Exposing original and duplicated regions using SIFT features and resampling traces*

David Vázquez-Padín[1] and Fernando Pérez-González[1,2,3]

[1] University of Vigo, Signal Theory and Communications Dept., Vigo, Spain
[2] GRADIANT, Vigo, Spain
[3] University of New Mexico, Dept. of Electrical and Computer Engineering, USA
{dvazquez,fperez}@gts.uvigo.es

Abstract. A common type of digital image forgery is the duplication of a region in the same image to conceal something in a captured scene. The detection of region duplication forgeries has been recently addressed using methods based on SIFT features that provide points of the regions involved in the tampering and also the parameters of the geometric transformation between both regions. However, considering this output, there is not yet any information about which of the regions are originals and which are the duplicated ones. A reliable image forensic analysis must provide this information. In this paper, we propose to use a resampling-based method to provide an accurate way to distinguish the original and the tampered regions by analizing the resampling factor of each area. Comparative results are presented to evaluate the performance of the combination of both methods.

Keywords: Image forensics, region duplication, resampling estimation, SIFT

1 Introduction

Today, digital images are widely used to inform, communicate and interact with people, above all, through the Internet. Due to the huge proliferation of visual information, a lot of image editing tools were designed initially to enhance the quality of digital images, but in the meantime these tools also allow their manipulation, alteration and even the creation of realistic synthetic images. So, nowadays, we often have to deal with cases where an image cannot be considered as an undeniable proof of occurrence of an event. For instance, very recently, we

* Research supported by the European Union under project REWIND (Grant Agreement Number 268478), the European Regional Development Fund (ERDF) and the Spanish Government under projects DYNACS (TEC2010-21245-C02-02/TCM) and COMONSENS (CONSOLIDER-INGENIO 2010 CSD2008-00010), and the Galician Regional Government under projects "Consolidation of Research Units" 2009/62, 2010/85 and SCALLOPS (10PXIB322231PR), and by the Iberdrola Foundation through the Prince of Asturias Endowed Chair in Information Science and Related Technologies.

(a) Original image. (b) Tampered image.

Fig. 1. Real example of a tampered image (*on the right*) shown in the BP Web site by copying and moving parts of the original image (*on the left*). Courtesy of The Washington Post.

have seen during the BP oil crisis, how the image shown in Fig. 1(a) was doctored on the BP Web site by filling the blank screens with other parts of the original photo (see the result in Fig. 1(b)).

As a mean to prove the authenticity or verify the integrity of an image and cope with these manipulations, a lot of techniques have arisen in the past few years that can be classified as active or passive. Active approaches require a known signal that is embedded in the image to detect forgeries, while passive approaches, also known as blind, work in the absence of any prior information of the original image. Currently, in the context of passive techniques there are several methods that exploit the intrinsic properties of an image [1], allowing for instance: the identification of the source or the origin of an image; the detection of lighting inconsistencies, double compressed images or region duplications; and also the detection and estimation of inconsistencies in the resampling factor of an image. In this paper, we will focus on the detection of duplicated regions and the estimation of the resampling factor on such regions.

Specifically, in this work we combine these two different but complementary forensic tools to get better results and to provide a more accurate forensic analysis of tampered images. The main idea is to mitigate the drawbacks of each technique by using the characteristics of the other. For example, by detecting a cloned region with one of the existing algorithms (e.g. [2] or [3]), it is viable to estimate the geometric relation between the original area and the cloned one, but it is not possible to know which of the two regions is the original and which is the clone. However, by estimating the resampling factor of each zone[4] using any of the methods in [4], [5] or [6], we can differentiate both regions as the original and the duplicated one, since their resampling factors will be different. In the other hand, if the cloned area has not been resized, the resampling estimator cannot help to infer such manipulation (since the resampling factors are equal), but using the region duplication detector this problem is solved.

[4] We are supposing that the copied region has been spatially transformed.

The pros and cons of each technique are discussed in more detail in the next section. In Section 3, the model used is described focusing on how we propose to combine both techniques to improve performance. Experimental results carried out with this image forensic scheme are summarized in Section 4. Finally, Section 5, provides the conclusions and further work.

2 Motivation

In the context of passive image forensics techniques, there does not exist a common framework to analyze images and detect forgeries, i.e. there is not a universal tool that can explicitly determine all the modifications or transformations applied to an image. Instead of that, there is a bunch of tools that exploit some of the inherent characteristics of an image, and in doing so, try to detect the alterations such image has been subject to.

