GRIFFITH UNIVERSITY

An Overview of Reversible Data Hiding

Author

Awrangjeb, M

Published

2003

Conference Title

International Conference on Computer and Information Technology (ICCIT)

Version

Accepted Manuscript (AM)

Rights statement

© 2003 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

Downloaded from

http://hdl.handle.net/10072/392039

Link to published version

http://www.iccit.org/

Griffith Research Online

https://research-repository.griffith.edu.au

This paper was published at ICCIT 2003, 19-21 Dec, Jahangirnagar University, Bangladesh, pp 75-79

An Overview of Reversible Data Hiding

Mohammad Awrangjeb
Department of Computer Science
School of Computing
National University of Singapore
Singapore 117543
E-mail: mohamma1@comp.nus.edu.sg

ABSTRACT

Reversible data hiding is the technique that allows embedding (hide) data inside an image and later the hidden data can be retrieved as required and the exact copy of the original image is found. One of the most important requirements of reversible data hiding is that the distortions to the original signal should be such that artifacts are not visible. Because these distortions create problems in some fields such as medical, astronomical, and military images due to legal reasons. Another requirement is to have higher embedding capacity. The reversible data hiding is an emerging field for content authentication of images where the authentication information (say hash) is embedded inside the image. The higher the capacity the more information can be embedded inside the image. In this paper we present a review of reversible watermarking techniques proposed so far and give suggestions how to get reversible data hiding technique with higher embedding capacity and invisible artifacts.

Keywords: reversible, lossless, watermarking, authentication, embedding, capacity

1. INTRODUCTION

Reversible data hiding is mainly used for the content authentication of multimedia data like images, videos, electronic documents etc. because of its emerging demands in various fields such as law enforcement, medical imagery, astronomical research, etc. One of the most important requirements in this field is to have the original image during judgment to take the right decision. Cryptographic techniques based on either symmetric key or asymmetric key methods cannot give adequate security and integrity for content authentication. Because the main problem with the cryptographic techniques is that they are irreversible. Some authors use synonyms distortion-free, lossless, invertible, erasable watermarking for reversible data hiding. The lossless watermarking, a branch of fragile watermarking, is the process that allows exact recovery of the original image by extracting the embedding information from the watermarked image, if the watermarked image is deemed to be authentic, that means no single bit of the watermarked image is changed after embedding the payload to the original image. This technique embeds secret information with the image so that embedded message is hidden, invisible and fragile. Any attempt to change the watermarked image will make the authentication fail. Mehmet et al. [1] classify the reversible data hiding techniques into two types. In the first type of algorithms [2,3], during encoding a spread spectrum signal corresponding to the information payload is superimposed on the host signal. During decoding the payload (watermark signal) is subtracted from the watermarked image in a restoration step. In the second type of algorithms [1,4,5], some features of the original image are replaced with the watermark payload. The original portions of the image that will be replaced by watermark payload are compressed and passed as a part of the embedded payload during embedding. During the decoding process this compressed payload-part is extracted and decompressed. Thus the original image is achieved by replacing the modified portions with this decompressed original features. The algorithms of first type offer visible artifacts and lower capacity. On the other hand, algorithms of second type offer better visible quality and higher capacity than the first type; though, the first type algorithms offer little bit robustness that the second type algorithms do not show.

This paper presents a review of reversible data hiding techniques proposed so far in the literature. We explain some efficient algorithms with their advantages and disadvantages regarding the visible quality and capacity offered by them. In section 2 we represent the general principle of reversible data hiding, in section 3 we represent some recently proposed lossless watermarking algorithms. In section 4 we present some results, discussions comparing the algorithms, and our remarkable suggestions. In section 5 we conclude the paper.

