A survey on Reversible Watermarking Techniques for Image Security

Swati Gupta¹, Dr. Raju Baraskar², Dr. Shikha Agrawal³

¹Student, Department of Computer Science and Engineering, UIT-RGPV, BHOPAL-462036
²Assistant Professor, Department of Computer Science and Engineering, UIT-RGPV, BHOPAL-462036
³Assistant Professor, Department of Computer Science and Engineering, UIT-RGPV, BHOPAL-462036

Abstract - The image is significantly used in applications such as medical area, research area, image conferencing, a military image system, online transactions, digital signatures, passwords etc. Watermarking techniques along with some Cryptography are used to protect the confidentiality of images. Watermarking is basically a process of injecting data into an image in such a way that it can depict the authenticity of those possessing it. The digital information hidden inside an image is imperceptible to the user but can be easily detected by a computer or various digital devices. Watermarking Technique has proved to be a powerful technique for image security and a lot of research has been made over the years to how to embed the watermarks more efficiently and to recover the watermark from the image without any distortion in the original image. These types of technique where watermark can be fully extracted from the image and along with the restoration of the cover are popularly known as Reversible Watermarking techniques. Due to the rapid development of watermarking techniques, a concluding review of recent research in this field is highly desirable. The major focus of this survey is on reversible watermarking techniques. Comparison of other reversible watermarking techniques is also tabularized.

Key Words: Watermarking Techniques, Reversible Watermarking, Histogram Modification, Tamper Localization, Image security

1 INTRODUCTION

The scope of Information Security evolves faster than the capability of computing. We can think of it as a set of techniques to protect the secrecy, integrity, and confidentiality of computer system and data against threats.

Secrecy of data means that only authorized people should be able to access the data. Integrity means that only authorized people should have the ability to use or modify the data. Availability means authorized people always have access to the data. With the advent of the Internet, Computers have become interconnected allowing us to communicate easily and instantly across the globe. The image is widely used in many applications like government scanned documents, in medical diagnosis where any loss in confidential data can lead to serious copyright issues or even reputation of the organization. Therefore it is necessary to keep the confidential information out of reach of malicious users who intend to break the security of the system. One way to ensure security is to hide a watermark inside an image. Various watermarking algorithms are being used which provides efficient embedding of the watermark. Also, the watermark can be easily recovered from the image for authentication purposes. Besides watermarking, there are two other techniques which are used to protect the digital content which is Steganography and Cryptography [1, 2, 3, 4, 5, 6].

Steganography like watermarking is used to hide the secret information in the cover image. But in steganography, if the existence of a secret message is found, it fails while watermarking aims to make any changes in secret message or information impossible.

Cryptography aims to protect the communication channel. It provides a protocol to encrypt the information in such a way that the user with the decryption key would be able to access it. It helps in preventing any changes or modifications by any intruders in the communication channel. It is achieved using hash functions and public key cipher. Many cryptographic algorithms like DES, Blowfish, RSA ensures a high level of confidentiality. A combination of the above techniques is also used for achieving high security.

1.1 Basic Concept

A digital watermark is an impression within an image which may be visible or invisible to the user. They are embedded to discern the ownership of the copyright of such image [17, 18, 19]. Digital watermarking has proved to be a promising technology for augmenting the security of images. It is used to identify forged notes in the bank and to track copyright violation. A watermark is usually a pattern of bits, an image or some information extracted from the image such as edges, corners.

There are 2 types of watermark-visible and invisible watermarks.
A visible watermark is basically a semi-transparent text or image laid over the original image while an invisible watermark is embedded in such a manner that it is imperceptible to the user but can easily be recognized by the computer or digital devices.

