

A FEATURES DECOUPLING METHOD FOR MULTIPLE MANIPULATIONS IDENTIFICATION IN IMAGE OPERATION CHAINS

Jiaxin Chen¹, Xin Liao^{*1,3}, Wei Wang², Zheng Qin¹

¹ College of Computer Science and Electronic Engineering, Hunan University, China.

² National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, China.

³ State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, China.

ABSTRACT

Recently, many forensic techniques have been developed to detect the use of a certain processing operation. When utilizing several manipulations to alter an image, artifacts left by manipulations that have been applied later can potentially disguise traces left by manipulations that were applied earlier. Therefore, the detection of manipulations become difficult. In this paper, we focus on identifying the manipulations in an image operation chain composed of multiple manipulations in a certain order. To address this issue, we analyze the relationship between manipulations identification and blind signal separation. Then, we propose a features decoupling method based on blind signal separation, which decouples the coupled features due to the superimposed processing artifacts and exploits the decoupled features to identify multiple operations. The experiments carried out on two image operation chains confirm the effectiveness of the proposed method.

Index Terms— Image forensics, multiple manipulations identification, image operation chain, features decoupling

1. INTRODUCTION

Nowadays, the large availability of image editing software makes it very easy to manipulate images. Therefore, it is hard to trust the authenticity of an image [1]. To determine whether an image has undergone a specific operation, researchers have designed many state-of-the-art techniques, such as resizing [2], contrast enhancement [3], and JPEG compression [4, 5].

Since most existing works aim at detecting a certain operation and these works cannot be applied to determine other operations, several positive efforts have been made to develop methods that can identify various operations simultaneously [6]-[9]. Most of these techniques assume that only one operation may be applied to alter the original images. In a realis-

tic scenario, multiple operations are needed when completing a forgery. It is well known that multiple falsifications may lead to superimposed processing artifacts, which means the later operations may affect the traces of prior operations [10]. When multiple manipulations are utilized, different types and orders will lead to varying artifacts, making operations detection difficult. Thus, existing methods based on the single operation hypothesis cannot be effectively employed to determine the manipulations in an image operation chain.

Recently, there have been some works considering multiple falsifications [11]-[16]. In [15], an information theoretical framework was proposed to quantify the detectability of manipulations in operation chains. Considering the order of operations, Boroumand et al. [16] investigated a convolutional neural network (CNN) for detecting four processing classes when an image downsampled and again JPEG compressed.

As multiple falsifications could result in the superimposed processing artifacts and two chains of the same n operations using different orders will leave different traces. Thus, detecting manipulations in operation chains to search the history of image falsification is still a challenging problem. In this paper, we formulate the problem of identifying manipulations into a blind signal separation problem. Then, we propose a features decoupling method to determine the operations. Finally, the identification of operations in two processing chains is examined to illustrate the effectiveness of our method.

2. IMAGE MANIPULATIONS IDENTIFICATION FROM THE VIEW OF BLIND SIGNAL SEPARATION

Blind signal separation is a useful method for recovering the unknown source signals from the observed mixed signals under the premise that the source signals are independent [17].

One of the most widely used examples of blind signal separation is the cocktail party problem [18], which describes how to separate the independent voice signal of each person speaking at the same time in a cocktail party. There are n people in the cocktail party and the voice signal of each person is regarded as a source signal. It is usually assumed that there are m sensors placed in every corner at the party to record the

This work is supported by the National Natural Science Foundation of China (Nos. 61972142, 61972395, 61772191), Hunan Provincial Natural Science Foundation of China (No. 2020JJ4212), Beijing Natural Science Foundation of China (No. 4192058), Key Lab of Information Network Security, Ministry of Public Security (No. C20611).

* Corresponding author: Xin Liao, xinliao@hnu.edu.cn.

