Regular Paper

A Histogram and GLCM-based Approach for Image Copy-Move Forgery Detection

Songpon Teerakanok[1,a)] Tetsutaro Uehara[2,b)]

Received: November 23, 2018, Accepted: June 11, 2019

Abstract: In this paper, a novel copy-move forgery detection (CMFD) method for the digital image using the histogram and GLCM-based rotation-invariant feature descriptor is proposed. In developing an efficient CMFD method, there are two fundamental challenges needed to be addressed: accuracy and processing time. To achieve this goal, a fast and straightforward histogram-based, and GLCM-based local features are combined to increase the uniqueness and accuracy of the detection while also maintaining the computational cost, resulting in relatively fast detection mechanism suitable for practical use. The detection mechanism, firstly, performs keypoint detection using SURF-based keypoint detection method. The local GLCM-based and histogram-based features for each block are then calculated and combined using the convolution method. All generated features are then sorted and compared. Finally, lines between matched features are drawn to express the relationship between the original and copy-move regions. Experimental results show that the proposed method outperforms some traditional methods in term of accuracy while also greatly reduces the computational complexity of the system compared to some existing techniques.

Keywords: copy-move forgery, keypoint, GLCM, SURF, histogram, CMFD

1. Introduction

Counterfeit or falsified information has become a crucial problem in today's digital information system. Today, it is not unusual for people to unconsciously cast doubt on what they heard or saw on news websites or the internet. An excellent and also a classic case of using the tampered digital image in the press was the photo of Iran's provocative missile test [2] appeared in several primary and well-known news websites, i.e., Los Angeles Times, The Chicago Tribune, BBC News and The New York Times. The mentioned digital image was tampered by duplicating and adding a picture of missile to the target image to emphasize the frightfulness of the event. Another good example concerning media tampering in our daily life is clickbait. Clickbait [3] is a term referring to web contents created to attract users attention and trying to gain advertising revenue from the number of user clicks. Clickbait usually involves altered digital images which exploit users' curiosities. **Figures 1** and **2** show real examples of image forgeries in our daily life.

Given an example of crime investigation system in Japan, after receiving permission to arrest the suspect from the court, the police have a total of 72 hours to investigate or analyze all collected evidence before deciding whether to release the suspect or not. Practically, these 72 hours, to the police, are not long since they have too much data and evidence to process, especially in

Fig. 1 The photo of Iranian missile test appeared in the press: original (left), tampered (right) (original photo from Nizza, N., and Lyons, P.J. [2]).

Fig. 2 A real example of using forged media, i.e., clickbait, on the internet (from Taboola advertising company).

[1] Graduate School of Information Science and Engineering, Ritsumeikan University, Kusatsu, Shiga 525–8577, Japan
[2] College of Information Science and Engineering, Ritsumeikan University, Kusatsu, Shiga 525–8577, Japan
[a)] songpon.te@cysec.cs.ritsumei.ac.jp
[b)] t-uehara@fc.ritsumei.ac.jp

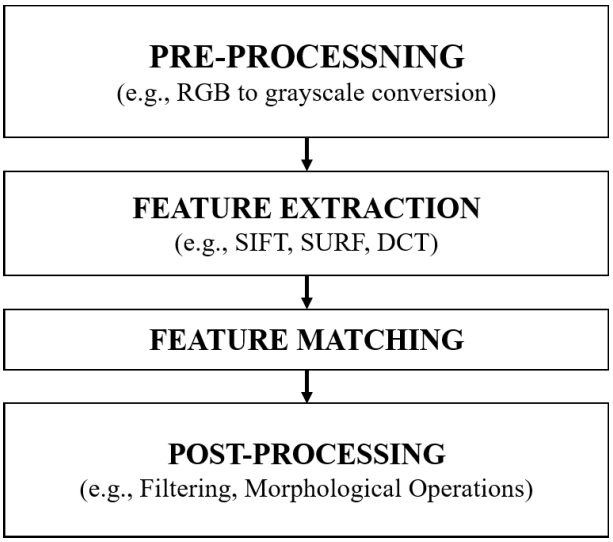

Fig. 3 The most common process pipeline/framework of CMFD.

cybercrime cases. Therefore, a fast and accurate approach for verifying the authenticity of every piece of digital evidence can provide significant support to reduce the time spent and allow the entire investigation processes to go smoothly.

To overcome the problems of forged media and counterfiet information, there are a number of detection methods and techniques proposed in the literature. In this paper, our main contribution is focused on the detection of copy-move forgery (CMF). Regarding copy-move forgery detection (or CMFD), there are many structures and frameworks proposed and studied in the last few years. Most of the frameworks, however, also share some fundamental components as shown in **Fig. 3**. The CMFD pipeline consists of 4 main stages: pre-processing, feature extraction, matching, and post-processing.

First, pre-processing is a stage where some conversion, transform, or decomposition techniques are performed. The goal of the pre-processing stage is to prepare and represent data in a way that makes the subsequent feature extraction stage more efficient. The most simple, yet important methods of pre-processing involve grayscale conversion, where RGB pixels are converted into a grayscale image with a range of 0 to 255, and color space conversion (e.g., RGB to HSV [4], or RGB to YCbCr [5]). Decomposition techniques, for example, wavelet decomposition [6], or principal component analysis (PCA) [7], are belonged to this stage.

Next, feature extraction is then performed. Feature extraction is one of the most critical stages which will determine the overall accuracy of the system. The goal of this step is to create a set of short yet meaningful data vectors (so-called "feature descriptors") to represent each part of the target digital image. There are a large number of techniques for extracting feature vectors from the digital image [8], [9].

An extremely robust and very well-known technique in this category is the Scale-Invariant Feature Transform (SIFT), first introduced by Lowe [10], [11], [12]. While SIFT provides efficient feature description technique which is robust against the various challenging transformation including scaling, and rotation, it also suffers from the high computational cost/complexity. Although the detection results from SIFT may seem to be very efficient in term of accuracy, SIFT, however, may not be a preferable choice for using in real criminal investigations where a massive number of digital images are needed to be verified before using in the court-of-law. Many studies attempt to speed-up SIFT keypoint detection and feature extraction processes. One of the well-known and most successful method proposed so far is Speeded-up Robust Feature (SURF) presented by Bay et al. [13]. Moreover, some interesting techniques also utilize the pre-processing stage in order to speed up the following SIFT feature extraction process.

After extracting of feature descriptors in the previous stage, feature matching processes are performed. This process involves searching for matched patches or segments of the target image with similar feature descriptors. The matching process is also a crucial stage determining the overall detection speed of the CMFD system. Lastly, stage 4 represents post-processing, where raw matched detection results are filtered or processed to enhance and produce the final detection results with the highest quality.

In this paper, we propose a new CMFD scheme using Gray-Level Co-Occurrence Matrix (GLCM) [15] and histogram-based feature description technique. This method is an upgraded version of our previous work [1] with significant changes and improvements. The main contribution of our proposed method is introducing a novel technique used in the feature extraction process. Therefore, the details regarding pre-processing and post-processing will not be discussed in this paper. We, however, also point out some potential pre-processing and post-processing methods for further developments.

The proposed method consists of 3 main steps: 1) keypoint detection using SURF, 2) rotation-invariant feature extraction using GLCM and histogram local features, and 3) feature matching process, which are presented in the following Sections 2, 3, and 4 respectively. Section 5 presents and discusses the experimental results in details. Section 6 explains the further works and directions of our research. Finally, a brief conclusion is drawn in the following Section 7.

