CrossMark

Detecting median filtering via two-dimensional AR models of multiple filtered residuals

Jianquan Yang[1,2] · **Honglei Ren**[1] · **Guopu Zhu**[1,3] (ID) ·
Jiwu Huang[4] · **Yun-Qing Shi**[5]

Abstract Median filtering, being an order statistic filtering, has been widely used in image denoising and recently also in image anti-forensics and anti-steganalysis. In the past few years, several methods have been developed for median filtering detection. However, it is still a challenging task to detect median filtering in JPEG compressed images. In this paper, we propose a novel method to solve this challenging task. We first generate median filtered residual (MFR), average filtered residual (AFR) and Gaussian filtered residual (GFR) by calculating the differences between an original image and its filtered images. Then, we propose to use two-dimensional autoregressive (2D-AR) model to characterize MFR, AFR and GFR separately, and further combine the 2D-AR coefficients of these three residuals

---

Jianquan Yang and Honglei Ren contributed equally to this work

✉ Guopu Zhu
gp.zhu@siat.ac.cn

[1] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, GD 518055, China

[2] Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen, GD 518055, China

[3] State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100093, China

[4] College of Information Engineering, Shenzhen University, Shenzhen, GD 518055, China

[5] Department of Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ 07102, USA

⦾ Springer

into a set of features. Finally, the extracted feature set is fed into a support vector machine classifier for training and detection. Extensive experiments have demonstrated that compared with existing methods, the proposed one can achieve a considerable improvement in detecting median filtering in heavily compressed images.

1 Introduction

Nowadays, digital images are fairly prevalent on the Internet. However, among these images, there are often forged ones, which might result in harming individual and public interests, such as, injuring someone's reputation, and misleading public opinion. Therefore, as an effective technique to unveil the forgeries in digital images, image forensics has drawn increasing attention over the past more than ten years [11, 30, 33]. So far, extensive studies have been conducted in the field of image forensics, including the detection of median filtering [4, 23, 37], the estimation of JPEG compression history [9, 14, 38], the identification of source camera [18, 26], the localization of tampered regions in forgery images [7, 25, 31], and so on.

This paper focuses on the study of median filtering detection. Median filtering is an order statistic operation. Because of its capability of removing image noise meanwhile preserving image detail [17], median filtering has been widely studied and applied in image denoising. In recent years, much attention has also been paid to the detection of median filtering. This is mainly due to the following reasons. First, one goal of image forensics is to detect the use of image editing and retouching in digital photos [21, 33]. Since median filtering is one of the basic tools for image editing and retouching, there is a need to detect median filtering applied to images. Second, as a nonlinear image operation, median filtering can effectively conceal the artifacts caused by some other operations, but without much further deterioration of image quality. Therefore, median filtering has been employed as a tool to attack image forensics, such as to hide the trace of image resampling [22] or image compression [35]. In addition, it could be used to degrade the performance of steganalysis methods [19, 23]. Median filtering detection is a simple but effective way to counter these anti-forensics and anti-steganalysis. Third, in the scenario of image tampering, it may happen that a portion of an image is pasted on another image to create a tampered image. If the two original images are filtered in different ways, i.e., one is smoothed with median filter, while the other is not, then the tampered region can be localized by a block-by-block detection of median filtering in the tampered image, as done in [4, 19, 37]. That is to say, median filtering detection is also useful to expose image tampering.

Several studies have been conducted on the detection of median filtering. In [23], Kirchner and Fridrich proposed a feature to detect median filtering by measuring streaking artifacts[1] in digital images. They found that the number of pixels whose first-order

---

[1]Streaking artifacts, first analyzed by Bovik [2], refer to the phenomenon that the probability of two adjacent pixels being the same increases greatly after median filtering, which can thus be evaluated by using first-order difference images.

differences are zero, denoted as h_0, is likely to increase relative to that of pixels whose first-order differences are one, denoted as h_1. Hence, they introduced the ratio h_0/h_1 as a detection feature. Experimental results have shown that the streaking-based feature is effective for uncompressed images, but not robust to JPEG compression. In order to reliably detect median filtering in JPEG compressed images, the authors of [23] also proposed to apply SPAM, a set of steganalysis features proposed in [29], for median filtering detection. It is shown that the application of SPAM can achieve better performance in detecting median filtering in JPEG compressed images. Almost simultaneous to the work reported in [23], Cao et al. [5] independently proposed a detection feature similar to the above streaking-based feature. A year later, Yuan [37] made several interesting observations on median filtered images. These are, firstly, the pixels in a small block tend to have the same value as the median of the block; secondly, the gray value of the center pixel of a block occurs more frequently in the block than before filtering; and thirdly, median filtering makes the quantity of gray values in a small block decrease. Based on these observations, Yuan proposed a set of features, named as median filtering forensics features (MFF), including the distribution of the block median, the occurrence of the block-center gray value, the quantity of gray levels in a block, and so on. Experiments have shown that MFF can improve the performance of detecting median filtering in low-quality images. Recently, Chen et al. [4] proposed two sets of new features extracted from the difference domain of images. The first set of features, namely global probability feature set, was derived from the cumulative histograms of the first- and second-order difference images; and the second set of features, namely local correlation feature set, was computed based on the correlation between adjacent pixels in the difference images. They combined these two sets of features as the final global and local feature set (GLF). According to the experiments reported in [4], GLF can achieve good performance in detecting median filtering in low resolution and highly compressed images. Later, Kang et al. [19] found that median-filtered residuals (MFR) have a higher capability of suppressing both image textures and block artifacts than difference images. Then, they made a contribution to median filtering detection by extracting features from MFR. This is quite different from the earlier works, which usually extract features from the difference domain of images. Another contribution of [19] is to apply one-dimensional autoregressive (1D-AR) model to describe the difference of MFRs between median-filtered and non-filtered images. More recently, Zhang et al. [39] proposed to use high-order local ternary pattern (LTP), which is an improved local binary pattern, for median filtering detection. Their results have shown that LTP is more robust to JPEG compression than SPAM and MFF. Lai et al. [24] proposed to detect the presence of median filtering based on the local binary patterns in both the spatial and DCT domains. Qiu et al. [32] reported that spatial rich model (SRM) [13], originally proposed for universal steganalysis, can effectively detect median filtering when the test images have a large size (in their tests, the size of the images varies from 480×640 to 4288×4752) and are not compressed. However, the dimension of SRM features is too high. Different from the above-mentioned methods that use handcrafted features for detection, Chen et al. [3] are the first to introduce convolutional neural network (CNN) to the field of median filtering detection. Their proposed CNN-based method can learn discriminative features automatically from image samples, and outperforms the MFF [37], GLF [4], and 1D-AR [19] methods in detecting median filtering in 32×32 and 64×64 images. From an interesting point of view, Pasquini et al. [28] performed a theoretical analysis on the deterministic property of median filter, and then proposed a novel approach to detect the presence of median filtering in 1D signal, which can guarantee 0% false negative in theory. It is also worth to note that a few studies [6, 10, 12, 36, 40] have been done on anti-forensics and countering anti-forensics of median filtering.

In this paper, inspired by the work of Kang et al. [19], we propose a novel method for median filtering detection. We regard not only the median filtered residual (MFR) but also the average filtered residual (AFR) and Gaussian filtered residual (GFR) as forensic fingerprints. Considering that each of the filtered residuals should be a two-dimensional stationary random process, we propose to use 2D-AR model instead of 1D-AR model to characterize the three kinds of the filtered residuals separately. The AR coefficients of the three filtered residuals of an image are then used together as features to distinguish between median filtered and non-median filtered images. Our experimental results indicate that the proposed method outperforms existing methods in detecting median filtering in JPEG compressed images.

The rest of the paper is organized as follows. In Section 2, we present the proposed detection method in detail. Database and experimental methodology are described in Section 3. Then, in Section 4, experimental results are reported and discussed. Finally, conclusions are summarized in Section 5.

2 Proposed detection method

The proposed detection method is a learning-based method. It can be divided into two parts: the feature extraction, and the classifier training and testing. The second part is very similar to that applied in previous works. Therefore, in this section we focus on the description of the feature extraction. Before describing it, we first briefly introduce the median filtering applied to images.

2.1 Median filtering

Median filtering is an order statistic operation, which is implemented by replacing the gray value of each pixel in an image with the median of its neighboring pixels. Given an image $\mathbf{I}$ to be filtered, the process of median filtering can be formulated as

$$\hat{\mathbf{I}}(i, j) = \text{median} \{\mathbf{I}(i + r, j + s), \ (r, s) \in W\},$$
$$\text{for} \quad (i, j) \in (1, 2, \ldots, M) \times (1, 2, \ldots, N), \tag{1}$$

where $\mathbf{I}(i, j)$ denotes the pixel value of the image $\mathbf{I}$ at coordinate (i, j), W denotes the set of coordinates within a square window centered at $(0, 0)$, and M and N denote the image height and width, respectively. Since 3×3 and 5×5 windows are the most widely used for median filtering, this paper concentrates on the detection of 3×3 and 5×5 median filterings, as done in previous works [4, 19, 23, 37].

