Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 653164, 6 pages
http://dx.doi.org/10.1155/2015/653164

*Research Article*

# A Feature-Based Forensic Procedure for Splicing Forgeries Detection

**Irene Amerini,[1] Rudy Becarelli,[1] Roberto Caldelli,[1,2] and Matteo Casini[1]**

[1] *Media Integration and Communication Center (MICC), Università degli Studi di Firenze, 50134 Firenze, Italy*
[2] *National Interuniversity Consortium for Telecommunications (CNIT), 43124 Parma, Italy*

Correspondence should be addressed to Roberto Caldelli; roberto.caldelli@unifi.it

Received 4 June 2015; Revised 11 November 2015; Accepted 20 December 2015

Academic Editor: Fazal M. Mahomed

Nowadays, determining if an image appeared somewhere on the web or in a magazine or is authentic or not has become crucial. Image forensics methods based on features have demonstrated so far to be very effective in detecting forgeries in which a portion of an image is cloned somewhere else onto the same image. Anyway such techniques cannot be adopted to deal with splicing attack, that is, when the image portion comes from another picture that then, usually, is not available anymore for an operation of feature match. In this paper, a procedure in which these techniques could also be employed will be shown to get rid of splicing attack by resorting to the use of some repositories of images available on the Internet like Google Images or TinEye Reverse Image Search. Experimental results are presented on some real case images retrieved on the Internet to demonstrate the capacity of the proposed procedure.

## 1. Introduction

When looking at a digital image, it is quite common to wonder if it is original or has been counterfeited in some way. Such a doubt is basically determined by the easiness with which digital images can be manipulated to change their content and especially their visual meaning. The contexts where doctored pictures could be involved are very disparate; they could be used in a tabloid or in an advertising poster or included in a journalistic report but also in a court of law where digital (sometimes printed) images are presented as crucial evidences for a trial in order to influence the final judgement. So, especially in the last case, reliably assessing image integrity becomes of fundamental importance. *Image forensics* specifically deals with such issues by studying and developing technological tools which generally permit determining, by only analyzing a digital photograph (i.e., its pixels), if that asset has been manipulated or even which could have been the adopted acquisition device (such an issue is not relevant to the topic of the present paper). Moreover, if it has been established that something has been altered, it could be important to understand in which part of the image itself such a modification occurred, for instance, if a person or a specific object has been covered, if an area of the image has been cloned, if something (i.e., a face or a weapon) has been copied from another different image, or, even more, if a mixture of these processes has been carried out. Among the different attacks that can be carried out to modify an image, two are surely the most important. The first one is the splicing attack which is performed when a portion of an image has been cut out and, after having been adapted (e.g., zoomed in or out, filtered), is inserted into another one to build a new "fake image." The second one is the *copy-move attack* which is basically a splicing attack but what is crucial is that the clipped image portion is pasted somewhere else within the same image. On the other side, regarding forgeries individuation, three are the principal classes of detectors studied so far: those based on double JPEG compression [1–3] adopted to reveal splicing attack, those based on inconsistent shadows [4], and finally those based on local features descriptors (mainly SIFT—Scale Invariant Feature Transform) [5–8] usually used to get rid of copy-move attack. A complete overview of forensic methods for tampering detection is well introduced in [9]. In particular, features-based methods (based on SIFT or others feature descriptors) have been demonstrated so far to be effective against copy-move attack and constitute one of the most promising techniques addressing this issue because they are resistant to JPEG compression, scaling, rotation, and affine transformations and also to digital/analog/digital conversion [10, 11]. Because of their effectiveness, though this kind of techniques has been presented to operate in a copy-move attack scenario, we have decided to propose and investigate, in this paper, a framework in which they can be utilized to cope with situations in which a splicing operation has occurred. The basic problem is that usually only the final forged photo is available for inspection to the forensic analyst and the source image(s) is often unobtainable. Because of that, the similarity matching procedure, which is the cornerstone of the feature-based methods, could not take place. However, in practical scenarios, this is not completely true. In fact, it happens that the forensic analyst is asked to give an assessment on a digital image which is, in some way, related to an image dataset. This could be the case, for example, of a certain person under judgment, whose hard disk or pen drive or, more generally, his social network profile has been confiscated and a set of his images is under investigation. The circumstance in which contents coming from multiple web-source images can be combined, for example, in montages or splicing, to create what is called an image composition is very frequent. As an example, we may consider pornographic compositions by involving *vip* pictures (such as celebrities or politicians) with the purpose of public shaming or sometimes for bullying. It can be comprehended that, in such an operative scenario, it can happen that the source images, used to generate a fake content, potentially belong to a huge available collection created by the images shared by users on the web. It is easy to understand that it is necessary to set up a method that can be adopted to determine if a set of near-duplicates of the to-be-checked image exists among the images indexed, for example, on the web and then assess, within such retrieved collection, a degree of similarity between the image under test and each photo of the collection. Furthermore it should be possible not only to detect the image as being a composition, but also to recover the source image(s) used to create it. Succeeding in detecting such links and in automating the retrieval and forgery detection procedures could help investigation activities. To do this, it is necessary to accurately evaluate and compare a set of images, as well as precisely localizing the common and the uncommon content between images. So our idea, starting from the method in [7], consists of building a framework to conveniently address the splicing attack issue by resorting to the use of repositories of images available on the Internet (Google Images Search, TinEye Reverse Image Search, etc.), ranking the images with respect to a *degree of similarity*, selecting a set of near-duplicates candidates, and then automatically detecting and localizing the tampered regions.

