Survey on Retargeting Techniques used in Multimedia Forensic of Mobile devices

Shashikant S. Dumbare
ABHA Gaikwad-Patil College of Engineering,Nagpur

Sulbha Patil
ABHA Gaikwad-Patil College of Engineering

ABSTRACT
This is a survey paper on Multimedia forensic in mobile device. By observing numbers of multimedia forensic technique, This paper focus on Seam carving technique. Seam carving is an adaptive multimedia retargeting technique to resize multimedia data for different display sizes.This technique has found promising applications in media consumption on mobile devices such as tablets and smartphones. However, seam carving can also be used to maliciously alter image content and when combined with other tampering operations, makes tampering detection very difficult by traditional multimedia forensic techniques. In this paper, we study the problem of seam carving estimation and tampering localization using very compact side information called forensic hash. The forensic hash technique bridges two related areas, namely robust image hashing and blind multimedia forensics, to answer a broader scope of forensic questions in a more efficient and accurate manner. We show that our recently proposed forensic hash construction can be extended to accurately estimate seam carving and detect local tampering.

General Terms
Algorithms, Security

Keywords
Forensic hash, Seam carving, Visual words, SIFT

1. INTRODUCTION
In recent years, mobile devices with multimedia capturing capability, and social networks that provide media sharing and streaming services, are rapidly emerging. This has significantly increased multimedia generation and consumption over the internet, and brought great social impact. However, the digital nature of multimedia allows easy modification of its content. Multimedia data can be intentionally altered to create a forgery and convey a different meaning. For example, objects can be removed from or inserted into an image, and multiple pieces of content may be combined into a new creation. As such, it is critical to provide detailed evaluation on trustworthiness of online multimedia information. In this paper, we develop forensic techniques to detect one increasingly popular image processing operation called seam carving and estimate its parameters using carefully designed side information.

Seam carving is a specific type of content aware image resizing, also known as image retargeting techniques [7], proposed by Avidan et al. [1]. It resizes an image in a way adaptive to its content by removing seams, which are eight-connected paths of low energy, from the image while keeping salient objects intact. Details on the seam carving algorithm can be found in [1]. Due to its capability of preserving salient objects and its aspect ratio after resizing, seam carving has found promising applications in showing images and videos on smaller displays, such as mobile phones.

The objective of seam carving is better image resizing, but as demonstrated in [1], seam carving can also be used to intentionally remove objects from the image. Such tampering brings challenges to forensic tasks. Traditional blind multimedia forensics try to detect potential tampering of digital images/videos without proactive aids such as signature attachment or embedded watermark [2]. This is accomplished by analyzing intrinsic traces, such as inconsistencies in signal characteristics, left by the processing operations [10]. However, the adaptive nature of seam carving makes it difficult to identify any traces or inconsistencies unique to the operation.

Recent work by Sarkar et al. [9] and Fillion et al. [3] detect whether an image has undergone seam carving or not by using a machine learning framework based on intuitive features extracted from the image. The accuracy of seam carving detection in the above mentioned works is around 80-90% for large amount of seam carving (e.g., number of seams larger than 30% of original image size), and becomes lower for smaller amount of seam carving. In contrast to answering only a binary question of whether the image is seam carved or not using blind forensic techniques, in this paper, we explore using compact side information to not only detect seam carving but also estimate the amount and location of seam carving. Such detailed information can help us better evaluate the trustworthiness of the image.

The idea of using side information to assist multimedia forensic analysis is explored in [5, 6]. The carefully designed side information is called forensic hash in the sense that it is as compact as traditional image hash [11] but provides forensic capability beyond binary authentication.

Compared to blind multimedia forensics, forensic hash needs to be securely attached to the image during transmission or transmitted through another trusted channel, but it is designed to answer broader scope of forensic questions with improved e_ciency and accuracy.

In this paper, we use the forensic hash based on visual words representation of SIFT features [6] and extend it for seam carving detection and estimation. Experiments show that forensic hash can accurately estimate the amount of seam carving and their approximate locations. With the estimation results, we further explored reconstruction of original image from the seam carved image for tampering detection. Such a forensic analysis can provide a detailed trustworthiness evaluation of the image in terms of which part can be trusted and which part might be tampered.