2.2 Feature extraction

In [19], Kang et al. proposed to extract features only from the median filtered residual (MFR). More generally, the proposed method extracts features from multiple filtered residuals, which are defined as

$$\mathbf{R}_t(i, j) = \mathbf{I}(i, j) - \hat{\mathbf{I}}_t(i, j), \tag{2}$$

where $\mathbf{I}$ and $\hat{\mathbf{I}}_t$ denote the original and filtered images, respectively, and the subscript t denotes the index of the filtered image. As a specific example, in this paper we let $t \in$ {"m", "a", "g"}, which represent median, average and Gaussian filterings, respectively. The window size of 3×3 can achieve a low computational complexity, and was adopted in [19].

For simplicity and a fair comparison with [19], in this paper the window size of the three kinds of filterings is also set as 3×3, and the mean and variance of the Gaussian filtering are set as 0 and 0.5, respectively. Therefore, as described above, not only the MFR used in [19], but also the average filtered residual (AFR) and the Gaussian filtered residual (GFR) are applied as the forensic fingerprints of median filtering. Note that compared with the first- and second-order differences of images, the filtered residuals have better capability of removing the interference caused by image edges and textures. Besides, the three kinds of filtered residuals are different from each other, thus it is possible to extract different distinguishing features from different kinds of filtered residuals.

In [19], 1D-AR model was utilized to fit the statistical properties of MFR in the horizontal and vertical directions. However, a digital image is a 2D signal, and median filtering is usually performed by sliding a square filtering window over the entire image. This results in that each pixel in a filtered image has strong correlations with its neighboring pixels in all directions. Therefore, 1D-AR model is not good enough to fit the statistical properties of MFR. In this paper, we consider each of the filtered residuals as a 2D stationary random process. Based on this, we propose to apply 2D-AR model to characterize the filtered residuals, and to extract the coefficients of the 2D-AR model as our features.

The 2D-AR model $AR(p, q)$ of each filtered residual $\mathbf{R}_t$ ($t \in \{$"m", "a", "g"$\}$) can be given by

$$\mathbf{R}_t(i, j) = \sum_{\substack{k=0 \\ k+l \neq 0}}^{p} \sum_{l=0}^{q} \mathbf{a}_t(k, l) \mathbf{R}_t(i - k, j - l) + \boldsymbol{\varepsilon}_t(i, j), \tag{3}$$
$$\text{for} \quad (i, j) \in (1, 2, \ldots, M) \times (1, 2, \ldots, N),$$

where $\mathbf{a}_t$ and ε_t denote the AR coefficient and prediction error matrices of $\mathbf{R}_t$, respectively, p and q ($p, q \in \mathbb{N}^+$) denote the orders of the 2D-AR model in the row and column directions, respectively, (k, l) indexes the elements of the matrix $\mathbf{a}_t$ of size $(p + 1) \times (q + 1)$, and $\mathbf{a}_t(0, 0)$ is specified as 1. Note that owing to the row and column symmetry of 3×3 and 5×5 median filterings, in this paper we use the AR model $AR(p, q)$ with $p = q$ to characterize the filtered residuals.

It is difficult to estimate the AR coefficient matrix $\mathbf{a}_t$ directly from (3). But fortunately, $\mathbf{a}_t$ can be obtained via the correlation method introduced by Kashyap [20]. According to the correlation method, the estimation of $\mathbf{a}_t$ is reduced to solving the following linear equation sets

$$\boldsymbol{\rho}_t(m, n) = \sum_{\substack{k=0 \\ k+l \neq 0}}^{p} \sum_{l=0}^{q} \mathbf{a}_t(k, l) \boldsymbol{\rho}_t(m - k, n - l), \tag{4}$$
$$\text{for} \quad (m, n) \in (0, 1, \ldots, p) \times (0, 1, \ldots, q), \ m + n \neq 0,$$

where $\boldsymbol{\rho}_t$ denotes the autocorrelation matrix of each filtered residual, which is calculated by

$$\boldsymbol{\rho}_t(m, n) = \begin{cases} \dfrac{\displaystyle\sum_{i=1}^{M-m} \sum_{j=1}^{N-n} \mathbf{R}_t(i, j) \mathbf{R}_t(i + m, j + n)}{(M - m) \times (N - n)} & \text{if} \quad m \geq 0, n \geq 0, \\[6mm] \dfrac{\displaystyle\sum_{i=1}^{M-m} \sum_{j=1-n}^{N} \mathbf{R}_t(i, j) \mathbf{R}_t(i + m, j + n)}{(M - m) \times (N - n)} & \text{if} \quad m > 0, n < 0, \end{cases} \tag{5}$$

and $\boldsymbol{\rho}_t(-m, -n) = \boldsymbol{\rho}_t(m, n)$.

Next, we investigate the statistical difference of the AR coefficients $\mathbf{a}_t(k, l)$ among MFR, AFR and GFR. Figure 1 shows the element values (except for $\bar{\mathbf{a}}_t(0, 0)$) of the average matrices $\bar{\mathbf{a}}_t$ of MFR, AFR and GFR. Each $\bar{\mathbf{a}}_t$ was averaged over the matrices $\mathbf{a}_t$ of 6000 images, and in Fig. 1 was rearranged into a vector by scanning it in zigzag order. Note that considering the trade-off between feature dimension and discriminability, this paper uses the $AR(p, q)$ model with $p = q = 6$ to calculate the AR coefficents $\mathbf{a}_t(k, l)$. It is shown in Fig. 1 that the average AR coefficients $\bar{\mathbf{a}}_t(k, l)$ among MFR, AFR and GFR are obviously different, whether for the original non-filtered images or for the median filtered images. Moreover, by comparing the plots in Fig. 1a with those in Fig. 1b and c, we can see that not only for MFR but also for AFR and GFR, the average AR coefficients of the non-filtered images take quite different values from those of the median filtered images. The above findings indicate that the AR coefficients have different statistical characteristics among the three different kinds of filtered residuals, and the AR coefficients of any kind of the filtered residuals have the ability to distinguish the median filtered from the non-filtered images.

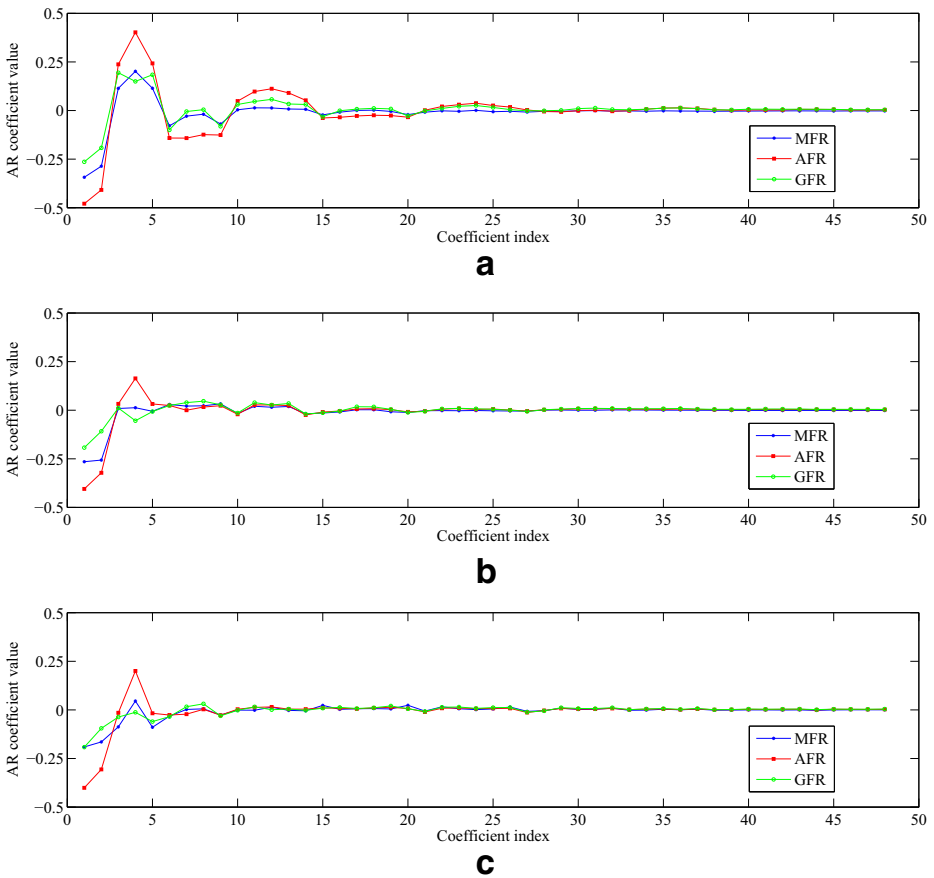**Fig. 1** Comparison of the average AR coefficients $\bar{\mathbf{a}}_t(k, l)$ among MFR, AFR and GFR. The filtered residuals are calculated from **a** original non-filtered images, **b** 3×3 median filtered images, and **c** 5×5 median filtered images. The model $AR(6, 6)$ was used to calculate the AR coefficients, and each $\bar{\mathbf{a}}_t$ was averaged over 6000 images and sorted in zig-zag order

Hence, it is reasonable to combine the three sets of the AR coefficients extracted from MFR, AFR and GFR as distinguishing features. Figure 1 also shows that the AR coefficient values of both the median filtered and non-filtered images almost vanish when their indices are larger than about 30. This means the AR coefficients $\mathbf{a}_t(k, l)$ have almost no distinguishability when $k + l > 6$. Therefore, only the AR coefficients $\mathbf{a}_t(k, l)$ with $k + l \leq 6$ and $k + l \neq 0$ (i.e., except for $\mathbf{a}_t(0, 0)$) are chosen as features.