The main objective of this paper is to provide a novel image forensic tool to reach better results in terms of estimation accuracy of digital forgeries, by combining two different techniques that complement the needs of each other. As it was previously mentioned, one of the techniques allows the detection of region duplication forgeries where a part of an image itself is copied, probably geometrically transformed and pasted into another part of the same image to conceal something. The second technique, provides a way to statistically detect and estimate the resampling factor of an image block which gives information about the type of spatial transformation locally applied.

The complementary behavior of both techniques can be established from the analysis of advantages and drawbacks of each one, as it is summarized below.

2.1 Advantages/drawbacks of region duplication detectors

Starting from the first approach for detection of region duplication based on an exhaustive search and analysis of correlation properties of the image [1], until the most recent methods proposed in [2] or [3] capable of estimating the geometric transformation applied between the duplicated regions; the main shortcoming of all these techniques, supposing that they are able to find the duplicated regions, is the impossibility to identify which are the original regions and which are the duplicated ones.

For example, Fig. 2(a) represents the possible output of any of these methods, highlighting two duplicated regions (tagged with **1** and **2**). Taking only into account the provided output, can we assert that the region labeled as **1** is the source and the region labeled as **2** is the duplicate, from a mathematical point of view? The answer is negative, as these methods only provide a match between different pixel areas. Even being able to estimate the geometric relation between both regions (with the method proposed in [2] or [3]), it is not possible to distinguish, in a mathematical sense, the original region from the cloned patch. The more suitable solution to provide this information is to compute the resampling

(a) Region duplication detection. (b) Resampling analysis.

Fig. 2. Examples of drawbacks of each technique. *On the left*, the detected regions are highlighted and tagged with **1** and **2**. *On the right*, the tampered region is highlighted and each analyzed block is denoted by a white border box.

factor of each region and also of the neighborhood and check if both are consistent. By analyzing this relation between the resampling factors, we can identify the tampered regions of the image.

Taking into account the advantages of the region duplication detectors, these methods are able to detect copy-move forgeries[5], while resampling detectors fail looking for inconsistencies in the resampling factor. Besides, the most recent proposed methods based on SIFT ([2] and [3]), allow a very fast analysis of an entire image, in terms of computation time. As a counterpart, they have also an important limitation in terms of detection performance since it is only possible to extract reliable keypoints from peculiar points of the image.

2.2 Advantages/drawbacks of resampling detectors

The detection of resampling traces and the estimation of the resampling factor (or equivalently, the spatial transformation applied to an image block) are closely related and have been studied in several works [4–6]. Although these methods provide good results in controlled scenarios, when they are evaluated in more realistic situations, their performance get worse. For instance, looking for a more efficient forensic analysis in terms of computation time, these methods usually process an image using non-overlapped blocks of a fixed size (e.g. 128×128 pixels). However, with high probability, the location of a tampered region will not be aligned with the grid defined by these blocks, as it is shown in Fig. 2(b). Thus, in such cases, the detection of the tampered region will fail, since the number of resampled samples included in each block is small with respect to the number of original samples.

An important handicap of these methods is the impossibility to detect copy-move forgeries, since the resampling factor of the whole image remains constant.

[5] A copy-move forgery is considered when the duplicated region is not spatially transformed, just translated.

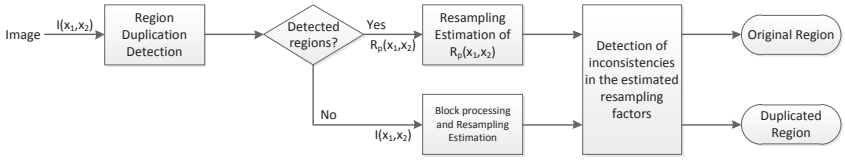


Fig. 3. Block Diagram of the proposed image forensic analysis tool.

We have just seen before that this problem can be easily solved by using a region duplication detector. Additionally, the processing of each block of the image, looking for inconsistencies in the resampling factor, is highly time-consuming.

Hence, once we have seen the positive characteristics and also the negative ones of each technique, it can be expected that the combination of both ideas will provide better performance and also a more complete and accurate forensic analysis of tampered images.

3 Model Description

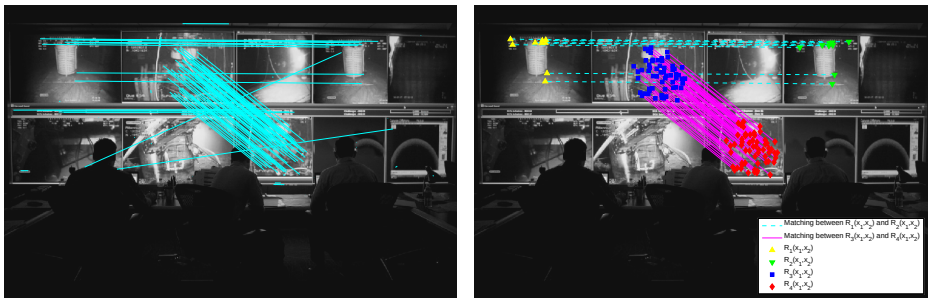
In order to overcome the problem related to the identification and differentiation of the original regions and the tampered ones using a region duplication detector and to avoid the previously mentioned misdetections of the resampling detectors, the proposed approach uses a combination of both techniques.