2. FORENSIC HASH CONSTRUCTION
In this section, we brief review the concept of forensic hash and its construction based on visual words representation of SIFT features. Details can be found in [6]. The research on forensic hash is related to robust image hashing and traditional

blind multimedia forensics. Their relation is illustrated in Fig. 1. Blind forensic techniques

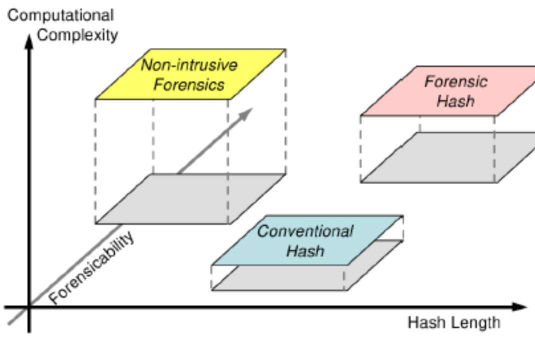


Figure 1: Motivation of forensic hash

answer forensic questions without requiring any signature attachment, but typically incur high computational cost. Traditional image hashing uses short signature for efficient image authentication, but only provide a binary answer on image integrity. A good design of forensic hash, on the other hand, encodes compact side information from the original image as in conventional hashing, but can provide enhanced

forensic capability in a more computationally efficient way than blind forensics.

For modularity and extensibility, a forensic hash may contain several components, as illustrated in Fig. 2, each targeting different operations, complementing each other, and working in a synergistic way. For example, it is shown in [5, 6] that alignment is a necessary step to enable accurate image integrity evaluation using block-based features.

The forensic hash used here is a compact representation of stable SIFT features in an image [6]. SIFT [4] with high contrast values are robust to a wide range of image operations. The characteristic scale and dominant orientation of SIFT points can be used to estimate geometric transforms such as rotation and scaling. To compactly represent high

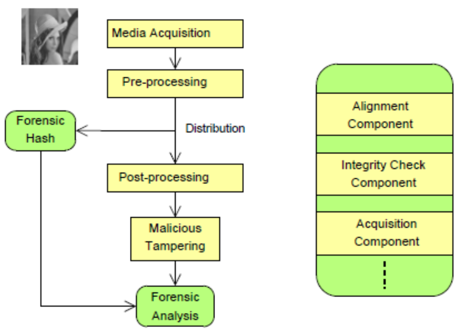


Figure 2: Flowchart of media processing and modular design of forensic hash

dimensional SIFT descriptors, [6] used visual words representation to store only the visual word ID rather than the full descriptor. More speci_cally, the generation of forensic hash involves the following steps: _rst the top k stable SIFT points are selected and their descriptors are hierarchically quantized based on a pre-trained vocabulary tree to get the visual word IDs; then each point is represented by a vector of 5 parameters: visual word ID, x and y positions in the image, its characteristic scale and dominant orientation, denoted as (id; x; y; _; _). For a vocabulary tree with 1000 visual words and an image size of 1024x1024, each vector would take around 50 bits.

When distributing the image, the forensic hash can be attached to the image along with a digital signature signed by the trusted content provider, such as news agencies or big media websites. Given an image claimed to be from certain content provider, any receiver can retrieve the public key from the content provider and verify that the hash is authentic. To estimate geometric transform that a received image may have undergone, its stable SIFT points and their 5-parameter vectors are computed. We first find corresponding SIFT points between the received image and the original image by looking for visual word IDs that have single occurrence in both images. These point pairs are denoted by (p1; ~p1), (p2; ~p2), (pn; ~pn). Each matching pair gives an estimate of the scaling factor and rotation angle. Since there can be mismatches among the point pairs, the scaling factor and rotation angle are estimated using robust estimation algorithms such as RANSAC.

The advantage of such a forensic hash construction is that it provides accurate and robust geometric transform estimation and enables further forensic analysis such as tampering localization, as demonstrated in [6].

