Reversible Data Hiding of Images Based on Plaintext Domains

Yu Fu[1*]

[1]Institute of international education, Chongqing college of mobile communication, Hechuan, Chongqing, China

Abstract. Information hiding technology can embed secret information into images, video, audio and other media. Authorized users can get embedded data from the medium of implicit data, so as to realize the transmission of information. Traditional information hiding only focuses on covert transmission, making it impossible for unauthorized persons to detect the existence of hidden information and does not require the carrier to remain undamaged after information extraction. But reversible data hiding technology can not only realize covert communication, but also realize lossless restoration of the original carrier image after extracting the hidden information, therefore, it has extremely important value in the field of high fidelity, This article systematically reviews the development of image reversible information hiding technology, Four methods of lossless compression, difference expansion, histogram shifting, and prediction-error expansion are introduced, the advantages and limitations of various methods were summarized, and a comparative analysis was conducted. The future development direction was discussed, providing new ideas for the future development of this technology.

1 Introduction

In the digital age, various multimedia such as audio and images have become the main ways of information transmission, among them, images are one of the most widely used and important information carriers. Images are the core information carriers in fields such as medical diagnosis, military reconnaissance, and copyright tracing, so global attention has been focused on the encryption and security of data. Traditional information hiding only focuses on covert transmission, while the core of reversible image data hiding technology is to maintain the fidelity of the image after completing the embedding of secret information, making it difficult for attackers to detect the hidden information. Reversible data hiding (RDH), as an important branch of information hiding technology, is mainly divided into two major categories: plaintext domain and ciphertext domain. The reversible hiding technology of image information originates from early information hiding techniques. Because irreversible data hiding can permanently damage the carrier and cannot meet the

* Corresponding author's email: ffffayee@outlook.com

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

high-fidelity requirements of carriers such as medical images and military images, RDH technology emerged. And image is one of the most common digital media, thus it is widely used for information storage and transmission. The color and gray of adjacent pixels in the image have a strong correlation, which makes it difficult for the naked eye to identify whether the image contains hidden information, so as to protect the encrypted information. At the same time, RDH technology is confronted with unique technical requirements and challenges, for example visual sensitivity, robustness and security. The research on RDH has been driven by the information security demands in the era of big data. This paper conducts a systematic review of the development history of reversible data hiding technology in plaintext domains, and summarizes the representative methods and core ideas at different development stages of this technology, and analyzes the advantages and limitations of each method. This paper will analyze the traditional RDH method, discuss the technological breakthroughs in recent years and look forward to the future development direction, so as to provide new ideas for the development of this technology in the future.

2 Image RDH method

2.1 Lossless Compression

Image compression can be classified into two major categories based on the degree of information retention: lossy compression and lossless compression. Lossy compression allows for a certain degree of information loss to some extent, while lossless compression has higher requirements and stronger confidentiality compared to lossy compression. Since information loss is not allowed, the compression efficiency is difficult to improve.

The earliest RDH method was mainly based on lossless compression technology. Its purpose was to collect the low significant bits of the original image for lossless compression, and then the low significant bits of the original image would be jointly replaced by the secret information to be embedded and the compressed low significant bits. However, due to the poor local correlation of low significant bits, it cannot accommodate a large amount of embedding space. If multiple low significant bit planes are accepted, it will cause significant embedding distortion.

Lossless image compression uses the redundancy of the image, by accurately identifying and eliminating the redundant information in the image, so as to achieve file volume compression. The first is to eliminate spatial redundancy. By taking advantage of the highly similar characteristics of adjacent pixels in color and brightness in the image, the difference data between the reference pixel and adjacent pixels are recorded by using difference coding and other technologies, instead of storing the complete pixel information repeatedly, so as to greatly reduce data redundancy. The second is to optimize coding redundancy. For the problem that the coding length of the original pixel value does not match its occurrence frequency, Huffman coding, LZW coding and other algorithms can be used to allocate shorter coding for pixels with high frequency and longer coding for pixels with low frequency, so that data storage is more efficient.

At the same time, lossless image compression has some limitations. 1）Only limited compression can eliminate the inherent redundancy of data. Since the core of lossless compression is not to lose information, it cannot discard non critical information like lossy compression. 2）It has higher requirements for hardware and time. During the compression and decompression process, complex algorithms such as Huffman coding and LZW coding need to be utilized to perform a large amount of data calculation and processing. When dealing with high-definition videos and large-capacity documents, more CPU resources will be consumed and a longer time cost will be required. 3）The applicable scenarios are limited and cannot meet the scenarios with extremely high requirements on file volume, such as network transmission of high-definition short videos, mobile phone storage of massive photos, etc., which are still large, so it is difficult to balance the storage cost and transmission efficiency

2.2 Difference Expansion

Tian was the first to propose the Difference Expansion (DE) method, which is based on image redundancy. By enlarging the difference between adjacent pixels and freeing up space in the difference to embed additional bit information [1]. The purpose is high-capacity information hiding embedding. By modifying the difference between adjacent pixel pairs of the image and expanding it, secret information such as watermarks is embedded in the difference. The maximum embedding rate in a single pass is 0. 5bpp, but its capacity can be further increased through optimization such as double-layer embedding. At the same time, it can ensure high fidelity and reversibility of the image. When embedding, it can slightly modify the pixel difference, control the visual distortion of the image, and restore the original image without loss.