In Fig. 3 we represent in block diagram form the steps involved in the proposed forensic analysis of an image. As a first step we use a region duplication detector to extract the original and the cloned regions, but if the method is not able to find any tampered region, it is necessary to analyze the entire image by processing blocks and looking for inconsistencies in the resampling factor of each block. Nevertheless, if the region duplication detector is capable of finding the duplicated regions, then the resampling-based method is just applied to estimate the resampling factor of each area. Finally, according to the results obtained in the previous stages, the system determines and differentiates the original regions from the tampered ones.

Next, we describe the specific methods considered for each technique to provide a possible practical implementation of the proposed forensic analysis tool.

3.1 A SIFT-based method for region duplication detection

As it was previously introduced, there are several recently published methods based on the matching of image features and keypoints (e.g. [2] and [3]), that provide good results for the detection of duplicated regions. In this paper, we consider the method proposed by Amerini et al.



(a) Matched keypoints. (b) Clustered and detected regions.

Fig. 4. Steps followed for the detection of cloned areas. *On the left side*, solid lines represent the matching between keypoints and, *on the right side*, different markers are used to identify the clustered data and solid/dashed lines link the related regions.

Following the steps described in [2] we start with one of the color space component of a sampled image $I(\boldsymbol{x}) = I(x_1, x_2)$ with size $N_1 \times N_2$ pixels, where $0 \leq x_1 \leq N_1 - 1$ and $0 \leq x_2 \leq N_2 - 1$. We apply the algorithm proposed by Lowe in [7] to reach a set $\mathcal{X}$ of n keypoints: $\mathcal{X} = \{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \mid \boldsymbol{x}_k = (x_1, x_2)\}$; with their respective SIFT descriptors: $\mathcal{D} = \{\boldsymbol{d}_1, \ldots, \boldsymbol{d}_n\}$, where each $\boldsymbol{d}_k$ is a 128-dimensional vector. Since the descriptors of a duplicated region will look like those of the original area, we want to identify the nearest neighbor of each descriptor to find a possible match of similar keypoints. Thus, a vector that contains the Euclidean distance between each pair of descriptors is computed for each descriptor $\boldsymbol{d}_k$, obtaining a set

$$\mathcal{S} = \{s_l = \|\boldsymbol{d}_i - \boldsymbol{d}_j\|_2 \mid j \in \{1, \ldots, n\}, j \neq i\}$$

that will be sorted in ascending order, for convenience. The matching between keypoints is satisfied if the ratio between the distance of the closest neighbor s_1 and that of the second-closest one s_2 is less than a threshold Υ, i.e.

$$\frac{s_1}{s_2} < \Upsilon.$$

For instance, considering a threshold $\Upsilon = 0.6$ and applying this procedure to the BP tampered image shown in Fig. 1(b), we get the result depicted in Fig. 4(a). Once the set of matched keypoints $\mathcal{X}_m$ is obtained, it is necessary to cluster these data in such a way as to be able to distinguish the different matched regions.

For clustering on the spatial location of the matched points, an agglomerative hierarchical clustering is used as it is proposed in [2]. Considering that we have at least two matched areas, the result of this process provides $P \geq 2$ different sets of matched points $\mathcal{M}_p$, so $\mathcal{X}_m = \mathcal{M}_1 \cup \cdots \cup \mathcal{M}_P$, and this allows the definition of the different duplicated regions.

Continuing with the BP doctored image, we illustrate in Fig. 4(b), the four set of points that determine the two different tampered regions matched with

Table 1. Followed steps by the method proposed in [4].

For each frequency pair (α_1, α_2) of the $N \times N$ 2-D FFT grid:
1. From the data $R_i(x_1, x_2)$, build up a cyclostationary vector $\hat{\boldsymbol{c}}_{xx}$, for a set of lags $\boldsymbol{\tau}$.
2. Estimate the covariance matrix $\hat{\boldsymbol{\Sigma}}_{xx}$.
3. Compute the test statistic $\mathcal{T}_{xx} = N^2 \hat{\boldsymbol{c}}_{xx}^H \hat{\boldsymbol{\Sigma}}_{xx} \hat{\boldsymbol{c}}_{xx}$.
4. Set Γ based on the probability of false alarm P_F.
5. Consider a cyclic frequency pair if $\mathcal{T}_{xx} > \Gamma$.
From the set of cyclic frequencies estimate the transformation.

the solid and dashed lines. Note that some outliers have been removed after the clustering process.