3. SEAM CARVING ESTIMATION

Below, we formally describe the seam carving estimation algorithm using forensic hash. We denote the original image as I and its forensic hash h = {v1, v2,…Vk}, where vi is the parameter vector of the ith stable SIFT point in I. Image I is transmitted and undergoes seam carving and potentially additional geometric transforms and tampering operations such as cut-and-paste. We denote the received image as ~I and its top stable SIFT points as $\tilde{\mathrm{h}} = \{\tilde{\mathrm{v}}_1, \tilde{\mathrm{v}}_2, \cdots, \tilde{\mathrm{v}}_k\}$. Given h and ~h, we match their SIFT points based on their visual word IDs and denote the matched points as $(p_1, \tilde{p}_1)$, $(p_2, \tilde{p}_2)$,$\cdots$,$(p_n, \tilde{p}_n)$. Before seam carving estimation, we _rst need to make sure~I is on the same scale and orientation as I. This is achieved by estimating the rotation angle _ and scaling factor _ using the matched points, as described in [6]. The image~I is then transformed to be $S(R(\tilde{\mathbf{I}}, -\theta), 1/\delta)$, thus aligned to I. Here R(.,.) is the rotation operator and S(.,.) is the scaling operator. For simplicity, below we still denote the transformed image as $\tilde{\mathbf{I}}$.

To locate the vertical seams removed from the image, we sort the matched points based on their x-coordinates and then compare the distances between every adjacent matched pairs. An illustrative example is given in Fig. 3, where xi, xi+1, xi+2 are the x-coordinates of three adjacent points from original image I and $\tilde{x}_i, \tilde{x}_{i+1}, \tilde{x}_{i+2}$ are the x-coordinates of corresponding points in the resized image $\tilde{\mathbf{I}}$.We denote the distance between two adjacent points in I as di = xi+1 □ xi and the distance between corresponding points in ~I as $\tilde{d}_i = \tilde{x}_{i+1} - \tilde{x}_i$, then the number of vertical seams removed in the horizontal range [~xi; ~xi+1] in ~I can be computed as $\Delta C_i = d_i - \tilde{d}_i$. A positive ΔC_iindicates seam removal and a negative one indicates seam insertion. For the example given in Fig. 3, we can see that there are seam insertions in the range $[\tilde{x}_i, \tilde{x}_{i+1}]$ and seam removal in the range $[\tilde{x}_{i+1}, \tilde{x}_{i+2}]$. After considering all adjacent matched pairs, we obtain the estimation results $\{\Delta C_i, \tilde{x}_i, \tilde{x}_{i+1}\}$where $i \in \{1, \cdots, n\}$. Furthermore, in order to estimate the number of seams removed before the _rst point ~p1 and that after the last

point ~pn, we can include the size of the original image into the forensic hash and obtain the complete estimation results as $\{\Delta C_i, \tilde{x}_i, \tilde{x}_{i+1}\}$ where $i \in \{0, \cdots, n+1\}$,
$\Delta C_0 = x_1 - \tilde{x}_1$, $\Delta C_{n+1} = (w - x_n) - (\tilde{w} - \tilde{x}_n)$. w and ~ w are the width of images I and~I, respectively. To locate horizontal seams, we sort the matched points along y-coordinates and follow the same procedure.

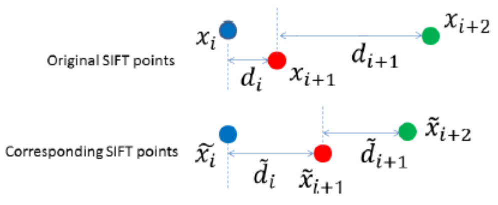


Figure 3: Vertical seam estimation using x-coordinates of matched SIFT point pairs

To give an example, we show the original image and its 50 vertical seams to be removed in Fig. 4(a). The resized image and its stable SIFT points with contrast value larger than 0.05 are shown in Fig. 4(b).

The green circles are SIFT points matched with the original SIFT points encoded in the forensic hash and red ones are those not matched due to seam emoval. By comparing the original and new distances of every adjacent pairs of matched points, we estimated that there are 30 vertical seams removed in the horizontal range of [1,9] in the resized image, 2 seams in the range of [9,30], 17 seams in the range of [285,363], and 1 seam in the range of [404,418]. Compared with ground truth knowledge, we have correctly estimated all 50 seams with no false alarm.

(a) Original image and its 50 vertical seams

(b) Image after seam carving and its stable SIFT points

Figure 4: Illustration of seam carving estimation (This figure is better viewed in color)

It should be noted that such capability of estimating the amount and location of removed seams does not require modifying the hash construction in [6]. This shows that a good design of forensic hash can be used to answer a broad scope of forensic questions. The estimation here provides the regions in the received image where seam carving has occurred. This information is very helpful to reconstruct the original image and enable further forensic analysis such as tampering localization, as will be demonstrated in the next section.

