TAMPER DETECTION BASED ON REGULARITY OF WAVELET TRANSFORM COEFFICIENTS

Yagiz Sutcu, Baris Coskun

Polytechnic University
Electrical & Computer Engineering Dept.
Brooklyn, NY, 11201, USA

Husrev T. Sencar, Nasir Memon

Polytechnic University
Computer & Information Science Dept.
Brooklyn, NY, 11201, USA

ABSTRACT

Very very Powerful digital media editing tools make producing good quality forgeries very easy for almost anyone. Therefore, proving the authenticity and integrity of digital media becomes increasingly important. In this work, we propose a simple method to detect image tampering operations that involve sharpness/bluriness adjustment. Our approach is based on the assumption that if a digital image undergoes a copy-paste type of forgery, average sharpness/blurriness value of the forged region is expected to be different as compared to the non-tampered parts of the image. The method of estimating sharpness/bluriness value of an image is based on the regularity properties of wavelet transform coefficients which involves measuring the decay of wavelet transform coefficients across scales. Our preliminary results show that the estimated sharpness/bluriness scores can be used to identify tampered areas of the image.

Index Terms— Digital Forensics, Regularity, Forgery Detection, Image Authentication

1. INTRODUCTION

Despite the availability of extremely powerful technologies in both generating and processing digital images, there is a severe lack of techniques and methodologies for validating the authenticity of digital images. Due to this asymmetry, digital images appear to be the source of a new set of legal disputes and problems rather than being a solution. Furthermore, combined with the ease with which image processing tools can be obtained and used to modify images in *indistinguishable* ways, verifying the integrity of digital images proves to be a challenging task. This in turn undermines the credibility of digital images presented as news items, as evidence in a court of law, as part of a medical record or as financial documents since it may no longer be possible to distinguish whether an introduced image can be considered as the original, or a (maliciously) modified version.

Recognizing the complexity of the problem, various digital watermarking techniques have been proposed as a means for authenticating images that are most likely to undergo various types of processing. In this approach to problem, a fragile watermark is embedded into the original image to create a marked image which is later extracted to determine if marked images has been tampered and to give the localization information as to which part of the image has been tampered, e.g., [1][2] [3]. While this approach enables detector to establish the degree of authenticity and integrity of a digital object, it practically requires that the watermark was embedded during the creation of the digital object. This limits watermarking to applications where the digital object generation mechanisms have built-in watermarking capabilities, and therefore it cannot be offered as a general solution to the problem of authentication. Consequently, alternative approaches, that do not require much prior knowledge or processing of the original image, needed to be considered.

Another approach to verify integrity of digital images is inspired from the use of cryptographic hash functions for data authentication. The crux of this class of techniques is in the design of a, so called, robust perceptual hash function. Since a digital media content might have many different digital representations, robust hash functions are designed to produce the same hash value as long as the input has not been perceptually modified. Mihcak and Venkatesan [4] proposed such function based on iterative geometric filtering. Another method is proposed by Fridrich [5] wherein a robust hash is generated by first dividing an image into blocks, projecting each block onto pseudo-randomly generated smooth basis functions and then appropriately quantizing the resulting values. In [6], Venkatesan et al. proposed another robust image hashing scheme based on random quantization of the statistics of wavelet coefficients. However, Coskun and Memon [7] showed that, these robust hash functions do not have satisfactory diffusion capabilities meaning that the hash value remains similar as the perceptual information is slowly changed.

Another promising class of techniques that aim at detecting image tampering is based on the assumption that although image tampering might cause no visual artifacts or anomalies, it will nevertheless affect the underlying statistics of the image. Furthermore, one may safely assume that the process

of image manipulation will very often involve a sequence of processing steps to avoid the appearance of illicit human intervention. Typically, a tampered image (or parts of it) would have undergone some common image processing operations, like scaling, rotation, brightness adjustment, compression, etc., to produce visually consistent images. To detect such anomalies, Bayram et al. [8] compiled more than 100 features that are sensitive to various common image processing operations and constructed classifiers to detect images that have undergone such processing. Similarly, Ng and Chang examined bicoherence characteristics of images to detect photomontages. Fridrich et al. in [9], based on correlation procedures, proposed method for detecting forgeries created by copying and pasting parts of an image over other parts. Based on the observation that image resizing operation introduced pixel-wise correlations in an image Popescu et al. [10] proposed a procedure to detect image resizing. Later, Johnson et al. [11] proposed a method based on inspecting inconsistencies in lighting conditions and Assuming the camera (or a number of images taken by the camera) is available, Lukas et al. in [12] proposed a technique to detect and localize tampering by analyzing the inconsistencies in the sensor pattern noise extracted from an image. Along the same direction, Swaminathan et al. [13] used inconsistencies in color filter array interpolation to detect tampered parts of an image.

