the residual noise is normally distributed with zero expectation and  $\sigma^2$  variance. It is noted that  $n_{min}$  denotes the minimum value of the noise and  $n_{max}$  the maximum value, assuming that  $n_{min} \neq n_{max}$ .

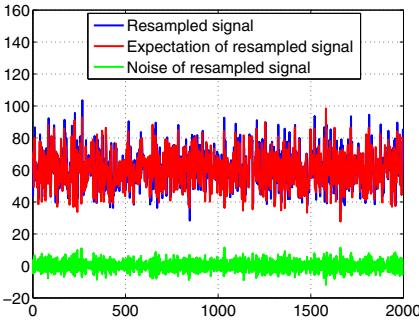
Then, by using Bayes' rule, let us immediately denote the conditional probability of each noise sample from the resampled signal by:

$$\Pr\{n_i \in \mathcal{H}_1 \mid n_i\} = \frac{\Pr\{n_i \mid n_i \in \mathcal{H}_1\} \Pr\{n_i \in \mathcal{H}_1\}}{\sum_{k=0}^1 \Pr\{n_i \mid n_i \in \mathcal{H}_k\} \Pr\{n_i \in \mathcal{H}_k\}}, \tag{8}$$

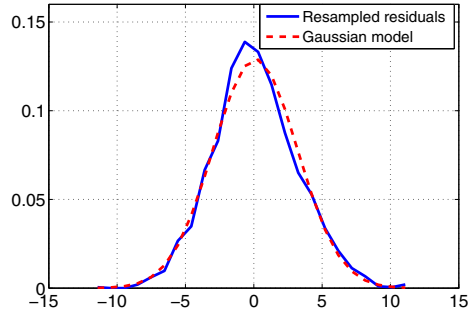
where it is assumed that the priors  $\Pr\{n_i \in \mathcal{H}_0\}$  and  $\Pr\{n_i \in \mathcal{H}_1\}$  are both equal to  $\frac{1}{2}$ . Then (7) is rewritten as follows:

$$\begin{cases} \Pr\{n_i \mid n_i \in \mathcal{H}_0\} = \frac{1}{n_{max} - n_{min}} \\ \Pr\{n_i \mid n_i \in \mathcal{H}_1\} = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-n_i^2}{2\sigma^2}\right). \end{cases} \tag{9}$$

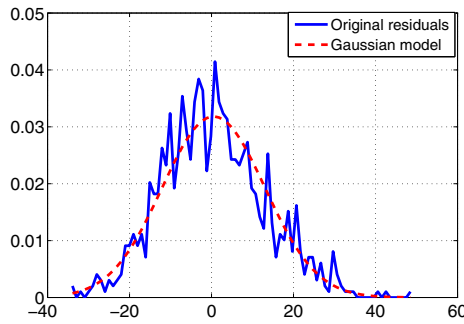
<sup>1</sup>In fact, it is very difficult to assume the accurate distribution model of the original signal without any prior information. In this context, we arbitrarily assume the original signal approximately follows the Uniform distribution in the designed test, which is the limitation of our proposed algorithm.



(a) Resampled signal, its estimated expectation and residual noise.



(b) Empirical results from the noise of resampled signal in Figure 2a and theoretical results from its Gaussian model fitting.



(c) Empirical results from the noise of original signal and theoretical results from its Gaussian model fitting.

**Fig. 2** Illustration of resampled signal with RF = 2 and statistical distributions of the residual noise (resampled and original version)

where in the practical detection, the variance  $\sigma^2$  can be replaced by its estimate  $\hat{\sigma}^2$  by using MLE as follows:

$$\hat{\sigma}^2 = \frac{1}{I - 1} \sum_{i=1}^I (n_i - \frac{1}{I} \sum_{i=1}^I n_i)^2, \tag{10}$$

where  $I$  denotes the total number of the pixels.

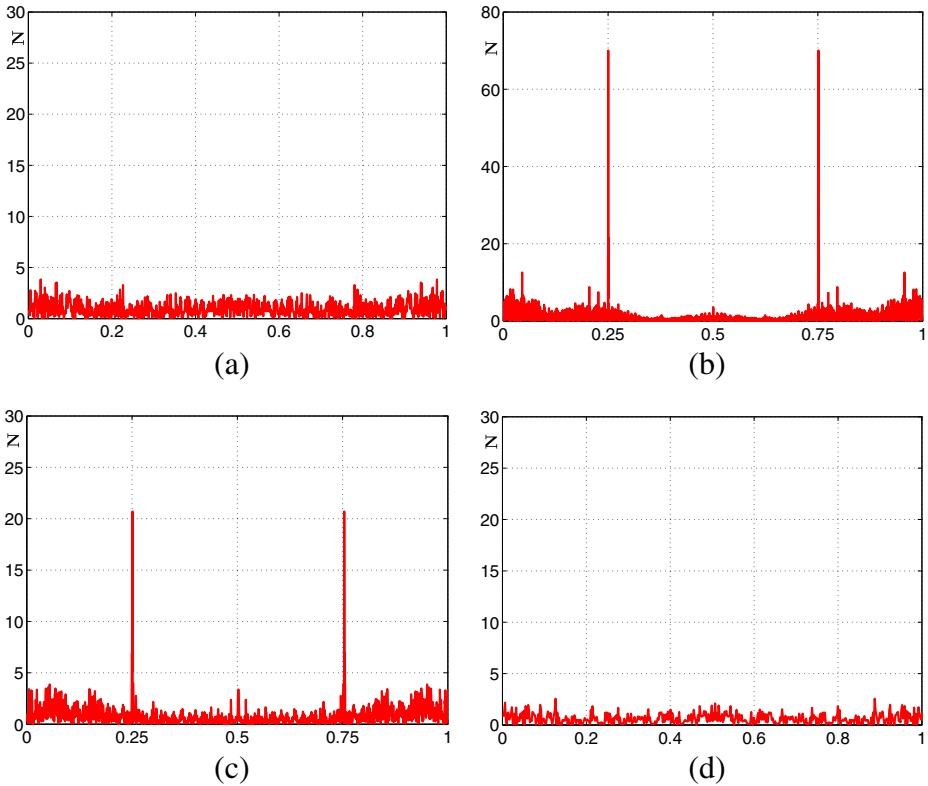
It is proposed to use the frequential representation of the probability of resampled residuals, denoted  $\mathbf{N}$  and formally defined as follows:

$$\mathbf{N} = |\mathbf{h} * |f[\text{Pr}\{\mathbf{n} \in \mathcal{H}_1 | \mathbf{n}\}]|, \tag{11}$$

where a vector  $\mathbf{h} = [1/4, 1/4, -1, 1/4, 1/4]$  denotes a 1-D high-pass filter,  $f[\cdot]$  represents the calculation of 1-D Discrete Fourier Transform (DFT). It should be noted that the high-pass filter aims at removing the lower frequential components for better visually displaying. Due to the periodical artifact caused by resampling operation, the spike-like peak value appears in the frequency domain which is the most distinctive artifact of the resampled signal (see Fig. 3b and c).

