Hindawi Publishing Corporation
Research Letters in Signal Processing
Volume 2007, Article ID 47549, 5 pages
doi:10.1155/2007/47549

Research Letter

Image Authentication Using Added Signal-Dependent Noise

Xin Cindy Guo and Dimitrios Hatzinakos

Department of Electrical and Computer Engineering, University of Toronto, Toronto, Canada M5S 1A8

Correspondence should be addressed to Xin Cindy Guo, cguo@comm.utoronto.ca

Received 16 September 2007; Accepted 24 October 2007

Recommended by Mark Liao

Image authentication has applications in security systems, photo forensics, and photo journalism. This paper presents an image authentication scheme using added signal-dependent noise. Imperceptible noise is embedded into the image at the time of acquisition according to the film grain noise model. During authentication, the image is divided into key-dependent overlapping blocks and the parameters of the embedded noise are extracted. The variance of the extracted parameters can be used to show the authenticity of an image. Test results indicate that the proposed algorithm is robust against content-preserving modifications such as JPEG compression and at the same time is capable of detecting malicious tampering.

1. INTRODUCTION

Due to the increasing popularity and accessibility of image manipulation softwares, digital manipulations of multimedia content have become difficult to detect. Determining the authenticity of digital images has widespread applications in security systems, forensics, and photo journalism.

Digital watermarking and digital signature have been proposed as image authentication methods in recent years. Most early digital watermarking schemes such as those based on spread-spectrum [1], discrete cosine transform (DCT) [2], and wavelet transform [3] can be defeated if small geometric distortions were introduced. Since the images might have gone through unavoidable modifications such as JPEG compression, additive noise and small geometric distortions that do not change the content of the images, recent algorithms have been focusing on the robustness of watermarking and signature generation schemes.

Bas et al. [4], Lu et al. [5], and Monga and Evans [6] developed feature-points-based watermarking embedding schemes. Instead of blindly inserting a predetermined mark into the images, they chose specific areas marked by the feature points that could not be destroyed in certain content-preserving modifications. This method, however, did not work well for images lacking edges or discontinuities. Recent digital signature schemes [7–10] based on extracting unique features from the images were robust against compression, noise, and distortions. However, they did not have good localization properties, and hence venerable against local distortions such as removal or insertion of foreign objects.

The main drawback of the current schemes is that they require complex computation. Therefore, watermarks or signatures cannot be attached to the images at the time when they are taken because most cameras are not equipped with such hardware. Detecting digital forgeries without digital watermark or signature had been studied by Popescu and Farid [11, 12]. Their idea was based on the fact that forgeries usually require interpolation and resampling of the insertion object. Their results, however, maybe not be accurate when compression and filtering are introduced. Lukas et al. proposed an image authentication method based on the camera sensor pattern noise [13]. The system was first trained by observing a number of images taken by one camera, and then the pattern noise of the image in question was extracted and compared to the observed noise pattern.

We propose an image tampering detection technique by introducing imperceptible signal-depended noise at the time of the image acquisition. During authentication, noise statistics is analyzed and any inconsistencies will point to possible tampering within the image. The main advantage of the algorithm is that it is easy to implement and it is robust against certain content-preserving modifications such as compression and additive noise.

Section 2 introduces the proposed algorithm. Testing results are presented in Section 3, and concluding remarks and future directions are discussed in Section 4.

2. METHOD

2.1. Signal-dependent noise

Signal-dependent noise is added to the original image $I(x, y)$ via the following formula:

$$R(x, y) = I(x, y) + kI^p(x, y)n(x, y), \quad (1)$$

where k and p are constants, and $n(x, y)$ is zero mean, unit variance Gaussian noise. This model is used for film grain noise generation. The parameter p, which depends on the film, is commonly set to 0.5 [14].

It can be shown that k can be estimated using the following equations [14]:

$$\sigma_r^2 = \sigma_i^2 + k^2 E[I], \quad (2)$$

$$c_3^r = c_3^i + 3k^2\sigma_i^2, \quad (3)$$

$$c_4^r = c_4^i + 6k^2 c_3^r - 15k^4\sigma_i^2, \quad (4)$$

where σ_r^2, c_3^r, and c_4^r are the variance, skewness, and kurtosis of the noise-embedded image $R(x, y)$ and σ_i^2, c_3^i and, c_4^i are of those original image $I(x, y)$. To calculate k, the statistics of the original image is assumed known a priori.