The 2D-AR coefficients of the MFR $\mathbf{a}_m(k, l)$, for $k + l \leq 6$ and $k + l \neq 0$, can be divided into three parts: the horizontal coefficients, the vertical coefficients and the rest coefficients (the fifteen coefficients labeled in blue-filled circles as shown in Fig. 2). The horizontal and vertical 2D-AR coefficients of the MFR are closely correlated to the 1D-AR coefficients of the MFR [22], whereas the rest 2D-AR coefficients of the MFR are the additional coefficients that 1D-AR cannot model. In order to validate the effectiveness of these additional 2D-AR coefficients, we visualize them by projecting them onto a 2D plane with linear discriminant analysis (LDA), as shown in Fig. 3. It is seen that the additional 2D-AR coefficients of the original non-filtered images, the 3×3 median filtered images, and the 5×5 median filtered images are clearly distinguishable from each other, indicating that the 2D-AR model is more suitable to describe the statistical properties of the filtered residuals than the 1D-AR model of [19].

2.3 Summary of the proposed method

Based on the above description, we summarize the proposed method as follows.

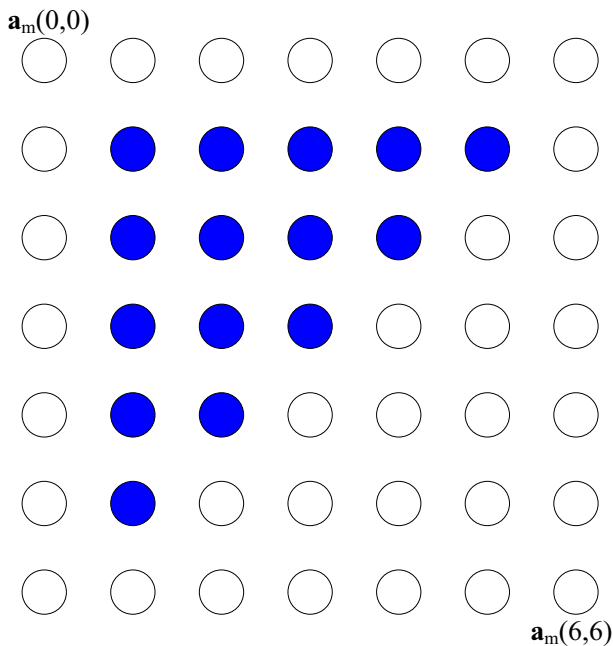1. Generate the MFR, AFR and GFR for each input image according to (2).

Fig. 2 Coordinates of the AR coefficients $\mathbf{a}_m(k, l)$ with $k + l \leq 6$ and $kl \neq 0$ (marked by *blue-filled circles*)
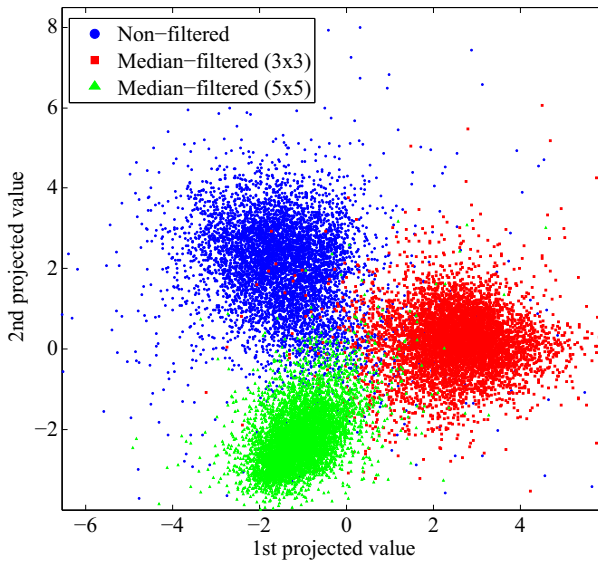
Fig. 3 Visualization of the AR coefficients $\mathbf{a}_m(k, l)$ with $k + l \le 6$ and $kl \ne 0$. The *blue*, *red*, and *green points* represent the AR coefficients of the original non-filtered, the 3×3 median filtered, and the 5×5 median filtered images, respectively. Here, the fifteen AR coefficients as labeled in Fig. 2 were projected onto the 2D plane with LDA, and each of the *blue*, *red*, and *green point* sets consists of 6000 sample points

2. Calculate the three AR coefficient matrixes $\mathbf{a}_t$ ($t \in$ {"m", "a", "g"}) for the MFR, AFR and GFR, respectively, according to (4) and (5).

3. For each of the three AR coefficient matrices $\mathbf{a}_t$ ($t \in$ {"m", "a", "g"}), choose the elements $\mathbf{a}_t(k, l)$ with $k + l \le 6$ and $k + l \ne 0$ as a set of features. Then the three sets of features are combined into the proposed feature set.

4. Train SVM classifier with the proposed feature sets extracted from median filtered and non-median filtered images. Finally, the trained classifier is used for median filtering detection.

Note that, as mentioned above, the order of the $AR(p, q)$ model is set as $(p, q) = (6, 6)$ in this paper, thus there are 27 features extracted from each $\mathbf{a}_t$, and totally 81 features in the proposed feature set.

The time complexity of the proposed method can be considered from two aspects, namely the feature extraction and the classifier training. The feature extraction of the proposed method consists of two parts: the generation of the filtered residuals and the computation of 2D-AR coefficients. The time complexities of the two parts are dominated by image filtering and autocorrelation computation, respectively. Both image filtering and autocorrelation computation can be implemented to run efficiently. Therefore, the feature extraction of the proposed method is not very time-consuming. It is known that the time complexity of SVM classifier training is related to the dimension of features. The dimension of the proposed features is not large compared with existing features. Therefore, the classifier training of the proposed method is also not time-consuming.

3 Database and experimental methodology

In order to evaluate the performance of the proposed method, we build a database of 6000 images. These images are taken from three different image databases, namely NRCS [27], BOWS2 [15] and BOSS [1], each of which provides 2000 images. Each image in our database is centrally cropped to a size of 512×512 pixels, and then converted to an 8-bit grayscale image. 3×3 and 5×5 median filterings are performed on each of the cropped grayscale images to generate 2×6000 median-filtered images. For testing the robustness of the proposed method, the filtered and non-filtered images are further central-cropped into sizes of 256×256, 128×128, and 64×64, and all of the cropped grayscale images are then compressed with quality factors (QFs) 90, 70, 50, and 30, respectively.

Five existing methods, including SPAM [23], MFF [37], GLF [4], 1D-AR [19], and LTP [39], are implemented for comparison. All these compared methods applied the SVM classifier for training and detection. In order to fairly compare our proposed feature set with the above five ones, the SVM classifier is also chosen in this paper, as done in [4, 19, 23, 37, 39]. 40% of the images are used for training, and the remaining 60% are used for testing. A five-fold cross validation is performed in the training to search the two hyper-parameters of the SVM classifier over the following grid:

$$(C, \gamma) \in \left\{ (2^i, 2^j) | i \in \{-5, -4, \cdots, 4\}, j \in \{-13, -12, \cdots, 13\} \right\}. \tag{6}$$

Note that the feature dimensions of SPAM, MFF, GLF, 1D-AR, LTP, and the proposed method are 686, 44, 56, 10, 2048 and 81, respectively. The time consumption of SVM classifier training is related to feature dimension. 1D-AR has only 10 features, which makes the training of its SVM classifier very fast. MFF, GLF, and the proposed method have less than 100 features, and the trainings of their SVM classifiers can be performed in an acceptable time. Compared to the other four methods, SPAM and LTP have a large number of features, so their classifier trainings are quite time-consuming.

The performance of each classifier is mainly evaluated by the minimum average decision error under equal probability of the positive and negative samples (i.e., the median-filtered and non-median filtered images), which is expressed by

$$P_e = \min \left(\frac{P_{fp} + 1 - P_{tp}}{2} \right), \tag{7}$$

where P_{fp} and P_{tp} denote the false positive and true positive rates, respectively. Besides, in some cases we also plot receiver operating characteristic (ROC) curves to illustrate the tradeoff between the false positive and true positive rates.

4 Experimental results

Extensive experiments have been carried out to investigate the detection performance of the proposed method on the above-described database.

4.1 Performance evaluation of different sets of the proposed features

As described in Section II, the proposed 2D-AR features consist of three sets of features extracted from MFR, AFR and GFR, respectively. In this subsection, we present

experiments to evaluate the performance of each of the three sets of features, as well as the superiority of the combination of the three sets over either set alone. The experiments are conducted on the images with various image sizes and different QFs of JPEG compression. Table 1 lists the detection results obtained by using the three sets of features from MFR, AFR, and GFR (denoted as 2D-AR-MFR, 2D-AR-AFR, and 2D-AR-GFR, respectively) and the combination of the three feature sets (denoted as 2D-AR-Comb). The detection results of the 1D-AR features [19] are also reported for comparison. In Table 1 and the following tables, "MF3 (or MF5) vs ORI" denotes that the positive and negative samples used for the tests are the 3×3 (or 5×5) median-filtered images and the original non-filtered images, respectively, and "UnC" denotes that the test images are not compressed. It can be seen from Table 1 that all of the three sets of the proposed features, i.e., 2D-AR-MFR, 2D-AR-AFR, and 2D-AR-GFR, have the capability to detect median filtering, especially when the test images have a large size and are not (or slightly) compressed. Moreover, 2D-AR-Comb is superior to any of the three sets of the proposed features in all the tests. This indicates that not only the features extracted from MFR, but also those from AFR and GFR contribute to the superior performance of the proposed feature set. It is also shown in Table 1 that 2D-AR-MFR outperforms 1D-AR in all the tests, which validates that 2D-AR model is more suitable than 1D-AR model to fit the statistical property of the filtered residuals.