Watermarking schemes reported in literature [7,8,9,10,11,12,13,14] can be classified as-

a) **Robustness**: In a fragile watermarking method, any slight change in watermark will cause the watermark to break. A Semi-fragile scheme is designed to survive some changes. And a robust watermarking scheme can withstand common image processing operations.

b) **Imperceptibility**: It is the important property of watermark which is used to measure the performance of the particular watermarking technique. Imperceptibility is good if the watermarked image and the cover image is identical. Loss of image quality after watermarking indicates poor imperceptibility.

c) **Embedding Capacity**: It is the other factor which determines the performance of watermarking scheme used. It is the measure of total no of information bits that can be embedded in the cover image

### 1.1.1 Phases of Watermarking-

Watermarking process consists of 3 steps as shown in Figure-1-

a) **Creation of Watermark**: The watermarked created must be unique to the user. It is a special code or information which is inserted into the image.

b) **Watermark Embedding and Transmission**: The embedding algorithm embed the watermark into the image using different watermarking techniques such as LSB substitution, histogram shifting etc. The image is then transmitted to another person over the internet where it can be attacked by malicious users or hackers. The possible modifications can be adding of noise or cropping of image etc.

c) **Detection and Verification**: The detection algorithm are used to extract the watermark from it. And the watermark may be present if the image is unmodified. The watermark then is extracted and verified. The verification algorithm verifies the authenticity and integrity of the image and the owner.

![Figure-1: Reversible Watermarking Process](image)

### 2. RESEARCH METHOD

Honsinger et al.[20] proposed one of the first reversible watermarking schemes. It was based on the addition of modulo 256. Because the original image can be recovered directly without any loss, a digital signature of the image can be embedded in the image itself and used to verify the authenticity of the image.

Macq[21] developed a new approach by combining the patchwork algorithm and modulo addition to achieve reversibility. It involves one-to-one mapping of a digital image onto another image. Although Honsinger et al. and Macq watermarking techniques [20,21] are robust, the watermarked images suffer from salt and pepper noise which makes them not acceptable for medical imagery noise generated is mainly due to use of modulo addition technique.

Later Fridich et al.[22] proposed a reversible watermarking technique which doesn’t use any modulo additions. The idea was to embed the watermark in the cover image by compressing the least significant bits of the cover image. For the uncompressed image formats, it used the RS method which embeds the message bits in the group of pixels. For compressed images such as...
JPEG, it is based on compressing the LSBs of a selected quantized DCT coefficient from all the blocks. But the embedding capacity was found to be quite low. Fridrich et al. [28] then proposed another approach to improving the imperceptibility and embedding capacity of the image. Reversible watermarking techniques have significantly gained importance over past years and a series of techniques have been developed in this field. In recent years many watermarking algorithms have been introduced. Their major focus was to gain imperceptibility versus capacity trade-off.

Zheng et al. [23] presented a survey on rotation, scaling and translation invariant image watermarking algorithms and later reviews of different reversible techniques were also reported.

Guo [24] proposed a technique for halftone images based on reversible watermarking. It introduced a technique named embeddable cells selection (ECS) for embedding data in variable amounts and also achieving good quality.

Feng et al. [26] categorized the reversible watermarking techniques as data compression, difference expansion, and histogram shifting. Also, a single technique was also proposed and major challenges are discussed.

Pan et al. [27] classified several reversible watermarking techniques into two-additive and substitution. Another classification based on watermarking domain-spatial or frequency domain was presented by Caldelli et al. [25]. They also proposed another review according to the properties of watermark-fragile, semi-fragile and robust. Above research, works were appreciable in this field but they discuss only a few of the work. An emergent concern in the field of the reversible watermarking technique is that these techniques are not able to identify any modifications done in the watermark which are useful in medical applications. Singular value decomposition (SVD) provides a way to detect a different attacked area of the watermarked image and recovers the original image even after a high tampering rate. The SVD based reversible watermarking has recently been reported in large numbers. Also, there has been a significant increase in the applications of reversible watermarking techniques. Therefore a recent review of new reversible watermarking techniques is necessary.

In this survey paper, the newly evolved reversible watermarking techniques are discussed. We classified the reversible watermarking techniques according to Feng et al. [26]. Also, Tamper localization based watermarking is added into this classification. So our classification consists of four groups i.e. data compression based, histogram shifting based, difference expansion based and SVD based. We discussed each of these groups while our major focus is on Tamper localization based. It is because of the better performance of the Singular Value Decomposition (SVD) based technique against other techniques. It is also resistant to a number of attacks such as text insertion, copy paste attack etc.

2.1 Data Compression based reversible watermarking

To achieve reversibility we need to hold some extra information along with the watermark. This requires more time and space. One solution is to embed the watermark by compressing some part of the cover image. This approach was reported in many reversible watermarking schemes.