2. SURF Keypoint Detection

To detect a set of good interest points within the target digital image, there are several ways to achieve this. Introduced by Lowe [10], Scale-Invariant Feature Transform (SIFT) is the most well-known and considered one of the most robust keypoint detection methods in this field. Using fundamental of SIFT keypoint detection, there are a number of keypoint detection techniques proposed in the literature. SIFT, however, has its main drawback regarding processing speed and high computational complexity due to the use of the difference of Gaussian (DoG) in computing Gaussian pyramid to achieve scale-space extrema detection. SIFT scale-space extrema detection is what makes its keypoint detection robust and accurate. However, it also makes SIFT not suitable for some practical application where instantaneous or near real-time processing speed is needed.

Inspired by SIFT, Speeded-Up Robust Features (SURF) was proposed by Bay et al. [13] to provide a much faster and computationally efficient alternative to SIFT. While Lowe's approach

Fig. 4 Keypoint detection using SURF (original image from Ref. [14]).

tries to approximate Laplacian of Gaussian (LoG) with DoG during the scale-space detection process, which has very high computational complexity, SURF achieves the same goal of LoG approximation using a simple box filter. Utilizing integral images can significantly lessen the computational cost of box filter convolution resulting in a speeded-up keypoint detection method. SURF, in addition, can also be applied with parallel processing to achieve better processing speed. In our work, SURF keypoint detection method is adopted for fast and efficient searching of interest points in the target digital image. The keypoint detection is done by utilizing the determinant of the Hessian matrix. The Hessian matrix is used to measure change around each point; the point with maximum determinant value is then selected as the location of the keypoint (see Ref. [13] for further details and concrete mathematical expression). **Figure 4** shows the use of SURF keypoint detection in our detection mechanism.

3. Feature Extraction Using GLCM and Histogram-based Descriptor

In this section, we generate feature vectors (or so-called "descriptors") from keypoints obtained from the previous stage. In this step, Gray-Level Co-Occurrence Matrix (GLCM) and histogram-based local features are computed and combined using convolution. To provide a better understanding of the GLCM-based local feature generation, we first briefly introduce GLCM in the following Section 3.1. The feature extracting procedures are then described in the next Section 3.2.

3.1 Grey-Level Co-occurrence Matrix (GLCM)

Grey-Level Co-occurrence Matrix [15] (or GLCM, for short) is an excellent statistical approach used in expressing spatial relationships between each pixel and its neighbor or surrounding pixels. GLCM is usually used in the field of texture analysis. Some commonly used statistical values derived from the co-occurrence matrix are homogeneity, energy, correlation, and contrast [16], [17], [18].

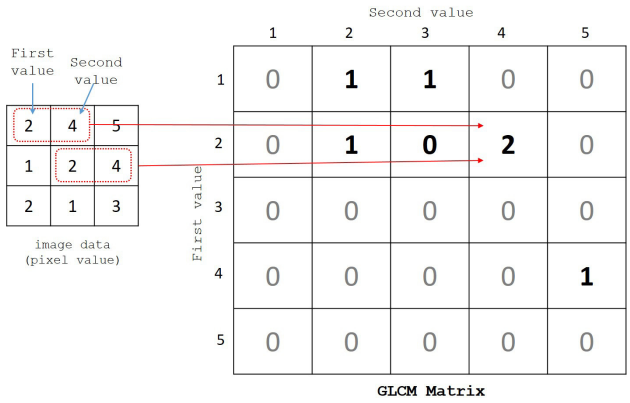

Fig. 5 Example of GLCM with (0, 1) offset.

$$G_{\Delta x,\Delta y}(i,j) = \sum_{x=1}^{n}\sum_{y=1}^{m}\begin{cases}1, & \text{if } I(x,y)=i \text{ and} \\ & I(x+\Delta x, y+\Delta y)=j \\ 0, & \text{otherwise}\end{cases} \quad (1)$$

Equation (1) shows the mathematical expression of GLCM, where i and j are referred to indexes of elements in the output matrix. Furthermore, x and y represent pixel location of the target digital image, and Δx and Δy are the spatial relationship parameters (so-called "offset"). Performing GLCM of the same target image I with different offset parameters (for example, $G_{1,-1}$ and $G_{1,1}$) may yield entirely different results. For better understanding, **Fig. 5** shows a practical example of GLCM.

The final result, i.e., co-occurrence matrix obtained from GLCM, contains meaningful statistical and spatial relationship information expressing uniqueness of the original image. Therefore, the co-occurrence matrix can further be used to create robust feature vectors for copy-move forgery detection.

3.2 Feature Extraction Process

For each keypoint detected by SURF, we then create the corresponding feature vector using the following procedures.

(1) For each key-point k, first, we choose the circle area $C_{k,r}$ around each keypoint within the radius r in the target grayscale image.

(2) With obtained pixels data from each circular area $C_{k,r}$, we then normalize pixels information by subtracting each pixel within the circular area with the local minimum pixel intensity and then quantize all pixel value into a discrete set of integers with the range of $[1, n]$, where n represents the number of quantizing level. With both normalization and quantization, we can reduce the effect of brightness adjustment significantly.

$$\begin{aligned} N_k &= C_{k,r} - min(C_{k,r}) \\ Q_k &= Quantize_n(N_k) \end{aligned} \quad (2)$$

(3) In this step, we will create GLCM-based feature descriptors. Using Q_k, we compute GLCMs of each keypoint with offset parameters: $(-1,-1)$, $(-1,0)$, $(-1,1)$ and $(0,1)$, resulting in $G_{-1,-1}$, $G_{-1,0}$, $G_{-1,1}$ and $G_{0,1}$ respectively. Each GLCM computed is an $n \times n$ matrix, where n is the number of graylevels used during the previous quantization process.

(4) Next, we sum all GLCMs into one final matrix, T. Let i

and j be the index of each row and column within T respectively. The GLCM-based feature vector $F_k(j)$ is then created by computing the mean of all elements in each column of T. Equation (3) shows the process of creating the GLCM-based feature vector.

$$F_k(j) = \frac{\sum_{i=1}^{n} T(i,j)}{n} \tag{3}$$

In Eq. (3), j is the number of each corresponding column in matrix T.

(5) After creating the GLCM-based feature of the keypoint k, we then compute the contrast feature γ from the previous co-occurrence matrix T. A mathematical expression of the contrast feature [16], [17] is shown in Eq. (4).

$$\gamma = \sum_{i,j=1}^{n} P_{i,j}(i-j)^2, \quad P_{i,j} = \frac{T(i,j)}{\sum_{p,q=1}^{n} T(p,q)} \tag{4}$$

(6) Next, we also create a histogram-based feature by subtracting each element of $C_{k,r}$ with its center pixel intensity. The subtracted data is used to created h-bin histogram H_k. In this work, the number of bin was empirically determined which can be changed or adjusted for better accuracy or processing speed.