The above results show that none of the above techniques can offer a definitive solution by themselves. Ultimately, a solution to the complex problem of image forensics require incorporation of all these methods together with many new techniques. With this perspective, in this paper, we aim at contributing to existing arsenal of image tamper detection techniques by exploiting the fact that blurriness/sharpness adjustment is a common form of processing performed during tampering. Since in the forged image parts the sharpness (blurriness) characteristics are expected to be different than in the non-tampered parts, by measuring this difference tampered regions of an image can be localized. For this purpose, we deploy a method based on regularity properties of wavelet transform coefficients. Such regularity based techniques have been previously used in various applications such as image interpolation and image quality estimation [14, 15]. Motivated by these results, in this work we extend this approach to the context of digital image forensics. We show that regularity of the wavelet transform coefficients can be used to estimate the overall sharpness/blurriness of the edges, which can in turn be used to detect image tampering. In this work, we demonstrate the potential of the method and in the final version, we will provide performance results obtained by applying the method to tampered image dataset used in [8].

The rest of the paper is organized as follows. In Section 2, we give a brief summary of regularity analysis in wavelet domain and explain the details of the proposed method. Experimental results are presented in Section 3 and conclusions with future efforts are discussed in the last section, namely, Section 4.

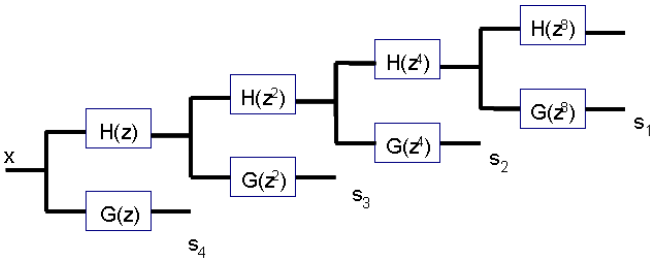

Fig. 1. Filterbank implementation of UDWT

2. ESTIMATING SHARPNESS/BLURRINESS

2.1. Regularity in Wavelet Domain

Visual smoothness of a function can be mathematically expressed by the Lipschitz exponent (also called the Hölder exponent) which essentially indicates the number of continuous derivatives that a function possesses. The Lipschitz exponent can be defined in the frequency domain as the largest α such that

$$\int_{-\infty}^{\infty} |F(w)|(1+|w|^{\alpha})dw < \infty \qquad (1)$$

holds. Here $F(w)$ is the Fourier transform of $f(t)$. The main problem with this definition is the fact that, it requires the knowledge of closed-form expression for the function f and this fact eliminate the applicability of this definition for measuring the regularity of digital media. Similar to Fourier methods, the Lipschitz regularity of a function can be determined by analyzing wavelet transform coefficients. Unlike Fourier transform, compactly supported wavelet basis make it possible to locate the irregularity of the function and provide information about local regularity.

Wavelet transform is a very powerful signal processing tool due to its energy compaction capability and realizability by appropriate filterbank structures. The undecimated discrete wavelet transform (UDWT) can be considered a discretized version of continuous-time wavelet transform and the coefficients of UDWT of a signal can be computed by projecting that signal onto a set of basis functions which are the translated and dilated versions of a mother wavelet:

$$\psi_{k,l}(x) = \psi(2^k x - l) \qquad (2)$$

where k is the scale and l is the offset values which are integers. Filterbank implementation of UDWT is illustrated in Figure 1.

Relation between the regularity of a signal and wavelet transform coefficients can be summarized as follows: Let $\mathbf{S}$ be the set of index pairs (k,l) such that for some $\epsilon > 0$, an interval $(x_0-\epsilon, x_0+\epsilon) \subset support(\psi_{k,l})$. A signal has local

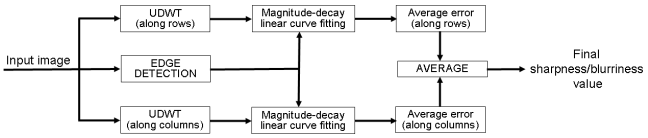

Fig. 2. Illustration of the proposed sharpness/blurriness estimation method

Lipschitz exponent α in the neighborhood $(x_0 - \epsilon, x_0 + \epsilon)$ if there exists a finite constant C such that the wavelet transform coefficients $w_{k,l} = \langle f, \psi_{k,l} \rangle$ satisfy

$$max_{(k,l)\in \mathbf{S}} |w_{k,l}| \leq C 2^{-k(\alpha + \frac{1}{2})} \tag{3}$$

for all scales k and all translations l [16]. In (3) strong edges achieve near equality. However, since this will not be true for blurred edges, the degree of inequality in (3) can be used to quantify the edge sharpness. When this method of sharpness/blurrines measurement is applied to different regions of an image, it provides a means to check whether the observed differences in the estimated values are within an acceptable range.