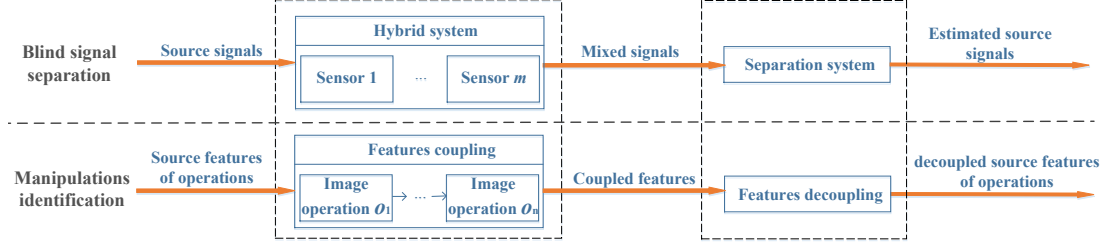


Fig. 1. Illustration of the relationship between manipulations identification and blind signal separation.

voice signals, where $m \geq n$. When n people speaking simultaneously, sensors will receive mixed signals. The adoption of the blind signal separation algorithm could derive the estimated source signal of each person from mixed signals.

Similarly, as illustrated in Fig. 1, in the image manipulations identification problem, supposing there are n tampering operations $\{o_1, o_2, \dots, o_n\}$. Each operation applied to forge an image can be regarded as a person in the cocktail party, and the operation feature extracted from the image is viewed as the voice signal of a person. Since the voice signal of each person is independent, the source feature of each operation is also independent. When the image undergoes multiple manipulations, subsequent operations could affect the features of previous operations. The features extracted from the image are coupled operation features that are similar to the mixed voice signals recorded by the sensors. If the features can be decoupled, the manipulations detection will become easier. Based on this analysis, we found that the detection process is quite analogous to blind signal separation. Namely, as the estimated source signals are isolated from the observed mixed ones through the separation system, decoupled source features of these operations can be obtained via features decoupling. Finally, we can exploit these decoupled features to detect the manipulations in the image operation chain composed of $\{o_1, o_2, \dots, o_n\}$.

Inspired by the view of blind signal separation, we propose a features decoupling method to identify processing operations in diverse image operation chains.

3. FEATURES DECOUPLING METHOD AND MANIPULATIONS IDENTIFICATION STRATEGY

Once an image has been modified by several manipulations, the inherent statistical properties within the original image will inevitably be distorted so that it is difficult to preserve them well after multiple falsifications. Moreover, performing a variety of tampering operations on an image will result in superimposed processing artifacts and make the feature extracted from the image a coupled feature, which brings trouble to the multiple operations identification. Based on this analysis, we investigate the way to decouple the coupled features.

Let $I = \{I_1, I_2, \dots, I_m\}$ and $O = \{o_1, o_2, \dots, o_n\}$ represent the set of tampered images and operations, where

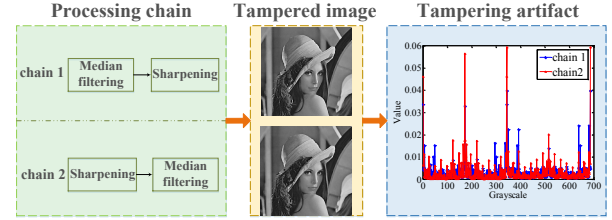


Fig. 2. Illustration of two image processing chains with the same two operations utilizing different order.

$m \geq n$. Supposing that the n operations are used to alter the image I_i , the artifacts of operation o_j is superimposed on the artifacts of operation o_i , where $j > i$. Fig. 2 demonstrates two processing chains composed of the same operations. If the order of median filtering and sharpening in chain 1 is inverted, this will correspond to another different chain 2. It can be observed that the visual effects of the images processed by the two operation chains are the same, but divers and unique artifacts are left in the images. Thus, when forging images with multiple manipulations, the artifacts of the manipulations are coupled in a nonlinear way.