Compared with the method based on lossless compression, when the same embedding capacity is obtained, the peak signal-to-noise ratio (PSNR) of the dense image obtained by the technology is higher, which means that the embedding distortion is reduced. And the theoretical embedding rate of the DE method is 0.5 bit/ pixel. However, this method has a relatively large modification range for some pixels, which is prone to causing the problem of boundary pixel value overflow. To solve this problem, it is necessary to build a location map to identify the scalability of pixel pairs: where "1" represents that there is no risk of overflow after the expansion of pixel pairs, which can be used to embed information; "0" means that overflow or underflow will occur after expansion, and information cannot be embedded [2]. The receiving end needs to extract the auxiliary information first, and then decode the secret information based on the information. However, for complex texture images, the volume of the compressed location map may exceed the total embedding capacity, resulting in the failure of the DE method. Even if the compression effect of the positioning map is ideal, it will still occupy some payload, thereby affecting the overall performance of the method. The sparse bitmap is lossless compressed, and then embedded into the carrier as auxiliary information and secret information. The receiving end needs to first extract the auxiliary information and then decode the secret information based on this information. However, for complex texture images, the volume of the compressed location map may exceed the total embedding capacity, resulting in the failure of the DE method. Even if the compression effect of the positioning map is ideal, it will still occupy some payload, thereby affecting the overall performance of the method. Therefore, enhancing the sparsity of topographic maps and reducing the volume scale of compressed topographic maps is a practical and feasible technical approach to optimizing this method. Zheng et al. pointed out that in the methods proposed by Tian Yu in relevant literature, when performing numerical expansion on large amplitude differences, significant distortion is introduced [3]. Assuming a threshold, the smaller the threshold, the fewer the number of differences available for expansion, and the smaller the available capacity obtained after expansion. Increasing the difference will increase the embedding distortion. Zheng et al. proposed to combine the theory of image bit-plane decomposition and extend the DE algorithm to the bit-plane domain, performing DE operations in the remaining lower-order bit-planes. This approach can not only compress significant differences but also increase the embedding capacity and reduce embedding distortion.

2.3 Histogram Shifting

Ni et al. Proposed a histogram shifting (HS) method [4]. The basic principle of this method is that, relying on the peak points in the image gray histogram, that is, the gray level and zero point with the highest frequency of pixels, all pixel values between the two points are moved left or right, subsequently, a pixel-free blank area is generated within the interval defined by the peak point and the zero point, and it is used as a carrier to achieve the embedding of secret information. The maximum embedding amount of HS algorithm is different due to the restriction of the peak point characteristics of the image itself; Its peak signal-to-noise ratio (PSNR) shows a decreasing rule as the embedding amount increases. However, even when the maximum embedding amount is reached, the PSNR can still be maintained at a relatively high level. The processed image has no obvious distortion. Although this algorithm has outstanding overall performance, it still has obvious shortcomings in the dimension of embedding capacity.

To improve performance via a prediction error histogram with reduced variance, Sachnev et al. developed a rhombus predictor [5]. The experimental results show that, compared with the traditional predictors, the prediction performance of this predictor is more advantageous; However, in order to achieve complete reversibility, it is necessary to introduce a two-stage embedding mechanism on this basis. First, all pixels of the image are divided into two types of checkerboard distribution according to the gray and white attributes. In each embedding stage, only one type of pixel is selected to carry the embedding information, while the other type of pixel remains in the original state to assist the prediction process.

Li et al. also proposed an improved PEE method combining pixel value ranking prediction, which effectively reduced the invalid translation of some pixel values [2]. The core process is as follows: First, the image is divided into pixel blocks of consistent size, and then the pixel values within each block are sorted in ascending order, then the following core processing steps are divided into two parts: The first part is the prediction stage, which specifically uses the second largest number in the target pixel set as the reference benchmark to predict the maximum value and the second decimal as the reference benchmark to predict the minimum value. The second part is the subsequent processing stage. Only the pixel values with an absolute prediction error equal to 1 are used to achieve the expansion of the embedded secret bit. For the pixel values with an absolute prediction error greater than 1, the bidirectional translation operation is uniformly implemented.