To verify the effectiveness of our proposed algorithm, it is proposed to compare the original signal with the resampled one by using different RFs. As Fig. 3 illustrates, the





**Fig. 3**  $N$  (frequential representation of the probability) comparison of residual noise between original and resampled signal in Fourier domain, **(a)** original signal, **(b)** resampled signal by a factor  $\frac{p}{q} = \frac{4}{1}$ , **(c)** resampled signal by a factor  $\frac{p}{q} = \frac{4}{3}$ , **(d)** resampled signal by a factor  $\frac{p}{q} = \frac{1}{2}$

resampled signal (see Fig. 3b and c) exposes the obvious peaks in the frequency domain, but the original signal (see Fig. 3a) has nothing. Moreover, surprisingly, the resampled period is accurately estimated by:

$$\hat{T} = \frac{1}{f_p} \tag{12}$$

where  $\hat{T}$  denotes the estimated period or interpolated rate of the resampled signal, which equals to the up-sampled factor  $p$ .  $f_p$  is the normalized frequency, corresponding to the position of the first peak in Fourier domain. By counting the number of the local peaks with a threshold  $\tau_1$ , we can design a test of detecting a resampled signal.

It should be noted that our proposed algorithm can not differentiate between two resampled signals by using the same up-sampled factor. For instance, the peaks associated with a resampled factor  $\frac{4}{1}$  (see Fig. 3b) appears similar to those by a resampled factor  $\frac{4}{3}$  (see Fig. 3c). In addition, in the case of the resampled signal by an up-sampled factor  $p = 1$ , where it has no obvious peaks (see Fig. 3d), it can not be distinguished with the original signal. In this case, each sample of the resampled signal exposes no periodic linear correlation with its neighbouring ones. Thus, our proposed algorithm is invalid.

### 3 Two-dimensional signal resampling and its artifact

In this section, inspired by extracting the artifact of the 1-D resampled signal (11), we can extend the algorithm from a 1-D signal to a 2-D image straightforward. Compared with a 1-D resampled signal, a 2-D resampled image has more complex linear correlation. A 1-D signal is interpolated by one direction, but a 2-D image is interpolated by the horizontal and vertical direction simultaneously. When a 2-D image is resampled only by the horizontal or vertical direction, it will cause the image visual distortion, which is not discussed in this context.

#### 3.1 Problem statement of two-dimensional signal resampling

Similarly, a procedure of resampling a 2-D image has three steps: up-sampling by the horizontal and vertical direction; interpolation by two directions; down-sampling by two directions. Then similar to the linear combination of 1-D signal, each pixel intensity of the resampled 2-D image correlates to its surrounding ones. Next, let us denote an original grey-level image as  $\mathbf{X} = \{x_{i,j} \mid i \in \{1, \dots, I\}, j \in \{1, \dots, J\}\}$ , where  $I$  and  $J$  denote the height and the width of the matrix  $\mathbf{X}$ . The case of color image with three channels can be obtained straightforward.

Let us denote the linear correlation between each pixel sample and its neighbouring ones by the following:

$$x_{i,j} = \sum_{p=-L}^L \sum_{q=-L}^L \alpha_{p,q} x_{i+p,j+q} \tag{13}$$

where  $\alpha_{p,q}$  denotes a weighted factor with the case of  $\alpha_{0,0} = 0, (2L + 1)^2 - 1 (L \in \mathbb{N}^+)$  the number of used neighbouring samples,  $x_{i,j} \mid i \in \{2, \dots, I - 1\}, j \in \{2, \dots, J - 1\}$  denotes a pixel intensity. Enlarging  $L$  the range of neighbouring indeed improves the accuracy of describing the linear correlation, but at the cost of high computation. For simplicity,  $L = 1$  is a good tradeoff between complexity and accuracy. Then let us rewrite a resampled image  $\mathbf{X} = \{\mathbf{x}_j\}$ , the vector  $\mathbf{x}_j = \{x_{2,j}, \dots, x_{I-1,j}\}^T$ , where  $j \in \{2, \dots, J - 1\}$ . Immediately, it is proposed to use a linear parametric model to deal with the 2-D image. To this end, the vector  $\mathbf{x}_j$  is defined:

$$\mathbf{x}_j \sim \mathcal{N}(\boldsymbol{\mu}_j, \sigma_j^2 \mathbf{I}_{I-2}) = \boldsymbol{\mu}_j + \mathbf{n}_j \tag{14}$$

where  $\boldsymbol{\mu}_j = (\mu_2, \dots, \mu_{I-1})^T$  denotes the expectation, each Gaussian-distributed column vector  $\mathbf{n}_j = (n_2, \dots, n_{I-1})^T$  takes variance  $\sigma_j^2 \mathbf{I}_{I-2}$  and  $\mathbf{I}_{I-2}$  the identity matrix of size  $(I - 2) \times (I - 2)$ . Furthermore,  $\boldsymbol{\mu}_j$  can be described with the following linear parametric model:

$$\boldsymbol{\mu}_j = \mathbf{A}_j \boldsymbol{\alpha}_j \tag{15}$$

where  $\mathbf{A}_j$  is a known full rank matrix of size  $(I - 2) \times 8$ , with  $I > 10$ , and  $\boldsymbol{\alpha}_j = \{\alpha_{p,q}\}$  where index  $(p, q) \in \{-1, 0, 1\}$  excluding  $(0,0)$  is a  $8 \times 1$  vector describing the expectation of  $\mathbf{x}_j$ . Then, similar to (6), let us define the estimation of  $\mathbf{n}_j$  as

$$\hat{\mathbf{n}}_j = \mathbf{x}_j - \hat{\mathbf{x}}_j = \mathbf{x}_j - (\mathbf{A}_j (\mathbf{A}_j^T \mathbf{A}_j)^{-1} \mathbf{A}_j^T) \mathbf{x}_j. \tag{16}$$

Note that in the present paper, the chosen linear parametric model dealing with the 2-D image is the following matrix  $\mathbf{A}_j$ :

$$\mathbf{A}_j = \begin{bmatrix} \mathbf{x}_j^{(-1,-1)} & \mathbf{x}_j^{(-1,0)} & \mathbf{x}_j^{(-1,1)} & \mathbf{x}_j^{(0,-1)} & \mathbf{x}_j^{(0,1)} & \mathbf{x}_j^{(1,-1)} & \mathbf{x}_j^{(1,0)} & \mathbf{x}_j^{(1,1)} \end{bmatrix} \tag{17}$$

where a matrix  $\mathbf{X}^{(p,q)} = \{\mathbf{x}_j^{(p,q)}\}$ , each column vector  $\mathbf{x}_j^{(p,q)} = \{x_{2+p,j+q}, \dots, x_{I+p-1,j+q}\}^T$ , where  $j \in \{2, \dots, J - 1\}$  and index  $(p, q) \in \{-1, 0, 1\}$  not including  $(0,0)$ .