4. RECONSTRUCTING ORIGINAL IMAGE AND DETECTING TAMPERING

Knowing where and how many seams have been removed is an important _rst step to evaluate image trustworthiness. Places with seam removal are less trustworthy than regions without seam carving. In this section, we explore how to further use such information to adaptively resize a received image to align with the original image and thus enable tampering localization through block-wise feature comparison.

There are several ways to resize the seam carved image. Without any knowledge about the seam carving amount and location, a na□ve resizing option is to resize the image or insert seams in the image. Such strategies cannot align the two images accurately and may make block-based comparison unreliable. In contrast, knowledge of the seam carving estimation can guide the reconstruction process by constraining the resizing or seam insertion to only those regions that have undergone seam carving and with only the necessary amount. More speci_cally, with estimation result f_Ci; ~xi; ~xi+1g, i.e., _Ci seams have been removed in the horizontal range [~xi; ~xi+1] of image ~I, we increase the width of~I by _Ci through either rescaling the vertical strip [~xi; ~xi+1] or inserting _Ci seams into the same range. Since the shape of a seam is irregular and its pixels may not all

fall into a vertical strip, we insert seams that have at least half of its pixels in the speci_ed range. Similarly for resized image due to seam insertion, we can reconstruct the original image by removing seams from the resized image.

To illustrate the accuracy of such reconstruction and alignment, we show an example below. For the same image in Fig. 4(b), we apply simple seam insertion without any constraints and seam insertion with constraints based on the estimation results to resize the image to its original size. The di_erence between the resized results and the original image

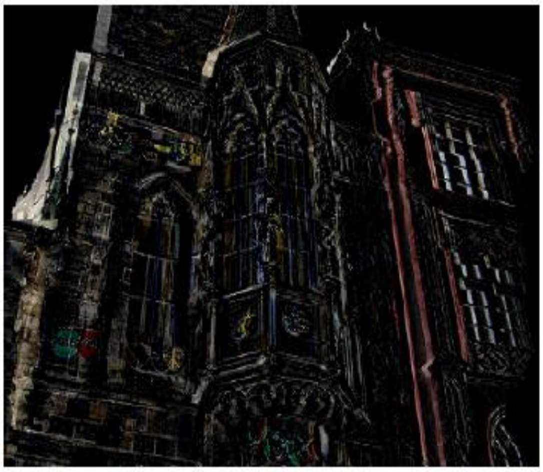
are shown in Fig. 5(a) and Fig. 5(b), respectively, where the di_erence is shown as a color image and black color indicates zero error. We can see that the constrained seam insertion can provide accurate alignment between the resized image and original image, while simple insertion without any constraints

mis-aligned the two images, causing large errors in many places of the image. The PSNR is 14.8 dB for simple insertion and 22.8 dB for constrained insertion.

Since the original image is not available during forensic analysis, compact block-based features can be encoded into the forensic hash as an integrity check component for tampering localization. The block feature used here is edge pixel direction histogram, quantized to four directions and has

size of 1 byte per block, as described in [6]. Using block size of 32 by 32, the average block feature distance is 65.05 for simple insertion and 15.02 for constrained insertion. The accurate alignment achieved by adaptive reconstruction that utilizes seam carving estimation result is very important, as it enables tampering localization using block-based features, which are not robust to misalignment. As can be seen from Fig. 5(a), misalignment may cause the block-based comparison to consider untampered regions as tampered.

With the knowledge of seam carving amount and locations, we resize each given range using simple scaling or seam insertion. For example, if seam carving estimation reveals that a vertical strip with horizontal range [10,29] has been carved 20 seams, we can resize this strip to a new width of 40 through scaling or inserting 20 new seams that pass through this vertical strip.



(a) Naïve reconstruction error

(b) Constrained reconstruction error

Figure 5: Di_erence between original and reconstructed image using seam insertion without and with constraints

For large resizing amount, constrained scaling and constrained seam insertion both produce a blurred result and seam insertion may introduce additional distortion along edges. For small resizing amount, seam insertion is a better option than scaling in the the sense that it can avoid blurness. Another case that seam insertion produces better reconstruction than scaling is when the removed seams have irregular shapes or going diagonal directions.