The remaining part of this paper is structured as follows. In Section 2, the proposed procedure is discussed, while Section 3 is devoted to the description of the core algorithm. Section 4 contains experimental results and Section 5 concludes the paper.

## 2. The Proposed Procedure

On the basis of the previous consideration, in this section, a general procedure in which the authenticity of a forged image is verified by resorting to the images indexed on the web is presented. Two are the main phases to be analyzed: first, web crawling, ranking, and image set selection and second, forgery detection and localization. It is easy to understand that all the cases described before in which an image subset is already available (e.g., recovered from a social network profile, a hard disk, and a SD card) can be catalogued as subcases of this general one by omitting the first of these two procedural phases.

*2.1. Crawling and Selecting Phase.* When dealing with user-generated content distributed online, forgeries could be created starting from a content available on different media sharing platforms. A typical example is a splicing forgery operated to substitute the face of a person with that of a celebrity. It is then possible to search for copies of the entire image (i.e., versions of the same image differing because of processing operations or pictures of the same scene captured from a slightly different point of view) or of a detail of the image under analysis (e.g., face, body). This search can be either performed via web crawling or in a dataset under analysis. In particular, in our method, given a to-be-checked image, we perform an image search on one of the repositories on the web; the image search engines, like Google Images and TinEye Reverse Image Search, have been chosen to collect a set of near-duplicates candidates. Google Images and TinEye are reverse image search engines that resort to image identification technologies rather than keywords, metadata, or watermarks to perform the retrieval; they regularly crawl the web adding new images in the dataset. In this kind of search engines, an image is given as an input to the system instead of a keyword and the output is a similar matching image linked to the image source (i.e., search by image). After the search is executed, a list of sorted results is produced. By default, all of these retrieved images (for instance, at most by establishing a threshold on the number of page rank; see Section 4 for a specific threshold setting) could be passed to the successive step of forgery detection and localization, but, to reduce the amount of comparisons to be done within the second phase of the procedure, a selection functionality to skim the raw results has been envisaged and various solutions are still under analysis. Partially supervised solutions can also be considered when the system is not able to discern automatically. However, one of the main selection solutions is based on the assumption that, generally, the input image is itself found in the first positions/pages of the obtained results; so by applying a difference operator (e.g., PSNR and SSIM) and comparing it with a predefined threshold (such a choice directly impacts the capacity of the system to reveal small modified zones), all the images that do not show significant differences are discarded by the system. Then the target of the following phase is to compare the image under observation with each of the selected candidates to find differences or similarities between them.