2.2. Proposed algorithm

2.2.1. Noise embedding

In order to increase the security of the algorithm, the image is divided into random, overlapping tiles according to a private key. The statistics, that is, the variance, skewness, and kurtosis, of each subimage is calculated and stored either in the header or in a separate, encrypted file which would be transmitted to the receiver.

Random tiling is used by Monga and Evans [6] to generate localized feature points and also by Venkatesan et al. [9] to calculate the statistics of wavelet coefficients. The proposed method uses random division similar to that of Venkatesan's algorithm; however, the statistics alone cannot authenticate an image.

After the computation of the image statistics, noise is added to the original image via (1).

2.2.2. Image authentication

The received image is partitioned using the same private key, and the variance, skewness, and kurtosis of each block are calculated. k is then estimated using (2)–(4). It should be noted that one equation is sufficient in calculating k. The effects of using second-, third-, and fourth-order statistics will be discussed in later sections.

Since the image might have gone through content-preserving modifications such as compression and filtering, the estimated k value would not necessarily be the same as the one used in embedding. However, the k's that are estimated from all the blocks should be consistent. This is because most content-preserving changes usually affect all pixels in an image rather than certain pixels locally.

(a) (b)

Figure 1: (a) Original Lena, (b) modified Lena.

Table 1: Estimation of k using Lena, Baboon, and Peppers.

	Second order	Third order	Fourth order
Lena	0.097 ± 0.033	0.110 ± 0.067	0.105 ± 0.047
Baboon	0.086 ± 0.031	0.123 ± 0.042	0.106 ± 0.045
Peppers	0.100 ± 0.037	0.097 ± 0.042	0.103 ± 0.034

Figure 2 shows the estimated k values from 16 random tiles for the original Lena and modified Lena in Figure 1. In Figure 2(a), the estimated values are somewhat evenly spread around the true value 0.1. In Figure 2(b), one of the estimated values is very different from the rest, indicating attacks occurred in this region.

3. RESULTS

3.1. Experiment setup

The algorithm is tested on over 20 images, including commonly used Lena, baboon, and peppers. Each image is scaled to 512×512 in size and divided into 16 overlapping regions. For colour images, the noise is embedded in the Y channel. In all test cases, k is set to 0.1. The peak signal-to-noise ratio (PSNR) of the noisy image to that of original is 47 dB on average. The added noise is imperceptible to human eyes.

Table 1 shows the mean and variance of the estimated k from 16 blocks. As an example, results for only 3 of the test images are presented. The mean value is very close to the true value for k which is 0.1. The variances of the estimations are very small.

The average embedding time for a 512×512 image is about 0.5 seconds on a Pentium(R) 4 machine. The time for computing and comparing k from 16 blocks is 0.3 seconds on average.

3.2. Performance in content preserving modifications

The algorithm is tested on none invasive modifications such as JPEG compression, additive Gaussian noise, and local denoising. JPEG compressed images are created by Stirmark