### 3.2 Exposing two-dimensional signal resampling

By using the linear parametric model, the residual noise can be extracted from an inspected image. Let us denote  $\mathbf{N} = \{\mathbf{n}_2, \dots, \mathbf{n}_{J-1}\}$  residual noise in which each residual column vector  $\mathbf{n}_j = \{n_{i,j}\}, i \in \{2, \dots, I - 1\}$ . Thus, the problem of detecting between the non-resampled and resampled image can be formulated by the following two hypotheses:

$$\begin{cases} \mathcal{H}_0 = \{n_{i,j} \sim \mathcal{U}(-255, 255)\}, \forall i = (2, \dots, I - 1), \\ \mathcal{H}_1 = \{n_{i,j} \sim \mathcal{N}(0, \sigma_j^2)\}, \forall i = (2, \dots, I - 1), \end{cases} \tag{18}$$

where  $j \in \{2, \dots, J - 1\}$ , under hypothesis  $\mathcal{H}_0$ , the residual noise of the non-resampled image follows the Uniform distribution;<sup>2</sup> under hypothesis  $\mathcal{H}_1$ , the residual noise of the resampled image is normally distributed with zero mean and  $\sigma_j^2$  variance. Besides, it should be noted that due to the pixel intensity of a gray-level image ranging from 0 to 255, residual noise (difference between the pixel’s intensity and its expectation value) falls in the range from -255 to 255. Then, by using Bayes’ rule, let us immediately denote the conditional probability of each residual sample:

$$\Pr\{n_{i,j} \in \mathcal{H}_1 | n_{i,j}\} = \frac{\Pr\{n_{i,j} | n_{i,j} \in \mathcal{H}_1\} \Pr\{n_{i,j} \in \mathcal{H}_1\}}{\sum_{k=0}^1 \Pr\{n_{i,j} | n_{i,j} \in \mathcal{H}_k\} \Pr\{n_{i,j} \in \mathcal{H}_k\}}, \tag{19}$$

where it is assumed that the priors  $\Pr\{n_{i,j} \in \mathcal{H}_0\}$  and  $\Pr\{n_{i,j} \in \mathcal{H}_1\}$  are both equal to  $\frac{1}{2}$ . Then (18) is rewritten as follows:

$$\begin{cases} \Pr\{n_{i,j} | n_{i,j} \in \mathcal{H}_0\} = \frac{1}{510}, \\ \Pr\{n_{i,j} | n_{i,j} \in \mathcal{H}_1\} = \frac{1}{\sqrt{2\pi\hat{\sigma}_j^2}} \exp\left(\frac{-n_{i,j}^2}{2\hat{\sigma}_j^2}\right), \end{cases} \tag{20}$$

where the variance  $\hat{\sigma}_j^2$  can be estimated by using MLE as follows:

$$\hat{\sigma}_j^2 = \frac{1}{I-3} \sum_{i=2}^{I-1} (n_{i,j} - \frac{1}{I-2} \sum_{i=2}^{I-1} n_{i,j})^2, \tag{21}$$

where  $I$  denotes the total number of pixels in each column  $j$ .

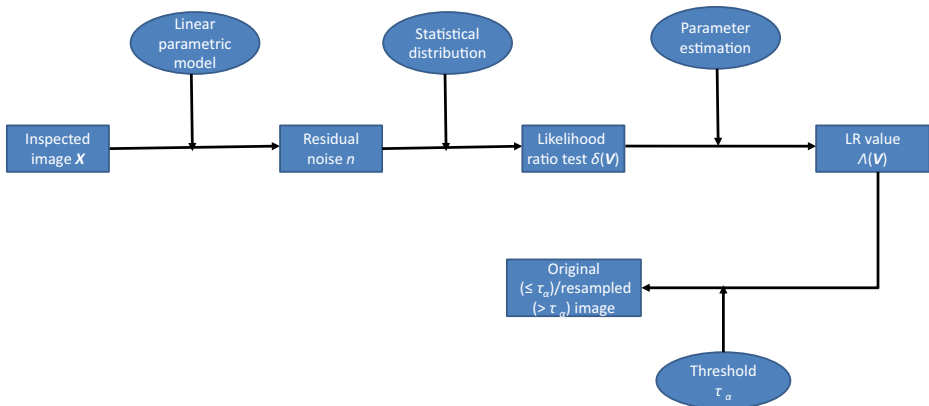
Then it is proposed to use the frequential representation of the probability of residual noise, denoted  $\mathbf{N}^{2D}$  and formally defined as follows:

$$\mathbf{N}^{2D} = |\mathbf{H} * |\mathcal{F}[\mathbf{P}_{map}]||, \tag{22}$$

where  $\mathbf{H}$  denotes a 2-D high-pass filter,  $\mathcal{F}[\cdot]$  represents the calculation of 2-D DFT.  $\mathbf{P}_{map} = \{\Pr\{n_{i,j} \in \mathcal{H}_1 | n_{i,j}\}\}, i \in \{2, \dots, I - 1\}, j \in \{2, \dots, J - 1\}$  denotes the so-called probability map.

A 2-D resampled image exposes bright spots surrounding the central spot (similar to peaks in the 1-D signal) distributed symmetrically in Fourier domain, which is the critical

<sup>2</sup>In fact, it is very difficult to accurately define the distribution of the residual noise from a non-resampled image. Here, it is proposed to arbitrarily approximate its distribution using the Uniform distribution inspired by the reference [21].



**Fig. 4** A flowchart of the proposed algorithm

resampling artifact of a resampled image. Due to the fact that the interpolated pixel obtained from its neighbourhood pixels, the periodic characteristic is uncovered unavoidably in the frequency domain. Therefore, the symmetric bright spots appear in Fourier domain when an image is resampled. Based on the assumed distributions (18), together with hypothesis theory, we can establish a practical LRT of detecting a resampled image. The specific discussion will be extended in the following section.

### 4 Design of the practical LRT for resampling detection

Prior to our proposed algorithm in this paper, some state-of-the-art methods confine their ideas to detection of the peak values (bright spots) in Fourier domain. Namely, by counting the number of bright spots, the designed test detects if the inspected image is resampled. Therefore, that proposed test has to estimate the suitable threshold for selecting bright spots existing in the Fourier domain (see [14, 21]). However, since some post-camera<sup>3</sup> operations such as compression possibly generate some bright spots which are not caused by image resampling. That will to some extent impact the detection accuracy of that prior-art test.