From the points in a region $\boldsymbol{x}_q \in \mathcal{M}_q$ and the corresponding matched points $\boldsymbol{x}_r \in \mathcal{M}_r$, we can estimate the geometric transformation applied between the two matched areas:

$$\begin{bmatrix} \boldsymbol{x}_q^T \\ 1 \end{bmatrix} = \boldsymbol{H}_{qr} \begin{bmatrix} \boldsymbol{x}_r^T \\ 1 \end{bmatrix},$$

where $\boldsymbol{H}_{qr}$ represents an affine homography. By using the Random Sample Consensus (RANSAC) algorithm, a Maximum Likelihood estimation of the affine homography $\boldsymbol{H}_{qr}$ can be carried out.

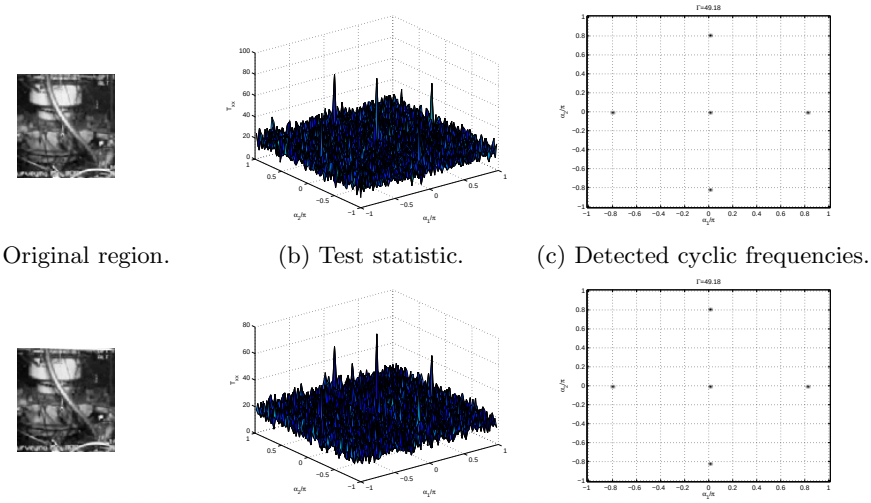
Now, suppose that from the SIFT-based method we obtain $P = 2$ identified regions $R_1(x_1, x_2)$ and $R_2(x_1, x_2)$ and also the estimation of the relation between both $\hat{\boldsymbol{H}}_{12}$, then, using this information, can we demonstrate that $R_1(x_1, x_2)$ is the original area and $R_2(x_1, x_2)$ is the duplicated one, or vice-versa? The method explained below will help to answer this question.

3.2 A resampling-based method to reveal tampered regions

An appropriate way to determine if a matched region is the source or the duplicated one, is to use a resampling estimator that gives a measure of the applied spatial transformation, based on the intrinsic properties of the image pixel region. If the SIFT-based method is not able to find any duplicated region, then we can use any of the proposed methods [4],[5] or [6] to make an exhaustive analysis, processing all the blocks of the image and looking for inconsistencies in the resampling factor.

However, we are more interested in the case when the SIFT-based method does provide the detected cloned regions. So, considering that we get two regions $R_1(x_1, x_2)$ and $R_2(x_1, x_2)$ and taking into account that these regions are generally non-square, for the identification of the original and the duplicated one, we will use the method proposed in [4], which takes a block of the image and applies a statistical test for the evaluation of the presence of almost cyclostationarity in the analyzed block. The steps followed by the method are summarized in Table 1.

Since this method works with square blocks, we have to adapt the detected regions to get a square form. A simple way to do that is to take a square region



(a) Original region. (b) Test statistic. (c) Detected cyclic frequencies.

(d) Duplicated region. (e) Test statistic. (f) Detected cyclic frequencies.

Fig. 5. Application of the two-dimensional statistical test to one of the pair of matched regions in the BP image: $R_3(x_1, x_2)$ (*top row*) and $R_4(x_1, x_2)$ (*bottom row*).

that includes the area under analysis and pad with zeros all the pixels that are not contained in the region $R_i(x_1, x_2)$. The zero-padding approach is probably a suboptimal solution, but doing this we can estimate the resampling factor for each region $R_i(x_1, x_2)$. One of the objectives of this paper is also to study the performance of this method in such scenario.