Li et al. introduced the concept of "complexity" on the basis of the PEE method, which can accurately distinguish the pixel features of different texture regions in the image, and then allocate the prediction errors corresponding to pixels of different complexities to multiple histograms for reversible modification [2]. So this method, known as mhm, not only improves the adaptive ability of the target embedding capacity, but also more reasonably filters the extended position parameters, reduces redundant pixel adjustments, and finally obtains a higher quality encrypted image by reducing the distortion caused by invalid translation. And the performance of this method is determined by the extended position parameters. The core objective is to minimize the distortion caused by invalid translation while meeting the target embedding capacity. Given the embedded capacity demand, it is necessary to select the optimal parameters for each sub prediction error histogram through the "capacity allocation" strategy to minimize distortion. In subsequent studies, Some people proposed to select multiple pairs of extended positions in a single histogram to increase the embedding capacity, and respectively optimized the parameter selection algorithm[2]; Wang et al. transformed the multi-parameter selection problem into a rate-distortion optimization problem, solved it with the help of genetic algorithms, and proposed a multi-feature complexity measurement method at the same time, which enhanced the correlation between prediction error and complexity, thereby more accurately distinguishing image textures and obtaining an MHM partition close to the optimal one [6]. In order to achieve the same goal, Weng et al. calculated new features, combined with the K-means clustering algorithm to divide multiple sub histograms, and improved the vertical and horizontal intersection algorithm to search the best embedding point of each sub histogram [7]; Chang et al. Combined two-dimensional histogram modification with multi histogram modification to further improve the performance of the method [8].

2.4 Prediction-Error Expansion

Thodi and Rodriguez first proposed the prediction error expansion (PEE) method in 2004 [9]. Compared with the de method which only considers the correlation between two adjacent pixels, the pee method further incorporates more local correlation between pixels, and embeds secret bits by replacing the pixel difference with prediction error, thus significantly improving the embedding capacity. Also the PEE algorithm can be applied to a very wide range of modules, and there are also many related studies. Feature selection and feature modification are two steps included in a complete PEE algorithm. The embedding features mentioned here are predictive residuals, which are equivalent to dividing the differences. "Bad" differences are not conducive to embedding. In a system mainly based on PSNR, the smaller the variation of pixels by the algorithm, the better the performance and the more accurate the prediction results. The core criterion for a "good" algorithm is to make the prediction residuals as close to zero as possible. Simultaneously,this stage completes the reconstruction of the embedded features, converting the pixel distribution of the original image into the distribution pattern of the predicted residuals. Given the high similarity between the residual distribution and the Laplacian distribution, the representation of image information is more streamlined, thereby achieving an efficient entropy reduction effect. However, compared with the DE method, the embedding rate of the PEE method can theoretically reach 1 bit/ pixel, but the missing points of pixel value overflow and high embedding distortion are still relatively obvious. In order to solve this problem, thodi and Rodriguez proposed a method combining HS with PEE. The prediction error histogram is a 0-centered Laplacian like distribution, which is steeper and has smaller variance than the pixel statistical histogram [10]. Compared with the traditional HS method, the embedding performance is better. The fusion scheme of HS and PEE integrates the core advantages of both methods, which can not only minimize the adjustment range of pixel values to the greatest extent, but also significantly increase the data embedding capacity. To solve the problem of pixel value overflow, since the maximum change amount of a single pixel is only 1, it is only necessary to preprocess the original boundary pixel values 255 and 0 to 254 and 1. Although the positioning map still needs to be embedded in the image as auxiliary information, fewer pixel modifications improve the sparsity of the overflow positioning map, thereby reducing the impact of pixel value overflow to a certain extent.

3 Comparative Analysis and Discussion

Image reversible data hiding technology, with its core feature of "information embedding - lossless recovery", has established irreplaceable application value in key fields such as sensitive data protection and image traceability authentication, and has formed a technical system of multi-method integration and multi-scenario adaptation. From the early basic framework based on lossless compression to the advanced scheme combining deep learning and adaptive embedding strategy, the technology has made continuous breakthroughs in the balance of embedding capacity, visual quality and robustness, laying a solid foundation for cross domain applications.