(7) Finally, we create the final feature vector υ_k of the keypoint k by using the following Eq. (5).

$$\upsilon_k = (\gamma \frown F_k) * H_k \tag{5}$$

As shown in Eq. (5), the final feature vector for each keypoint is created by computing discrete convolution between two data arrays: $(\gamma \frown F_k)$ and H_k, which are $(n+1)$ and h-element arrays respectively. The $(\gamma \frown F_k)$ denotes the concatenation between the contrast feature γ and the GLCM-based feature F_k. Since the discrete convolution between any a and b-element arrays will result in a $(a+b-1)$-length array. Hence, the current length of each feature vector υ_k is equal to $(n+1)+h$.

4. Feature Matching

With detected keypoints and their corresponding features, CMF tampered area of the target digital image can now be efficiently detected. The following steps explain the feature matching process in detail.

(1) With feature vectors derived from detected keypoints, we then form a feature matrix M of size $m \times (n+1)+h$, where m and $(n+1)+h$ are the total number of keypoints and the length of each feature vector respectively. The k-th row of this feature matrix M_k represents a previously derived feature vector υ_k. **Figure 6** shows the structure of the feature matrix.

(2) For each row in M, we perform the inner product between M_k and the transpose version of M (M^t). The arccosine of the inner product results is then computed. Finally, sorting the previous outcome will result in a sorted array S_k which can be used for making the final decision.

$$S_k = sort(\arccos(M_k \times M^t)) \tag{6}$$

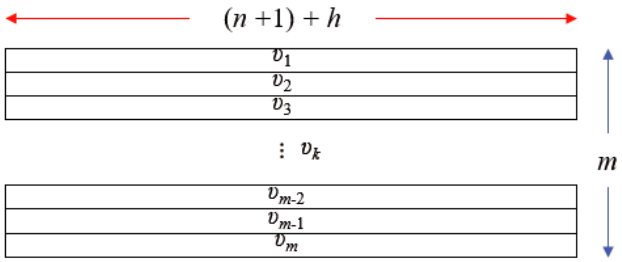

Fig. 6 Overall structure of the feature matrix.

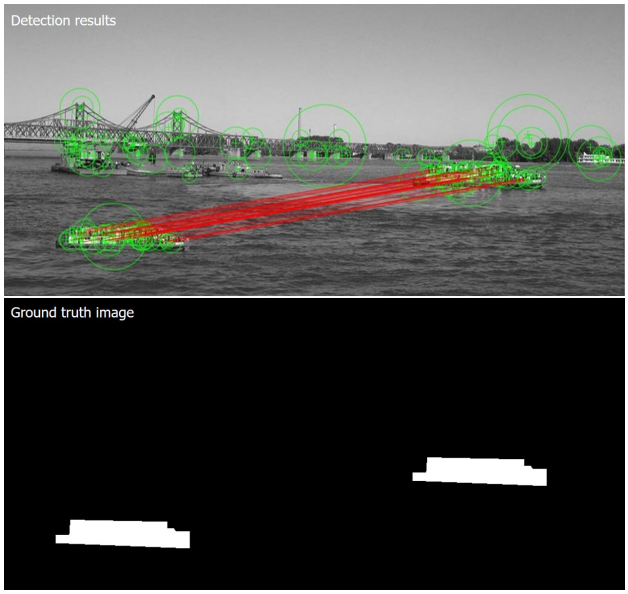

Fig. 7 An example of detection results obtained from our feature matching process (original image from Ref. [14]) (top: detection results, bottom: the ground truth image)

(3) Finally, to determine whether the image at keypoint k was tampered or not, we compared the second and third element of the sorted array S_k. The first element in S_k is generally a value created using inner product of M_k and itself which is usually meaningless in this CMF detection problem. The second and third elements in S_k, however, generally should share the same level of value. In case the difference between the second and third elements of S_k is greater than the predetermined threshold, the image at keypoint k can be considered as CMF tampered.

5. Experiment Results & Discussion

In this section, the detection results and some further discussion are presented. First, the experimental results against the public dataset are provided in Section 5.2, following with some further discussions in Sections 6.1 to 6.5.

5.1 Testing Environments

The experiments were conducted using Matlab v.9.4.0.813654 (R2018a) on Intel Xeon Processor E3-1240 (3.50 GHz) computer with 48 GB RAM running 64-bit Windows 10 OS. In this experiment, 50 randomly selected forged digital images (without scaling) from Ardizzone et al. public dataset [14] (i.e., $D0$ and $D1$) are used to evaluate the accuracy of the proposed CMFD techniques. Also, the proposed technique was also tested against 50

original images without tampering from *D*3 dataset [14] for better evaluation. Regarding parameters, the proposed method is tested using radius of the circular area $C_{k,r}$ equals to 8 pixels ($r = 8$), 64 GLCM quantizing levels ($n = 64$), 64-bin histogram ($h = 64$). These numbers, however, are empirically determined and are subjected to change for better performance.

Lastly, the experiment results were compared to the other 3 important CMFD techniques: 1) SURF-based CMFD proposed by Bo et al. [19], 2) SIFT-based detection method, and 3) CMFD technique using dyadic wavelet transform (DyWT) combined with SIFT proposed by Anand et al. [20].

5.2 Experimental Results

The F_1 score (so-called "F-score" or "F-measure") [21] was adopted to measure the efficacy of each detection mechnaism. The following Eq. (7) describes the F_1 score:

$$F_1 = \frac{2}{\frac{1}{p} + \frac{1}{r}} \tag{7}$$

The value p and r represent precision and recall respectively. In general, precision yields the rate of how many selected keypoints lead to the actual forgeries (true positive keypoint, TP for short). Moreover, false negative keypoint (FN, for short) refers to the keypoint where the algorithm correctly predict that it was not related to CMF. Out of the total number of TP and FN within the target digital image, the recall shows how many TP was discovered by the detection algorithm. Let TTP, TFP, and TFN be the total number of true positive, false positive, and false negative keypoints detected respectively. During the experiment, locations of all matched keypoints were compared against the ground truth image in order to determine precision and recall. The following Eq. (8) shows the mathematical expression of precision and recall.

$$\begin{aligned} precision(p) &= \frac{TTP}{TTP + TFP} \\ recall(r) &= \frac{TTP}{TTP + TFN} \end{aligned} \tag{8}$$

In addition to the use of F_1, we also use the accuracy score to compare the efficiency of each detection mechanism. Let k be the total number of keypoints within the target image and TTN be the number of keypoints where the algorithm correctly predict that it was led to CMF.

$$ACC = \frac{TTP + TTN}{k} \tag{9}$$

Figures 8 and **9** present the visual example of CMF detection using the proposed method compared with three other ap-