From m tampered images, we will extract m coupled features, which is denoted as $\mathbf{F} = [f_1, f_2, \dots, f_m]^T$. Note that the dimension of f_i is d and $\mathbf{F} \in \mathbb{R}^{m \times d}$. In our work, we use an unknown coupling matrix \mathbf{A} ($\mathbf{A} \in \mathbb{R}^{m \times n}$) and nonlinear coupling function g to denote the coupling way. The features coupling process can be presented as follows,

$$\mathbf{F} = g(\mathbf{AS}), \quad (1)$$

where $\mathbf{S} = [s_1, s_2, \dots, s_n]^T$ is the set of intrinsic operation feature corresponding to processing operation $O = \{o_1, o_2, \dots, o_n\}$ and $\mathbf{S} \in \mathbb{R}^{n \times d}$.

Considering a nonlinear decoupling function p and a decoupling matrix \mathbf{W} , where $\mathbf{W} \in \mathbb{R}^{n \times m}$. According to Eq. (1), the relationship among the decoupled operation features $\mathbf{Y} = [y_1, y_2, \dots, y_n]^T$ and them can be expressed as follows,

$$\mathbf{Y} = \mathbf{W}p(\mathbf{F}) = \mathbf{W}p(g(\mathbf{AS})). \quad (2)$$

If $p = g^{-1}$ is the inverse function of the nonlinear coupling function g and $\mathbf{W} = \mathbf{A}^+$ represents the pseudo inverse matrix of \mathbf{A} , then

$$p(g(\mathbf{AS})) = g^{-1}(g(\mathbf{AS})) = \mathbf{AS}, \quad (3)$$

$$Y = Wp(g(AS)) = WAS = A^+AS = S. \quad (4)$$

Our task is to find the decoupling function p and the decoupling matrix W . Therefore, the decoupled operation features Y can be derived, which are effective resources to detect tampering operations.

In order to de-correlate the coupled features from m images, we use the classical Sphering method [19] to whiten F ,

$$Z = VF, \quad (5)$$

where $Z = [z_1, z_2, \dots, z_m]^T$ denotes the whitened features matrix and V represents a whitening matrix. Note that the matrix V is calculated as follows,

$$V = \Lambda^{-\frac{1}{2}} U^T, \quad (6)$$

where Λ represents the diagonal matrix with the eigenvalues of the covariance matrix of F as diagonal elements, and U denotes the matrix with the eigenvector corresponding to each eigenvalue as a column.

In the following analysis, the whitened features Z are applied as input data. Because the decoupling is performed on the whitened feature z_i rather than the coupled feature f_i , we use p' and W' to represent the decoupling function and matrix corresponding to the whitened features matrix Z . From Eq. (2), we can get the decoupled features Y ,

$$Y = W'p'(Z). \quad (7)$$

In this paper, we utilize the TanH function as the nonlinear function in the features decoupling process to decouple the whitened features Z , which has great decoupling performance.

$$p'(z_i) = \frac{e^{z_i} - e^{-z_i}}{e^{z_i} + e^{-z_i}}. \quad (8)$$

Then, a fixed-point algorithm [20] based on the fourth-order cumulant will be used to calculate the optimal decoupling matrix W' ,

$$\begin{cases} W'_i(t+1) = E[p'(Z)(W'^T_i(t)p'(Z))^3] - 3W'_i(t), \\ W'_i(t+1) = \frac{W'_i(t+1)}{\|W'_i(t+1)\|}, \end{cases} \quad (9)$$

where t denotes iteration times and W'_i is the i^{th} row in the matrix W' . In addition, $E[\cdot]$ and $\|\cdot\|$ represent the mean value of the sampled values of $p'(Z)$ and matrix norm, respectively.

According to Eqs. (7-9), we can obtain the decoupled features $Y = [y_1, y_2, \dots, y_n]^T$, where y_i represents the corresponding processing operations o_i in the set of manipulations and provides evidence for identifying manipulations.

Applying the features decoupling method to identify multiple manipulations, and the implementation strategy is divided into the training and testing stages. Notice that the fixed-point algorithm requires the number of mixed signals is not less than the number of source signals [20]. However, a given

image can only be regarded as one observation channel. For the manipulations detection in operation chains, we need to obtain m ($m \geq n$) observation channels by performing an appropriate image expansion. In this work, we rotate the image 30 degrees by a bilinear interpolation algorithm and crop it to the same size as the given image. Then the rotation is performed on the previously rotated image to generate m images. Due to the difference in application form and purpose, using rotation will not increase the number of operations applied in the chain, and the condition $m \geq n$ can be satisfied.