Table 1 Detection results of the three sets and the whole of the proposed features

Size	QF	MF3 vs ORI					MF5 vs ORI				
		1D-AR	2D-AR				1D-AR	2D-AR			
		[19]	MFR	AFR	GFR	Comb	[19]	MFR	AFR	GFR	Comb
512	UnC	0.65	0.38	0.97	1.15	**0.13**	0.33	0.18	1.04	1.08	**0.18**
	90	1.49	1.10	2.04	2.00	**1.06**	0.49	0.45	1.90	1.85	**0.36**
	70	3.60	3.31	5.47	5.54	**2.44**	1.38	1.31	3.22	3.08	**1.07**
	50	6.18	5.11	7.67	7.17	**3.86**	2.21	1.93	4.04	4.25	**1.78**
	30	9.64	6.97	10.00	9.94	**6.21**	2.92	2.60	5.14	5.25	**2.33**
256	UnC	1.11	0.67	1.65	1.78	**0.33**	0.78	0.47	1.57	1.63	**0.26**
	90	2.33	2.11	3.40	3.53	**1.46**	1.03	0.81	2.72	2.42	**0.64**
	70	5.82	5.32	7.58	7.50	**4.56**	2.67	2.03	4.96	4.94	**1.99**
	50	9.31	7.43	10.83	10.64	**6.46**	3.90	3.27	6.42	5.89	**3.13**
	30	13.28	10.50	13.72	13.44	**9.83**	5.26	4.19	7.53	7.36	**3.85**
128	UnC	2.49	1.61	2.43	2.53	**0.85**	1.79	0.93	2.10	2.18	**0.60**
	90	4.49	4.04	5.82	5.94	**3.11**	2.83	1.83	4.28	4.32	**1.39**
	70	10.19	8.67	11.90	11.71	**7.47**	5.58	4.26	7.33	7.71	**3.64**
	50	14.06	12.32	15.17	14.90	**10.71**	7.28	5.97	9.21	9.07	**5.36**
	30	18.00	14.83	18.31	17.76	**14.72**	9.38	7.36	11.11	10.36	**6.69**
64	UnC	4.97	3.76	3.25	3.32	**1.40**	3.56	2.33	3.11	3.03	**1.49**
	90	9.26	7.81	10.00	9.99	**6.14**	6.15	3.79	6.89	6.57	**3.17**
	70	15.86	13.99	17.22	17.07	**12.82**	10.79	8.26	11.13	10.99	**6.69**
	50	19.81	17.63	20.86	19.97	**16.89**	13.33	11.25	12.86	12.31	**9.04**
	30	23.24	22.04	23.07	22.94	**20.71**	15.68	12.82	15.04	14.61	**11.15**

Here, these detection results are reported with the minimum error P_e (%). The best results are marked in bold

4.2 Performance comparison with existing methods

A set of experiments have been performed to compare the proposed method with existing methods, including SPAM [23], MFF [37], GLF [4], 1D-AR [19], and LTP [39]. The comparative results are given in Table 2. We can see that the six compared methods generally perform better in "MF5 vs ORI" than in "MF3 vs ORI". This is not surprising since 5×5 median filtering leaves more significant filtering artifacts and thus is more easily detected than 3×3 median filtering. We can also see that, in general, the performance of these detection methods becomes worse as the image size decreases. The same happens as the QF of the JPEG compression goes down. These are consistent with the following two facts: first, the smaller the image size is, the less the information can be used for feature extraction; and second, the compression can disturb the artifacts left by the median filtering, and the more severe the compression is, the larger the disturbance becomes. It is also shown in Table 2 that for the uncompressed images, all of the six compared methods perform quite well in the median filtering detections. LTP [39] and SPAM [23] are clearly superior to the other four compared methods, but have much larger dimensions of features, especially the dimension of the LTP features is more than 2000. The proposed 2D-AR-Comb method has a detection

Table 2 Comparison between the proposed and five existing methods in distinguishing the median filtered images from the original non-filtered images

Size	QF	MF3 vs ORI						MF5 vs ORI					
		SPAM [23]	MFF [37]	GLF [4]	1D-AR [19]	LTP [39]	2D-AR -Comb	SPAM [23]	MFF [37]	GLF [4]	1D-AR [19]	LTP [39]	2D-AR -Comb
512	UnC	0.03	0.10	0.10	0.65	**0.01**	0.13	0.04	0.03	0.18	0.33	**0.01**	0.18
	90	1.18	4.97	3.36	1.49	**1.00**	1.06	0.40	3.94	2.78	0.49	**0.29**	0.36
	70	5.58	10.68	5.10	3.60	5.56	**2.44**	1.68	6.07	3.53	1.38	2.04	**1.07**
	50	9.53	14.10	6.40	6.18	8.76	**3.86**	2.46	7.21	4.11	2.21	3.21	**1.78**
	30	14.44	16.01	8.56	9.64	12.29	**6.21**	4.28	7.68	4.44	2.92	4.96	**2.33**
256	UnC	0.06	0.18	0.18	1.11	**0.01**	0.33	0.06	0.10	0.24	0.78	**0.01**	0.26
	90	1.60	6.15	5.24	2.33	1.89	**1.46**	0.69	4.94	4.03	1.03	**0.60**	0.64
	70	8.28	12.89	7.93	5.82	7.56	**4.56**	2.61	8.24	5.46	2.67	3.26	**1.99**
	50	12.33	15.26	9.49	9.31	10.74	**6.46**	3.92	9.04	5.25	3.90	4.57	**3.13**
	30	16.97	18.21	11.65	13.28	15.68	**9.83**	6.35	9.44	5.97	5.26	6.99	**3.85**
128	UnC	**0.10**	0.40	0.35	2.49	**0.10**	0.85	0.11	0.29	0.72	1.79	**0.06**	0.60
	90	3.35	9.57	7.24	4.49	3.64	**3.11**	**1.14**	6.90	5.51	2.83	1.61	1.39
	70	12.67	16.11	10.76	10.19	11.03	**7.47**	4.65	10.46	6.83	5.58	6.24	**3.64**
	50	16.74	18.21	12.86	14.06	15.99	**10.71**	6.81	11.28	7.39	7.28	8.06	**5.36**
	30	20.83	21.29	17.22	18.00	20.46	**14.72**	9.44	12.38	8.53	9.38	10.57	**6.69**
64	UnC	**0.17**	0.56	0.75	4.97	0.31	1.40	**0.19**	0.75	0.76	3.56	0.26	1.49
	90	7.08	13.88	10.99	9.26	7.39	**6.14**	**2.93**	9.79	7.89	6.15	4.26	3.17
	70	18.21	19.60	15.04	15.86	18.10	**12.82**	8.00	13.29	9.25	10.79	9.69	**6.69**
	50	21.76	22.32	18.53	19.81	22.14	**16.89**	10.78	14.07	10.43	13.33	12.47	**9.04**
	30	25.24	24.88	22.19	23.24	25.49	**20.71**	13.07	15.94	11.53	15.68	15.58	**11.15**

Here, the detection results are reported with the minimum error P_e (%). The best results are marked in bold

error of less than 1.00% when the size of the test images is larger than 64×64. Whereas, for the compressed images, most of the compared methods no longer perform very well, particularly when the image size is small and "MF3 vs ORI" is tested. In contrast to the other compared methods, the proposed method is relatively robust to JPEG compression. 2D-AR-Comb achieves the best performance among the six methods in all cases that the QF of the compression is not larger than 70, and also in some cases when the QF takes 90. The superiority of 2D-AR-Comb over the other five methods is evident when the test images are heavily compressed. For example, when the image size and the QF are 512×512 and 50, respectively, and "MF3 vs ORI" is tested, 2D-AR-Comb has a detection error of 3.86%, which is 37.5% lower than that of 1D-AR (i.e., the best result of the other five methods). The ROC curves of some representative tests are also provided in Fig. 4. It can be seen that these results represented by the ROC curves are consistent with those reported in Table 2.