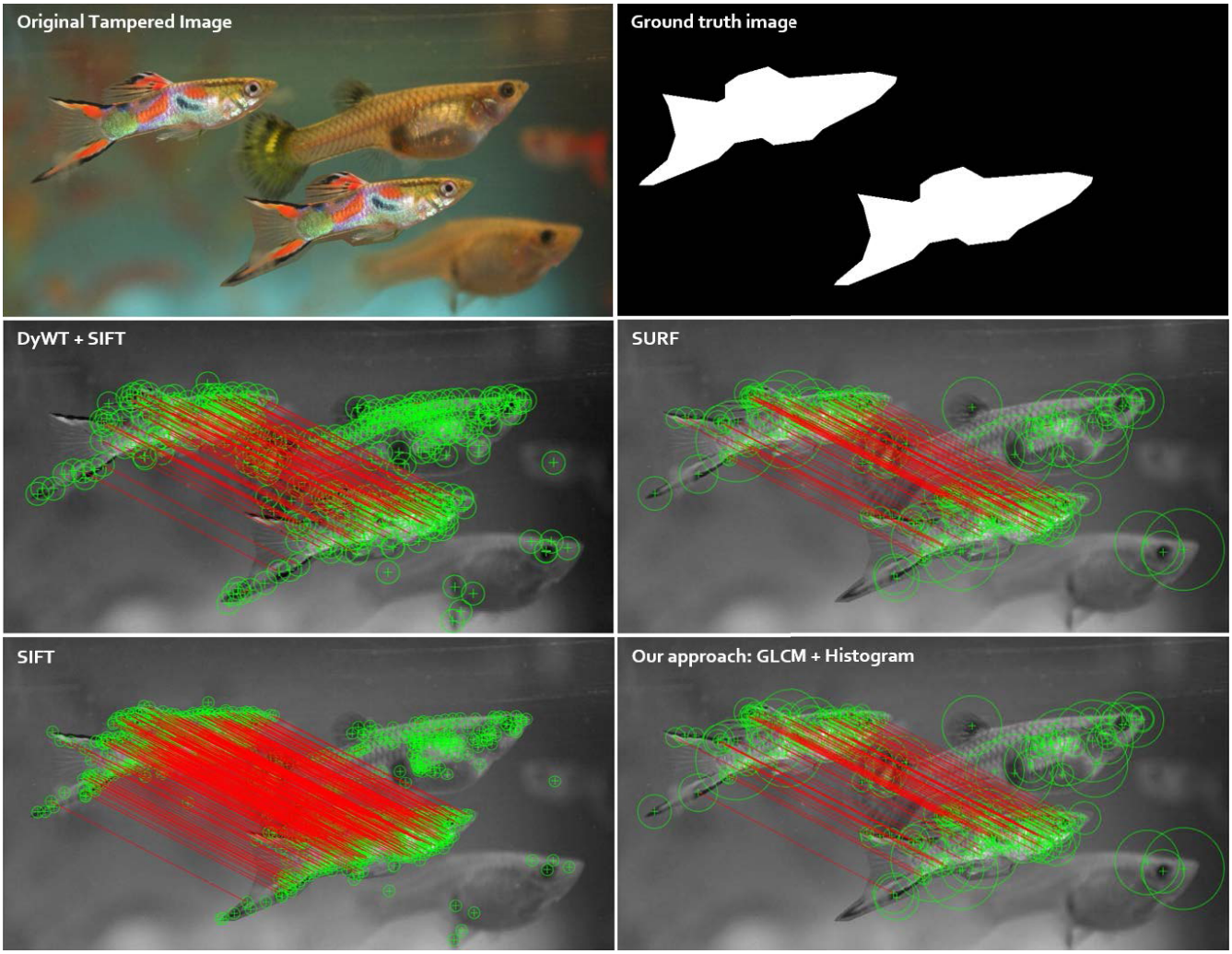

Fig. 8 Comparison of CMF detection techniques (original photo from Ardizzone et al. [14]).

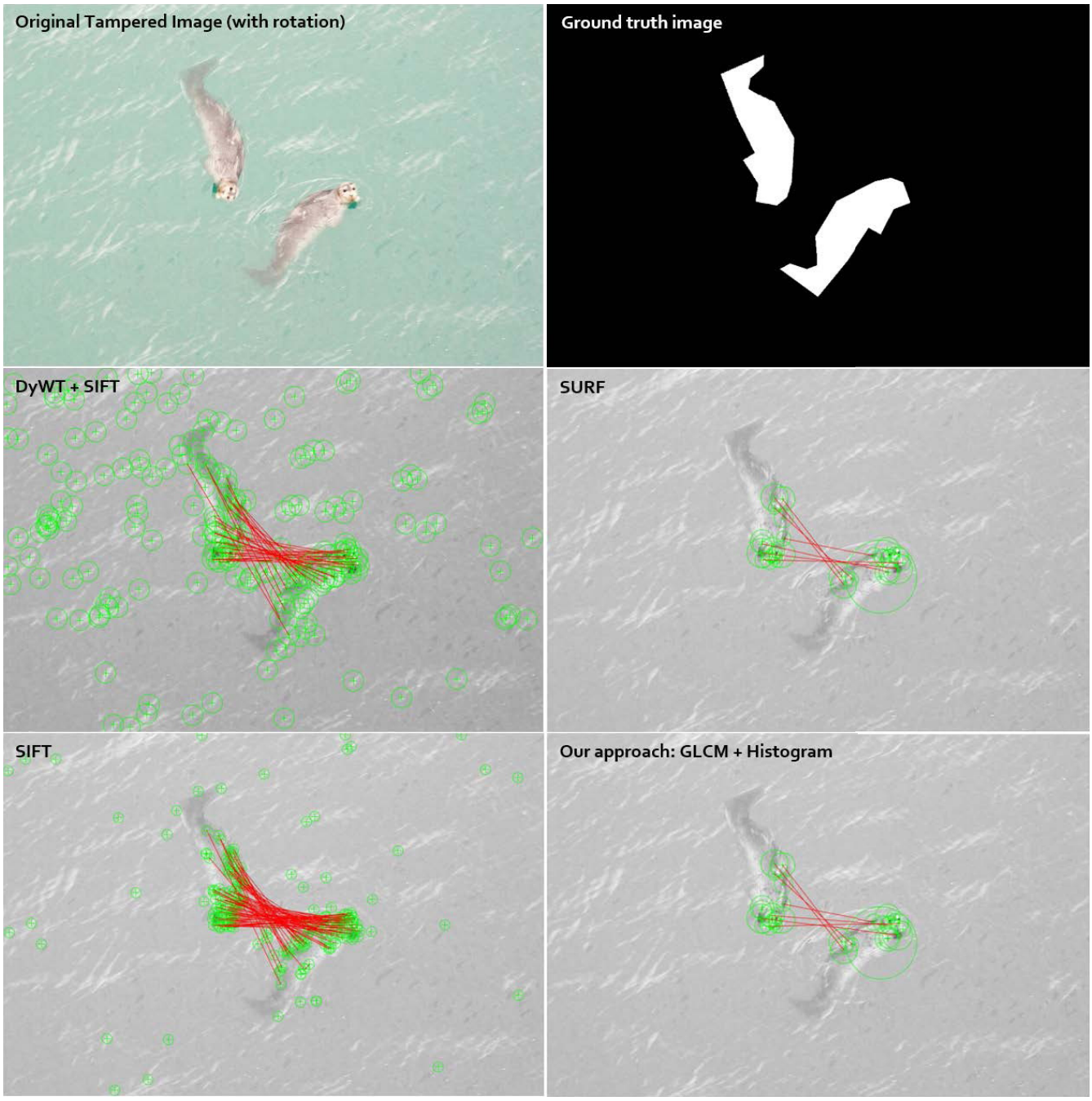

Fig. 9 CMF Detection against rotated tampered image (original photo from Ardizzone et al. [14]).

proaches. We can see at a glance that the method having keypoint detection algorithm based on SIFT can detect more keypoints covering the broader area of tampered regions than the methods based on SURF. The same goes for results on the rotated tampered image. The SIFT-based approaches are, however, cost significantly higher amount of computational cost. SURF-based approaches, however, not only show the acceptable amount of accuracy and F_1 scores during the experiments, but it can also significantly outperform SIFT-based approach in term of speed (time consumption). **Tables 1** and **2** present the experiment results against tampered and authentic images respectively.