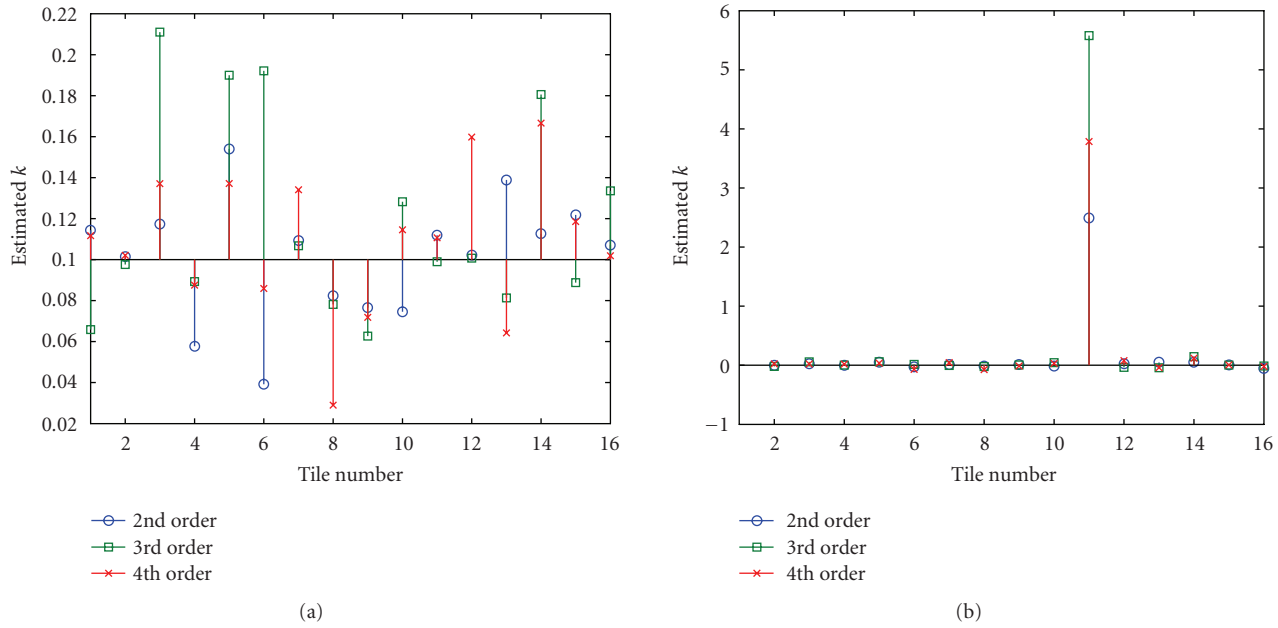

Figure 2: Comparison of estimated k from 16 overlapping tiles. (a) Estimated k from original Lena, (b) estimated k from modified Lena.

Figure 3: Examples of object insertion; (a) five icons, (b) eight icons.

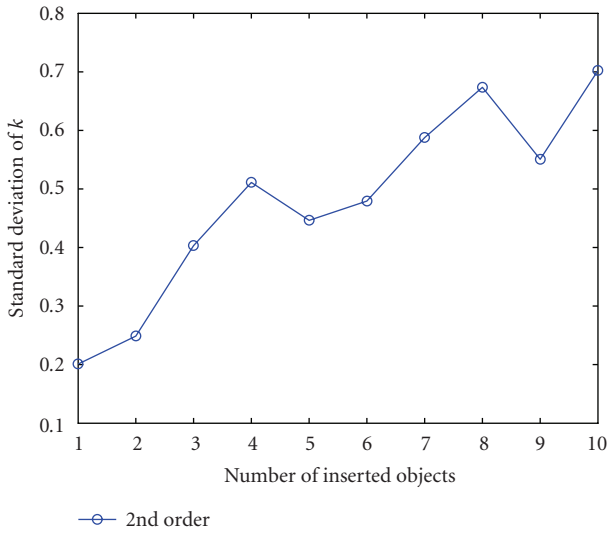

Figure 4: Variance of k for object insertion.

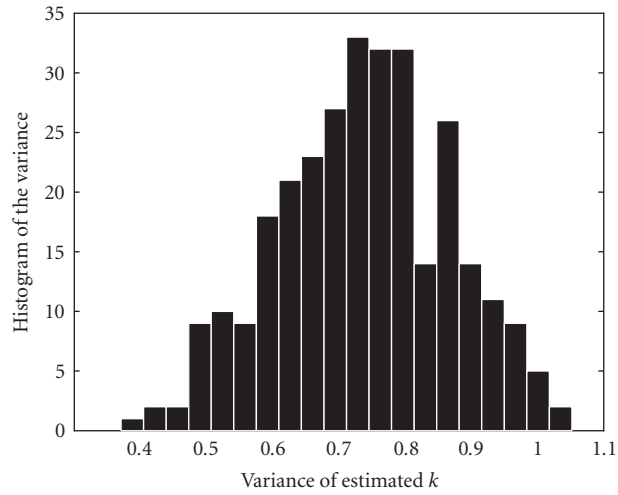

Figure 5: Histogram of the variance of estimated k using different keys.

software [15]. The lowest compression quality factor used is 60, which corresponds to a PSNR of about 27 dB. Gaussian noise of variance between 2 and 10 is added to the images directly. The worst PSNR in this case is around 28 dB. In local denoising, the noise of a selected area is suppressed using Photoshop CS. Four areas of size 50×50, 100×100, 150×150, 200×200 pixels are used.

