A Secure Image Hashing Technique for Forgery Detection

Tanmoy Kanti Das[1] and Piyush Kanti Bhunre[2]

[1] Academy of Technoogy, India
dastanmoy@gmail.com
[2] National Institute of Science and Technology, India
kbpiyush@gmail.com

Abstract. Nowadays most of the multimedia contents are in digital form. With the increased use of powerful computer and image processing software, along with wide availability of digital cameras have given rise to huge numbers of doctored images. Several forgery detection algorithms are available. However, these techniques do not address the issue from cryptographic point of view. As a result, even if an image or video is identified as doctored, most of the time it is not possible to track the actual offender. Here, we present a perceptual hash function which can be used for both detection of forged images as well as tracking of forgers.

1 Introduction

Analog photos and video images have always been accepted as a "proof of occurrence" of the depicted event. For that very reason, courts have set high standard to ensure the integrity of those images. Advent of digital images raises additional concerns, because the images can so easily be manipulated and in many occasions, forged images are used to influence the naive people. Although digital watermarks have been proposed as a tool to provide authenticity of images, it is a fact that most of the images that are captured today do not contain any watermark. And we expect this situation will not change in immediate future. Hence, it is required to develop techniques those can detect the tampering of digital images. Some of the well known digital image tempering techniques can be found in [2]. In light of these problems, the subject of digital forensics has been developed to find the answers to the following questions [5].

- Is this an original image or manipulated image?
- What is the processing history of the image?
- What parts of the image has undergone processing and up to what extent?
- Was the image acquired by the device as claimed by the producer?
- Did this image originate from a source X as claimed?

These are just a few questions that are routinely faced by forensic experts and law enforcement agencies. Most of the existing research in this area is based on image processing techniques and lack a proper cryptographic framework. And it is well known that, once an image processing based forgery detection methodology is developed, the forgers will find new ways to circumvent it. Here, we propose a new perceptual hashing algorithm which use cryptographic framework for both authentication of digital images and tracking of the forgers.

R. Natarajan et al. (Eds.): ICDCIT 2015, LNCS 8956, pp. 335–338, 2015.

2 Image Hashing Technique

In general, cryptographic hash functions or message authentication functions are used to ensure data integrity. However, these functions are key dependent and sensitive to change in every bit of information. We know that minor changes in image information (i.e. pixel values) do not change the image visually. For example, one can generate some image I' form original image I by applying lossy image compression over I and both I' and I remain visually indistinguishable. In this scenario, we want the hash function to produce same hash values for I and I' as long as these images remain visually indistinguishable. Several image hash functions [1,6] were proposed in the existing literature. But they mostly depend on complex image processing techniques. Here we propose a new image hashing technique which is inspired by the ideas presented in [4]. Our technique is based on *wavelet transform* and uses basic statistical features like *mean, standard deviation, kurtosis, skewness* etc. to generate the hash value.

Before we proceed further, let us first discuss about wavelet transform using an example. Consider the following one dimensional signal $I = [11, 5, 7, 15]$ consisting of four samples. After applying Haar wavelet transform [7], the coefficients look like $I_{wav}^{L=1} = [\frac{11+5}{2}, \frac{7+15}{2}, \frac{11-5}{2}, \frac{7-15}{2}] = [8, 11, 3, -4]$. In the next level of wavelet transform, low frequency coefficients (here, first two coefficients) are subjected to further processing. Thus, $I_{wav}^{L=2} = [\frac{8+11}{2}, \frac{8-11}{2}, 3, -4] = [9.5, -1.5, 3, -4]$. As, in this example, there is only one low frequency component, we can not proceed further. To extend these ideas to images, we consider an image as a 2D signal and apply wavelet transform separately, first along the rows and then along the columns. In each level of wavelet transform four different bands are generated and they are denoted as LL_n, HL_n, LH_n, HH_n, where n is the level number. Let us now describe the hashing algorithm.