Regarding the data presented in Tables 1 and 2, all numerical results are average values calculated from the raw data; therefore, some small calculation errors or inconsistencies may be presented due to the approximation and rounding. According to the results, the comparison shows that our approach outperforms both SURF and the combination between DyWT and SIFT (DyWT+SIFT,

Table 1 Performance against CMF tampered images.

Methods	p	r	F_1	ACC	Time(s)
SIFT	0.63	0.71	0.65	94.68	7.92
DyWT+SIFT [20]	0.48	0.41	0.43	90.06	1.71
SURF [19]	0.57	0.59	0.55	89.64	0.81
Our approach	0.72	0.60	0.64	92.36	2.98

Table 2 Performance against original authentic images.

Methods	TNR	FPR	ACC	Time(s)
SIFT	96.978	3.021	96.978	11.617
DyWT+SIFT [20]	95.810	4.189	95.810	1.834
SURF [19]	93.957	6.043	93.957	0.784
Our approach	97.252	2.747	97.252	2.968

for short) in term of precision, recall, F_1, and accuracy score. F_1 and accuracy of our developed approach are still lower than SIFT which is known for it very accurate detection results. The proposed method, however, also yield the best performance while having the highest TNR (true negative rate) and also the lowest FPR (false positive rate) when being tested against a set of 50 original authentic images in Ref. [14]. Focusing on the computational cost, although the proposed method can perform slower than SURF and DyWT+SIFT, the processing time of our approach is acceptable; therefore, with the higher level in efficiency, it can be used in practice.

5.3 Performance Analysis and Discussion

Although the obtained detection results from keypoint-based CMFD techniques (including our proposed method) may look accurate and practically usable, the results, however, still need some further post-processing steps (e.g., boundary tracing operations [39], RANSAC [24], or morphological operations [26], [36], [37]) to make the final results complete or closer to the ground truth image. Furthermore, keypoint-based approaches also have one main drawback regarding the incapability of dealing with low-contrast areas of the image (e.g., sea, sky, desert). Comparing to the exhaustive search-based CMFD techniques utilizing sliding windows or overlapping blocks which a massive number of feature descriptors are generated from all windows or overlapping blocks, the exhaustive search-based approaches may have an advantage over keypoint-based approach in term of accuracy and more precise localization of the tampered regions due to the massive number of keypoint generated from every pixel of the target image.

On the other hand, keypoint-based methods have significant advantages over the traditional exhaustive search-based approaches in term of computational performance and speed. Given an example of a CMF tampered image with the dimension of $1{,}024 \times 768$ pixels as shown in **Fig. 10**. The CMFD system will produce a total of 773,797 feature descriptors using an exhaustive search-based approach (Mahmood et al. [37], for example) and only 255 descriptors using our proposed method.

Let us assume that the exhaustive matching algorithm with the complexity of $O(n^2)$ proposed by Mahmood et al. [37], in 2018, is used to search for CMF regions. Searching for CMF regions, the detection algorithm will need to compare and match feature descriptors for approximately $(773{,}797)^2$ times and $(255)^2$ times using the traditional exhaustive feature description technique and

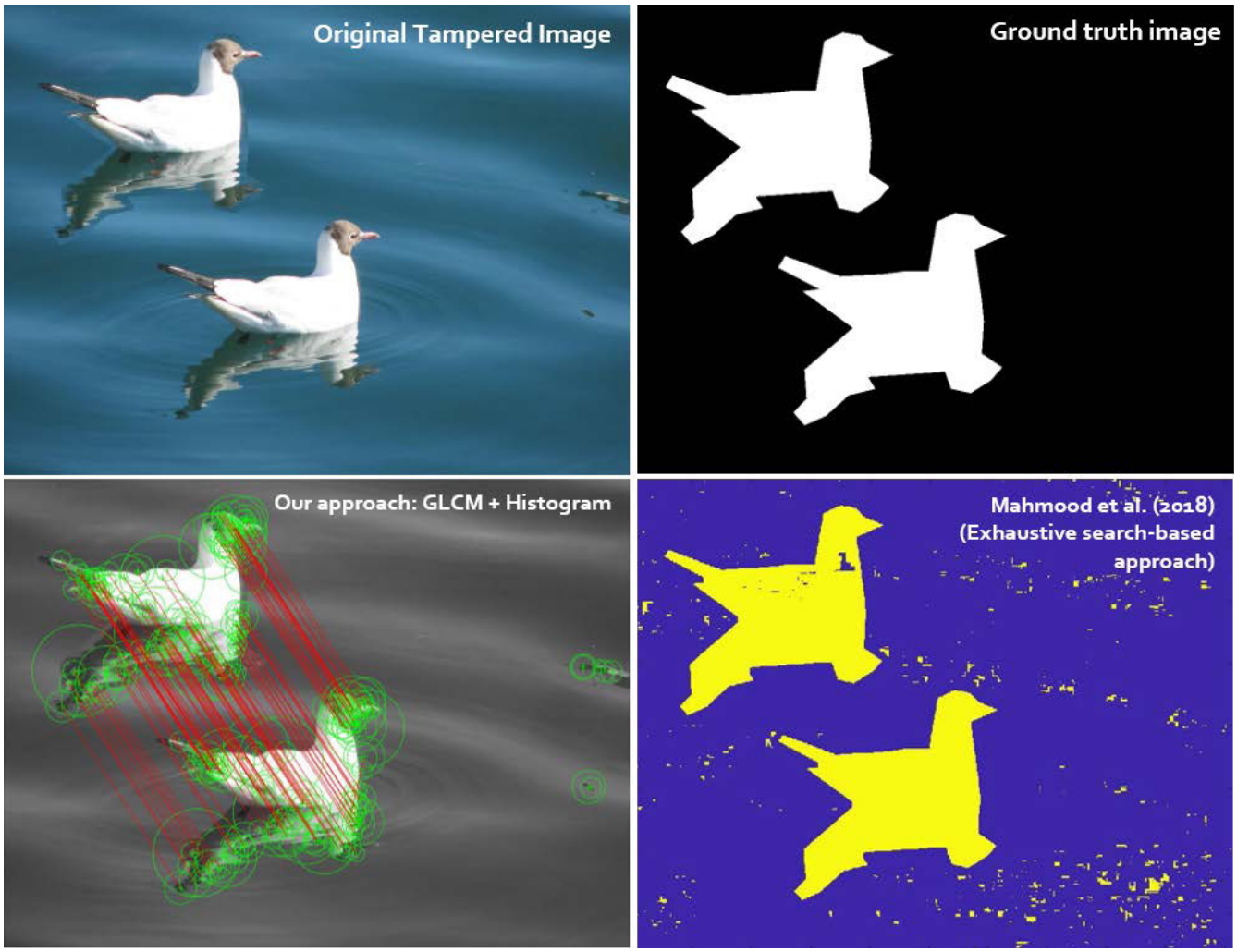

Fig. 10 Comparison between the proposed method and the traditional exhaustive search-based feature description approach (original photo from Ardizzone et al. [14]).