4.3 Discrimination between median filtering and other manipulations

For median filtering detection, a good classifier should have the ability to distinguish the median filtered images not only from original images but also from those processed by some other image manipulations, such as, Gaussian filtering, average filtering, and image resizing. In order to assess this ability of the proposed method, a set of experiments have also been conducted, including "MF vs AVE", "MF vs GAU", "MF vs D-SCA" and "MF vs U-SCA". Note that "MF vs AVE (or GAU, D-SCA, and U-SCA)" denotes that the positive and negative samples are the median filtered images and those processed by an average filtering with a window size of 3×3 (or a Gaussian filtering with a standard deviation of 0.5, a downscaling with a factor of 0.9, and an upscaling with a factor of 1.1). In practice, we may not know the kind of manipulation to be distinguished from median filtering, as discussed in [19]. To simulate a more practical scenario, we created a set of negative samples, denoted as "ALL", which consists of equal amounts of the original images, the Gaussian filtered images, the average filtered images, the downscaled images, and the upscaled images. Then the experiments of "MF vs ALL" were also conducted. For brevity, we only report the results for the case that the test images are compressed with a QF value of 70. It is interesting to note from Table 3 that GLF [4] performs quite well (it obtains the second best performance in general) in 3×3 median filtering detection, but has no superiority over most of the compared methods in 5×5 median filtering detection. Note that here 3×3 (or 5×5) median filtering detection refers to the classification between 3×3 (or 5×5) median filtering and other manipulations including Gaussian filtering, average filtering and image resizing. In contrast to GLF, the proposed 2D-AR-Comb method always achieves the best results among the six compared methods in all the tests. Moreover, the improvement of the proposed method is often significant when compared with SPAM [23], MFF [37], 1D-AR [19], and LTP [39] in detecting 3×3 median filtering. The good performance of the proposed method can also be seen from Fig. 5, where the ROC curves of some representative tests are reported.

4.4 Discrimination between 3×3 and 5×5 Median filterings

To identify the parameter of an image manipulation is another important issue for image forensics. Hence, we have conducted a set of experiments to examine the ability of the proposed method in identifying the window size of median filtering (i.e., in distinguishing between 3×3 and 5×5 median filterings). In this set of experiments, the positive and
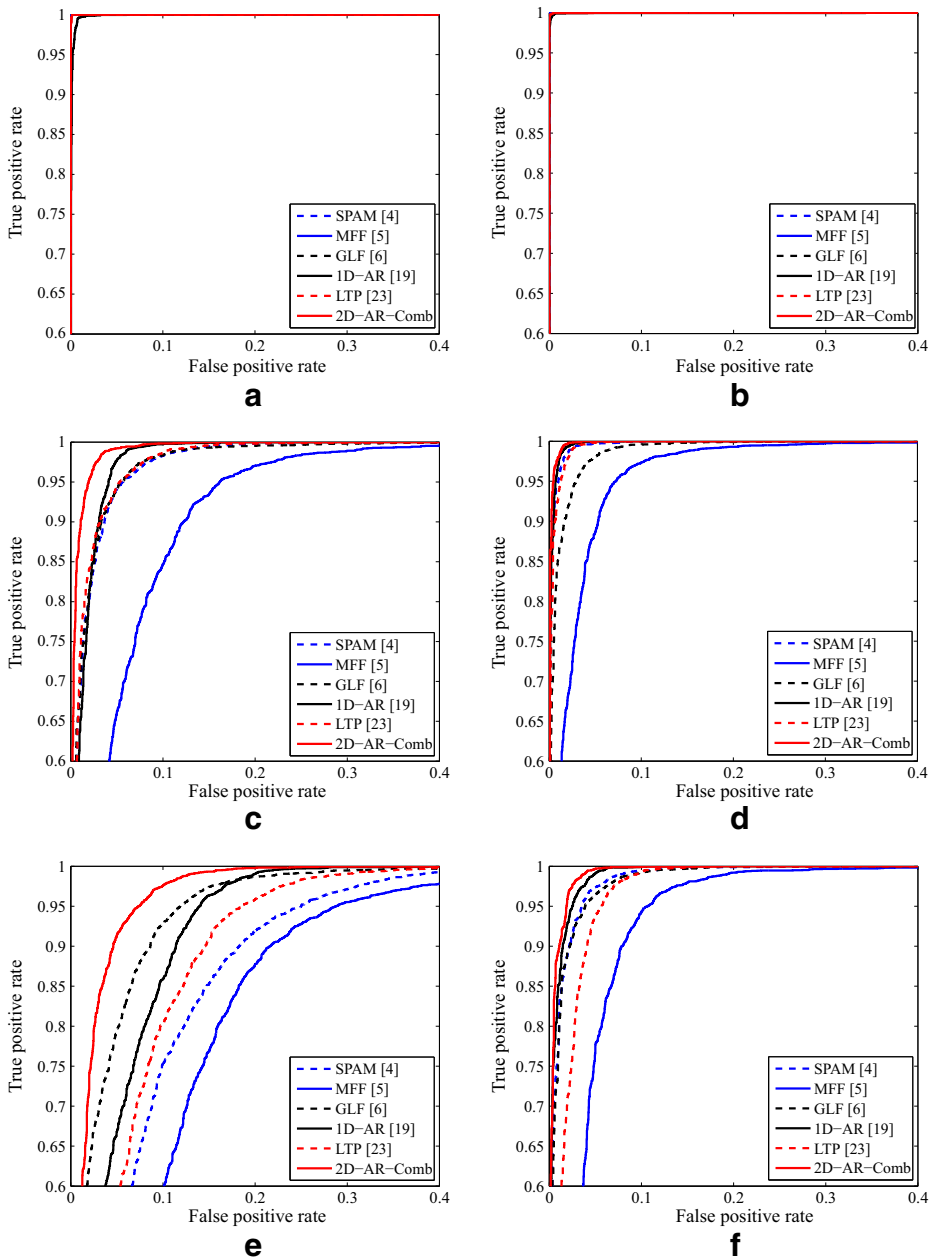
Fig. 4 ROC curves of the proposed and five existing methods for distinguishing the median filtered images from the non-filtered images. **a** MF3 vs ORI, and QF="UnC"; **b** MF5 vs ORI, and QF="UnC"; **c** MF3 vs ORI, and QF=70; **d** MF5 vs ORI, and QF=70; **e** MF3 vs ORI, and QF=30; **f** MF5 vs ORI, and QF=30. Here, all of the test images have a size of 512×512

Table 3 Comparison between the proposed and five existing methods in distinguishing between median filtering and other manipulations

	Size	MF3						MF5					
		SPAM	MFF	GLF	1D-AR	LTP	2D-AR	SPAM	MFF	GLF	1D-AR	LTP	2D-AR
		[23]	[37]	[4]	[19]	[39]	-Comb	[23]	[37]	[4]	[19]	[39]	-Comb
MF	512	4.56	7.69	3.07	8.15	4.88	**1.29**	1.83	4.56	1.75	2.74	2.11	**0.46**
vs	256	7.06	12.08	4.72	10.96	7.31	**2.29**	2.57	7.43	2.58	5.57	4.00	**1.14**
AVE	128	11.24	18.51	7.14	15.60	11.72	**5.07**	5.46	13.43	4.89	11.25	7.43	**2.63**
	64	18.00	25.43	12.04	21.96	19.42	**10.38**	12.90	19.99	9.25	18.42	14.11	**6.50**
MF	512	8.76	14.00	5.60	5.93	7.51	**3.43**	2.00	7.93	3.54	1.31	2.40	**1.06**
vs	256	12.18	18.22	8.31	9.33	10.63	**4.72**	2.86	10.69	4.81	3.62	2.51	**2.10**
GAU	128	18.81	24.60	13.19	14.93	15.32	**9.79**	5.58	14.69	7.92	5.38	7.12	**4.10**
	64	26.74	30.76	19.60	21.92	24.06	**17.31**	10.00	18.90	11.74	10.22	12.00	**7.82**
MF	512	8.03	13.63	5.31	4.85	6.40	**2.88**	2.04	7.53	3.69	1.32	2.35	**0.93**
vs	256	10.24	15.97	7.46	7.10	9.35	**4.56**	2.42	9.47	4.36	2.25	3.51	**1.93**
D-SCA	128	16.36	21.96	11.33	11.61	13.72	**8.25**	4.58	12.99	6.26	4.61	6.25	**3.57**
	64	23.49	26.97	18.03	19.17	22.43	**15.53**	9.18	16.67	10.71	9.97	10.54	**7.44**
MF	512	7.64	13.44	5.18	7.54	6.56	**3.06**	1.75	8.25	3.17	1.24	2.08	**1.17**
vs	256	11.94	18.54	8.29	11.74	10.21	**5.69**	2.86	12.21	4.94	2.68	3.62	**1.99**
U-SCA	128	19.82	25.83	12.82	19.03	16.79	**10.74**	5.88	15.96	7.44	5.51	7.22	**4.53**
	64	26.94	32.61	20.53	27.32	24.82	**19.22**	10.86	20.13	12.28	10.76	12.47	**8.36**
MF	512	9.56	14.08	5.72	8.82	7.33	**3.14**	2.56	9.11	3.58	2.76	3.50	**1.35**
vs	256	12.26	18.43	8.47	12.44	11.01	**5.79**	4.01	12.51	5.43	4.46	5.53	**2.24**
ALL	128	19.24	27.69	13.15	18.38	17.94	**10.00**	7.61	18.18	8.54	8.67	9.25	**5.13**
	64	28.35	34.31	20.25	26.33	26.99	**18.50**	13.50	23.96	13.22	15.85	16.75	**9.83**

Here, the detection results are reported with the minimum error P_e (%), and the test images were compressed with QF = 70. The best results are marked in bold

negative samples are the 3 × 3 and 5 × 5 median filtered images, respectively, and all of the filtered images are compressed with QF = 70, 50. The detection results are presented in Table 4 and Fig. 6. It is shown that the proposed method performs very well in these tests, and still achieves the lowest detection error among the six compared methods.