2. THE GENERAL PRINCIPLE

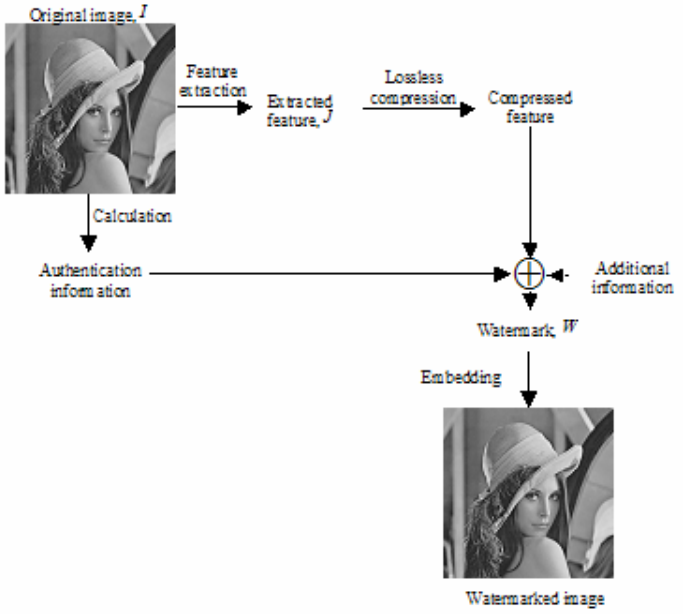


Fig. 1: General Principle (Embedding)

We represent here the general principle of lossless data hiding techniques from [8]. The general principle of reversible data hiding is that for a digital object (say a JPEG image file) *I*, a subset *J* of *I* is chosen. *J* has the structural property that it can be

easily randomized without changing the essential property of *I*, and it offers lossless compression itself to have enough space (at least 128 bit) to embed the authentication message (say hash of *I*). During embedding *J* is replaced by the authentication message concatenated with compressed *J*. If *J* is highly compressible only a subset of *J* can be used.

During the decoding process authentication information together with compressed *J* is extracted. This extracted *J* (compressed) is decompressed to replace the modified features in the watermarked image; hence the exact copy of the original image is found. Fig. 1 shows the graphical representation of the general principle of reversible data hiding. In this figure 'Additional information' means we can add here information to be conveyed as necessary. The decoding process is just the reverse of the embedding process. Therefore, so we do not present a separate figure for it.

3. REVERSIBLE DATA HIDING TECHNIQUES

In this section we represent some reversible data hiding algorithms proposed so far in the lossless watermarking literature. We also investigate related advantages and disadvantages of each algorithm.

3.1. Lossless Compression and Encryption of Bit-Planes

Fridrich et al. propose this algorithm in [3]. Space to hide data is found by compressing proper bit-plane that offers minimum redundancy to hold the hash (authentication information). Lowest bit-plane offering lossless compression can be used unless the image is not noisy. In completely noisy image some bit-planes exhibit strong correlation. These bit-planes can be used to find enough room to store the hash. Hash length is generally 128 bit using MD5 algorithm [6]. The algorithm starts lossless compression from 5th bit-plane and calculates redundancy by subtracting compressed data size from number of pixels. The authors use the JBIG lossless compression method [7] to compress the bit-planes.

During embedding the algorithm first calculates the hash of the original image, finds the proper bit-plane, and add the hash with the compressed bit-plane data. Then it replaces selected bit-plane by concatenated data. For more security the concatenated hash with compressed data is encrypted using symmetric key encryption based on 2-dimensional chaotic maps [9]. This algorithm takes variable sized blocks and gives the encrypted message as long as the original message, so no padding is needed. Other public or symmetric key algorithms can be used, but they require padding to embed the encrypted message and hence increase distortion. During decoding after key bit-plane selection the data is decrypted and hash is separated from the compressed original bit-plane data. The bit-plane is replaced by the decompressed data; hence the exact copy of the original image is found. The hash of the reconstructed image is calculated and compared with the extracted hash; if both are same the image in question is authentic.

The advantages of this algorithm are – (i) high capacity, (ii) security is equivalent to the security provided by cryptographic authentication, and (iii) can be applied for the authentication purposes of JPEG files, complex multimedia objects, audio files, digitized hologram, etc. The disadvantages are – (i) noisy image forces the algorithm to embed information in higher bit-plane when the distortions are higher and easily visible, (ii) single bit-plane in a small image does not offer enough space to hide hash after compression, so two or more bit-planes are required and the artifacts must be visible, and (iii) capacity is not high enough to embed large payload.