In the training stage, we first collect M original images, and the corresponding $c+1 = \sum_{i=1}^n A_n^i + 1$ kinds of training images (including the original images and images forged by A_n^k kinds of operations, where $1 \leq k \leq n$) will be created by applying the n operations to modify the original images. For each training image, we expand it into m training images by image rotation. Then, we extract coupled features from the m images and employ the features decoupling method to obtain n decoupled features. Finally, cascading the n features into a new feature vector and use it to train a multi-class classifier.

In the testing stage, as in the processing of the training stage, a testing image may be forged by n operations. We first expand the image to m images via image rotation. Then, the coupled features of these m images are extracted. By applying the proposed features decoupling method, the decoupled operation features are derived and cascaded as a feature vector. The feature then is fed to the obtained multi-class classifier to get the manipulations classification results.

By performing the above strategy with supervised learning, the identification of manipulations in an image processing chain can be realized.

4. EXPERIMENTAL RESULTS

4.1. Experimental Setup

According to our features decoupling method, we need to find a feature based on artifacts leftover from manipulations. As is well known, there exists consistency between adjacent pixels of the image, while operations may alter spatial correlations among neighboring pixels. In our experiments, we give priority to a typical universal feature, i.e., subtractive pixel adjacency matrix (SPAM) [6, 21], to detect different operations.

To assess the effectiveness of our features decoupling method for performing manipulations identification in processing chains, we take the following image operation chain containing two different operations as an example.

- 1) The chain consists of median filtering and resizing.
- 2) The chain consists of median filtering and sharpening.

Note that our method is generalizable because of its ability to eliminate the coupling effect of multiple manipulations and the use of universal features. For a given image, the following five hypotheses for possible processing history of the image are considered.

Table 1. Comparison of the accuracy (%) with state-of-the-art methods under different median filtering and resizing parameters. (*A*: median filtering, *B*: resizing)

Parameters	Methods	H_0	H_1	H_2	H_3	H_4	AVG
$w = 3$ $s = 1.4$ $QF_1 = 80$ $QF_2 = 95$	Method in [9]	75.10	22.20	54.10	82.20	93.50	65.42
	Method in [15]	12.00	8.40	11.60	10.50	60.40	20.58
	Method in [16]	99.90	85.90	99.20	96.70	99.70	96.28
	SPAM w/out fd	50.10	93.40	34.10	13.20	3.30	38.82
	Cascade w/out fd	82.80	81.30	73.10	65.90	33.20	67.26
	Our method	99.00	96.80	99.30	97.70	98.40	98.24
$w = 5$ $s = 0.7$ $QF_1 = 70$ $QF_2 = 85$	Method in [9]	44.50	51.80	37.00	91.90	51.10	55.26
	Method in [15]	69.10	8.80	9.90	10.60	9.10	21.50
	Method in [16]	96.80	47.00	90.60	97.50	72.80	80.94
	SPAM w/out fd	15.20	14.20	64.10	81.60	29.10	40.84
	Cascade w/out fd	5.20	0.00	99.90	83.00	99.00	57.42
	Our method	97.70	91.70	97.80	93.70	97.90	95.76

$$\begin{aligned}
H_0 &: \text{It is double compressed with quality factors } QF_1 \text{ then } QF_2, \\
H_1 &: \text{It is double compressed with quality factors } QF_1 \text{ then } QF_2 \text{ interleaved by } A, \\
H_2 &: \text{It is double compressed with quality factors } QF_1 \text{ then } QF_2 \text{ interleaved by } B, \\
H_3 &: \text{It is double compressed with quality factors } QF_1 \text{ then } QF_2 \text{ interleaved by } A \text{ then } B, \\
H_4 &: \text{It is double compressed with quality factors } QF_1 \text{ then } QF_2 \text{ interleaved by } B \text{ then } A,
\end{aligned} \tag{10}$$

where *A* and *B* represent the tampering operations.