4.5 Performance comparison on more image databases

We have also conducted a set of experiments to compare the proposed method with the existing methods on three more image databases, namely, UCID [34], RAISE2k [8], and DRESDEN [16]. The UCID database has 1338 images with size of 384 × 512; the RAISE2k database has 2000 images with size of 4928 × 3264 or 4288 × 2848; and the DRESDEN database has 1388 images with size of 3872 × 2592 or 3008 × 2000. We centrally cropped each images of these databases into the size of 256 × 256. Then, the cropped images were used as the test images. The detection error rates of the six compared methods are listed
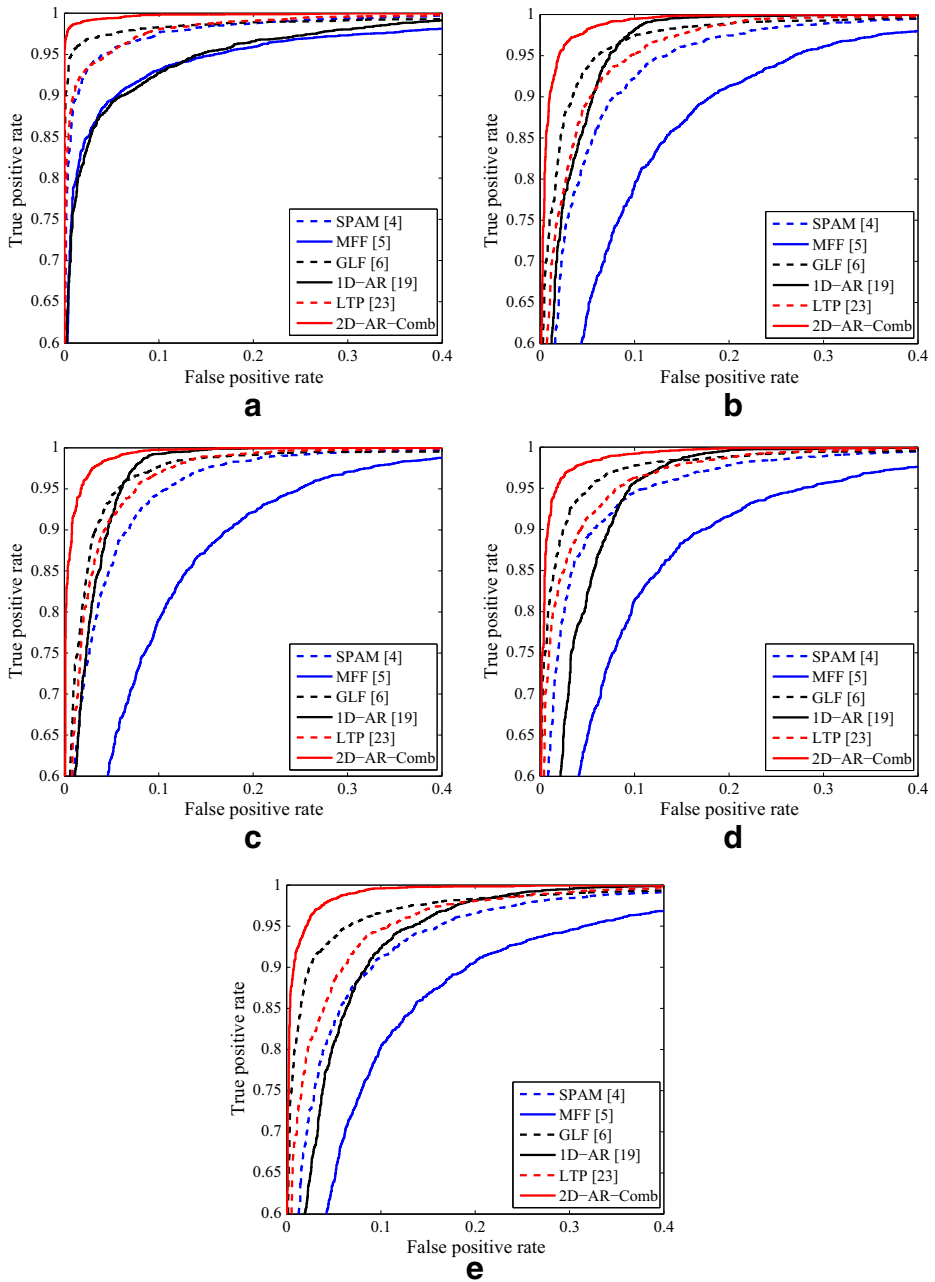
Fig. 5 ROC curves of the proposed and five existing methods for distinguishing between median filtering and other manipulations. **a** MF3 vs AVE; **b** MF3 vs GAU; **c** MF3 vs D-SCA; **d** MF3 vs U-SCA; **e** MF3 vs ALL. Here, all of the test images have a size of 512×512, and were compressed with QF=70

Table 4 Comparison between the proposed and five existing methods in distinguishing between 3×3 and 5×5 median filterings

	Size	QF=70						QF=50					
		SPAM [23]	MFF [37]	GLF [4]	1D-AR [19]	LTP [39]	2D-AR -Comb	SPAM [23]	MFF [37]	GLF [4]	1D-AR [19]	LTP [39]	2D-AR -Comb
MF3	512	3.42	19.39	10.99	1.92	3.79	**1.82**	4.97	19.14	10.39	3.18	5.97	**2.68**
vs	256	5.24	20.42	14.26	4.10	6.35	**3.65**	8.11	18.99	13.75	6.39	8.90	**4.90**
MF5	128	9.72	23.22	17.36	7.58	12.00	**6.93**	13.56	22.71	17.29	10.57	15.17	**9.18**
	64	16.43	26.50	21.68	14.33	18.68	**12.29**	19.72	25.62	21.32	17.74	22.01	**16.57**

Here, the detection results are reported with the minimum error P_e (%), and the test images were compressed with QF = 70, 50. The best results are marked in bold

in Table 5. It is shown that the compared methods perform better on the UCID database than on the RAISE2k and the DRESDEN databases. This may be explained as follows. The images from UCID database have a much smaller size than those from the RAISE2k and the DRESDEN databases, which results in that the 3×3 (or 5×5) local neighbour-hoods of the UCID images usually appear less smooth than those of the RAISE2k and the DRESDEN images. Therefore, after median filtering, the filtering artifacts left on UCID images are usually more obvious than those on the RAISE2k and the DRESDEN images, hence the median filtering is more easily detected in the UCID images than in the RAISE2k and the DRESDEN images. In addition, Table 5 also shows that the proposed method still outperforms the other compared methods in most of the test cases.

4.6 Image forgery detection via median filtering detection

Finally, we present an experiment to illustrate the application of median filtering detection in image forgery detection. One scenario of image forgery can be described as follows:
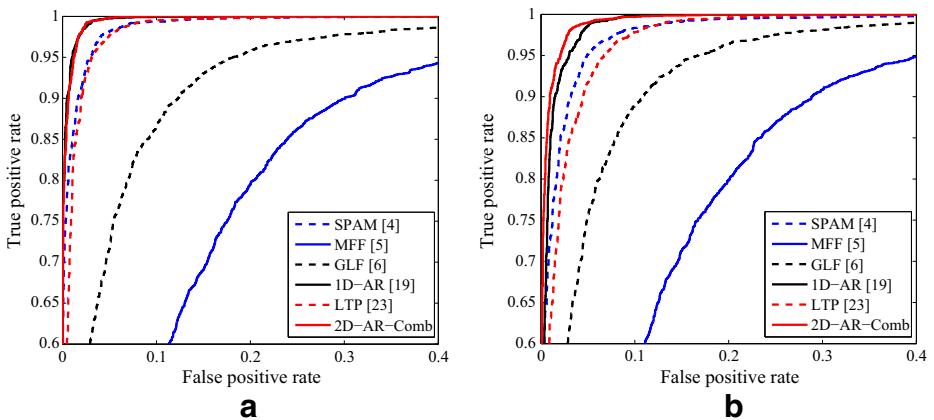

Fig. 6 ROC curves of the proposed and five existing methods for distinguishing between 3×3 and 5×5 median filterings. **a** MF3 vs MF5, and QF=70; **b** MF3 vs MF5, and QF=50. Here, all of the test images have a size of 512×512, and were compressed with QF=70

Table 5 Performance comparison between the proposed and five existing methods on UCID, RAISE2k, and DRESDEN databases