Yang et al. [29] proposed a technique in the transform domain. It is a bit shift based reversible watermarking technique. It achieves good embedding capacity but suffers from overflow/underflow problems.

Celik et al. [30] developed a popular compression technique. It involves quantizing the pixel values in the cover image and compressing the remainder using a codec with the watermark information. This whole data is added to the quantized image to get the watermarked image.

Xuan et al. [31] used a companding function based on integer wavelet transform to compress the values. It resulted in an increase in the capacity of data to be embedded. However, it increased the auxiliary data. To improve the capacity of this technique, Memon et al. [32] proposed a threshold optimization technique.

Arsalan et al. [33] combined Xuan et al.'s [31] companding function with a genetic algorithm to develop a robust reversible watermarking technique. It showed great performance in imperceptibility versus capacity trade-off. However, the time complexity was quite high.

2.2. Histogram Modification based reversible watermarking

Watermarking algorithms based on this approach make use of zero or minimum points of the histogram of an image and modifies the pixel values to embed data into the image. Many works in this field have been reported in the literature.

Vleeschouwer et al. [41] introduced circular interpretation based reversible watermarking. In this technique, an image is divided into no of blocks which are further divided into zones. A histogram is calculated for each zone and bin shifting is
performed. However, distortion occurs due to the shifting of the highest and lowest bin. Later an improved version of this work is also introduced by him in bijective transformation [42].

In 2006, Ni et al.[43] proposed an excellent reversible watermarking technique based on histogram shifting. From the histogram of the cover image, a pair of peak and zero points is selected and only pixel value between them is modified. Different algorithms based on this approach were developed to increase the embedding capacity. Lin et al [44] proposed an approach which utilizes histogram of difference image for embedding of data. The difference image is obtained by subtracting two adjacent pixels of the cover image. The image is divided into no of blocks and difference image is created for each block. But the above method required more space to store the extra information. By examining Vleeschouwer et al.'s[41] approach Ni et al.[43] concluded that it uses modulo addition for operations which suffer from salt and pepper noise. He proposed another method free of salt and pepper noise. Gao et al [46] method improved all the shortcomings of Ni et al.'s[43] work.

Tsai et al.[47] proposed a prediction based method in which a difference between a basic pixel and every pixel in the block is calculated. However, the method reduces the embedding capacity. Kim et al.[48] developed a method which uses spatial correlation between sub-images and embedding is done by modifying the histogram.

Kamran et al.[49] improved the performance of Gao et al.'s[46] method by introducing the concept of down sampling. It involves obtaining sub-sampled versions from down sampling and generating blocks for embedding which is done by histogram modification. A location map is also embedded in the watermarked image.

Boulahia et al.[62] proposed a histogram shifting modulation scheme that identifies parts of the image that can be watermarked using a classification process. It consists of 2 HS modulations: Pixel Histogram Shifting and Dynamic Error Pixel Histogram Shifting. However, the scheme was not much robust to simple modifications.

In 2015, Yadav and Naskar[63] proposed a reversible watermarking scheme based on histogram bin shifting by using a tamper localization approach. This approach helps in finding the tampered region which can be selectively rejected from the image thereby reducing the transmission overhead on the communication channel. It uses a minimum and maximum point of the histogram for embedding. At the receiverside, the watermark is extracted and the image is restored and a hash is computed. If the hash matches original image is accepted and in case of a mismatch, the image is rejected due to authentication failure. The scheme proposed is a good technique for optimizing the number of transmission overheads. In the future, more focus will be given on recovering the watermark from the image even after modification which helps in locating every single detail of image which is important for medical purposes. The existing schemes are good for content authentications but self-recovery of the image itself will solve a number of problems and challenges in image security. So tamper based reversible watermarking is a domain we have classified in our survey and also explains the research work in this field and challenges and issues in coping with the security of images.

2.3 Prediction Error based reversible watermarking

It utilizes prediction error to achieve reversibility. The current pixel is predicted by the correlation among its adjacent pixels. The prediction error is a special type of Difference Expansion technique. The difference expansion technique involves selecting a pair of pixels from the image and inserts information into the pixel difference of pairs by maintaining its average value. The insertion is only performed for the first two blocks. A location map is used for determining these blocks and for ensuring recovery of information from extracted data. Different category of reversible watermarking schemes was introduced by Tian[51] known as Difference Expansion. In this technique, there is no loss of data due to compression and decompression. Also, its embedding capacity is high and performs better than other techniques. Thodi et al.[52] extended this scheme by introducing prediction-error(PE) expansion. Various reversible watermarking work is reported in the literature based on prediction error.