As we have stated before, a resampling detector cannot differentiate the original source from the duplicated versions if a copy-move forgery is performed. That is exactly what happens with the tampered regions, labeled as $R_2(x_1, x_2)$ and $R_4(x_1, x_2)$ in Fig. 4(b). In fact, applying the statistical test to the matched regions $R_3(x_1, x_2)$ and $R_4(x_1, x_2)$, we obtain the same resampling factor ($\hat{\rho} \approx 5/3$) in both cases, as we can see in Fig. 5. Thus, in this particular scenario, the resampling-based method only identifies the scaling factor applied to the image, but it is not able to distinguish the source region from the clone (since the resampling factor is the same).

However, considering that the pasted regions are geometrically adapted to the scene, then to determine which parts of the image are the copies and which one of those is the source, it is enough to analyze the neighborhood of each region. So, taking a square block that only includes the adjacent neighbor pixels of each region $R_i(x_1, x_2)$ (removing the pixels that belong to the area under analysis), the resampling factor of the neighborhood can be estimated. Finally, for the classification of the regions, we know that an original region will have the same resampling factor in the neighborhood and inside the region, but the tampered regions will have different values in both cases.

(a) (b) (c) (d) (e) (f)

Fig. 6. Different masks used to test the performance of the proposed forensic tool.

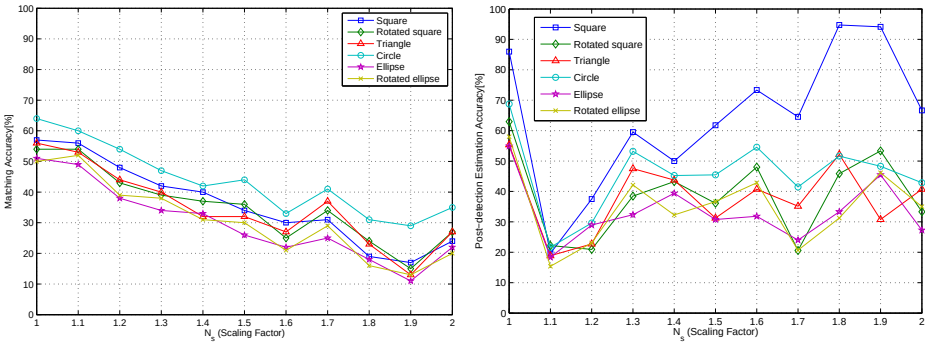
4 Experimental results

For the evaluation of this image forensic scheme, we use 100 images from a personal image database composed by several realistic scenarios with different indoor and outdoor scenes. All the images in this collection have been captured in a RAW format by a Nikon D60 digital camera and have been converted into uncompressed TIFF images in the RGB color space. The original resolution of each image was 3872×2592, but due to the increase of computational complexity, all the images were cropped to 1024×1024 pixels. The resampling factor of each color channel is equal to 2, due to the color filter array (CFA) interpolation performed inside the camera. This fact will be taken into consideration along the application of the resampling-based method and for simplicity we will only process the green component of the RGB color space.

To test the performance of the proposed scheme (Fig. 3), as a first step we evaluate the SIFT-based method and the resampling-based method separately, and then we combine both to see how the results of the forensic analysis improve. In order to get more realistic forgeries in our experiments, we use six different patterns for the duplicated areas, that are depicted in Fig. 6. We use these masks to copy a region located at a random position of an image, then we scale this region by one of the scaling factors N_s in the set $\{1, 1.1, \ldots, 2\}$ and, finally, we paste the duplicated region in a new random location on the same image. The random position of both regions is the same for all masks in order to make a fear comparison, but this one changes for different scaling factors and for each image. Since the tampered regions tend to be relatively small, we have made the experiments in such a way that the resampled region fits always in a 128×128 block.

For the SIFT-based method we use a threshold $\Upsilon = 0.6$, we remove false positive matching points if their distance is less than 10 (i.e. $\|\boldsymbol{x}_i - \boldsymbol{x}_j\|_2 < 10$) and once we get the hierarchical clustering we remove the outliers of each region if their distance to the mean point of their corresponding region is higher than 3 times the variance of the points in the considered region. The implementation of the SIFT algorithm used in the following experiments has been taken from [8] and for the RANSAC homography estimation we have used the functions available from [9].

The configuration of the resampling-based method is almost the same as the one used in [4] (i.e. we use a spectral window to smooth the periodogram of size 11×11 and a set of $K = 9$ lags), but we do not use the threshold Γ to detect



(a) SIFT-based method. (b) SIFT-based method.