*2.2. Detection and Localization of Common or Uncommon Parts.* All the images that have been identified as near-duplicate of the test image are compared by means of the proposed method (see Section 3 for details) to find differences or similarities; then, if a forgery exists, it is detected and then spatially localized. In particular, given a pair of images, the objective is to detect the differences between them by segmenting the parts that are not in common when the patch involved in the forgery is small with respect to the background or otherwise revealing the common parts and cosegmenting the rest. The latter objective is achieved by generalizing the *Copy-Move Forgery Detection* technique presented in [7] by separating simultaneously the common parts (cosegmentation of a pair of images [12, 13]). In particular, the test image is registered onto its near-duplicate in order to compensate for geometrical transformations such as cropping and resizing. This is done by using SIFT matching and J-Linkage clustering [7] by adjusting some settings like clustering threshold, number of affine transformations, and so on (such settings will be described in Section 3). In Figure 1, an intermediate result in which the matched SIFT keypoints are evidenced is proposed; some false alarms that later will be eliminated are visible too. A correlation operation is performed between the near-duplicate image and the registered version of the test one by obtaining a correlation map. Basically, two different cases can be distinguished by taking into account that the forgery is done by inserting an image portion which is relatively small with respect to the whole image size. The former (*fake-background case*) is the one in which such a portion is taken from an original image and inserted in a fake context: so in this case the correlation map is taken for successive localization. The latter (*fake-foreground case*) is the one in which the portion is extracted from an image and pasted in an original context: in this case, on the contrary, the inverse correlation map is considered for successive localization. To better understand these two circumstances, the reader can refer to Section 4 for the cases named *Naomi Campbell* and *Barack Obama*, respectively. Obviously, for each of the near-duplicate images, a different localization result is obtained: at the moment, a predefined number of result images (usually at most 10) are provided as output of the entire procedure. Anyway, in the next future works, operations of refinement and/or merging among the result images can be envisaged to improve the final answer. Anyway it is important to underline that the basic aim of the proposed methodology is to retrieve and evidence discrepancies, if any, between images under comparison, with the goal of providing a support to image authenticity assessment. As it often happens in image forgery detection, the process to draw conclusions is then left to the final analysis of an end-user that is able to contextualize the image itself.

## 3. The Core Algorithm

In the following we review the core algorithm of the proposed method based on [7]. Such a method is subdivided into two steps in which the first one is devoted to SIFT feature extraction and to the keypoint matching, while the second one is in charge of performing clustering and localization of the doctored regions. In the following subsections, these two steps are briefly described.

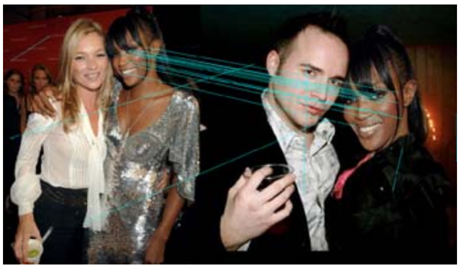


Figure 1: An intermediate result of the procedure: the matched SIFT keypoints between cloned regions.

*3.1. SIFT Computation and Keypoint Matching.* The initial step of the core algorithm resorts to SIFT features; in fact, they are quite well-known as being robust to rotation, scaling, and affine transformations and so they are well-suited for the forgeries recognition in near-duplicate images as has been debated in [5, 6]. Stable local extrema, detected in the scale space, are taken as keypoints and a feature vector is computed, for each of them, from a pixel neighborhood around the individuated point. Let $\mathscr{K} = \{\mathbf{k}_1, \ldots, \mathbf{k}_n\}$ be the set of $n$ interest points taken from an image $I$, $\mathbf{k}_i = \{\mathbf{c}_i, \mathbf{f}_i\}$ being the vector containing the coordinates $\mathbf{c}_i = (x, y)$ of the keypoint and $\mathbf{f}_i$ the feature descriptor of the zone around the keypoint (such descriptor is composed of 128 elements that represent the histogram of local gradient orientations). Match keypoints can be achieved by straightforwardly fixing a global threshold where the Euclidean distances among descriptors are compared; because of the high dimensionality of the feature space, this method can result in a low accuracy, some descriptors being much more discriminative than other ones. On this basis, it has been proposed in [14] that, given a keypoint, also the distance with the second keypoint of the ranked list, not with the first one only, is to be taken into account; in particular, the ratio between the distance with the candidate match and the distance with the second similar point (i.e., the so-called *2NN test*) is suggested to be considered. If this ratio is lower than a threshold $\tau$ (usually equal to 0.55 or 0.60), a match is declared:

$$\frac{d_1}{d_2} < \tau, \quad \text{where } \tau \in (0, 1). \tag{1}$$

However, such an approach works well if a region is copied once, but if it is copied multiple times, which could happen in image forensic applications, performances decrease. So, to deal with this circumstance, a generalization of the matching technique, proposed in [6] and named *g2NN test*, has been used. Such a generalization consists in an iteration of the *2NN test* by computing the ratio $d_i/d_{i+1}$ until this ratio is under the predefined threshold $\tau$. The value in which the procedure ends being $l$, each keypoint with distance ratio within the range $\{d_1, \ldots, d_l\}$ $(1 \leqslant l < n)$ is labeled matching with the keypoint under inspection. By using the *g2NN* strategy on all the keypoints, a set of $s$ matched pairs $\mathscr{P} = \{\mathbf{p}_1, \ldots, \mathbf{p}_s\}$, where $\mathbf{p}_i = (\mathbf{k}, \mathbf{k}')$, to be input in the following step is obtained.

*3.2. J-Linkage Clustering and Cloned Regions Localization.* J-Linkage clustering algorithm [15] does not operate in the spatial domain of the matched points but conversely in the geometric transformation domain. In particular, an adaptation of the J-Linkage algorithm has been introduced in [7] to provide a solution to the classical drawbacks spatial clustering generates: first, the inability to separate cloned regions that are very close to each other and, second, the difficulty to identify a patch as unique when it contains nonuniform spatially distributed keypoints. J-linkage clustering randomly samples the input matched pairs and consequently estimates $m$ affine geometric transformations. A preference set vector (PS) is associated with each pair and it indicates which geometric transformations of available $m$ are preferred according to a fixed threshold. The PS vector is consequently a binary vector and represents each pair in a *conceptual space* $\{0, 1\}^m$. Since the matched pairs belonging to the original and to the duplicated patches share similar transformations, consequently they will show similar conceptual representations. Each preference set vector is assigned to a cluster within a hierarchical agglomerative clustering procedure in which, cyclically, the two clusters with smallest distance in the conceptual space are merged. Finally, every cluster will get at least one transformation shared by all its matched pairs. Outlier transformations that fit with a number of elements under a fixed threshold are discarded, while, if one or more transformations are detected, a copy-move operation is revealed. The coordinates of the matched pairs previously selected, $\{(\mathbf{c}, \mathbf{c}')_1, \ldots, (\mathbf{c}, \mathbf{c}')_{z+1}\}$, constitute the values for the geometric transformation estimation. In particular, affine transformations have been considered in order to model geometric distortions such as scaling, rotation, and shearing between the original and the copied patches. An affine transformation has six degrees of freedom and, consequently, can be computed by resorting to three noncollinear matched pairs. Such a computation is performed by using the normalized Direct Linear Transformation (DLT) algorithm for affine homography [16] in which, given a set of correspondences $(\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_{(z+1)})$ and $(\mathbf{c}'_1, \mathbf{c}'_2, \ldots, \mathbf{c}'_{(z+1)})$, the algorithm tries to minimize the objective function written in

$$\sum_{i=1}^{z+1} \left\| \mathbf{c}'_i - \mathbf{H}\mathbf{c}_i \right\|^2 . \quad (2)$$