Database	QF	MF3 vs ORI						MF5 vs ORI					
		SPAM [23]	MFF [37]	GLF [4]	1D-AR [19]	LTP [39]	2D-AR -Comb	SPAM [23]	MFF [37]	GLF [4]	1D-AR [19]	LTP [39]	2D-AR -Comb
UCID	UnC	0.12	0.12	0.12	1.81	**0.06**	0.56	0.25	**0.12**	0.31	1.12	0.25	0.93
	90	1.12	4.42	3.86	1.99	**1.06**	1.25	1.12	3.36	2.80	1.49	**0.87**	1.00
	70	7.41	7.72	6.16	3.42	6.66	**2.37**	2.62	5.60	4.61	2.37	3.80	**1.56**
	50	12.14	10.83	6.54	3.99	10.09	**3.67**	4.23	6.04	4.98	3.05	5.29	**2.30**
	30	13.89	11.39	7.72	5.73	14.13	**4.86**	5.35	5.98	4.98	4.61	6.91	**2.43**
RAISE2k	UnC	0.42	0.42	0.88	5.17	**0.25**	0.79	0.33	0.42	0.63	1.01	**0.29**	0.50
	90	3.79	11.63	7.04	7.75	3.21	**2.38**	1.67	6.63	3.88	1.04	2.13	**0.75**
	70	11.46	23.04	12.46	10.88	10.75	**8.21**	3.92	10.96	5.08	2.96	4.13	**2.67**
	50	20.50	31.17	17.67	17.79	18.33	**12.33**	6.38	15.04	8.21	5.50	8.50	**4.33**
	30	31.04	35.33	24.63	26.38	28.54	**20.00**	11.46	20.08	11.13	9.42	12.67	**7.42**
DRESDEN	UnC	0.11	**0.06**	0.17	6.22	0.17	0.45	**0.06**	0.11	0.11	1.90	**0.06**	0.22
	90	**0.17**	8.72	3.13	9.78	0.28	1.23	**0.06**	5.64	1.06	1.96	**0.06**	0.73
	70	**5.70**	19.94	11.06	16.31	10.45	8.32	**0.95**	9.89	5.47	3.52	2.68	1.96
	50	10.06	20.16	11.23	16.98	11.68	**8.99**	3.52	11.40	6.03	6.93	4.53	4.53
	30	32.23	32.07	25.14	27.60	31.90	**20.61**	11.73	17.37	16.37	13.69	15.87	**11.51**

The detection results are reported with the minimum error P_e (%). The best results are marked in bold

Given two images, one of which is a median filtered image and the other is a non-filtered image, as shown in Fig. 7a and b, respectively. A forged image is generated by pasting a region cut from the median filtered image into the non-filtered image and then compressing it with QF=70, as shown in Fig. 7c. To identify the cut-and-paste region in Fig. 7c, we first divide the forged image in Fig. 7c into blocks of size 64×64. Then, the SVM classifiers that have been trained in Subsection IV-B are used to make binary decision on whether each of the blocks is a median filtered block or not. According to the outputs of the classifiers for all of the blocks, we generate six binary maps for the six compared methods, respectively, as shown in Fig. 8. For reference, the ground truth of the forged region is already given in Fig. 7d. The miss detection rates (i.e., the percentage of wrongly classifying median filtered blocks as non-median filtered ones) of SPAM, MFF, GLF, 1D-AR, LTP, and the proposed method are 26.92%, 14.93%, 31.42%, 24.53%, 25.25%, and 22.81%, respectively. The false alarm rates (i.e., the percentage of wrongly classifying non-median filtered blocks as median filtered ones) of the six methods are 14.07%, 17.29%, 8.12%, 8.72%, 9.92%, and 6.95%, respectively. On the whole, the error rates of the six methods are 20.49%, 16.11%, 19.77%, 16.63%, 17.58%, and 14.88%, respectively. It is shown that among the six compared methods, the proposed method has the second lowest miss detection rate and the lowest false alarm rate, which leads to the lowest error rate in detecting the image forgery. Note that the detection errors mainly occur at the top left corner of the forged image for all the six compared methods. The reason might be that the top left corner of the forged image is a low-contrast region so that the statistical traces left by median filtering are too weak to be detected.
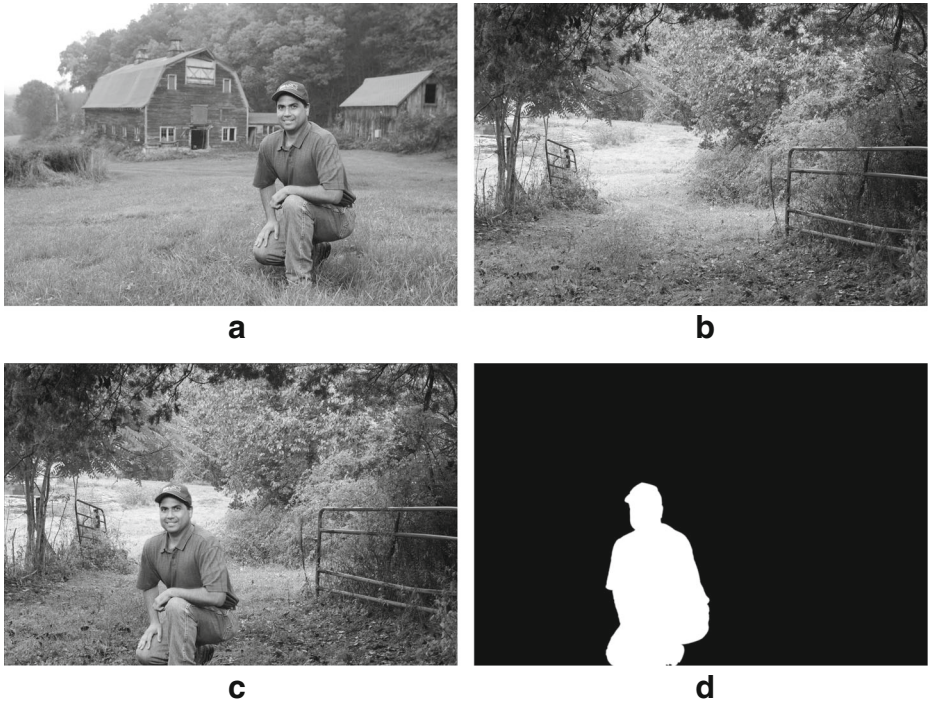
Fig. 7 Original and forged images. **a** First original image (3×3 median-filtered); **b** Second original image (non-filtered); **c** Forged image; **d** Ground truth of the forgery. Here, all of the original and forged images has a size of 1800×2700
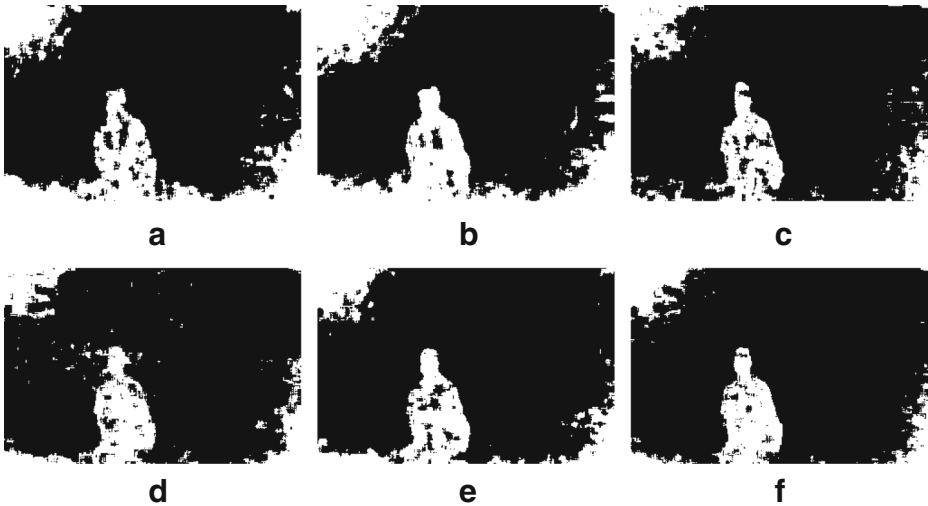

Fig. 8 Forgery detection results of the proposed and five existing methods. **a** SPAM [23]; **b** MFF [37]; **c** GLF [4]; **d** 1D-AR [19]; **e** LTP [39]; **f** Proposed method

5 Conclusion

Inspired by the work of Kang et al. [19], in this paper we have proposed a novel method for median filtering detection. The main contributions of this paper are as follows:

First, considering that the filtered residuals are two-dimensional stationary random processes, we have proposed to apply 2D autoregressive model instead of 1D-AR model to characterize the filtered residuals. According to our experiments, 2D-AR model is more effective than 1D-AR model to fit the filtered residuals.

Second, we found that, in addition to median filtered residual (MFR), average filtered residual (AFR) and Gaussian filtered residual (GFR) also exhibit different characteristics between median filtered images and non-median filtered images. Hence, we use not only MFR but also AFR and GFR as fingerprints for median filtering detection. Our experiments have validated that the use of more kinds of filtered residuals can evidently improve the detection performance.

Extensive experiments have been conducted to compare the proposed method with five existing methods. The results show that the proposed method performs quite well in all the tests. It is often superior to the existing methods in detecting median filtering in heavily compressed images. Although only grayscale natural images have been tested in out experiments, other types of images, such as medical images, should be modified in a similar way after median filtering, and hence leave similar statistical traces. With appropriate modification, the proposed method should be applicable to other types of images. Our future work includes investigating the discriminability of other residuals to detect median filtering, and countering the anti-forensics of median filtering detection. Moreover, it is feasible to combine the proposed feature set with other existing feature sets, such as the 1D-AR feature set [19], to achieve improved performance, which could also be an interesting future work.