3.2. Reversible Data Hiding at Low Pixel-Levels

Mehmet et al. [1] propose a reversible data hiding technique that uses prediction based conditional entropy coder utilizing static portions of the input signal as side-information to improve the compression efficiency. Hence the lossless data embedding capacity is increased. This spatial domain method is the modification of generalized LSB embedding technique and uses very simple signal features: lowest levels of raw pixels. It follows the general principal [8] of lossless embedding.

The algorithm searches the whole image to have the first L (say, L = 4) lowest levels of pixel values. It compresses these pixel values using CALIC lossless image compression algorithms [16] and check whether it gives enough space (128-bit for hash). If the given capacity is lower than expectation the algorithm increases L and continues searching. Once it finds enough capacity, it concatenates the hash with compressed pixel values. The concatenated bit-string is converted into L-ary symbols to replace the lowest L-levels of pixel values. The decoding process is just the reverse of the embedding phase.

The advantages are – (i) simple algorithm, and (ii) higher capacity can be found with the increase of embedding level L. The disadvantages are – (i) capacity depends on image structure, smooth images give higher capacity than irregular textured images, and (ii) artifacts are visible with the increase of embedding level L. Though the algorithm gives a very high capacity, it gives incredible distortions to the original image.

3.3. Circular Interpretation of Bijective Transformations

Macq et al. proposed an original circular interpretation of bijective transformations as a solution to fulfill all quality and functionality of lossless watermarking. In first work Macq [15] proposes an additive method that is criticized by him in [17] for having 'salt and pepper' visual artifacts due to wrapped around pixels. In [14, 17] they propose a modification that decreases the problem – wrapped around pixel.

It essentially follows the idea of patchwork algorithm [12]. In that method each bit of payload (message) is associated with a group of pixels. Each group of pixels is divided into two-pseudo random set of pixel-zones *A* and *B*. The histogram of each zone is mapped to a circle and the position of each gray scale is defined in corresponding position on the circle by a weight proportional to its frequency in the histogram. The position of the center of mass (represented by vectors $\vec{v}_a$ and $\vec{v}_b$ for zone *A* and *B* respectively) in the resulting distribution of weights for each zone corresponds to the average luminance

value of that zone. Since, zones *A* and *B* are pseudo randomly selected it is highly probable that vectors $\vec{v}_a$ and $\vec{v}_b$ are very close to each other (average luminance values for zones *A* and *B* are almost same) before embedding. Slight rotations of vectors $\vec{v}_a$ and $\vec{v}_b$ in opposite directions allows to embed one bit of information. So, during embedding phase based on the bit being embedded their luminance values are incremented or decremented (vectors $\vec{v}_a$ and $\vec{v}_b$ are rotated). The vector $\vec{v}_a$ rotates clockwise (to embed a '1') or anti-clockwise (to embed a '0') or the vector $\vec{v}_b$ rotates clockwise (to embed a '0') or anti-clockwise (to embed a '1'). During extraction phase the bit is inferred from the sign of smallest angle between vectors $\vec{v}_a$ and $\vec{v}_b$, i.e. difference between the mean values of zones *A* and *B*. In fact, the angle between vectors $\vec{v}_a$ and $\vec{v}_b$ at the receiver end provides direction of rotation during embedding and enables bit retrieval and reversibility at the pixel levels, i.e. to reconstruct the original image.

The advantages of this algorithm are – (i) visible quality is improved in watermarked image, and (ii) capacity is high. The disadvantages are – (i) there are some groups of pixels that do not offer themselves to embed message: so expected capacity decreases, (ii) additional information should be conveyed with payload to enable reversing for some problematic blocks at the receiver end: so effective capacity decreases, and (iii) complex algorithm.

3.4. Based on Integer Wavelet Transform

In Xuan et al. [13] propose a lossless data hiding having large capacity based on integer wavelet transform. It hides authentication information and bookkeeping data into a middle bit-plane of integer wavelet coefficients in high frequency sub-bands. The histogram modification or integer modulo addition is used to prevent gray scale overflowing during data embedding. The method uses second-generation wavelet transform IWT [11].