Such a linear method allows managing well the $m$ affine transformation hypotheses, $m$ having a rather high value (in our experiments $m = 500$) due to the random sampling of the pairs. Once the homography has been determined, rotation and scaling parameters can be then calculated and translation is obtained by means of the centroids of the two matched clusters. When an image is recognized as forged, the system tries to achieve accuracy as possible localization of the duplicated regions. By applying the estimated transformation $\mathbf{H}$ to the entire image, all pixels of the original region are linked to those of the duplicated one:

$$\mathbf{R}^D = \mathbf{H}\mathbf{R}^O . \quad (3)$$

After that, a warped image is obtained in which the original region will overlap the duplicated one. A block-wise correlation measure (between the two images), based on Zero Mean Normalized Cross-Correlation (ZNCC), is computed and a binary thresholding is performed to localize cloned regions.

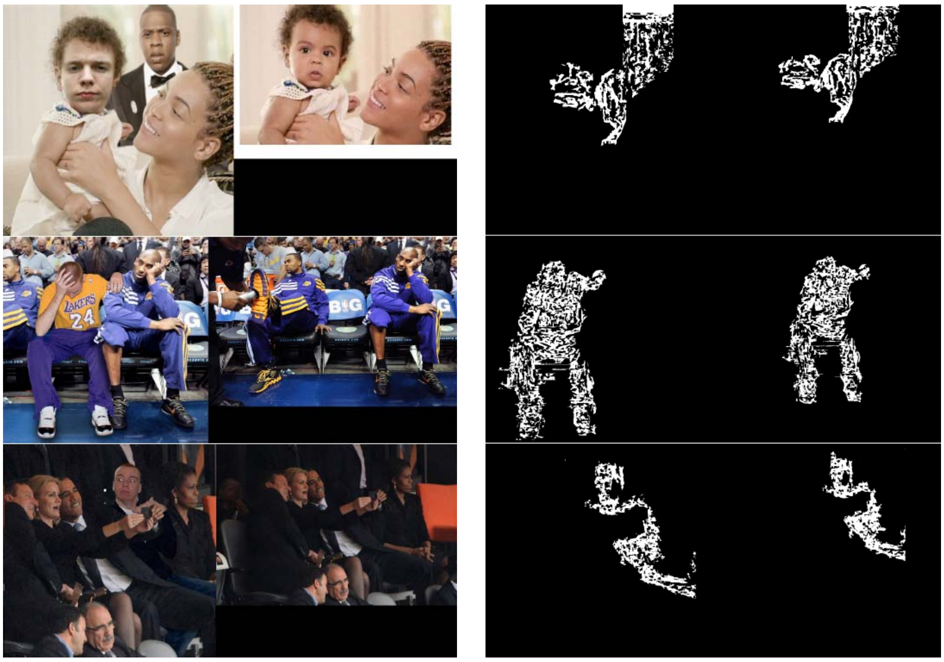
## 4. Experimental Results

This section shows some experiments and the consequent obtained results to demonstrate the validity of the proposed method. We analyzed different popular images with various size appearing on the web and/or social networks (mostly last year); most of them are taken from Instagram (https://instagram.com/peejet/) or Pinterest (https://www.pinterest.com/everetthiller/); others have been gathered from international online newspapers (e.g., http://www.dailymail.co.uk/home/index.html) and from web sites, like Celeb Jihad (http://www.celebjihad.com/), that generally create photomontages to rubbish celebrities.