1. Pre-process the original image O to get I.
2. Compute wavelet transform of the image I upto n^{th} level. So there will be $n \times 3 + 1$ bands. Exclude band LL_n from further processing.
3. For each band, compute *mean, median ,mode, range, standard deviation, kurtosis, skewness* and represent the result in a matrix form. So, there will be 7 columns, each representing one statistical feature and there will be $n \times 3$ rows.
4. We convert all the values obtained in the last step to a 3 bit integer number and apply gray coding [3] to get a bit sequence.
5. The bit sequence is decoded using $(7, 3)$ Reed-Solomon code to get the hash value.
6. Encrypt the hash value H using the private key of the owner X using RSA algorithm or any other suitable public key algorithm to form the digital signature D_X. Now X can publish the image along with the digital signature D_X.

Here, we always choose LL band for next level of wavelet transform. Though one can choose any of the available bands i.e. LL, LH, HL, HH for next level of wavelet transform. In fact, the choice of band for the next level of wavelet transform can be made *key* dependent to introduce randomness. In this scenario, one cannot compute the hash value without the knowledge of the key. Thus the hash value becomes a keyed-hash message authentication code (**HMAC**). Now, after receiving an image I along with the digital signature D_X from X, one can can check whether the image is authentic or not, using the following steps:

1. Generate the hash value H' using the received image I
2. Decrypt the hash value D_X to get H using public key of X.
3. If *normalized hamming distance* between H and H' is less than threshold, then
   (a)Image is authentic as the hash values match.
   (b)X is the owner of the image as we can decrypt D_X using the public key of X.
4. Else the image is not authentic.

Security Analysis

Security of image hashing technique is not well defined and an active area of research. In this context, Swaminathan et al. proposed a security metric based on differential entropy of the hash value in their paper [6]. In simple terms, one can describe differential entropy is the amount of effort an adversary has to put to compute the correct image hash without the knowledge of the key. So, larger value of differential entropy is better for security. Our algorithm when used in HMAC mode performs very well in this regard.

Suppose we compute the k features $M_1^{(p)}, M_2^{(p)}, M_3^{(p)}, \cdots, M_k^{(p)}$ from a wavelet band at p^{th} level, where $p = 1, 2, \cdots, n$. At p^{th} level, one of the wavelet bands $LL^{(p)}, LH^{(p)}, HL^{(p)}, HH^{(p)}$ is chosen at random for the computation of the wavelet bands in the next level. In the proposed scheme, the wavelet bands are chosen with equal probabilities. Note that, as the wavelet bands are chosen randomly, the computed features will also take random values. Let us first consider the probability distribution of the i^{th} feature at the 2^{nd} level. The i^{th} feature $M_i^{(2)}$ have four possible values depending upon the choice of wavelet band at the first level to generate the wavelet bands at the 2^{nd} level and they are equally likely. Therefore the entropy of $M_i^{(2)}$ is $log(4)$. Hence the entropy of k random features, denoted by a vector $M^{(2)} = [M_1^{(2)}, M_2^{(2)}, \cdots, M_k^{(2)}]$ at the 2^{nd} level is $klog(4)$. The wavelet band that is chosen for next level of wavelet computation can be represented as follows.

$$B^{(p)} = \delta_{LL}^{(p)} LL^{(p)} + \delta_{LH}^{(p)} LH^{(p)} + \delta_{HL}^{(p)} HL^{(p)} + \delta_{HH}^{(p)} HH^{(p)} \quad (1)$$

where, $\delta_{LL}^{(p)}, \delta_{LH}^{(p)}, \delta_{HL}^{(p)}, \delta_{HH}^{(p)}$ are delta-functions associated with each wavelet band and its value can be either 0 or 1. The value of delta-function is 1 only when the corresponding wavelet band is chosen for next level of wavelet transform. Therefore, value of random variable $M_i^{(2)}$ can also be written as:

$$M_i^{(2)} = \delta_{LL}^{(1)} M_i^{(2)}(LL^{(1)}) + \delta_{LH}^{(1)} M_i^{(2)}(LH^{(1)}) + \delta_{HL}^{(1)} M_i^{(2)}(HL^{(1)}) + \delta_{HH}^{(1)} M_i^{(2)}(HH^{(1)}) \quad (2)$$

In level 3, the randomly chosen wavelet band, denoted by $B^{(2)}$, is further decomposed into four wavelet bands. The randomly chosen wavelet band and the extracted feature can be written as follows.