(a) Original image (b) Reconstructed image using constrained scaling (c) Reconstructed image using constrained seam insertion

In this case, scaling a vertical or horizontal strip will distort the image content. An example is given in Fig. 6. The original image and its removed seams are shown in Fig. 6(a). The reconstructed images using constrained scaling and constrained seam insertion are shown in Fig. 6(b) and Fig. 6(c), respectively. We can see that the removed seams in the lower center region of the original image moves in a diagonal direction, therefore, scaling will produce a blurred strip in the center of the resized image (Fig. 6(b)) while the seam insertion can avoid such distortion in the image content (Fig. 6(c)).

Reconstructing the exact original image from a seam carved image is a challenging and open problem, but for the purpose of aligning the resized image with the original image for tampering detection, both the constrained scaling and constrained seam insertion work well and we will show more experiments in the next section.

5. EXPERIMENT RESULTS

in this section, we perform several experiments to evaluate the performance of seam carving estimation and tampering detection using forensic hash. The image dataset used in the paper includes 200 images from Flicker with 40 different tags, such as beach, building etc. The image size is about 500x300. For each of the 200 images, we perform seam carving along its larger dimension and generate four resized

images whose modified dimension has 60%, 70%, 80%, and 90% of the original size, respectively. The total number of resized images is 800.

Robust geometric transform estimation To estimate seam carving, the modified image should be first aligned with the original image to the same orientation and scale. Such alignment can be achieved through geometric transform estimation using forensic hash. Here, we validate the robustness of such alignment against seam carving operation.

For an image undergone seam carving operation, a robust geometric transform estimation should report a scaling factor close to 1. The estimation results on the 800 resized images are shown in Table 1, which shows the absolute rotation angle estimation error and relative scaling.

Table 1. Robustness of forensic hash against seam carving

Resize factor	90%	80%	70%	60%
Rotation error	0.06	0.26	0.97	2.82
Scaling error	0.18%	0.76%	1.88%	3.37%

Seam carving estimation After geometric alignment, we can estimate the amount and position of removed seams as described in Section 3. By comparing with the ground truth seam carving amount, we evaluate the estimation accuracy using probability of correct detection (Pd) and probability of false detection (Pf), which are de_ned as follows:

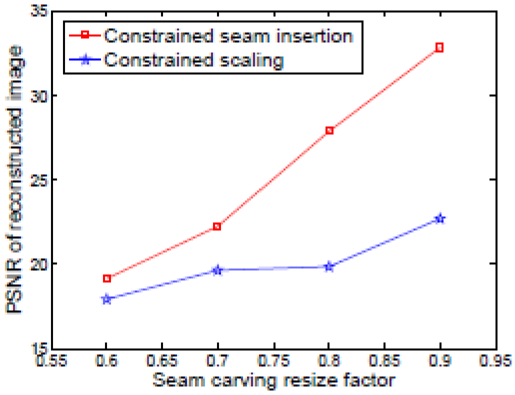
$$P_d = \sum_i \frac{\min(\Delta\hat{C}_i, \Delta C_i)}{\Delta C}, \ P_f = \sum_i \frac{\max(\Delta\hat{C}_i - \Delta C_i, 0)}{\Delta C}.$$

ΔC_i is the estimated seam carving amount in the range given by the ith matched SIFT pairs, ΔC_i is the actual carving amount in the same range, and ΔC_i is the ground truth value of total number of seams that have been removed. We perform estimation on the 800 images with deferent resize factors. The probability of correct detection and false detection at deferent hash lengths are shown in Table 2.

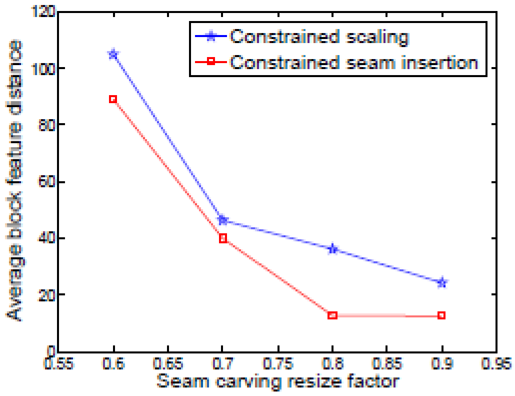
With hash length at around 50 bytes, the average probability of correct detection is 99.4% and average probability of false detection is 2%. We can also see that longer hash length may not improve the estimation accuracy. Actually, when more SIFT points are used, there is higher chance of mismatch of SIFT points and we see slight decrease in estimation accuracy.