In this section, it is proposed to solve the problem of resampling detection by designing a practical LRT, which has been verified to be a very effective algorithm in [23–26]. A flowchart of the proposed algorithm is illustrated in Fig. 4. By using a linear parametric model, residual noise can be extracted from the inspected image. Based on the statistical distributions of residual noise under two hypotheses, it is proposed to design a practical LRT. In this practical LRT, we first estimate the unknown parameter, and meanwhile calculate the LR, and finally detect the inspected image based on the given threshold.

For solving the statistical detection problem (18), it follows from the Neyman-Pearson lemma [16, Theorem 3.2.1] that the LRT is optimal in the sense described below. For definition, let

$$\mathcal{H}_\alpha = \left\{ \delta : \sup_{\sigma_j} \mathbb{P}_{\mathcal{H}_0} [\delta(\mathbf{N}) = \mathcal{H}_1] \leq \alpha \right\} \tag{23}$$

<sup>3</sup>The term *post-camera* is referred to as the operation after acquiring a digital image.

be the class of tests, solving problem (18), with an upper-bounded FAR  $\alpha$ . Here  $\mathbb{P}_{\mathcal{H}_j}[\cdot]$  is the probability under  $\mathcal{H}_j$ ,  $j \in \{0, 1\}$ , and the supremum over model parameter  $\sigma_j$  can be understood as the distribution parameter ensuring that the FAR  $\alpha$  can not be exceeded. Among all tests in  $\mathcal{H}_\alpha$  the LRT is the most powerful test, it maximizes the detection power:

$$\beta_\delta = \mathbb{P}_{\mathcal{H}_1}[\delta(\mathbf{N}) = \mathcal{H}_1]. \tag{24}$$

In the practical test, parameter  $\sigma_j$  is unknown which can be successfully estimated by using MLE (see (21)). Afterwards, we can establish our proposed practical LRT for detecting resampled images. Based on the probability density function (pdf) of hypothesis  $\mathcal{H}_0$  and  $\mathcal{H}_1$  (20), the Likelihood Ratio (LR) value for the  $i$ -th pixel-wise on  $j$ -th column vector is given by:

$$\Lambda_{i,j}(n_{i,j}) = \frac{\Pr\{n_{i,j}|n_{i,j} \in \mathcal{H}_1\}}{\Pr\{n_{i,j}|n_{i,j} \in \mathcal{H}_0\}} = \frac{510}{\hat{\sigma}_j \sqrt{2\pi}} \exp\left(\frac{-n_{i,j}^2}{2\hat{\sigma}_j^2}\right). \tag{25}$$

Then, it follows from the statistical independence of vectors  $n_{i,j}$  that the practical LRT for all the pixel intensities is given by:

$$\delta(\mathbf{N}) = \begin{cases} \mathcal{H}_0 & \text{if } \Lambda(\mathbf{N}) = \sum_{j=2}^{J-1} \sum_{i=2}^{I-1} \log(\Lambda_{i,j}) \leq \tau_\alpha \\ \mathcal{H}_1 & \text{if } \Lambda(\mathbf{N}) = \sum_{j=2}^{J-1} \sum_{i=2}^{I-1} \log(\Lambda_{i,j}) > \tau_\alpha \end{cases} \tag{26}$$

where the column index  $j \in \{2, \dots, J - 1\}$  in which  $J$  denotes the total number of columns, and the decision threshold  $\tau_\alpha$  is the solution of equation  $\mathbb{P}_0[\Lambda(\mathbf{N}) > \tau_\alpha] = \alpha$  to guarantee that  $\delta(\mathbf{N}) \in \mathcal{H}_\alpha$ .

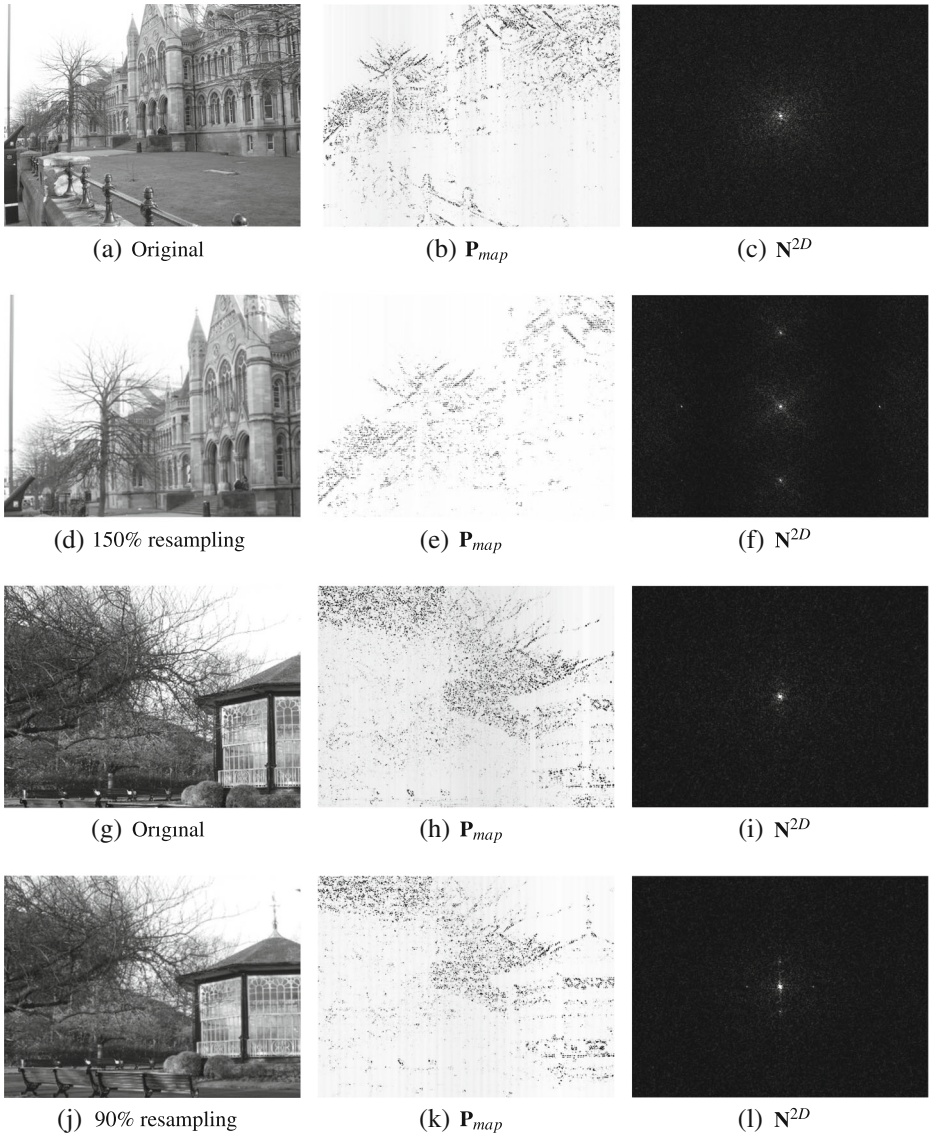
## 5 Experiments

### 5.1 Performance comparison on single resampled image

To verify the sharpness and effectiveness of our proposed methodology, numerical experiments are illustrated in this section. Prior to our experiments, it is proposed to randomly choose some images from the image database [30], in which all 1338 RAW format images with size  $512 \times 384$  are converted to TIFF format or JPEG format with different quality factors.