For detecting each operation chain, we use 2,000 images from the BOSSbase image set [22] to generate a training database and 1,000 images from the UCID database [23] to generate a testing database. These images are firstly converted into gray-scale images. A set of falsified images is created by employing the corresponding processing operations to these selected images based on Eq. (10). Finally, a total of 10,000 training images and 5,000 testing images are obtained.

All the experiments are conducted using a support vector machine (SVM) classifier. The polynomial kernel is utilized as the kernel function, where the degree and gamma in kernel function are 3 and 2.8, respectively. Besides, all the experiments utilize the same database as described above.

4.2. Identify Manipulations in Operation Chains

For verifying that our method is effective rather than the extracted image features are valid, a comparison experiment that does not use features decoupling is conducted. To demonstrate that our method is effective in decoupling rather than increasing the feature dimension, we add a comparison experiment where manipulations are detected by cascading coupled features extracted from the original image and its expanded image. Furthermore, the proposed method is compared with

Table 2. Comparison of the accuracy (%) with state-of-the-art methods under different median filtering and sharpening parameters. (*A*: median filtering, *B*: sharpening)

Parameters	Methods	H_0	H_1	H_2	H_3	H_4	AVG
$w = 5$ $\sigma = 0.8$ $\lambda = 1.5$ $QF_1 = 70$ $QF_2 = 85$	Method in [9]	96.10	98.50	96.90	82.60	42.90	83.40
	Method in [15]	48.50	23.90	86.30	70.00	37.60	53.26
	Method in [16]	98.60	99.00	98.50	92.40	53.80	88.46
	SPAM w/out fd	34.50	25.80	69.40	83.90	11.10	44.94
	Cascade w/out fd	73.60	84.10	73.90	89.20	62.10	76.58
	Our method	98.20	88.20	97.70	97.90	89.20	94.24
$w = 5$ $\sigma = 1.3$ $\lambda = 1$ $QF_1 = 85$ $QF_2 = 80$	Method in [9]	83.90	88.70	96.00	84.40	67.50	84.10
	Method in [15]	13.90	29.80	74.20	39.40	24.00	36.26
	Method in [16]	98.90	98.30	90.40	93.30	66.60	89.50
	SPAM w/out fd	42.10	29.40	46.30	91.70	1.60	42.22
	Cascade w/out fd	75.70	77.60	74.50	81.40	53.50	72.54
	Our method	97.20	85.60	96.80	95.30	82.00	91.38

three state-of-the-art methods, i.e., Bayar et al.'s method [9], Gao et al.'s method [15], and Boroumand et al.'s method [16].

Table 1 provides the comparison accuracies of operations identification in the chain consists of median filtering and resizing. It should be noted that "w/out fd" denotes the method without using features decoupling and "AVG" calculates average accuracy. Obviously, our method can identify manipulations with great accuracy and it is superior to the other comparison methods when an image is downscaled.

Table 2 demonstrates the comparison results of operations identification in the chain consisting of median filtering and sharpening. These results prove the effectiveness of our method. Moreover, when the image undergoes multiple operations, the average accuracy of our method is at least 8% higher than other state-of-the-art methods.

We have also verified in the chain consists of three operations. The average accuracy is 87.08%, which shows that our method can effectively identify more than two manipulations.

5. CONCLUSION

In this paper, we focus on identifying manipulations in image operation chains. We first formulate the manipulations identification problem into a blind signal separation problem. To the best of our knowledge, this is the first attempt to apply the blind signal separation technique to forensics in operation chains. Then, a features decoupling method is proposed for decoupling the observed coupled features, which can provide evidence to detect operations. Finally, we examine the identification of two different operation chains. The experimental results have shown the effectiveness of our method. As part of our future effort, we will try to extend the features decoupling idea to design a new detector for dealing with local tampering operations classification and forgery localization.