Some of them, in which famous people are involved, have been chosen to be proposed hereafter and to point out the effectiveness of the proposed framework. In the first stage, we employed the Google Image and TinEye image search to find the similar images for each of the test image. As stated in Section 2.1, when the test image itself is retrieved, this is discarded (the first 20 pages of results have been considered) and only those ones presenting differences with respect to the predefined threshold are passed to the detection/localization phase. This selection phase is anyway quite complex, because sometimes it can happen that different images (i.e., original source image) are not retrieved though they are really present on the web. In fact, in this case, it could help to assist the image search engine, for example, by refining the search with some keywords (this is possible in Google Image) or by performing the image search by using a subpart of the test image (e.g., the supposed authentic image area). Such functionalities have not been fully integrated within the proposed procedure yet and these circumstances have been managed in a semiautomatic manner: an interesting analysis of such an issue is presented in Table 1 where the ranking position of the selected near-duplicate in the list of results obtained with Google Images and TinEye is reported for some sample cases. It can be easily understood how diverse situations can occur and, above all, how important providing additional or specific (image subparts) information to improve the raw image search can be, as for the case "Barack Obama" in which just simply by adding the keyword "obama" the ranking position moves from 96th to 1st for Google Image. Both search engines give almost similar results in most cases, even if Google Images has a much larger indexed database than TinEye. So, after that, the second phase of the proposed procedure is launched to check out all the possible selected couples of photos (the image under test and the selected duplicate) looking for common/uncommon areas. The method is able to detect and then segment the "supposed forged" areas; some sample results obtained for the images considered within the experimental tests are pictured in Figures 2 and 3. In particular, in Figure 2 the cases named *Beyonce*, *Kobe Bryant*, and *Barack Obama* are presented: in Figure 2(a)

Table 1: Ranking position of the selected duplicate among the list of results. NF stands for *Not-Found*; in brackets is the position when a specific text keyword is added to the image search or when a portion of the test image is cut out and used for the search.

| Image | Google Images | TinEye |
|---|---|---|
| *Beyonce* | 1 | 6 |
| *Kobe Bryant* | NF (1 with "kobe bryant bench") | 19 |
| *Barack Obama* | 96 (1 with "obama") | 1 |
| *Rihanna* | 1 | 3 |
| *Naomi Campbell* | NF (34 with "naomi") | NF (NF by cutting out Naomi Campbell) |
| *Mario Balotelli* | 4 | 15 |
| *Jenny Thompson (Escort)* | NF (1 by cutting out Jenny Thompson) | NF (NF by cutting out Jenny Thompson) |



(a)

(b)

Figure 2: In (a), the test image is on the left and its near-duplicate is on the right. In (b), the segmentation of the uncommon parts revealing the splicing is presented.

the test image and the retrieved near-duplicate can already visually show the altered areas, while in Figure 2(b) the cosegmented masks of the uncommon zones can further help in determining the modifications. In Figure 3, another output of the procedure is given in terms of localization after the refinement of the binary map: the test images named *Rihanna*, *Naomi Campbell*, and *Balotelli-Thompson* are pictured. In the image *Rihanna*, it is interesting to point out that the system is able to also give evidence of the small forged areas in the people behind *Rihanna* (uncommon areas are evidenced in red). The image *Naomi Campbell* contains a portion coming from the original which is small with respect to the fake destination image but the procedure result is not worsened by such an issue. Finally, the image *Balotelli-Thompson* is interesting because there is a mixed composition generated from two diverse sources where the two subjects are singularly present (common areas are highlighted in blue).

## 5. Conclusion

A novel forensic procedure to detect and localize splicing forgeries by means of a feature-based technique has been presented. The use of automatic web search to retrieve near-duplicate images has been introduced to support the implementation of such image forensic methods. Future works will be dedicated to finding solution to better automatize the entire pipeline both for the ranking phase and for merging of the different achieved localization maps. Alternative methods for features extraction which performs better than SIFT onto flat areas could also be tested and implemented.