Coltuc et al.[53] proposed contrast mapping based reversible watermarking technique. It involves the transformation of a pair of pixels. The current pixel is modified to adjust prediction error but some pixels cause overflow/ problems. To increase performance, a location map was created to identify expandable locations.

Later Lu et al.[54] proposed an enhanced version of the above method. Hong et al.[55] presented a new approach based on contrast mapping. It was a modified version of Coltuc et al.’s [53] method. Here the embedding is performed after the image is divided into blocks.

Sachnev et al [56] introduced a PE based technique by using a sorting mechanism. Dragoi et al.[55] reported an enhanced version by performing adaptive prediction of pixels that are not present in smooth regions. Wu et al. developed a hybrid approach where embedding is done in a pair of pixels and the image is divided into blocks.
approach using Sachnev et al’s [56] method. Zhou et al.[58] presented a technique based on the optimization of parameters that regulates the volume of prediction error.

A different prediction error method was introduced by Li et al[64] which is known as Pixel value ordering(PVO). The image is first divided into some blocks and PVO is performed. The maximum and minimum prediction error is calculated for each block. These values are only used for embedding. It resulted in higher PSNR and imperceptibility but the embedding capacity was low.

A different type of technique which is based on estimating the missing pixels known as Image interpolation was also reported in the literature. It uses interpolation error instead of prediction error for embedding of data. Luo et al.[59] proposed this type of method which resulted as an efficient algorithm. However, it suffered from some problems. A modification to this method was given by Abadi et al.[60]. It was based on histogram shifting of pixel values to prevent underflow during embedding and resulted in high embedding capacity. Naheed et al.[61] developed a modified approach based on particles warm optimization for estimating missing pixels.

### 2.4 Tamper localization based reversible watermarking

Tamper localization based reversible watermarking techniques is able to identify any modifications or changes done in watermark making it susceptible to tamper location-based attacks.

Dhole et al.[39] proposed a watermarking technique out of tamper recovery which provides good tamper localization but it cannot handle Vector Quantization (VQ) attack. This problem arises due to blocking the independent nature of the scheme. It is found that watermarking without block dependency can simply suffer from damages such as VQ attack. To get rid of these attacks, it is necessary to introduce a chaotic pattern in the watermarking scheme. It helps in recovering the tampered region in the watermarked image. Later many research works on block dependency were reported in the literature.

Patra et.al[38] developed a fragile watermarking technique based on the Chinese remainder theorem(CRT). It showed improved computational complexity. Block size of $8 \times 8$ was chosen to maintain tamper localization. But the results were not that accurate due to the large block size.

A method known as Singular value decomposition (SVD) was used by Abdulaziz et al.[40] to figure out any transformation in the original image. It helps in image authentication and also provides a way to detect attacked areas of the watermarked image. It is very useful in medical applications.