2.2. Proposed Method

Estimating regularity of digital signals by measuring the decay of wavelet transform coefficients across scales is not a new topic [17][14]. The proposed method is also based on this phenomenon and can be summarized as follows:

- *Edge detection:* Employ an edge detection algorithm to determine edge locations of the given image.
- *Wavelet transform:* Take L-level UDWT of each and every row and column separately by using the above described filterbank implementation.
- *Linear curve fitting:* Edge locations are located by analyzing the edge image and corresponding maximum amplitude values of wavelet subband signals are determined. Then, a linear curve is fitted to the log of these maximum amplitude values.
- *Final sharpness/blurriness value:* The goodness of the linear curve fitting is the row-based (and column-based) sharpness/blurriness measure of the given image. The final sharpness/blurriness value of the given image is determined as the mean value of these two (row and column) values.

These steps of the proposed scheme are illustrated in Figure 2.

However, in the context of forgery detection, the proposed sharpness/blurriness measure will be calculated by dividing an image into non-overlapping regions. This provides the ability to observe the variation of proposed sharpness/blurriness measure over different parts of the image. Outliers and/or marginal deviations from that distribution will help to identify the forgery regions.

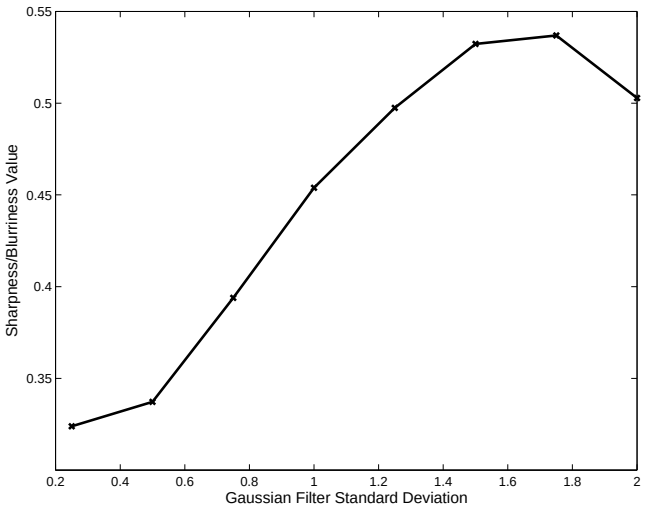

Fig. 3. Performance of the proposed blurriness estimation method

3. EXPERIMENTAL RESULTS

In this setup, we implemented 4-level UDWT decomposition in MATLAB using appropriate filterbank structure and the mean absolute error calculated after linear curve fitting step is taken as the goodness of fit, in other words, as a measure of blurriness. It is clear that the lower the mean absolute error, the more sharp the image will be.

Firstly, we tested the accuracy of the proposed blurriness determination method using 256x256 grayscale image (given in Figure 5(a)) and its blurred versions. Different blur amounts are obtained by filtering the image with 7x7 Gaussian filter with different standard deviations which are the multiples of 0.25 changing from 0 to 2. Results of our simulations are given in Figure 3.

As can be seen from the Figure 3, proposed blurriness measure captures the increasing value of blur introduced by Gaussian filtering reasonably well. However, when the value of the standard deviation of the Gaussian filter exceeds the value of 1.75, blurriness value calculated by the proposed method becomes unreliable. This is mainly because the edge detection process starts to get inaccurate for overly smoothed images.

Secondly, in order to test the validity of our regularity-based forgery detection scheme, we considered a digitally tampered image and applied the proposed method to many different parts of that image. The original and the tampered versions of that image are given in Figure 5. We considered 30 different regions cropped from the tampered version of the