Figure 5 illustrates the results from resampled images by using *bilinear* interpolation. Compared with the single bright spot from an original image (see Fig. 5c), the up-sampled image with RF equal to 150 %, has a series of bright spots symmetrically surrounding the central point in Fourier domain (see Fig. 5f). It is observed that with increasing the up-sampled factor, the distance between two spots horizontally or vertically is enlarged. This interesting phenomenon is very similar to the pattern of the distance between symmetric peaks in Fourier domain of the 1-D resampled signal. The results of down-sampled images are displayed in Fig. 5l. Compared to the results of up-sampled images, the bright spots surrounding the central point are not very obvious. Let us attribute that characteristic to that down-sampled images lose more information of linear correlation among neighborhood pixels than those of up-sampled images. Thus, it is more difficult to detect a down-sampled image.

Next, it is proposed to detect a rotated image. Similarly, it can be observed that surrounding bright spots are scattered in Fourier domain (see Fig. 6). Finally, Fig. 7 illustrates the experimental results from another affine transformation. Numerical experiments verify that our proposed algorithm indeed is capable of detecting a resampled image. Besides, we can



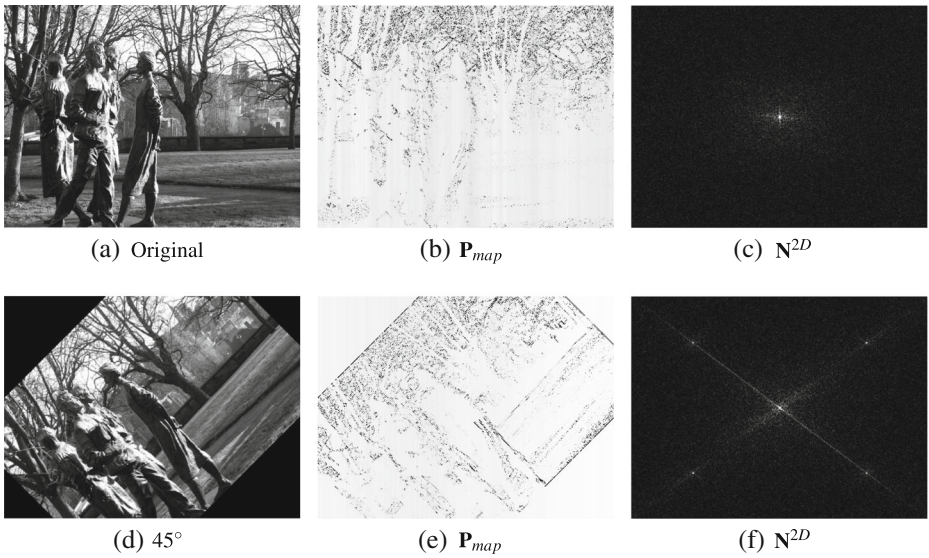
**Fig. 5** *Left-most column:* comparison between an original image and its corresponding resampled ones (bilinear interpolation); *middle column:* its corresponding  $P_{map}$ 's; *right-most column:* periodic resampling artifacts in Fourier domain

modify the high-pass filter  $H$  for enhancing the difference between the bright spot and its surrounding nuisance noise in order to make peak values more visible.

## 5.2 Performance comparison on batch of resampled images

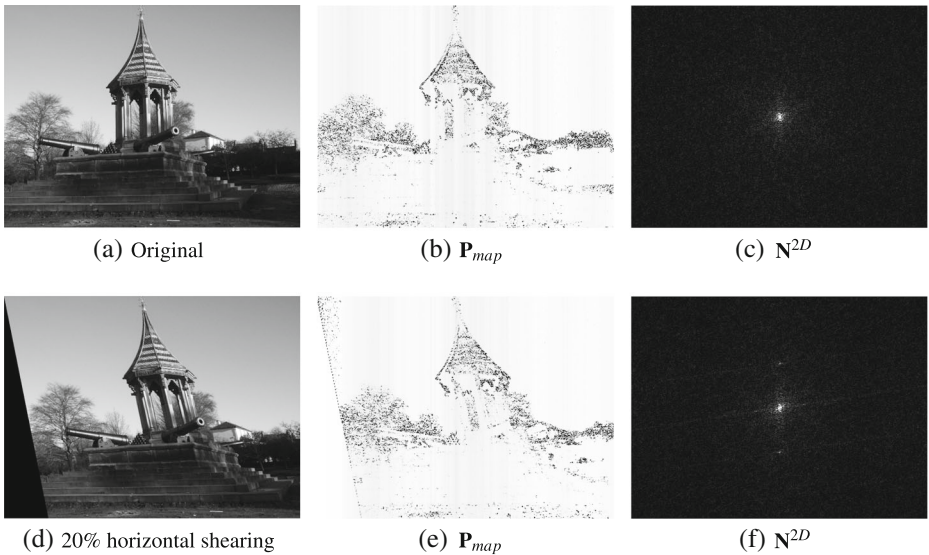
To verify our proposed test in a large scale of images, it is proposed to give the experimental performance of our detector for processing a batch of inspected images. We establish an



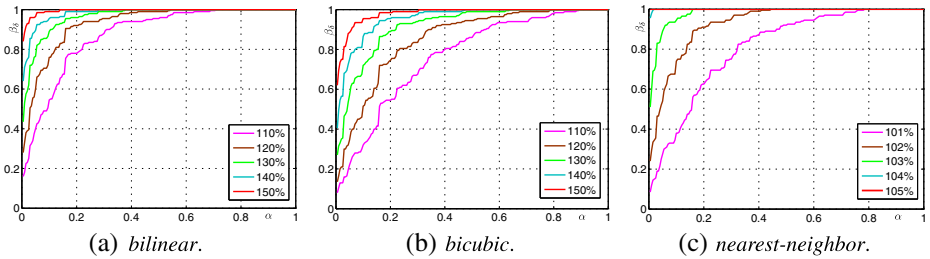


**Fig. 6** *Left-most column*: comparison between an original image and its corresponding resampled one (*bilinear* interpolation); *middle column*: its corresponding  $P_{map}$ 's; *right-most column*: periodic resampling artifacts in Fourier domain

image database of 200 uncompressed gray-level images, which is randomly chosen from Bossbase database [2]. In the following experiments, three practical interpolation methods for image resizing, that are *bilinear*, *bicubic*, *nearest-neighbor*, are used for verifying