The average mean value and standard deviation of the estimated k from the images are shown in Tables 2, 3, and 4. As expected, the accuracy of the estimation decreases as the image quality gets worse. However, the differences among the blocks of the same image remain relatively small, as evident in the standard deviations.

Table 2: Estimation of k in JPEG compressed images.

Quality	Second order	Third order	Fourth order
60	0.222 ± 0.120	0.287 ± 0.143	0.262 ± 0.135
70	0.237 ± 0.101	0.254 ± 0.155	0.261 ± 0.143
80	0.213 ± 0.094	0.203 ± 0.094	0.224 ± 0.115
85	0.180 ± 0.095	0.175 ± 0.105	0.211 ± 0.089
90	0.177 ± 0.074	0.181 ± 0.075	0.197 ± 0.081
95	0.194 ± 0.037	0.139 ± 0.061	0.205 ± 0.044

Table 3: Estimation of k in images with additive Gaussian noise.

Variance	Second order	Third order	Fourth order
2	0.197 ± 0.050	0.152 ± 0.070	0.132 ± 0.127
4	0.356 ± 0.059	0.204 ± 0.105	0.184 ± 0.164
6	0.540 ± 0.052	0.223 ± 0.119	0.221 ± 0.205
8	0.719 ± 0.065	0.252 ± 0.119	0.226 ± 0.213
10	0.898 ± 0.090	0.323 ± 0.127	0.313 ± 0.268

Table 4: Estimation of k in images with local denoising.

Size	Second order	Third order	Fourth order
50×50	0.203 ± 0.055	0.373 ± 0.133	0.276 ± 0.185
100×100	0.222 ± 0.111	0.366 ± 0.152	0.319 ± 0.229
150×150	0.231 ± 0.115	0.397 ± 0.152	0.333 ± 0.276
200×200	0.266 ± 0.196	0.379 ± 0.215	0.394 ± 0.297

The results obtained using (2)–(4) are fairly similar in the JPEG compression case. However, when additive Gaussian noise is introduced, the third- and fourth-order statistics give better estimations in k. This is because higher-order statistics is better at suppressing white Gaussian noise. The tradeoff, on the other hand, is the high bias. Estimation using second-order statistics, that is, (2), yields significantly smaller variances. This is a confirmation of the observation in [14]. In our application, accurate estimation of k is not necessary. What is important is how evenly spread out the estimated values are. Therefore, the use of second-order statistics alone is sufficient.

3.3. Performance in attacks

3.3.1. Content-changing attacks

An important aspect of image authentication algorithm is its ability to detect malicious, content-changing modifications. To systematically test the proposed scheme, 1 to 10 icons of size 32×32 pixels are randomly inserted into the test images. Figure 3 shows an example of the images used in this test. In the worst case, about 4% of the pixels are modified. This experiment focuses on small changes because the larger the modified area, the more changes in the pixels, and therefore the easier it is to detect.

Figure 4 shows the variance of k in the object insertion case. Comparing to Tables 2, 3, and 4, the standard deviations of the estimated k are significantly higher.

A threshold on the variance can be set to distinguish authentic images from the maliciously modified ones. In this paper, we choose the threshold to be 0.2, which results in an error rate of 8% using second-order statistics.

The attacking region can be traced by examining the estimated k value from different blocks. However, releasing the regions used by the algorithm is dangerous to security attacks.

3.3.2. Security attacks

As mentioned in Section 2.2.1, the security of the algorithm comes from random tiling of the image. Without the key, attackers are not able to locate the positions that the algorithm use to estimate k. This property is shown in Figure 5.

Figure 5 is the histogram of the variance of the estimated k using 300 different keys. The variance is larger than the predefined accepted range of 0.2. It shows that an attacker with the knowledge of the embedding algorithm will not be able to generate authentic images without the embedding key.