The authors find more bias between 1s and 0s starting from 2nd bit-plane to higher bit-planes of IWT coefficients. To make the watermarked image perceptually as same as the original image and to have high PSNR they tell to embed information into middle bit-plane and in the high frequency sub-bands respectively. To compress it they use arithmetic coding from [10]. The watermark payload concatenated with compressed data is embedded with a secret key. In extraction phase, the watermark (say, hash of original image) is extracted and the original image is reconstructed in the opposite manner. To prevent the gray-scale overflow either histogram modification or gray scale modification is used as pre-processing and post-processing during embedding and extraction phases respectively.

The advantages are – (i) high capacity, and (ii) use of secret key during embedding increases security. The disadvantages are – (i) often multiple bit-planes are required to have enough space when the artifacts become visible, and (ii) gray scale mapping (lowest and highest 16 gray levels to 15 and 240 respectively) may introduce visible artifacts into the watermarked image.

3.5. High Capacity Watermarking Based on Difference Expansion

In [18,19] Tian propose a high quality reversible watermarking method with high capacity based on difference expansion. Pixel differences are used to embed data, this is because of high redundancies among the neighboring pixel values in natural images.

During embedding – (i) differences of neighboring pixel values are calculated, (ii) *changeable* bits in that differences are determined, (iii) some differences are chosen to be expandable by 1-bit, so *changeable* bits increases, (iii) concatenated bit-stream of compressed original *changeable* bits, the location of expanded difference numbers (location map), and the hash of original image (payload) is embedded into the *changeable* bits of difference numbers in a pseudo random order, (iv) use the inverse transform to have the watermarked pixels from resultant differences. During watermark extraction – (i) differences of neighboring pixel values are calculated, (ii) *changeable* bits in that differences are determined, (iii) extract the *changeable* bit-stream ordered by the same pseudo random order as embedding, (iv) separate the compressed original *changeable* bit-stream, the compressed bit-stream of locations of expanded difference numbers (location map), and the hash of original image (payload) from extracted bit-stream, (v) decompress the compressed separated bit-streams and reconstruct the original image replacing the changeable bits, (vi) calculate the hash of reconstructed image and compare with extracted hash.

The advantages are – (i) no loss of data due to compression-decompression, (ii) also applicable to audio and video data, and (iii) encryption of compressed location map and *changeable* bit-stream of different numbers increases the security. The disadvantages include – (i) there may be some round off errors (division by 2), though very little, (ii) largely depends on the smoothness of natural image; so cannot be applied to textured image where the capacity will be zero or very low, and (iii) there is significant degradation of visual quality due to bit-replacements of gray scale pixels.

3.6. Reversible Data Hiding by Histogram Shifting

Ni et al. [20] utilizes zero or minimum point of histogram. If the peak is lower than the zero or minimum point in the histogram, it increases pixel values by one from higher than the peak to lower than the zero or minimum point in the histogram. While embedding, the whole image is searched. Once a peak-pixel value is encountered, if the bit to be embedded is '1' the pixel is added by 1, else it is kept intact. Alternatively, if the peak is higher than the zero or minimum point in the histogram, the algorithm decreases pixel values by one from lower than the peak to higher than the zero or minimum point in the histogram, and to embed bit '1' the encountered peak-pixel value is subtracted by 1. The decoding process is quiet simple and opposite of the embedding process. The algorithm essentially does not follow the general principle of lossless watermarking in [8].

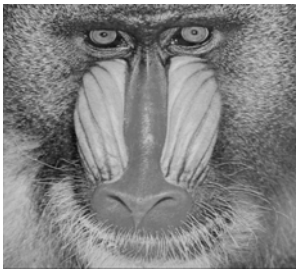
The advantages of this method are – (i) it is simple, (ii) it always offers a constant PSNR 48.0dB, (iii) distortions are quite invisible, and (iv) capacity is high. The disadvantages are – (i) capacity is limited by the frequency of peak-pixel value in the histogram, and (ii) it searches the image several times, so the algorithm is time consuming.

There are some more efficient algorithms have also been proposed. We refer our survey report of lossless watermarking in [21] to have complete research knowledge of reversible data hiding.