Fig. 7. Matching and post-detection estimation accuracy (in terms of percentage), obtained with the SIFT-based method for different masks and scaling factors.

the cyclic frequencies, since we just estimate the applied transformation (i.e. a scaling factor) from the cyclic frequency with highest magnitude, excluding the frequency at zero (DC).

4.1 Detection results using the SIFT-based method

Taking into account the described set of tampered images, we consider that the SIFT-based method matches correctly a tampered area if it is able to find at least four common points between the original and the duplicated region. Fig. 7(a) depicts the matching accuracy of this method in terms of percentage, showing the different results for each used mask and for the different values of the scaling factor N_s.

Next to this graph, Fig. 7(b) shows the (post-detection) estimation accuracy of the affine transformation applied between the previously matched regions, using the RANSAC method. Note that we are drawing the post-detection estimation accuracy, i.e. the estimation accuracy of the scaling factor applied between the correctly matched regions in the previous step (thus, it is clear that the represented percentage of accurate estimation is not relative to the 100 images of the database). In this case, since we cannot know which region is the original we get two possible estimations: $\hat{\boldsymbol{G}}_{12} \approx \boldsymbol{H}_{12}$ or $\hat{\boldsymbol{G}}_{12} \approx \boldsymbol{H}_{12}^{-1}$. We consider that the estimation is correct if any of both estimated affine transform satisfies $|\hat{\boldsymbol{G}}_{12}(1,1) - N_s| \leq 0.05$ or $|\hat{\boldsymbol{G}}_{12}(2,2) - N_s| \leq 0.05$, where $\hat{\boldsymbol{G}}_{12}(i,j)$ represents the element of the matrix $\hat{\boldsymbol{G}}_{12}$ located at the ith row and at the jth column. In this case, $\hat{\boldsymbol{G}}_{12}(1,1)$ and $\hat{\boldsymbol{G}}_{12}(2,2)$ represent the estimation of the scaling factor applied in each axis in the affine transformation.

As we can observe from the two graphs of Fig. 7, with the SIFT-based method it is easier to match and estimate copy-move forgeries than duplicated regions that have been geometrically transformed. However, from the estimation point of view, it is more difficult to estimate the homography for scaling factors near

one like $N_s = 1.1$ or $N_s = 1.2$, than for higher values. The matching accuracy is not very high, due to the lack of reliable keypoints in several images of the dataset (the number of keypoints per image was in the range $[250, 17500]$), but, as it was mentioned earlier, this is an intrinsic limitation of any SIFT-based method. With respect to the used masks, the intuitive idea that small areas are more challenging for detection and estimation purposes, comes up in both plots.

At this point we are just able to find matches between regions and estimate the relation between both, but we cannot indicate which is the source and which is the forged region.

4.2 Detection results using the resampling-based method

Before considering the union of the two methods, we evaluate the resampling-based method when it is applied to the whole image, processing block by block to find inconsistencies in the resampling factor ρ. As it was previously noticed, due to the CFA interpolation applied by the camera, we know that the resampling factor of each non-tampered block is $\rho = \rho_{\text{orig}} = 2$, and then the corresponding value to a scaled version by N_s will be $\rho = \rho_{\text{orig}} \times N_s$. Therefore, once we attain a different value from ρ_{orig} we tag the block under analysis as a digitally forged region. In this case, because the tampered regions have a similar size, we use a 128×128 block of analysis.

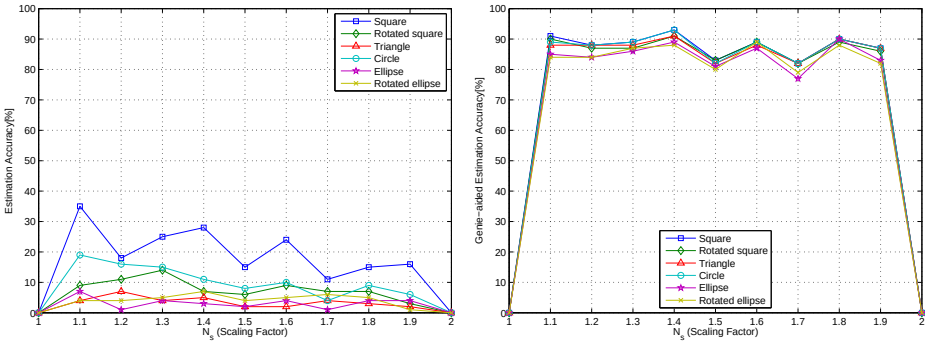
The classification of every single block is performed by analyzing the test statistic $\mathcal{T}_{xx}$ computed in each case. As we have said at the beginning of Section 4, the resampling factor is estimated from the cyclic frequency (α_1, α_2) with highest magnitude (excluding DC), and using the following relation:

$$\hat{\rho} = \max_{i \in \{1,2\}} \hat{\rho}_i = \max_{i \in \{1,2\}} \frac{2\pi}{|\alpha_i|}$$

where we have exploited the fact that, in this case, $\rho \geq 2$ since $1 \leq N_s \leq 2$. We consider that the detection of the tampered region is correct if we discover any inconsistency in the resampling factor (i.e. $\hat{\rho} \neq \rho_{\text{orig}}$) and the corresponding estimated resampling factor $\hat{\rho}$ satisfies $|\hat{\rho} - 2N_s| < 0.05$ and since we have the interference created by the CFA pattern we will also check if $|\hat{\rho}/(\hat{\rho} - 1) - N_s| < 0.05$ is satisfied.

Applying this approach to the tampered images of the database, we obtain the results shown in Fig. 8(a). As we have stated before, this method cannot detect copy-move forgeries, since there are not inconsistencies in the resampling factor along the whole image and that is the reason why the estimation accuracy is equal to zero at $N_s = 1$. Given the ambiguity inserted in the estimation, caused by the CFA pattern, we are not able to distinguish between a scaling factor $N_s = 1$ or $N_s = 2$, and that is why the estimation accuracy is also zero for $N_s = 2$. The rate of accurate estimation of the tampered region is not very high for all the masks used (in the best case we barely reach a 35%), so this method presents very bad performance when identifying forgeries.

Nevertheless, to demonstrate the generally good performance of the resampling estimator, we analyze the estimation accuracy in an ideal case where we



(a) Processing block by block. (b) Using exact matching (genie-aided).

Fig. 8. Estimation accuracy (in terms of percentage), obtained through the application of the resampling-based method in two different scenarios for several masks and scaling factors.

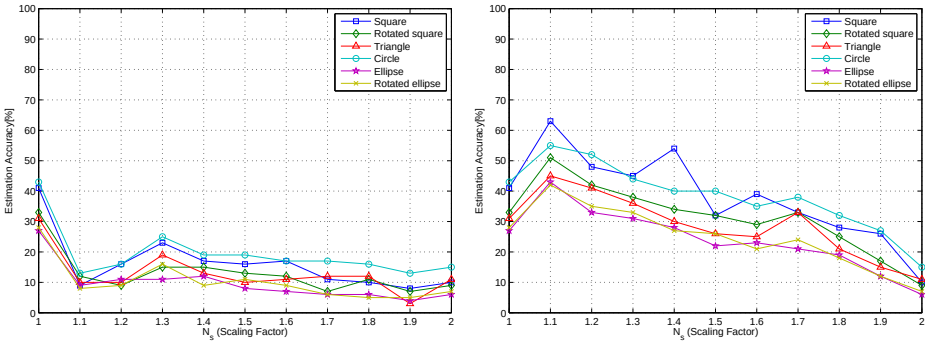
use the information supplied by a *genie* that tell us exactly the location of the original region and that of the tampered region (the application of a "genie-aided" detection is commonly used in communications to determine performance bounds). Thus, knowing exactly the location of both regions in the pixel domain and using the same criteria for the estimation of ρ, as in the previous scenario, we show in Fig. 8(b) the results of accurately estimate which region is the original and which is the duplicated. As it was said before, the correct distinction of the two regions when a spatial transformation has not been applied is not possible with the resampling-based method. However, the detection performance is very high (around a 90% for all the masks) if we compare it with the obtained when the image is processed block by block.

So, according to the results collected in this ideal case, the problem does not lie in the resampling estimator itself, but in the correct matching of the tampered area, and that is the reason why a SIFT-based method is needed.

4.3 Detection results combining both methods

As it was discussed along the paper, the combination of both methods provides a deeper and enhanced forensic analysis of the tampered regions (since we are able to identify which region is the source and which is the duplicated one) and it also brings a way to compensate the drawbacks of each method with the advantages of the other.

Certainly, since the SIFT-based method is not capable to find all the duplicated regions, mostly due to the unavoidable lack of reliable keypoints, combining both approaches we will get worse results than those depicted in Fig. 8(b) (i.e. the ideal "genie-aided" case where we perfectly match all the regions). However, with the use of the SIFT-based method, the detection of the tampered regions is more accurate than processing the image block by block, so we will get better



(a) SIFT-based method. (b) Proposed forensic tool.