References

1. Bas P, Filler T, Pevný T (2011) Break our steganographic system: the ins and outs of organizing BOSS. In: Information hiding. Springer, Berlin, pp 59–70
2. Bovik AC (1987) Streaking in median filtered images. IEEE Trans Acoust Speech Signal Process 35(4):493–503
3. Chen J, Kang X, Liu Y, Wang ZJ (2015) Median filtering forensics based on convolutional neural networks. IEEE Signal Process Lett 22(11):1849–1853
4. Chen C, Ni J, Huang J (2013) Blind detection of median filtering in digital images: a difference domain based approach. IEEE Trans Image Process 22(12):4699–4710
5. Cao G, Zhao Y, Ni R, Yu L, Tian H (2010) Forensic detection of median filtering in digital images. In: Proceedings IEEE Int. Conf. Multimedia and Expo, pp 89–94
6. Dang-Nguyen D-T, Gebru ID, Conotter V, Boato G, De Natale FGB (2013) Counter-forensics of median filtering. In: Proceedings IEEE Int. Workshop on Multimedia Signal Processing, pp 260–265
7. Dirik AE, Memon N (2009) Image tamper detection based on demosaicing artifacts. In: Proceedings IEEE Int. Conf. Image Processing, pp 1497–1500
8. Dang-Nguyen D-T, Pasquini C, Conotter V, Boato G (2015) RAISE: a raw images dataset for digital image forensics. In: Proceedings 6th ACM Multimedia Systems Conference, pp 219–224
9. Fan Z, de Queiroz RL (2003) Identification of bitmap compression history: JPEG detection and quantizer estimation. IEEE Trans Image Process 12(2):230–235

10. Fan W, Wang K, Cayre F, Xiong Z (2015) Median filtered image quality enhancement and anti-forensics via variational deconvolution. IEEE Trans Inf Forensics Secur 10(5):1076–1091

11. Farid H (2009) Image forgery detection. IEEE Signal Proc Mag 26:16–25

12. Fontani M, Barni M (2012) Image tamper detection based on demosaicing artifacts. In: Proceedings 20th European Signal Processing Conference, pp 1239–1243

13. Fridrich J, Kodovsky J (2012) Rich models for steganalysis of digital images. IEEE Trans Inf Forensics Secur 7(3):868–882

14. Fu D, Shi YQ, Su W (2007) A generalized Benford's law for JPEG coefficients and its applications in image forensics. In: Proceedings SPIE, Electronic Imaging, Security and Watermarking of Multimedia Contents IX, vol 6505, pp 1L1–1L11

15. Furon T, Bas P (2008) Broken arrows. EURASIP J Inf Secur 2008, p 3

16. Gloe T, Böhme R The Dresden Image Database for benchmarking digital image forensics. In: Proceedings 25th Symposium On Applied Computing (ACM SAC 2010), vol 2, pp 1585–1591

17. Gonzalez RC, Woods RE (2002) Digital image processing. Upper Saddle River NJ: Prentice-hall

18. Kang X, Li Y, Qu Z, Huang J (2012) Enhancing source camera identification performance with a camera reference phase sensor pattern noise. IEEE Trans Inf Forensics Secur 7(2):393–402

19. Kang X, Stamm MC, Peng A, Liu KJR (2013) Robust median filtering forensics using an autoregressive model. IEEE Trans Inf Forensics Secur 8(9):1456–1468

20. Kashyap RL (1984) Characterization and estimation of two-dimensional arma models. IEEE Trans Inf Theory 30(25):736–745

21. Kee E, Farid H (2011) A perceptual metric for photo retouching. Proc Natl Acad Sci 108(50):19907–19912

22. Kirchner M, Böhme R (2008) Hiding traces of resampling in digital images. IEEE Trans Inf Forensics Secur 3(4):582–592

23. Kirchner M, Fridrich J (2010) On detection of median filtering in digital images. In: Proceedings SPIE, Electronic Imaging, Media Forensics and Security II, vol 7541, pp 1–12

24. Lai Y-N, Gao T-G, Li J-X, Sheng G-R (2015) Forensic detection of median filtering in digital images using the coefficient-pair histogram of DCT value and LBP pattern. In: Proceedings Int Conf Intelligent, pp 421–432

25. Lin Z, He J, Tang X, Tang C-K (2009) Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis. Pattern Recogn 42(11):2492–2501

26. Lukas J, Fridrich J, Goljan M (2006) Digital camera identification from sensor pattern noise. IEEE Trans Inf Forensics Secur 1(2):205–214

27. (2002). Natural resources conservation service photo gallery, http://photogallery.nrcs.usda.gov/, United States Department of Agriculture

28. Pasquini C, Boato G, Alajlan N, De Natale FGB (2016) A deterministic approach to detect median filtering in 1D data. IEEE Trans Inf Forensics Secur 11(7):1425–1437

29. Pevný T, Bas P, Fridrich J (2010) Steganalysis by subtractive pixel adjacency matrix. IEEE Trans Inf Forensics Secur 5(2):215–224

30. Piva A (2013) An overview on image forensics, ISRN Signal Processing, vol. 2013

31. Popescu AC, Farid H (2004) Exposing digital forgeries by detecting duplicated image regions, Dept Comput Sci, Dartmouth College, Tech Rep TR2004-515

32. Qiu X, Li H, Luo W, Huang J (2014) A universal image forensic strategy based on steganalytic model. In: Proceedings 2nd ACM Workshop on Information hiding and multimedia security, pp 165–170

33. Rocha A, Scheirer W, Boult T, Goldenstein S (2011) Vision of the unseen: Current trends and challenges in digital image and video forensics. ACM Comput Surv 43(4):Article 26

34. Schaefer G, Stich M (2004) UCID: An Uncompressed color image database. In: Proceedings SPIE 5307, Storage and Retrieval Methods and Applications for Multimedia, vol 5307, pp 472–480

35. Stamm MC, Liu KJR (2011) Anti-forensics of digital image compression. IEEE Trans Inf Forensics Secur 6(3):1050–1065

36. Wu Z-H, Stamm MC, Liu KJR (2013) Anti-forensics of median filtering. In: Proceedings IEEE Int. Conf. acoustics, Speech, Signal Process, pp 3043–3047

37. Yuan H-D (2011) Blind forensics of median filtering in digital images. IEEE Trans Inf Forensics Secur 6(4):1335–1345

38. Yang J, Xie J, Zhu G, Kwong S, Shi YQ (2014) An effective method for detecting double JPEG compression with the same quantization matrix. IEEE Trans Inf Forensics Secur 9(11):1933–1942

39. Zhang Y, Li S, Wang S, Shi YQ (2014) Revealing the traces of median filtering using high-order local ternary patterns. IEEE Signal Process Lett 21(3):275–279

40. Zeng H, Qin T, Kang X, Liu L (2014) Countering anti-forensics of median filtering. In: Proceedings IEEE Int Conf acoustics, Speech, Signal Process, pp 2704–2708

Jianquan Yang received his B.S. degree in 2008 and M.S. degree in 2010, both from Sun Yat-Sen University, Guangzhou, China. He is currently working toward the Ph.D. degree with the Shenzhen College of Advanced Technology, University of Chinese Academy of Sciences. He is also an engineer with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. His research interests mainly include multimedia security and image processing.
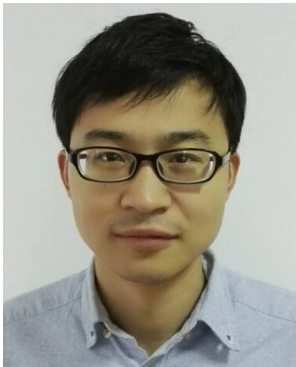
Honglei Ren received his B.S. and M.S. degrees from China University of Geosciences, Wuhan, China. He was a visiting student with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences. His research interests mainly include multimedia security and image processing.

Guopu Zhu received the B.S. degree from Jilin University, China, in 2002, and the M.S. and Ph.D. degrees from the Harbin Institute of Technology, China, in 2004 and 2007, respectively. He is currently a Professor with the Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences (CAS), China. His main research areas are multimedia security and image processing. He has authored or co-authored over 20 papers in SCI journals. He currently serves as an Associate Editor of the Journal of Information Security and Applications. He is a senior member of IEEE, and also is a member of the Youth Innovation Promotion Association, CAS.

Jiwu Huang received the B.S. degree from Xidian University, China, in 1982, the M.S. degree from Tsinghua University, Beijing, China, in 1987, and the Ph.D. degree from the Institute of Automation, Chinese Academy of Sciences, Beijing, in 1998. He is currently a Professor with the College of Information Engineering, Shenzhen University, Shenzhen, China. Before joining Shenzhen University, he was with the School of Information Science and Technology, Sun Yat-sen University, Guangzhou, China. His current research interests include multimedia forensics and security. He served as an Associate Editor of the IEEE TRANSACTIONS ON INFORMATION FORENSICS AND SECURITY and a member of Information Forensics and Security Technical Committee, IEEE Signal Processing Society. He was General Co-chair of the IEEE International Workshop on Information Forensics and Security 2013. He has been a Fellow of IEEE since 2015.

Yun-Qing Shi has joined New Jersey Institute of Technology, US, since 1987. He obtained his M.S. degree from Shanghai Jiao Tong University, China; his Ph.D. degree from University of Pittsburgh, US. His research interests include data hiding, forensics and information assurance, visual signal processing and communications. He is an author/coauthor of more than 300 papers, one book, five book chapters, and an editor of 10 books, 3 special issues and 13 proceedings. He holds 30 US patents. He serves as an associate editor of IEEE Transactions on Information forensics and Security, and an editorial board member of a few journals; he has served as an associate editor of IEEE Transactions on Signal Processing and IEEE Transactions on Circuits and Systems (II); also served as the technical program chair of IEEE ICME07, a co-technical chair of IWDW since 2006 and IEEE MMSP05, a co-general chair of IEEE MMSP02, a Distinguished Lecturer of IEEE CASS. He is a member of a few IEEE technical committees, and a Fellow of IEEE since 2005.