4. CONCLUSION

This paper introduces an image authentication algorithm based on added signal-dependent noise. The proposed scheme is simple and can be applied to the image at the time of acquisition. During authentication, the parameters of the embedded noise are extracted from key-dependent areas, and the statistics of the extracted parameter are examined. The algorithm is tested on both content-preserving modifications and malicious attacks. From the magnitude of the variance of the noise parameter, one can determine the authenticity of an image. Test results show that the proposed algorithm is robust against content-preserving changes such as JPEG compression, additive Gaussian noise, and local denoising. Also, when foreign objects are inserted into the image, the algorithm can successfully detect the anomalies. Future work includes design of an appropriate noise filter such as the one used in [14] so that the original image statistics can be accurately estimated from the observed image and hence eliminating the need to store or transmit the original statistics.

ACKNOWLEDGMENT

The completion of this research was made possible thanks to Bell Canada's support through its Bell University Laboratories R&D program.

REFERENCES

[1] I. J. Cox, J. Kilian, F. T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Transactions on Image Processing*, vol. 6, no. 12, pp. 1673–1687, 1997.

[2] C.-Y. Lin and S.-F. Chang, "A robust image authentication method distinguishing JPEG compression from malicious manipulation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 11, no. 2, pp. 153–168, 2001.

[3] W. Dietl, P. Meerwald, and A. Uhl, "Protection of wavelet-based watermarking systems using filter parametrization," *Signal Processing*, vol. 83, no. 10, pp. 2095–2116, 2003.

[4] P. Bas, J.-M. Chassery, and B. Macq, "Geometrically invariant watermarking using feature points," *IEEE Transactions on Image Processing*, vol. 11, no. 9, pp. 1014–1028, 2002.

[5] C.-S. Lu, S.-W. Sun, C.-Y. Hsu, and P.-C. Chang, "Media hash-dependent image watermarking resilient against both geometric attacks and estimation attacks based on false positive-oriented detection," *IEEE Transactions on Multimedia*, vol. 8, no. 4, pp. 668–685, 2006.

[6] V. Monga and B. L. Evans, "Perceptual image hashing via feature points: performance evaluation and tradeoffs," *IEEE Transactions on Image Processing*, vol. 15, no. 11, pp. 3452–3465, 2006.

[7] A. Swaminathan, Y. Mao, and M. Wu, "Robust and secure image hashing," *IEEE Transactions on Information Forensics and Security*, vol. 1, no. 2, pp. 215–230, 2006.

[8] J. Fridrich, "Visual hash for oblivious watermarking," in *Proceedings of the 12th Annual Symposium, Electronic Imaging, Security and Watermarking of Multimedia Content II*, vol. 3971 of *Proceedings of SPIE*, pp. 286–294, San Jose, Calif, USA, January 2000.

[9] R. Venkatesan, S.-M. Koon, M. H. Jakubowski, and P. Moulin, "Robust image hashing," in *Proceedings of IEEE International Conference on Image Processing*, vol. 3, pp. 664–666, Vancouver, Canada, September 2000.

[10] J. S. Seo, J. Haitsma, T. Kalker, and C. D. Yoo, "A robust image fingerprinting system using the rado transform," *Signal Processing: Image Communication*, vol. 19, no. 4, pp. 325–339, 2004.

[11] A. C. Popescu and H. Farid, "Exposing digital forgeries by detecting traces of resampling," *IEEE Transactions on Signal Processing*, vol. 53, no. 2, pp. 758–767, 2005.

[12] A. C. Popescu and H. Farid, "Exposing digital forgeries in color filter array interpolated images," *IEEE Transactions on Signal Processing*, vol. 53, no. 10, pp. 3948–3959, 2005.

[13] J. Lukas, J. Fridrich, and M. Goljan, "Detecting digital image forgeries using sensor pattern noise," in *Security, Steganography, and Watermarking of Multimedia Contents VIII*, vol. 6072 of *Proceedings of SPIE*, San Jose, Calif, USA, January 2006.

[14] P. Campisi, J. C.K. Yan, and D. Hatzinakos, "Signal-dependent film grain noise generation using homomorphic adaptive filtering," *Proceedings of IEE Vision, Image and Signal Processing*, vol. 147, no. 3, pp. 283–287, June 2000.

[15] F. A. P. Petitcolas, R. J. Anderson, and M. G. Kuhn, "Attacks on copyright marking systems," in *Proceedings of the 2nd International Workshop on Information Hiding (IH '98)*, pp. 218–238, Portland, Ore, USA, April 1998.