Fig. 9. Comparative results of the estimation accuracy for the SIFT-based method and the proposed forensic tool.

results than those included in Fig. 8(a). Finally, since the estimation of the resampling factor is not so dependent on outliers as it is the case for the estimate of the homography, we will also get better results than those comprised in Fig. 9(a), where we represent the estimation accuracy of the SIFT-based method when it is able to correctly match the two regions and also estimate their geometric relation. Explicitly, the estimation accuracy ploted in Fig. 9(a), corresponds to the product of the accuracy rates achieved in the matching step (Fig. 7(a)) and in the post-detection estimation step (Fig. 7(b)).

In Fig. 9(b) we can see the inferred estimation accuracy of the proposed forensic tool for different masks and scaling factors. If we compare this plot with the corresponding estimation accuracy obtained with the SIFT-based method alone (depicted in Fig. 9(a)), we can observe that with the scheme described in Fig. 3, performance is improved for almost all the scaling factors and masks considered. It is important to note that the resampling estimator takes as input the exact matching of the detected regions by the SIFT-based method, so the results provided can be considered as an upper performance bound of the estimation accuracy that we can attain with this approach.

Note also that, even with the combination of both methods, we are still not able to distinguish the original region from the tampered one when a copy-move forgery is carried out. Besides, in this particular case, occasioned by the CFA interpolation of the camera, we are neither able to identify the duplicated regions by a factor $N_s = 2$. Hence, the estimation accuracy should be strictly zero for the scaling factors $N_s = 1$ and $N_s = 2$ in Fig. 9(b). However, since with the SIFT-based method we are able to match the involved regions in the tampering and also their relation, then we add the estimation accuracy of this method in both cases, and that is the reason why we have the same values of estimation accuracy for the scaling factors $N_s = 1$ and $N_s = 2$ in both graphs of Fig. 9.

By comparing the estimation accuracy of the resampling-based method (processing block by block) with that obtained with the concatenation of both meth-

ods, we achieve an improvement of the exact classification of each region for all the scaling factors and masks considered. In addition, as it was expected, the best results are reached with those masks that have the largest area.

According to the results shown in this section, we can conclude that the forensic analysis scheme proposed in this paper provides a more accurate analysis since we can identify in an image where are located and which are the original regions and the tampered ones when a region duplication forgery is performed. Moreover, the performance in terms of estimation accuracy is increased with respect to the use, in an independent way, of the SIFT-based and the resampling-based methods.

5 Conclusions and Further lines

In this paper we have introduced a new scheme for image forensic analysis, by combining two complementary methods. The former, based on SIFT, is capable of finding duplicated regions and the latter, based on a resampling estimator, allows one to identify which region is the source and which is the tampered one. The proposed forensic analysis scheme provides better estimation results than considering each method separately.

As future research lines we will focus on the application of this method using JPEG compressed images trying to get similar results as in the case of uncompressed images. Another interesting question is the estimation of the resampling factor on a non-square area, since the zero-padding technique is not the optimal one.

References

1. Sencar, H. T., Memon, N.: Overview of state-of-the-art in digital image forensics. In: Part of Indian Statistical Institute Platinum Jubilee Monograph series titled Statistical Science and Interdisciplinary Research. World Scientific Press, (2008)
2. Amerini, I., Ballan, L., Caldelli, R., Del Bimbo, A., Serra, G.: Geometric Tampering Estimation by Means of a SIFT-based Forensic Analysis. In: Proceedings of 2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), pp. 1702–1705, (2010)
3. Pan, X., Lyu, S.: Region duplication detection using image feature matching. In: IEEE Transactions on Information Forensics and Security, vol. 5, no. 4, pp. 857–867, (2010)
4. Vázquez-Padín, D., Mosquera, C., Pérez-González, F.: Two-dimensional statistical test for the presence of almost cyclostationarity on images. In: Proceedings of 2010 17th IEEE International Conference on Image Processing (ICIP), pp. 1745–1748, (2010)
5. Mahdian B., Saic, S.: Blind authentication using periodic properties of interpolation. In: IEEE Transactions on Information Forensics and Security, vol. 3, no. 3, pp. 529–538, (2008)
6. Kirchner, M., Gloe, T.: On resampling detection in re-compressed images. In: Proceedings of 2009 First IEEE International Workshop on Information Forensics and Security (WIFS), pp. 21–25, (2009)

7. Lowe, D. G.: Distinctive image features from scale-invariant keypoints. In: Int. J. Comput. Vision, vol. 60, pp. 91–110, (2004)
8. SIFT Keypoint Detector, http://www.cs.ubc.ca/~lowe/keypoints
9. RANSAC algorithm, http://www.csse.uwa.edu.au/~pk/research/matlabfns