4. DISCUSSIONS AND SUGGESSIONS

For Lena image in Fig. 2(a), [1] gives a high embedding capacity of 9325 bytes when embedding level is 8 and PSNR is 38.0dB. With the increase of embedding level the capacity increases, but the distortions becomes more visible. For Baboon image in Fig. 2(b) it gives 1787 bytes of capacity with same embedding level and PSNR. Circular interpretation of bijective transformations in [17] gives 413 bytes of capacity for Lena image when image is divided into 4 by 4 blocks. For Lena image the algorithm in [13] gives about 10.5kB capacity with a PSNR of 36.64dB, for Tiffany image in Fig. 2(c) it gives a capacity of 11kB with 28.91dB PSNR, for Baboon its capacity is 1.82kB with a PSNR 32.76dB, and for F16 image in Fig. 2(d) its capacity is 11.5kB with 36.30dB PSNR. The lossless watermarking algorithm based on difference expansion [18,19] gives 12.33kB capacity with a PSNR 40.06dB. The algorithm in [20] gives embedding capacity 682 bytes for Lena image, 1.97kB for F16 image, 1097 bytes for Tiffany image, and 677 bytes for Baboon image with a constant 48.0dB PSNR.

(a) Lena



(b) Baboon

(c) Tiffany

(d) F16

Fig. 2: Test Images used in Watermarking Literature, all images are 512×512 8-bit gray scale image

The aims of the reversible data hiding are two folds: first is to make the visible distortions as low as possible so that the artifacts are not visible, second is to make the embedding capacity as high as possible. There is a trade of between distortions and embedding capacity. If we make the distortions low we can embed only a few data. On the other hand, we can get high capacity with low visible quality. The only way to achieve these two goals is to invent an algorithm that can make a better trade off between embedding capacity and visible artifacts.

According to the *human visual system* (HVS) we know that we can change (increase or decrease) a pixel value to a certain amount so that the change (distortion) is not noticeable to the human eye, i.e. the artifact becomes perceptually invisible. This certain amount of change to a pixel value is called the *just noticeable distortion* (JND) value of that pixel. If we can take the HVS into account when we embed information to the original image the watermarked image should be perceptually as same as the original image.

In order to gain high capacity we can consider several issues. First, the feature selection system should be such that the extracted features are highly lossless-compressible. Second, the lossless compression algorithm should be efficient. A compression algorithm does not compress all kinds of data with the same ratio. After selection of an efficient algorithm the extracted feature should be arranged such that they are highly compressible by the selected algorithm. Third, the embedding process should keep effective capacity (free capacity after embedding watermark: compressed hash and authentication information) as high as possible. For example, the reversible data-hiding algorithm by Mehmet et al. selects lowest L-levels of pixel values as features. After compression it converts the watermark data into L-ary symbols. So, if the extracted feature-size is *x* bytes, compressed feature-size is *y* bytes and watermark length is *z* bytes, the gained capacity is *x-y* bytes, but the effective capacity is $x - z\log_L 255$ bytes, instead of *x-z* bytes. The higher the effective capacity the higher the payload, i.e. additional necessary information can be added to the watermark. We can add confidential patient report, patient's personal information, referenced doctor's information etc. as payload to the medical images, which is an active research area to transfer medical image together with related information in *hospital information system* (HIS).

5. CONCLUSIONS

In this paper we have made a clear representation of recent reversible data hiding algorithms together with their advantages and disadvantages. We have shown comparisons of these algorithms and presented remarkable suggestions. Our suggestions include considering the HVS, allows changing each pixel value to a certain amount defined by the JND value of that pixel, which could make the watermarked image having perceptually better quality. Moreover, we have suggested finding an efficient lossless compression algorithm to compress the extracted features and also to embed the data to the original image such that the effective embedding capacity is high. We hope our remarkable suggestions would be helpful to invent an

algorithm that could make a better trade off between visible artifacts and embedding capacity.

REFERENCES

[1] M.U. Celik, G. Sharma, A.M. Tekalp., and E. Saber, "Reversible Data Hiding", In Proc. of International Conference on Image Processing, Rochester, NY, USA, Vol. 2, pp. 157-160, September 24, 2002.

[2] C.W. Honsinger, P.W. Jones, M. Rabbani, and J.C. Stoffel, "Lossless Recovery of an Original Image Containing Embedded Data", In US Patent no. 6278791, August 2001.

[3] J. Fridrich, M. Goljan, and D. Rui, "'Invertible Authentication", In Proc. of SPIE Photonics West, Security and Watermarking of Multimedia Contents III, San Jose, California, USA, Vol. 3971, pp. 197-208, January 2001.