### 3. COMPARISON AND ANALYSIS

Table 1 presents the comparison of different reversible watermarking schemes. It is observed that these techniques are suitable for content authentication purposes. Comparison of tamper localization based reversible watermarking techniques is presented in Table 2. They are useful in the security of medical images. It can identify even tampered location in the watermark.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Type</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Celik et al[30]</td>
<td>Compression based, Fragile</td>
<td>Better Performance than Fredich et al’s scheme</td>
</tr>
<tr>
<td>Xuan et al[31]</td>
<td>Compression based, Fragile</td>
<td>Enhanced capacity vs PSNR tradeoff.</td>
</tr>
<tr>
<td>Arsalan et al[33]</td>
<td>Compression based, Fragile</td>
<td>Enhanced capacity vs PSNR tradeoff</td>
</tr>
<tr>
<td>Vleeschouwer et al[41]</td>
<td>Histogram-based, Semi-fragile</td>
<td>Suffers from salt and pepper noise</td>
</tr>
<tr>
<td>Vleeschouwer et al[42]</td>
<td>Histogram-based, Semi-fragile</td>
<td>Suffers from salt and pepper noise</td>
</tr>
<tr>
<td>Ni et al.[43]</td>
<td>Histogram-based, Robust to compression</td>
<td>Partially reversible</td>
</tr>
<tr>
<td>Gao et al.[46]</td>
<td>Histogram-based, Robust to compression</td>
<td>Increased payload, High embedding capacity</td>
</tr>
<tr>
<td>Paper</td>
<td>Block Size</td>
<td>Technique</td>
</tr>
<tr>
<td>------------------------------</td>
<td>------------</td>
<td>----------------------------------------</td>
</tr>
<tr>
<td>Walton[34]</td>
<td>N/A</td>
<td>LSB modification</td>
</tr>
<tr>
<td>Yeung and Mintzer[35]</td>
<td>4*4</td>
<td>Chaotic pattern</td>
</tr>
<tr>
<td>Shao-Hui Liu et al[36]</td>
<td>4*4</td>
<td>Chaotic pattern</td>
</tr>
<tr>
<td>Sanjay Rawat et al[37]</td>
<td>4*4</td>
<td>Chaotic pattern</td>
</tr>
<tr>
<td>Patra et al[38]</td>
<td>8*8</td>
<td>Chinese remainder theorem</td>
</tr>
<tr>
<td>Dhole et al[39]</td>
<td>4*4</td>
<td>Mean value of the block</td>
</tr>
<tr>
<td>Abdulaziz et al[40]</td>
<td>4*4</td>
<td>SVD</td>
</tr>
</tbody>
</table>

Table 2: Comparison of tamper localization based Reversible watermarking technique
4. CONCLUSION

Reversible watermarking techniques are used in wide areas of applications in recent times. We have classified reversible watermarking techniques into four groups - (1) Compression based (2) Histogram modification (3) Expansion based (4) Tamper Localisation based. A combination of the above techniques is also found in the literature. But, it has been a challenging task to identify tamper locations in watermarks. SVD based fragile watermarking with block dependency offers more security and are able to detect the attacked area inside medical images. This paper provides a complete survey of reversible watermarking techniques for image security. Also, a detailed discussion with the comparative analysis is provided. The analysis provided for various existing proposals allow various users working in this domain to select one of the proposals with respect to its merits over the others. In the future, we will develop a new tamper based technique based on the discussion and analysis provided in this paper.

REFERENCES


[40] Abdulaziz Shehab, Mohamed Elhoseny, “Secure and Robust Fragile Watermarking Scheme for Medical Images”. IEEE, 2018


BIographies

Swati Gupta, pursuing Dual Degree Integrated Course in Computer Science and Engineering from University Institute of Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal (MP) India. She is currently a student in final year doing research on Reversible Watermarking Techniques. Her research areas include Image Processing and Image Security.

Dr Raju Barskar is an Assistant Professor in Department of Computer Science & Engineering at University Institute of Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal (MP) India. He obtained B.E. from SATI Vidisha(M.P.), then M.Tech. and Ph.D in Computer Science & Engineering from Maulana Azad National Institute of Technology, Bhopal. He has more than twelve years of teaching experience. His area of interest is Network Security, Vehicular Ad hoc networks, Image Processing, Parallel, Algorithm Pattern matching algorithm and Data Mining etc. He has published more than 20 research papers in different reputed international journals and 02 chapters. He is also member of various academic societies such as IEEE etc.
Dr Shikha Agrawal is an Assistant Professor in Department of Computer Science & Engineering at University Institute of Technology, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal (MP) India. She obtained B.E., M.Tech and Ph.D in Computer Science and Engineering from Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal. She has more than fifteen years of teaching experience. Her area of interest is Artificial Intelligence, Soft Computing and Particle Swarm Optimization and Database. She has published more than 40 research papers in different reputed international journals and 10 chapters. For her outstanding research work in Information Technology, she has been rewarded as “Young Scientist” by Madhya Pradesh Council of Science and Technology, Bhopal. Her other extraordinary achievements include “ICT Rising Star of the Year Award 2015” in International Conference on Information and Communication Technology for Sustainable Development (ICT4SD-2015), Ahmedabad, India and Young ICON Award 2015 in Educational category by Dainik News Paper Patrika, Bhopal, India. She got recognition of IEEE as a senior member. She is also a member of various academic societies such as IEEE, ISTE, CSI, ACM & CSTA.