6. REFERENCES

- [1] M. C. Stamm, M. Wu, and K. J. R. Liu, "Information forensics: an overview of the first decade," *IEEE Access*, vol. 1, pp. 167-200, 2013.
- [2] C. Pasquini, and R. Böhme, "Information-theoretic bounds of resampling forensics: new evidence for traces beyond cyclostationarity," *ACM Workshop on Information Hiding and Multimedia Security*, pp. 3-14, 2017.
- [3] G. Cao, Y. Zhao, R. Ni, and X. Li, "Contrast enhancement-based forensics in digital images," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 3, pp. 515-525, 2014.
- [4] Z. Fan, and R. L. de Queiroz, "Identification of bitmap compression history: JPEG detection and quantizer estimation," *IEEE Transactions on Image Processing*, vol. 12, no. 2, pp. 230-235, 2003.
- [5] Y.-L. Chen, and C.-T. Hsu, "Detecting recompression of JPEG images via periodicity analysis of compression artifacts for tampering detection," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 2, pp. 396-406, 2011.
- [6] X. Qiu, H. Li, W. Luo, and J. Huang, "A universal image forensic strategy based on steganalytic model," *ACM Workshop on Information Hiding and Multimedia Security*, pp. 165-170, 2014.
- [7] H. Li, W. Luo, X. Qiu, and J. Huang, "Identification of various image operations using residual-based features," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 28, no. 1, pp. 31-45, 2018.
- [8] B. Bayar, and M. C. Stamm, "A Deep learning approach to universal image manipulation detection using a new convolutional layer," *ACM Workshop on Information Hiding and Multimedia Security*, pp. 5-10, 2016.
- [9] B. Bayar, and M. C. Stamm, "Constrained convolutional neural networks: a new approach towards general purpose image manipulation detection," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 11, pp. 2691-2706, 2018.
- [10] M. C. Stamm, X. Chu, and K. J. R. Liu, "Forensically determining the order of signal processing operations," *IEEE International Workshop on Information Forensics and Security*, pp. 162-167, 2013.
- [11] V. Conotter, P. Comesaña, and F. Pérez-González, "Forensic detection of processing operator chains: Recovering the history of filtered JPEG images," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 11, pp. 2257-2269, 2015.
- [12] C. Pasquini, "Statistical and deterministic approaches for multimedia forensics," *Electronic Letters on Computer Vision and Image Analysis*, vol. 15, no. 2, pp. 16-18, 2016.
- [13] S. Milani, M. Tagliasacchi, and S. Tubaro, "Discriminating multiple JPEG compression using first digit features," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 2253-2256, 2012.
- [14] X. Liao, and Z. Huang, "A framework for parameters estimation of image operator chain," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 2787-2791, 2020.
- [15] S. Gao, X. Liao, and X. Liu, "Real-time detecting one specific tampering operation in multiple operator chains," *Journal of Real-Time Image Processing*, vol. 16, no. 3, pp. 741-750, 2019.
- [16] M. Boroumand, and J. Fridrich, "Deep learning for detecting processing history of images," *International Conference on Media Watermarking, Security, and Forensics*, pp. 2131-2139, 2018.
- [17] C. Jutten, and J. Herault, "Space or time adaptive signal processing by neural network models," *International Conference on Neural Network for Computing*, pp. 206-211, 1986.
- [18] T. Lee, A. Ziehe, R. Orglmeister, and T. Sejnowski, "Combining time-delayed decorrelation and ICA: towards solving the cocktail party problem," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 1249-1252, 1998.
- [19] J. V. Stone, "Independent component analysis: a tutorial introduction," *A bradford book*, 2004.
- [20] A. Hyvärinen, "A family of fixed-point algorithms for independent component analysis," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 3917-3920, 1997.
- [21] T. Pevný, P. Bas, and J. Fridrich, "Steganalysis by subtractive pixel adjacency matrix," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 2, pp. 215-224, 2010.
- [22] P. Bas, T. Filler, and T. Pevný, "Break our steganographic system – the ins and outs of organizing boss," *Proceeding of International Conference on Information Hiding*, pp. 59-70, 2011.
- [23] G. Schaefer, and M. Stich, "UCID-An uncompressed color image database," *SPIE Storage and Retrieval Methods and Applications for Multimedia*, pp. 472-480, 2004.