**Fig. 7** *Left-most column*: comparison between an original image and its corresponding resampled one (*bicubic* interpolation); *middle column*: its corresponding  $P_{map}$ 's; *right-most column*: periodic resampling artifacts in Fourier domain



**Fig. 8** ROC curves comparison by detecting up-sampled images where the legend denotes the resampling factor (RF)

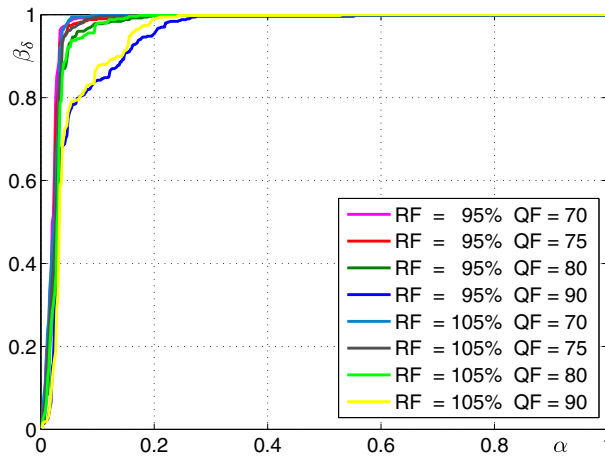
the effectiveness of our proposed algorithm. It should be noted that our proposed test can still detect other image re-scaling techniques like rotation or shearing. As Fig. 8 illustrates, by using *bilinear* or *nearest-neighbor* algorithm for resampling, our detector performs its ability to detect resampled images with small RFs. When *nearest-neighbor* is applied for interpolating, even resampled images with RF equal to 105 % can be easily detected at the cost of very small FAR. Compared with *bilinear* or *nearest-neighbor*, since *bicubic* partly destroys the correlation of pixels existing in the resampled image, the performance of the detector is degraded. Besides, with increasing RF, the performance of our detector is improved largely. Therefore the larger RF is, the more stronger correlation among pixels exists.

Apart from detecting resampled uncompressed images, our proposed detector is capable of detecting resampled compressed images. It is proposed to use dataset one consisting of 500 uncompressed non-resampled images and 500 compressed resampled images. After resampling original uncompressed images by using RF from 75 % to 150 %, all the resampled images are saved as JPEG format with Quality Factor (QF) from 50 to 90 (see Table 1).

**Table 1** Detection performance comparison at the given FAR  $\alpha_\tau = 0.05$  by detecting resampled once-compressed images in dataset one

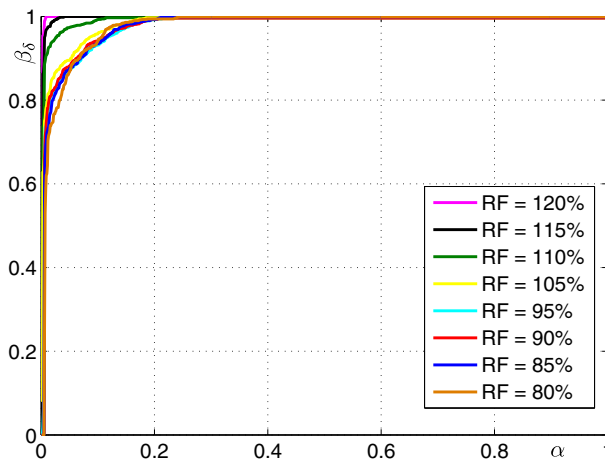
RF \ QF	50	55	60	65	70	75	80	85	90
150 %	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
140 %	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
130 %	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
120 %	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.99
110 %	1.00	1.00	1.00	1.00	1.00	1.00	0.99	0.97	0.92
95 %	1.00	1.00	1.00	1.00	1.00	0.99	0.98	0.93	0.84
90 %	1.00	1.00	1.00	1.00	1.00	0.99	0.98	0.94	0.85
85 %	1.00	1.00	1.00	0.99	0.99	0.99	0.97	0.92	0.81
80 %	1.00	1.00	0.99	0.99	0.98	0.97	0.96	0.91	0.80
75 %	0.99	0.99	0.99	0.98	0.97	0.95	0.92	0.85	0.70
Average	0.999	0.999	0.998	0.996	0.994	0.989	0.980	0.952	0.891





**Fig. 9** ROC curves comparison by detecting resampled double-compressed images in dataset two

By increasing QF, JPEG compression nearly does not impact our detection power. Besides, by using different RFs ( $> 100\%$  or  $< 100\%$ ), our proposed detector performs at the similarly equivalent level. In addition, it is proposed to use dataset two consisting of 500 compressed non-resampled images with QF 75 and 500 double-compressed resampled images. After resampling original compressed images by using RF 105% and 95%, all resampled images are saved as JPEG format with QF from 70 to 90 (see Fig. 9). Although double-compression operation to some extent impacts the detection accuracy, the proposed detector performs well by considering its detection rate at the given false alarm rate. Finally, it is proposed to design dataset three consisting of mixed 2000 images which includes 1600 JPEG images with different QFs and 400 uncompressed images. In our experiment, half of this mixed dataset is resampled with RF from 80% to 120% while the other half is not resampled. As Fig. 10 illustrates, at the very low FAR, the detection performance of the proposed



**Fig. 10** ROC curves comparison by detecting resampled compressed or uncompressed images in dataset three

**Table 2** Detection performance (true positive rate) comparison under two hypotheses between  $\mathcal{H}_0$  and  $\mathcal{H}_1$  at the given FAR  $\alpha_\tau$ ; the proposed test  $\delta$  (26) designed by using three algorithms; the interpolation algorithm is *bilinear*

Resampling factor	Linear model (15)	Fixed pattern [14]	EM algorithm [21]
(a) $\alpha_\tau = 0.05$			
200 %	1.00	1.00	0.71
150 %	0.95	0.86	0.43
140 %	0.81	0.72	0.34
130 %	0.69	0.61	0.28
120 %	0.57	0.51	0.24
110 %	0.41	0.39	0.19
95 %	0.52	0.46	0.07
90 %	0.54	0.45	0.07
85 %	0.53	0.43	0.08
80 %	0.53	0.42	0.09
Average	0.67	0.59	0.25
(b) $\alpha_\tau = 0.1$			
200 %	1.00	1.00	0.79
150 %	0.98	0.91	0.55
140 %	0.91	0.80	0.44
130 %	0.81	0.73	0.40
120 %	0.69	0.62	0.35
110 %	0.56	0.51	0.29
95 %	0.64	0.57	0.13
90 %	0.68	0.57	0.15
85 %	0.68	0.57	0.18
80 %	0.69	0.57	0.18
Average	0.76	0.69	0.35

detector achieves the relevant level. Besides, the resampled images with RF larger than one can be detected more accurately than the images with RF smaller than one. Because resampled images with larger RF has more correlated information existing among pixels than images with smaller RF. Meanwhile, resampling detection in the mixed dataset verifies that our designed test can detect resampled images with mixed compressed and uncompressed format.