$$B^{(2)} = \delta_{LL}^{(2)} LL^{(2)} + \delta_{LH}^{(2)} LH^{(2)} + \delta_{HL}^{(2)} HL^{(2)} + \delta_{HH}^{(2)} HH^{(2)} \quad (3)$$

$$M_i^{(3)} = \delta_{LL}^{(2)} M_i^{(3)}(LL^{(2)}) + \delta_{LH}^{(2)} M_i^{(3)}(LH^{(2)}) + \delta_{HL}^{(2)} M_i^{(3)}(HL^{(2)}) + \delta_{HH}^{(2)} M_i^{(3)}(HH^{(2)}) \quad (4)$$

From equation 1-4, it follows that at level 3, the i$^{\text{th}}$ feature can take 4^2 many different values due to different choices of δ's at level 1 and 2. Each of those values of the feature is equally likely. Hence the entropy for the i$^{\text{th}}$ feature in the 3$^{\text{rd}}$ level is $log(4^2)$ and the entropy for k independent features is $klog(4^2)$. Following

a similar argument, the entropy of a feature at n^{th} level will be $log(4^{n-1}) = (n-1)log(4)$. Then the entropy of k many independent features at n^{th} level is $k(n-1)log(4)$.

It is observed that the random vectors $M_i^{(p)}, p = 2, 3, \cdots, n$ are not independent. In fact, for any fixed i, it is obvious that the sequence of random variables $M_i^{(2)}, M_i^{(3)}, ...M_i^{(n)}$ will form a markov chain of order 1 and the conditional distribution of the random variable $M_i^{(p+1)}$ given $M_i^{(p)}$ is a discrete uniform distribution with probabilities $\frac{1}{4}$ for each of its four distinct values. Hence, the joint entropy of the i^{th} feature for all levels of the wavelet tree is as follows.

$$E(M_i^{(2)}, M_i^{(3)}, \cdots, M_i^{(n)}) = E(M_i^{(2)}) + E(M_i^{(3)}|M_i^{(2)}) + \cdots + E(M_i^{(n)}|M_i^{(n-1)})$$

$$\text{i.e., } E(M_i^{(2)}, M_i^{(3)}, \cdots, M_i^{(n)}) = log(4) + log(4) + \cdots + log(4) = (n-1)log4$$

Considering k many independent features, we can obtain the joint entropy of all features at all levels as $k \times (n-1) \times log(4)$ which is same as the entropy of the feature vector at the last level. This result shows that the joint entropy of the features is a linear function of both the level (p) of the wavelet tree and the number of features (k) is used for computation of the hash value. Now, considering $n = 6$ and $k = 7$, the *entropy* of our algorithm is 70. Though we cannot compare it directly with the results obtained by Swaminathan et. al. [6] as they have reported the *differential entropy*; however, the best value of differential entropy obtained by them is 16.39.

3 Conclusion

In this paper, we have proposed a secure and robust image hashing algorithm. The proposed technique possesses very good discriminating property and is very much sensitive to malicious image processing operations like object insertion. It is robust against the content preserving image processing operations such as JPEG compression, filtering, small rotation etc.. Nevertheless, further improvement is desired against the geometric operations such as rotation, scaling, translation.

References

1. Birajdar, G.K., Mankar, V.H.: Digital image forgery detection using passive techniques: A survey. Digital Investigation 10(3), 226–245 (2013)
2. Goldfarb, B.: Digital deception, `http://brucegoldfarb.com/larrysface/deception.shtml` (accessed on June 2014)
3. Gray, F.: Pulse code communication. U.S. Patent 2,632,058; (filed on November 13, 1947) (issued March 17, 1953)
4. Kailasanathan, C., Naini, R.S., Ogunbona, P.: Image authentication surviving acceptable modifications. In: Proc. of IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, Baltimore, MD (2001)
5. Sencar, H.T., Memon, N.: Overview of state-of-the-art in digital image forensics. In: Indian Statistical Institute Platinum Jubilee Monograph series titled 'Statistical Science and Interdisciplinary Research', pp. 325–348. World Scientific Press (2008)
6. Swaminathan, A., Mao, Y., Wu, M.: Robust and secure image hashing. IEEE Transactions on Information Forensics and Security 1(2), 215–230 (2006)
7. Vetterli, M., Kovacevic, J.: Wavelets and subband coding. Prentice-Hall (1995)