Table 2. Seam carving estimation performance

Hash length (bytes)	47	94	156	219
Prob. of correct detection	99.4%	98.6%	98.6%	98.3%
Prob. of false detection	1.99%	3.52%	4.26%	3.48%



(a) PSNR



(b) Average block feature distance

Reconstruction and alignment We perform reconstruction using both the constrained scaling and constrained seam insertion guided by the seam carving estimation results over the 800 images. The PSNR and block feature distance between the original image and the reconstructed image using a forensic hash of 47 bytes are shown in Fig. 7. Fig. 7(a) shows the PSNR of the reconstructed image at di_erent seam carving resize factors. We can see that constrained seam insertion consistently outperforms the constrained scaling, and the reconstruction quality is better if the seam carving amount is smaller, i.e., the resize factor is large. Fig. 7(b), compares the average block feature distances of the two methods, and again we can see constrained seam insertion has better performance.

Tampering localization By adaptively resizing the seam carved image, we can accurately align it with the original image for tampering localization. We illustrate one example below. The original image is shown in Fig. 8(a) and its tampered version is shown in Fig. 8(b). The tampered image has undergone seam carving to remove the central building and then cut-and-paste to insert a plane. Our seam carving estimation correctly identi_es that there are 125 missing seams in the center of the tampered image. After adaptive resizing using constrained scaling, the result of block-wise comparison of edge-direction histogram is shown in Fig. 8(c). Both the center region where the building has been removed and the inserted plane are correctly identi_ed as tampered, as covered by red blocks. Therefore, using forensic hash and additional block-based features, we can provide a detailed report on the trustworthiness of an image, in terms of which part can be trusted and which part might be tampered.

(a) Original image

(b) Tampered image

(c) Tampering detection result

Figure 7: Reconstruction and alignment performance

6. CONCLUSIONS

In this paper, we study the problem of detecting seam carving using compact side information called forensic hash.The adaptive nature of seam carving allows e_ective image tampering against traditional blind forensic techniques. However, we demonstrate that a very compact forensic hash around 50 bytes can reliably estimate both the amount and the location of seam carving, and further enable accurate alignment and tampering localization on a modi_ed image.

Such detailed information of trustworthiness provided by the forensic hash is important for better utilization of online multimedia information. More importantly, the proposed analysis for seam carving estimation can be extended to detect and estimate di_erent image retargeting operations [8] that involve multiple operators such as scaling, cropping, warping, and seam carving. We will consider such extension in the future work.

7. REFERENCES

[1] S. Avidan and A. Shamir. Seam carving forcontent-aware image resizing. ACM Trans. onGraphics, (Proceedings SIGGRAPH 2007), 26(3), 2007.

[2] E. Delp, N. Memon, and M. W. (eds). Special issue on forensics analysis of digital evidence. IEEE Signal Processing Magazine, 26(2), March 2009.

[3] C. Fillion and G. Sharma. Detecting content adaptive scaling of images for forensic applications. In Proc.SPIE: Media Forensics and Security, volume 7541, pages 7541{36, 2010.

[4] D. Lowe. Distinctive image features fromscale-invariant keypoints. International Journal of Computer Vision, 60(2):91{110, 2004.

[5] W. Lu, A. Varna, and M. Wu. Forensic hash for multimedia information. In SPIE Media Forensics and Security, 2010.

[6] W. Lu and M. Wu. Multimedia forensic hash based on visual words. In Proc. of IEEE ICIP, 2010. M. Rubinstein, D. Gutierrez, O. Sorkine, and

[7] A. Shamir. A comparative study of image retargeting. ACM Trans. on Graphics, Proceedings Siggraph Asia, 29(5), 2010.

[8] M. Rubinstein, D. Gutierrez, O. Sorkine, and A. Shamir. A comparative study of image retargeting. ACM Transactions on Graphics (Proc. SIGGRAPH Asia), 29(5), 2010.

[9] A. Sarkar, L. Nataraj, and B. S. Manjunath. Detection of seam carving and localization of seam insertions in digital images. In Proceedings of the 11th ACM workshop on Multimedia and security, pages 107{116, 2009.

[10] H. Sencar and N. Memon. Overview of state-of-the-art in digital image forensics. World Scienti_c Press, 2008.

[11] A. Swaminathan, Y. Mao, and M. Wu. Robust and secure image hashing. Information Forensics and Security, IEEE Transactions on, 1(2):215{230, 2006