Fig. 11 Best (top) and worst (bottom) experiment results (original photo from Ardizzone et al. [14]).

the proposed technique respectively. The processing time using the proposed method for the sample image in Fig. 10 is 1.41 s. The exhaustive approach, however, consumes roughly 66 hours (i.e., almost three days) to complete the detection under the same testing environment. Therefore, it is clearly shown that the exhaustive search-based approach is impractical for CMFD under time pressure condition.

Regarding detection failure in our approach, generally, the proposed method can efficiently detect anomalies within the target CMF tampered digital image without problem. However, in some cases where the keypoint detection mechanism fails to detect the keypoint within the tampered areas, the CMF detection of the target image will result in complete failure as shown in **Fig. 11**. Figure 11 shows results from the best and worst case scenarios using the proposed detection mechanism.

5.4 Theoretical Advantages

Although feature descriptors generated from the proposed method are inferior to SIFT-based descriptors in term of robustness and accuracy, the proposed method has advantages over SIFT for two main reasons. First, SIFT-based approaches use Laplaican of Gaussian (LoG) or Difference of Gaussian (DoG) in achieving scale-space extrema detection during the keypoint detection process. These processes are known for their high computational complexity and cost. The proposed method, in contrast, utilizes SURF-based keypoint detection mechanism which is a lot faster and more computational efficient. Although the SURF-based method may yield a lower number of detected keypoints, however, most of them can cover an entire image, i.e., keypoints from most of the objects within the target image are detected, which are generally sufficient for CMFD purpose. Second, since the number of keypoint being detected and used in our approach is smaller than SIFT-based approaches, the matching process will also consume less computational cost resulting in significant improvement in term of matching speed.

Comparing to SURF, the proposed method uses the same keypoint detection mechanism. However, the feature extraction processes are totally different. GLCM is known for its use in the field of texture classification; therefore, GLCM may have potential to become a good feature extraction technique for CMFD. After many studies and experiments, however, it is empirically determined that GLCM alone is still not sufficient for CMFD purpose due to the insufficient in detection accuracy.

Regarding utilizing of histogram, a histogram-based feature extraction is not something new. Histogram-based approach is also appeared in many feature extraction technique including the well-known SIFT [10] by using it to assign the orientation to each descriptor prior to generating each feature descriptor. With the

potential of histogram-based feature and GLCM-based feature, empirically, we found out that combining these two types of feature descriptors using discrete convolution method can produce robust feature descriptors with high discriminative power which are suitable for CMFD purpose.

6. Further Directions

In this section, we provide a discussion concerning the final goal of this research, adopting this technique in practice, and further plan for development.

6.1 Overall Picture

First of all, in this work, we originally designed a CMFD technique to help the forensic examiner in performing their digital investigation which is expected to speed up their processes and make CMFD become an easier task. With the problem of lacking number of forensic practitioners, the final goal of our research involves creating an automatic CMFD tool requiring less or no human interaction/final decision which can provide support to non-skilled personnel or general police officers in performing their basic investigation. Moreover, in the hand of experts, this tool is expected to provide significant supports during their investigation. **Figure 12** shows our current milestone for this development.

Utilizing machine learning or classification techniques (SVM [22], and Deep Learning [23], for example), the proposed CMFD method can be served as a core detection mechanism in the automatic detection system. Therefore, in further development, we first aim to apply some post-processing and false matching removal technique (e.g., RANSAC [24], Median filter [25], and morphological opening [26]) to efficiently localize the tampered region and enhance the quality of the final detection results. Finally, we plan to apply this technique to the classification/machine learning algorithm in creating the mentioned automatic CMFD system.

6.2 Enhancement of CMFD Using Pre-processing

To enhance the quality of the detection results or to speed up the entire process, pre-processing is an essential part of the CMFD system that is needed to be carefully designed. In this paper, our main contribution is not including the pre-processing process. However, there is also a simple pre-processing process used in our work: local normalization and quantization. Regarding local normalization, pixels in the area around each keypoint are first subtracted with their local minimum value and then quantize all pixels value into the range of $[1, n]$. By doing so, the effect of some brightness adjustment will be removed allowing the detection algorithm to perform more accurately.

Concerning other pre-processing techniques, there are a number of pre-processing technique that may be useful when applying to the proposed mechanism. Some of these interesting pre-processing techniques are segmentation (e.g., SLIC [27]), transformation-based techniques (e.g., Fast Fourier Transform (FFT) [28], or Discrete Wavelet Transform (DWT) [6]) and dimensionality reduction techniques (e.g., Principal Component Analysis (PCA) [7], etc.). By filtering or reducing dimensionalities of image data, these techniques can get rid of unnecessary information and are proved to be useful in many applications including CMFD. Some recent CMFD techniques utilizing these pre-processing techniques are presented in Refs. [29], [30], [31], [32].

6.3 False Match Removal

Generally, feature matching algorithms are not flawless. There is a moderate level of possibility for feature matching process to yield some false positive (or false match) results. There are many factors that might cause false positive results in detection; for example, low uniqueness (quality) feature extraction, or inefficient threshold value. Therefore, to enhance the quality of the final detection results, recent research has utilized false match removal techniques to achieve this goal. There are several techniques proposed and discussed so far. Some of them are proximity (distance) thresholding and the most commonly used Random Sample Consensus (RANSAC) technique.

Proximity (or distance) thresholding is a simple constraint which is applied during or after the matching process. During the matching of current keypoint K_i with another keypoint K_j $(i \neq j)$,

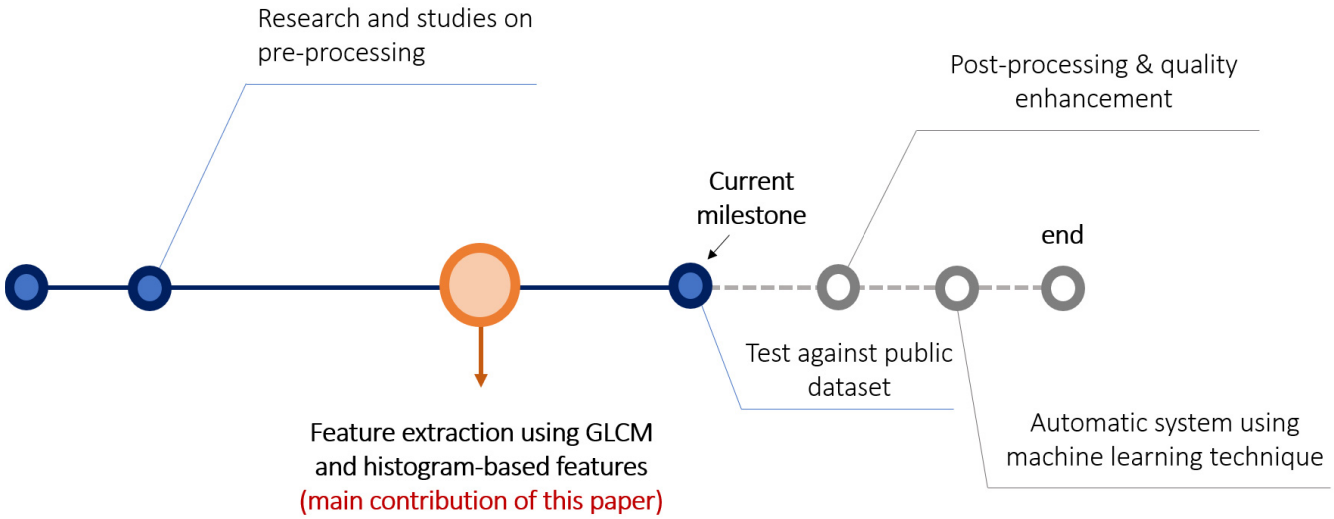

Fig. 12 Current milestone and further development plan of our research.

the matching process of these two keypoints will be skipped if the distance between them is lower than the pre-defined threshold. This simple process not only prevents the algorithm from matching the current keypoint with its neighbor which are likely to cause false matches but also speed up the matching process.