In practice, the periodic blocking artifact of the JPEG format image possibly collides with the periodic trace of the resampled image. Therefore, when a batch of resampled JPEG format images are inspected, it is very difficult to authenticate whether the periodic pattern is

**Table 3** Comparison of computation time (second) for each inspected image; the interpolation algorithm is *bilinear*

Resampling factor	Linear model (15)	Fixed pattern[14]	EM algorithm [21]
200 %	0.0996	0.0488	2.2354
150 %	0.0928	0.0449	1.9323
140 %	0.0827	0.0482	2.3860
130 %	0.0953	0.0485	2.0290
120 %	0.0937	0.0538	1.9387
110 %	0.1012	0.0516	2.0494
95 %	0.0808	0.0477	2.1660
90 %	0.0759	0.0430	1.8757
85 %	0.0717	0.0426	1.6390
80 %	0.0622	0.0379	1.5356
Average	0.0856	0.0467	1.9787

from the blocking artifact or the resampling artifact. Since the artifact is denoted as the only benchmark to detect a batch of images, those artifact-based detectors could not solve this tough problem. In our proposed algorithm, the practical LRT does not use the resampling artifact directly, but instead the ratio of the probability of residual noise, which effectively deals with the problem caused by JPEG compression.

### 5.3 Performance comparison with state-of-the-art competitors

Some state-of-the-art designed tests based on the bright spots, which are caused by the periodic characteristic of interpolation, have been proposed to detect a batch of resampled images (see [14, 21]). However, in this present paper, it is proposed to design a practical LRT for solving the problem of image resampling detection. To compare our proposed detector with the prior arts, in our following experiments, it is proposed to design three different LRTs based on estimated residual noise  $n_{i,j}$  of an image by using the following algorithms: our proposed linear parametric model (15); EM algorithm proposed in [21]; the fixed parameter pattern<sup>4</sup> proposed in [14]. Then, let us compare the detection performance of the designed practical LRT based on three different algorithms. We use all TIFF format images from the image database [30], which are resampled by using *bilinear* algorithm with 6 different RFs.

Table 2 illustrates the comparison of the detection performance at the give FAR  $\alpha_\tau = 0.05$  or 0.1. With increasing RF, the detection performance of three practical LRTs is improved. Moreover, in the case of using all 6 RFs, the practical LRT designed by the linear model outperforms the other two practical LRTs using prior-art algorithms [14, 21]. Besides, it should be noted that the performance of the practical LRT designed by our proposed linear model achieves 100 % at the given RF equal to 200 %, as well as the practical

<sup>4</sup>The fixed pattern can be described as  $\alpha_j = (0.5, 0.5, 0.5, 0.5, -0.25, -0.25, -0.25, -0.25)^T$ , where  $\alpha_j$  denotes the weighted factors (see (15)).

LRT established by using a fix pattern. We also compare the computation efficiency of the designed practical LRTs based on three algorithms, which is illustrated in Table 3. It can be observed that the LRT designed by the fix pattern performs best due to that the fixed pattern need not estimate the parameters. Since EM algorithm needs iterative computation, the LRT based on it performs less efficiently than two other designed LRTs. Hence, by considering the detection accuracy and computation efficiency, our proposed LRT based on the linear parametric model is the optimal choice.

## 6 Conclusion and discussion

This paper first studies the linear correlation of 1-D resampled signals. By using a linear parametric model, together with the Bayes' rule, we propose the algorithm of uncovering the resampling artifact. Numerical experiments verify that our proposed algorithm is capable of distinguishing between non-resampled signal and resampled one. Meanwhile, the up-sampled factor of a resampled 1-D signal can be estimated accurately. Moreover, it is proposed to use another linear parameter model and design a test of detecting a resampled image, whose artifact is described as bright spots in Fourier domain. In the case of detecting a batch of images, based on the probability of residual noise extracted from a non-resampled and resampled image respectively, it is proposed to design the practical LRT detector. To compare with LRTs designed by using the prior-art algorithms [14, 21], our proposed practical LRT performs best.

To the best of our knowledge, our designed LRT-based detector is the first test based on hypothesis testing theory for image resampling detection. It can achieve the maximal detection power at the prescribed FAR. Besides, our detector can deal with the problem of authenticating the resampled images from mixed compressed and uncompressed images. In this context, the Uniform distribution describing a non-resampled image is not very accurate. It can be replaced by the more accurate generalized model, which will be studied in the future works.

**Acknowledgments** This work is funded by the State Key Program of Zhejiang Province Natural Science Foundation of China under Grant No. LZ15F020003 and the Natural Science Foundation of China (No. 61602295) and the Natural Science Foundation of Shanghai (No. 16ZR1413100). The Ph.D thesis of Tong Qiao is funded by the China Scholar Council (CSC) and the region Champagne-Ardenne, IDENT project.

## References

1. Amerini I, Ballan L, Caldelli R, Del Bimbo A, Serra G (2011) A sift-based forensic method for copy-move attack detection and transformation recovery. *IEEE Trans Inf Forensics Secur* 6(3):1099–1110
2. Bas P, Filler T, Pevný T. (2011) Break our steganographic system — the ins and outs of organizing boss. In: Filler T (ed) 13th International Workshop on Information Hiding
3. Chen M, Fridrich J, Goljan M, Lukás J (2008) Determining image origin and integrity using sensor noise. *IEEE Trans Inf Forensics Secur* 3(1):74–90
4. Christlein V, Riess C, Jordan J, Angelopoulou E (2012) An evaluation of popular copy-move forgery detection approaches. *IEEE Trans Inf Forensics Secur* 7(6):1841–1854
5. Cogramme R, Zitzmann C, Fillatre L, Reirant F, Nikiforov I, Cornu P (2011) A cover image model for reliable steganalysis. In: Information Hiding, pp 178–192. Springer