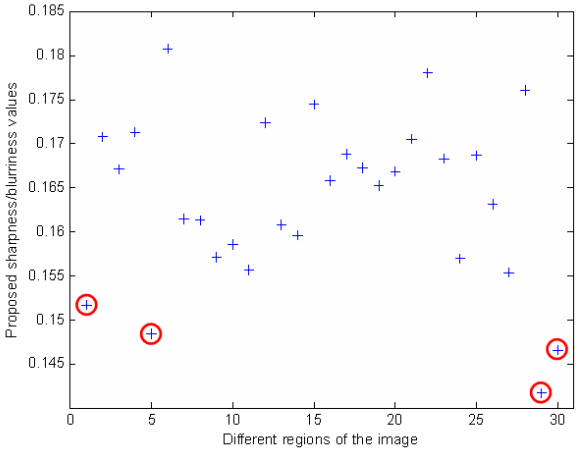

Fig. 4. Blurriness values calculated over different parts of the tampered image

(a) Original (b) Tampered

Fig. 5. Original and tampered image. Tampered regions are marked with white circles.

given image and 4 of them are the regions of forgeries. Figure 4 shows the results of our simulations where circled ones correspond to the tampered regions. As can be observed, blurriness values calculated from tampered regions of the image are the four lowest values. Since they differ from the measured blurriness values of the rest of the image regions, one can say that these regions has been replaced with different image .

4. CONCLUSION

In this paper, we propose a sharpness/blurriness measure based on regularity of wavelet coefficients. Our results show that when the devised method is applied to different regions of a purposefully blurred image, it can successfully estimate the degree of blurring (sharpening) the image has undergone. One promising application area for such a metric is in digital image forensics. Images that have undergone tampering are very likely to exhibit variations in blurriness (sharpness) characteristics. Therefore, our future work basically consists of testing the performance of the devised metric in a forensics setting, where we will include false-alarm and detection probability results concerning application of the method to tampered image dataset used in [8].

5. REFERENCES

[1] Fridrich J., "Image watermarking for tamper detection," *Proc. ICIP, International Conference on Image Processing*, vol. 2, pp. 404–408, 1998.

[2] Huang J., Hu J., Huang D., and Shi Y.Q., "Improve security of fragile watermarking via parameterized wavelet," *Proc. ICIP, International Conference on Image Processing*, vol. 2, pp. 721–724, 2004.

[3] Watanabe J., Hasegawa M., and Kato S., "A study on a watermarking method for both copyright protection and tamper detection," *Proc. ICIP, International Conference on Image Processing*, vol. 4, pp. 2155–2158, 2004.

[4] Mihcak M.K. and Venkatesan R., "New iterative geometric methods for robust perceptual image hashing," *Proc. of the Digital Rights Management Workshop*, November 2001.

[5] Fridrich J., "Robust bit extraction from images," *ICMCS 99, Florence, Italy*, June 1999.

[6] Venkatesan R., Koon S., Jakubowski M., and Moulin P., "Robust image hashing," *Proc. IEEE Int. Conf. on Image Processing*, 2000.

[7] Coskun B. and Memon N., "Confusion/diffusion capabilities of some robust hash functions," *Proc. CISS, Conf. on Information Sciences and Systems*, March 2006.

[8] B. Sankur S. Bayram, I. Avcibas and N. Memon, "Image manipulation detection," *Journal of Electronic Imaging – October - December 2006 – Volume 15, Issue 4, 041102 (17 pages)*, vol. 15(4), 2006.

[9] J. Fridrich, D. Soukal, and J. Luk, "Detection of copy-move forgery in digital images," *Proc. Digital Forensic Research Workshop, Cleveland, OH*, August 2003.

[10] A.C. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of resampling," *IEEE Transactions on Signal Processing*, vol. 53(2), pp. 758–767, 2005.

[11] M.K. Johnson and H. Farid, "Exposing digital forgeries by detecting inconsistencies in lighting," *Proc. ACM Multimedia and Security Workshop, New York*, pp. 1–9, 2005.

[12] Luk J., Fridrich J., and Goljan M., "Detecting digital image forgeries using sensor pattern noise," *Proc. of SPIE Electronic Imaging, Photonics West*, January 2006.

[13] M. Wu A. Swaminathan and K. J. Ray Liu, "Image tampering identification using blind deconvolution," *Proc. IEEE ICIP*, 2006.

[14] W.K. Carey, D.B. Chuang, and S.S. Hemami, "Regularity-preserving image interpolation," *IEEE Transactions on Image Processing*, vol. 8, pp. 1293–1297, 1999.

[15] R. Ferzli and L.J. Karam, "No-reference objective wavelet based noise immune image sharpness metric," *Proc. IEEE International Conference on Image Processing ICIP 2005*, vol. 1, pp. 405–408, September 2005.

[16] I. Daubechies, "Ten lectures on wavelets," *SIAM*, 1992.

[17] Rooms F., Pizurica A., and Philips W., "Estimating image blur in the wavelet domain," *ICASSP 2002*, 2002.