Table 1. Typical method characteristic

Method	Advantage	Limitation
Lossless compression	Data authenticity	Low compression ratio
Difference expansion	High embedding	Distortion problem
Histogram shifting	Low embedding distortion and high image quality	Low embedding capacity
Prediction-error expansion	Greatly improve compression efficiency	The compression effect on noisy images is poor

Table 1 sorts out the characteristics of four typical information hiding and image compression methods: lossless compression, difference expansion, histogram translation and prediction error expansion. The technical advantages and application limitations of each method show significant differentiation characteristics: Lossless compression takes the complete fidelity of data as the core technical advantage, which can realize the lossless restoration of compressed data, but its low compression rate limits its application in massive data storage scenarios. The difference expansion technology has a large capacity for information embedding and can meet the demand for hiding high-capacity data. However, it is prone to cause image distortion during the embedding process, which has a negative impact on the visual quality of the image. The histogram shifting method can control image distortion to the greatest extent and effectively guarantee the visual effect of the processed image. However, its information embedding capacity is obviously limited and it is difficult to adapt to the task of large capacity data hiding. The prediction-error expansion technology has significantly enhanced the efficiency of image compression, with outstanding compression performance. However, this method has a poor processing effect on noisy images and its anti-interference ability needs to be strengthened.

In the future, the iterative upgrading of related technologies can be promoted from three aspects: technology integration, optimization of shortcomings, and field expansion. And promote the research and development of hybrid technology and realize performance complementarity by integrating the technical advantages of different methods. For example, combine the efficient compression characteristics of prediction error expansion and the low distortion advantages of histogram translation to build a new processing model with high compression efficiency and high image quality. Also expanding towards intelligence and multi-field applications, an adaptive information embedding and compression model is constructed relying on machine learning algorithms. At the same time, the application scope of the technology is extended from static images to multimedia fields such as video and audio. On the basis of taking into account the capacity of information embedding, image visual quality and steganography security, the practical application value and environmental robustness of the technology are enhanced. By accurately responding to the needs of the times and overcoming the bottleneck of core technologies, image reversible information hiding technology will release greater application potential in more sensitive and general fields and provide more comprehensive technical support for information security.

4 Conclusion

In short, after more than 20 years of development, reversible data hiding technology in the plaintext domain has made significant progress, and its rate distortion performance is close to the theoretical upper limit. However, there is still room for improvement to further narrow the gap with the ideal theoretical limit.

At present, research in this field still mainly focuses on the trade-off among reversibility, embedding capacity and imperceptibility. To promote the reversible data hiding technology in plaintext domains from laboratory research to practical application, future research must go beyond the traditional paradigm of merely pursuing rate-distortion optimization. Computational efficiency, robustness against conventional image processing, and security against malicious attacks must be taken into account as core design elements.

In addition, build a comprehensive and scientific evaluation system to balance performance indicators and practical application needs. Only by adhering to this open and inclusive research direction, the reversible data hiding method in the plaintext domain achieve a win-win situation of theoretical breakthrough and practical value, so as to promote the steady development of the whole field.

References

1. J. Tian, Reversible data embedding using a difference expansion. IEEE Trans. Circuits Syst. Video Technol. 13, 890–896 (2003)
2. X.-S. Wu, M. Xu, T. Qiao, B.-M. Pan, X. Liao, A survey of reversible information hiding in image spatial domain. J. Image Graph. 27, 125–149 (2022)
3. Z. Wang, Y. Zhang, Reversible watermarking algorithm based on improved bit-plane decomposition and adaptive difference expansion, Signal Processing: Image Communication, 106, 116775 (2023).
4. Z. Ni, Y.-Q. Shi, N. Ansari, W. Su, Reversible data hiding. IEEE Trans. Circuits Syst. Video Technol. 16, 354–362 (2006)
5. V. Sachnev, H.J. Kim, J. Nam, S. Suresh, Y.-Q. Shi, Reversible watermarking algorithm using sorting and prediction. IEEE Trans. Circuits Syst. Video Technol. 19, 989–999 (2009)
6. J.-X. Wang, X. Chen, J.-Q. Ni, N.-X. Mao, Y.-Q. Shi, Multiple histograms-based reversible data hiding: framework and realization. IEEE Trans. Circuits Syst. Video Technol. 30, 2312–2328 (2020)
7. S.-W. Weng, W.-L. Tan, B. Ou, J.-S. Pan, Reversible data hiding method for multi-histogram point selection based on improved criss-cross optimization algorithm. Inf. Sci. 549, 13–33 (2021)
8. J. Chang, F. Ding, X.-L. Li, G.-P. Zhu, Hybrid prediction-based pixel-value-ordering method for reversible data hiding. J. Vis. Commun. Image Represent. 77, 103097 (2021)
9. D.M. Thodi, J.J. Rodriguez, Prediction-error based reversible watermarking. In Proceedings of the 2004 International Conference on Image Processing (IEEE, Singapore, 2004), pp. 1549–1552.
10. M.R. Islam, M. S. Hossain, Advanced expansion embedding techniques for high-capacity reversible watermarking, IEEE Transactions on Image Process, 30, 5678-5690 (2021).