6. Dempster AP, Laird NM, Rubin DB et al (1977) Maximum likelihood from incomplete data via the em algorithm. *J R Stat Soc* 39(1):1–38
7. Farid H (2009) Exposing digital forgeries from jpeg ghosts. *IEEE Trans Inf Forensics Secur* 4(1):154–160
8. Farid H (2009) Image forgery detection. *IEEE Signal Proc Mag* 26(2):16–25
9. Feng X, Cox IJ, Doërr G (2011) An energy-based method for the forensic detection of re-sampled images. In: 2011 IEEE International Conference on Multimedia and Expo (ICME), pp 1–6. IEEE
10. Feng X, Cox IJ, Doërr G (2012) Normalized energy density-based forensic detection of resampled images. *IEEE Trans Multimedia* 14(3):536–545
11. Fridrich AJ, Soukal BD, Lukáš AJ (2003) Detection of copy-move forgery in digital images. In: *Proceedings of Digital Forensic Research Workshop*. Citeseer
12. Gallagher A, Chen T (2008) Image authentication by detecting traces of demosaicing. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08*, pp 1–8. IEEE
13. Gallagher AC (2005) Detection of linear and cubic interpolation in jpeg compressed images. In: *The 2nd Canadian Conference on Computer and Robot Vision, 2005. Proceedings*, pp 65–72. IEEE
14. Kirchner M (2008) Fast and reliable resampling detection by spectral analysis of fixed linear predictor residue. In: *Proceedings of the 10th ACM workshop on Multimedia and security*, pp 11–20. ACM
15. Kirchner M (2010) Linear row and column predictors for the analysis of resized images. In: *Proceedings of the 12th ACM workshop on Multimedia and security*, pp 13–18. ACM
16. Lehman E R J (2005) *Testing statistical hypotheses*, 2nd Edn. Springer
17. Li J, Li X, Yang B, Sun X (2015) Segmentation-based image copy-move forgery detection scheme. *IEEE Trans Inf Forensics Secur* 10(3):507–518
18. Lukáš J, Fridrich J, Goljan M (2006) Detecting digital image forgeries using sensor pattern noise *Electronic Imaging 2006*, pp 60,720Y–60,720Y. International Society for Optics and Photonics
19. Mahdian B, Saic S (2008) Blind authentication using periodic properties of interpolation. *IEEE Trans Inf Forensics Secur* 3(3):529–538
20. Oppenheim AV, Schaffer RW, Buck JR et al. (1989) *Discrete-time signal processing*, vol 2. Prentice-hall, Englewood Cliffs
21. Popescu AC, Farid H (2005) Exposing digital forgeries by detecting traces of resampling. *IEEE Trans Signal Process* 53(2):758–767
22. Qiao T, Retraint F, Cogranne R (2013) Image authentication by statistical analysis. In: *Signal Processing Conference (EUSIPCO), 2013 Proceedings of the 21th European*, pp 1–5. IEEE, EURASIP
23. Qiao T, Retraint F, Cogranne R, Thai TH (2015) Source camera device identification based on raw images. In: *2015 IEEE International Conference on Image Processing (ICIP)*, pp 3812–3816. IEEE
24. Qiao T, Retraint F, Cogranne R, Zitzmann C (2015) Steganalysis of jsteg algorithm using hypothesis testing theory. *EURASIP J Inf Secur* 2015(1):1–16
25. Qiao T, Zitzmann C, Cogranne R, Retraint F (2014) Detection of jsteg algorithm using hypothesis testing theory and a statistical model with nuisance parameters. In: *Proceedings of the 2nd ACM workshop on Information hiding and multimedia security*, pp 3–13. ACM
26. Qiao T, Zitzmann C, Retraint F, Cogranne R (2014) Statistical detection of jsteg steganography using hypothesis testing theory. In: *2014 IEEE International Conference on Image Processing (ICIP)*, pp 5517–5521. IEEE
27. Qin C, Chang CC, Chen PY (2012) Self-embedding fragile watermarking with restoration capability based on adaptive bit allocation mechanism. *Signal Process* 92(4):1137–1150
28. Qin C, Chen X, Ye D, Wang J, Sun X (2016) A novel image hashing scheme with perceptual robustness using block truncation coding. *Inf Sci* 361:84–99
29. Ryu SJ, Lee HK (2014) Estimation of linear transformation by analyzing the periodicity of interpolation. *Pattern Recogn Lett* 36:89–99
30. Schaefer G, Stich M (2003) Ucid: an uncompressed color image database. In: *Electronic Imaging 2004*, pp 472–480. International Society for Optics and Photonics
31. Sencar HT, Memon N (2012) *Digital image forensics: There is more to a picture than meets the eye*. Springer
32. Stamm MC, Wu M, Liu K (2013) Information forensics: An overview of the first decade. *IEEE Access* 1(1):167–200
33. Thai TH, Cogranne R, Retraint F (2014) Camera model identification based on the heteroscedastic noise model. *IEEE Trans Image Process* 23(1):250–263

34. Wei W, Wang S, Zhang X, Tang Z (2010) Estimation of image rotation angle using interpolation-related spectral signatures with application to blind detection of image forgery. *IEEE Trans Inf Forensics Secur* 5(3):507–517
35. Xia Z, Wang X, Sun X, Liu Q, Xiong N (2016) Steganalysis of lsb matching using differences between nonadjacent pixels. *Multimedia Tools and Applications* 75(4):1947–1962
36. Zitzmann C, Cogranne R, Reirant F, Nikiforov I, Fillatre L, Cornu P (2011) Statistical decision methods in hidden information detection. In: *Information Hiding*, pp 163–177. Springer



**Tong Qiao** received the B.S. degree in Electronic and Information Engineering in 2009 from Information Engineering University, Zhengzhou, China, and the M.S. degree in Communication and Information System in 2012 from Shanghai University, Shanghai, China, and the Ph.D. degree in Systems Optimizations and Dependability in 2016 from University of Technology of Troyes, Laboratory of Systems Modeling and Dependability, Troyes, France. He is currently a lecturer at Hangzhou Dianzi University, School of Cyberspace. His current research interests focus on steganalysis and digital image forensics.



**Aichun Zhu** was born in Jiangsu, China, in 1986. He received the M.S. degree in engineering from China University of Mining and Technology, China, in 2012. He received the Ph.D. degree with the University of Technology of Troyes, France, in 2016. He is currently a lecturer at Nanjing Tech University, School of Computer Science and Technology. He is involved in human motion detection and tracking. His academic interests span computer vision and machine learning.



**Florent Retrait** received the M.Sc. degree in applied mathematics and the Ph.D. degree in image processing from the National Institute of Applied Sciences of Lyon, France, in 1994 and 1998, respectively. He is currently an Associate Professor with the University of Technology of Troyes. His research focuses on image modeling, statistical image processing, hypothesis testing theory, and anomaly detection and localization, with a main application to digital image forensics.