Regarding another very well-known and widely used technique, RANSAC is an iterative approach that is used to find outliers within the given set of data. In CMFD, RANSAC is deployed to detect the outliers (i.e., isolated matches) from the set of matched features obtained during the matching process. Some recent methods utilizing RANSAC for false match removal are proposed in Refs. [33], [34], [35], and the original paper of RANSAC by Fischler et al. [24]. Since RANSAC is an iterative method, it is also known for its slow processing speed. Therefore, RANSAC may not practically suitable for real-time processing or application which require fast processing speed. For further development, we plan to adopt RANSAC to remove false positive results and modify it to speed up the false match removal process.

6.4 Final Detection Results Quality Enhancement

To improve the quality of the final CMFD output, some quality enhancement techniques are needed. Morphological operations (e.g., dilation and erosion) are well-known and suitable for using in CMFD. Morphological opening (i.e., performing of image erosion and then follow by image dilation) is a very effective technique used in many research [36], [37], [38], for example. The morphological opening process will first remove unnecessary noise from the target detection image and then subsequently fill any small holes and gaps using image dilation process. Although this paper does not mainly focus on the quality enhancement of CMFD output, these enhancement techniques are also highly applicable and recommended for applying with our approach.

6.5 Remaining Challenges

To create an efficient automatic CMFD system, a good machine learning (ML) or classification method is needed to be considered. Since the automatic system requires no or less human interaction, the high level of detection accuracy is essential. Also, there is also a tradeoff between processing speed and detection accuracy. Thus, to practically use the automatic system in the real criminal investigation, the designed system should have high enough detection accuracy, while having relatively low computation complexity. Choosing an appropriate ML algorithm and solving the speed-accuracy tradeoff problem are still left as remaining challenges for our research.

7. Conclusion

In this paper, a CMF detection method using the histogram and GLCM-based rotation-invariant features are presented. Testing with the public dataset proposed in Ref. [14], it is shown that the developed outperform SURF [19] and DyWT+SIFT based approaches [20] in term of accuracy, precision, and recall which also resulted in the higher F1 scores. Although the proposed method performs slower than two techniques mentioned earlier, with the higher level of efficiency, it makes our approach an attractive way to overcome the problem of copy-move forgeries in practice.

References

[1] Teerakanok, S. and Uehara, T.: Copy-move Forgery Detection Using GLCM-Based Rotation-Invariant Feature: A Preliminary Research, *Proc. 42nd IEEE Annual Computer Software and Applications Conference* (*COMPSAC*), pp.365–369 (online), DOI: 10.1109/COMPSAC.2018.10259 (2018).

[2] Nizza, M. and Lyons, P.J.: In an Iranian Image, a Missile Too Many, The New York Times (online), available from ⟨https://thelede.blogs.nytimes.com/2008/07/10/in-an-iranian-image-a-missile-too-many⟩ (accessed 2018-03-15).

[3] Chakraborty, A., Paranjape, B., Kakarla, S. and Ganguly, N.: Stop Clickbait: Detecting and preventing clickbaits in online news media, *Proc. IEEE/ACM Intl. Conf. Advances in Social Networks Analysis and Mining* (*ASONAM*), pp.9–16 (online), DOI: 10.1109/ASONAM.2016.7752207 (2016).

[4] Loesdau, M., Chabrier, S. and Gabillon, A.: Hue and Saturation in the RGB Color Space, *Proc. Intl. Conf. Image and Signal Processing* (*ICISP*), pp.203–212 (online), DOI: 10.1007/978-3-319-07998-1_23 (2014).

[5] Color images (online), available from ⟨http://www.bk.isy.liu.se/courses/tsbk06/material/lect8-9b.pdf⟩ (accessed 2018-04-13).

[6] Bénéteau, C. and Fleet, P.J.V.: Discrete Wavelet Transformations and Undergraduate Education, *Notices of the AMS*, Vol.58, No.5, pp.656–666 (online), DOI: 10.1155/2017/7925404 (2017).

[7] Introduction to Principal Components and Factor Analysis (online), available from ⟨ftp://statgen.ncsu.edu/pub/thorne/molevoclass/AtchleyOct19.pdf⟩ (accessed 2018-07-11).

[8] Warif, N.B.A., Wahab, A.W.A., Idris, M.Y.I., Ramli, R., Salleh, R., Shamshirband, S. and Choo, K.-K.R.: Copy-move forgery detection: Survey, challenges and future directions, *J. Network and Computer Applications*, Vol.75, pp.259–278 (online), DOI: 10.1016/j.jnca.2016.09.008 (2016).

[9] Zhang, Z., Wang, C. and Zhou, X.: A Survey on Passive Image Copy-Move Forgery Detection *J. Information Processing Systems*, Vol.14, No.1, pp.6–31 (online), DOI: 10.3745/JIPS.02.0078 (2018).

[10] Lowe, D.G.: Distinctive Image Features from Scale-Invariant Keypoints, *Intl. J. Computer Vision*, Vol.60, No.2, pp.91–110 (online), DOI: 10.1023/B:VISI.0000029664.99615.94 (2004).

[11] Do, T.T., Kijak, E., Furon, T. and Amsaleg, L.: Understanding the security and robustness of SIFT, *Proc. ACM Intl. Conf. Multimedia* (*MM'10*), pp.1195–1198 (online), DOI: 10.1145/1873951.1874185 (2010).

[12] Lindeberg, T.: Scale Invariant Feature Transform, *Scholarpedia*, Vol.7, No.5, p.10491 (2012).

[13] Bay, H., Tuytelaars, T. and Gool, L.V.: SURF: Speeded Up Robust Features, *Proc. European Conf. Computer Vision* (*ECCV*), *Lecture Notes in Computer Science*, Vol.3951, pp.404–417 (online), DOI: 10.1007/11744023_32 (2006).

[14] Ardizzone, E., Bruno, A. and Mazzola, G.: Copy-Move Forgery Detection by Matching Triangles of Keypoints, *IEEE Trans. Information Forensics and Security*, Vol.10, No.10, pp.2084–2094 (online), DOI: 10.1109/TIFS.2015.2445742 (2015).

[15] Haralick, R.M., Shanmugam, K. and Dinstein, I.: Textural Features for Image Classification, *IEEE Trans. Systems, Man, and Cybernetics*, Vol.SMC-3, No.6, pp.610–621 (online), DOI: 10.1109/TSMC.1973.4309314 (1973).

[16] Mohanaiah, P., Sathyanarayana, P. and GuruKumar, L.: Image Texture Feature Extraction Using GLCM Approach, *J. Intl. Scientific and Research Publications*, Vol.3, No.5 (2013).

[17] Shijin, K.P.S. and Dharun, V.S.: Extraction of Texture Features using GLCM and Shape Features using Connected Regions, *J. Intl. Engineering and Technology*, Vol.8, No.6 (online), DOI: 10.21817/ijet/2016/v8i6/160806254 (2016).