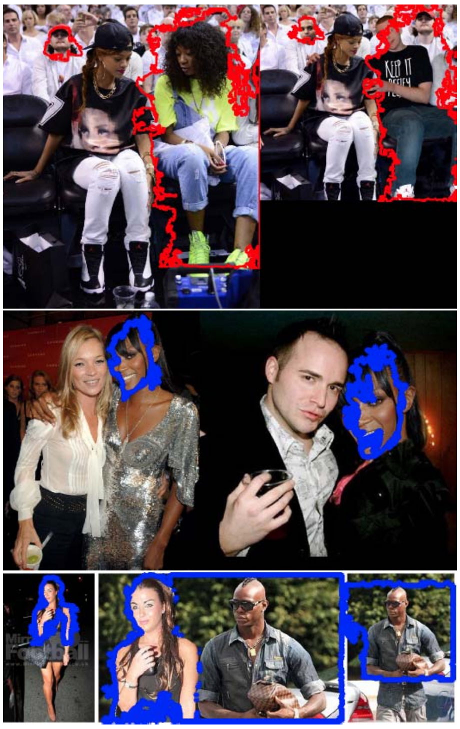


Figure 3: The localization results after the refinement of the binary map. The near-duplicate is on the left and the test image is on the right. In the last row, the test image is in the center with respect to the retrieved two near-duplicates.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

This work was partially supported by the SMARTVINO Project funded by the PRAF 2012–2015-1.2.e programme of the Tuscany Region (Italy).

## References

[1] L. Zhouchen, H. Junfeng, T. Xiaoou, and T. Chi-Keung, "Fast, automatic and fine-grained tampered JPEG image detection via DCT coefficient analysis," *Pattern Recognition*, vol. 42, no. 11, pp. 2492–2501, 2009.

[2] T. Bianchi and A. Piva, "Image forgery localization via block-grained analysis of JPEG artifacts," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 3, pp. 1003–1017, 2012.

[3] I. Amerini, R. Becarelli, R. Caldelli, and A. D. Mastio, "Splicing forgeries localization through the use of first digit features," in *Proceedings of the IEEE International Workshop on Information Forensics and Security (WIFS '14)*, pp. 143–148, Atlanta, Ga, USA, December 2014.

[4] E. Kee, J. F. O'Brien, and H. Farid, "Exposing photo manipulation with inconsistent shadows," *ACM Transactions on Graphics*, vol. 32, no. 3, article 28, pp. 1–12, 2013.

[5] X. Pan and S. Lyu, "Region duplication detection using image feature matching," *IEEE Transactions on Information Forensics and Security*, vol. 5, no. 4, pp. 857–867, 2010.

[6] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, and G. Serra, "A SIFT-based forensic method for copy-move attack detection and transformation recovery," *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 1099–1110, 2011.

[7] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, L. Del Tongo, and G. Serra, "Copy-move forgery detection and localization by means of robust clustering with J-Linkage," *Signal Processing: Image Communication*, vol. 28, no. 6, pp. 659–669, 2013.

[8] V. Christlein, C. Riess, J. Jordan, C. Riess, and E. Angelopoulou, "An evaluation of popular copy-move forgery detection approaches," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 6, pp. 1841–1854, 2012.

[9] 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.

[10] I. Amerini, R. Caldelli, A. Del Bimbo, A. Di Fuccia, L. Saravo, and A. P. Rizzo, "Copy-move forgery detection from printed images," in *Media Watermarking, Security, and Forensics*, vol. 9028 of *Proceedings of SPIE*, San Francisco, Calif, USA, February 2014.

[11] I. Amerini, R. Caldelli, A. Del Bimbo, A. Di Fuccia, A. P. Rizzo, and L. Saravo, "Detection of manipulations on printed images to address crime scene analysis: a case study," *Forensic Science International*, vol. 251, pp. e9–e14, 2015.

[12] D. S. Hochbaum and V. Singh, "An efficient algorithm for co-segmentation," in *Proceedings of the 12th International Conference on Computer Vision (ICCV '09)*, pp. 269–276, IEEE, Kyoto, Japan, October 2009.

[13] S. Vicente, C. Rother, and V. Kolmogorov, "Object cosegmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '11)*, pp. 2217–2224, Providence, RI, USA, June 2011.

[14] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, 2004.

[15] R. Toldo and A. Fusiello, "Robust multiple structures estimation with J-linkage," in *Computer Vision—ECCV 2008: 10th European Conference on Computer Vision, Marseille, France, October 12–18, 2008, Proceedings, Part I*, vol. 5302 of *Lecture Notes in Computer Science*, pp. 537–547, Springer, Berlin, Germany, 2008.

[16] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, Cambridge, UK, 2004.