[4] J. Fridrich, M. Goljan, and D. Rui, "Lossless Data Embedding - New Paradigm in Digital Watermarking", In Special Issue on Emerging Applications of Multimedia Data Hiding, Vol. 2, pp. 185-196, February 2002.

[5] J. Tian, "Wavelet-based Reversible Watermarking for Authentication", In Proc. Security and Watermarking of Multimedia Contents IV, Electronic Imaging 2002, Vol. 4675, pp. 679-690, 20-25 January 2002.

[6] R. Rivest, "The MD5 Message-Digest Algorithm", In DDN Network Information Center, http://www.ietf.org/rfc/rfc1321.txt, April 1992.

[7] K. Sayood, "Introduction to Data Compression", Morgan Kaufmann, 1996, pp. 87-94.

[8] J. Fridrich, M. Goljan, and D. Rui, "Lossless Data Embedding for all Image Formats", In Proc. SPIE Photonics West, Electronic Imaging, Security and Watermarking of Multimedia Contents, San Jose, California, USA, Vol. 4675, pp. 572-583, January, 2002.

[9] J. Fridrich, "Symmetric Ciphers Based on Two-Dimensional Chaotic Maps", On Int. Journal of Bifurcation and Chaos, 8(6), pp. 1259-1284, June 1998.

[10] Y. Q. Shi, and H. Sun, "Image and Video Compression for Multimedia Engineering", Boca Raton, FL: CRC, 1999.

[11] A. R. Calderbank, I. Daubechies, W. Sweldens, and B. Yeo, "Wavelet Transformations that Map Integers to Integers", In Proc. of Applied and Computational Harmonic Analysis, 1998, Vol. 5, No. 3, pp. 332-369.

[12] W. Bender, D. Gruhl, N. Morimoto, and A. Lu, "Techniques for Data Hiding", In IBM Systems Journal, 1996, Vol. 35, No. 3-4, pp. 313-336.

[13] G. Xuan, J. Chen, J. Zhu, Y.Q. Shi, Z. Ni, and W. Su, "Lossless Data hiding based on Integer Wavelet Transform", In Proc. of IEEE International Workshop on Multimedia Signal Processing. Marriott Beach Resort St. Thomas, US Virgin Islands, 9-11 December 2002.

[14] C. De Vleeschouwer, J. E. Delaigle, and B. Macq, "Circular Interpretation of Histogram for Reversible Watermarking", In Proc. of IEEE 4th Workshop on Multimedia Signal Processing, pp. 345-350, 2001.

[15] B. Macq, "Lossless Multi-Resolution Transform for Image Authenticating Watermarking", In Proc. of EUSIPCO, Tempere, Finland, Sept 2000.

[16] X. WU, "Lossless Compression of Continuous-Tone Images via Context Selection, Quantization, and Modeling", IEEE Transactions on Image Processing, Vol. 6, No. 5, pp. 656-664, May 1997.

[17] C. De Vleeschouwer, J. E. Delaigle, and B. Macq, "Circular Interpretation of Bijective Transformations in Lossless Watermarking for Media Asset Management", On IEEE Transactions on Multimedia, March 2003.

[18] J. Tian, "Wavelet Based Reversible Watermarking for Authentication", In Proc. Security and Watermarking of Multimedia Contents IV, Electronic Imaging 2002, Vol. 4675, pp. 679-690, 20-25 January 2002.

[19] J. Tian, "Reversible Watermarking by Difference Expansion", In Proc. of Workshop on Multimedia and Security, pp. 19-22, December 2002.

[20] Z. Ni, Y.Q. Shi, N. Ansari, and W. Su, "Reversible Data Hiding, In Proc. of International Symposium on Circuits and Systems, Bangkok, Thailand, Vol. 2, pp. 912-915, 25-28 May 2003.

[21] M. Awrangjeb, "A Survey Report: Content Authentication with Lossless Watermarking", http://comp.nus.edu.sg/~mohamma1/survey.pdf.

This paper was published at ICCIT 2003, 19-21 Dec, Jahangirnagar University, Bangladesh, pp 75-79

View publication stats