[18] Acuna, M.B.A.: Rotation-invariant texture classification based on greylevel co-occurrence matrices, PhD Thesis, University of Sao Paulo Polytechnic School (2013).

[19] Bo, X., Junwen, W., Guangjie, L. and Yuewei, D.: Image Copy-Move Forgery Detection Based on SURF, *Proc. Intl. Conf. Multimedia Information Networking and Security*, pp.890–892 (online), DOI: 10.1109/MINES.2010.189 (2010).

[20] Anand, V., Hashmi, M.F. and Keskar, A.G.: A Copy Move Forgery Detection to Overcome Sustained Attacks Using Dyadic Wavelet Transform and SIFT Methods, *Intelligent Information and Database Systems*, *LNCS*, Vol.8397, pp.530–542 (2014).

[21] Sasaki, Y.: The truth of the F-measure (online), available from ⟨https://www.toyota-ti.ac.jp/Lab/Denshi/COIN/people/yutaka.sasaki/F-measure-YS-26Oct07.pdf⟩ (accessed 2018-02-17).

[22] Ng, A.: CS229 Lecture notes - Support Vector Machines (online), available from ⟨http://neuralnetworksanddeeplearning.com/chap6.html⟩ (accessed 2018-10-01).

[23] Nielsen, M.: Deep learning (online), available from ⟨http://cs229.

stanford.edu/notes/cs229-notes3.pdf⟩ (accessed 2018-11-01).
[24] Fischler, M.A. and Bolles, R.C.: Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, *J. Graphics and Image Processing*, Vol.24, No.6, pp.381–395 (1981).
[25] Fisher, R., Perkins, S., Walker, A. and Wolfart, E.: Median Filter (online), available from ⟨https://homepages.inf.ed.ac.uk/rbf/HIPR2/median.htm⟩ (accessed 2018-05-18).
[26] OpenCV, Morphological Transformations (online), available from ⟨https://docs.opencv.org/3.4/d9/d61/tutorial_py_morphological_ops.html⟩ (accessed 2018-07-28).
[27] Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P. and Süsstrunk, S.: SLIC Superpixels, *EPFL Technical Report*, Vol.149300, pp.1–15 (2010).
[28] Weisstein, E.W.: Fast Fourier Transform, available from ⟨http://mathworld.wolfram.com/FastFourierTransform.html⟩ (accessed 2018-08-19).
[29] Zandi, M., Mahmoudi-Aznaveh, A. and Talebpour, A.: Iterative Copy-Move Forgery Detection Based on a New Interest Point Detector, *IEEE Trans. Inforfmation Forensics and Security*, Vol.11, No.11, pp.2499–2512, DOI: 10.1109/TIFS.2016.2585118 (2016).
[30] Jwaid, M.F. and Baraskar, T.N.: Study and Analysis of Copy-Move & Splicing Image Forgery Detection Techniques, *Proc. Intl. Conf. IoT in Social, Mobile, Analytics and Cloud* (*I-SMAC*), pp.1–6, DOI: 10.1109/ICCUBEA.2017.8463695 (2017).
[31] Fadl, S.M. and Semary, N.A.: Robust Copy-Move forgery revealing in digital images using polar coordinate system, *J. Neurocomputing*, Vol.265, pp.57–65, DOI: 10.1016/j.neucom.2016.11.091 (2016).
[32] Nithiya, R. and Veluchamy, S.: Key point descriptor based copy and move image forgery detection system, *Proc. Intl. Science Technology Engineering and Management* (*ICONSTEM*), pp.577–581, DOI: 10.1109/ICONSTEM.2016.7560959 (2016).
[33] Das, T., Hasan, R., Azam, M.R. and Uddin, J.: A Robust Method for Detecting Copy-Move Image Forgery Using Stationary Wavelet Transform and Scale Invariant Feature Transform, *Proc. Intl. Conf. Computer, Communication, Chemical, Material and Electronic Engineering* (*IC4ME2*), pp.1–4, DOI: 10.1109/IC4ME2.2018.8465668 (2018).
[34] Emam, M., Han, Q., Li, Q. and Zhang, H.: A robust detection algorithm for image Copy-Move forgery in smooth regions, *Proc. Intl. Conf. Circuits, System and Simulation* (*ICCSS*), pp.119–123 (online), DOI: 10.1109/CIRSYSSIM.2017.8023194 (2017).
[35] Jin, G. and Wan, X.: An improved method for SIFT-based copy-move forgery detection using non-maximum value suppression and optimized J-Linkage, *J. Signal Processing: Image Communication*, Vol.57, pp.113–125, DOI: 10.1016/j.image.2017.05.010 (2017).
[36] Li, Y.: Image copy-move forgery detection based on polar cosine transform and approximate nearest neighbor searching, *J. Forensic Science International*, Vol.224, pp.59–67, DOI: 10.1016/j.forsciint.2012.10.031 (2013).
[37] Mahmooda, T., Mehmood, Z., Shah, M. and Saba, T.: A robust technique for copy-move forgery detection and localization in digital images via stationary wavelet and discrete cosine transform, *J. Visual Communication and Image Representation*, Vol.53, pp.202–214 (online), DOI: 10.1016/j.jvcir.2018.03.015 (2018).
[38] Sharma, S. and Ghanekar, U.: A rotationally invariant texture descriptor to detect copy move forgery in medical images, *Proc. IEEE Intl. Conf. Computational Intelligence and Communication Technology* (*CICT*), pp.795–798 (online), DOI: 10.1109/CICT.2015.88 (2015).
[39] Yadav, N. and Kapdi, R.: Copy move forgery detection using SIFT and GMM, *Proc. Nirma University Intl. Conf. Eng.* (*NUiCONE*), pp.1–4 (online), DOI: 10.1109/NUICONE.2015.7449647 (2015).

Songpon Teerakanok received his B.E. and M.E. degrees in Computer Engineering, and Information Science and Engineering from Prince of Songkla University and Ritsumeikan University in 2013 and 2016, respectively. Currently, he is pursuing his doctoral degree in Information Science and Engineering at Ritsumeikan University. Regarding past research, he was a former research student at the Centre for Network Research (CNR) (Hatyai, Thailand) from 2009 to 2013. He is now also a current member of the Cyber Security Laboratory at Ritsumeikan University. His research interest covers Cryptography, Privacy, Location-based Service (LBS), and Digital Forensics.

Tetsutaro Uehara received his B.E., M.E., and D.Eng. degrees from Kyoto University in 1990, 1992, and 1996, respectively. He was an assistant professor on the Faculty of Systems Engineering, Wakayama University from 1996 to 2003. From 2003 to 2005, he was an associate professor at the Center for Information Technology of the Graduate School of Engineering, Kyoto University. From 2006 to 2011, he was an associate professor at the Academic Center for Computing and Media Studies, Kyoto University. From 2011 to 2013, he was a Deputy Director of Standardization Division in the Ministry of Internal Affairs and Communication, Japan. He has been a professor at College of Information Science and Engineering, Ritsumeikan University from 2013. He has also been the vice president of the Institute of Digital Forensics from 2017. His research interest covers Systems Security, Digital Forensics, Privacy, Education in Information Ethics, and Information System Management in Local Government.