Digital Watermarking

and its applications in

Transaction Tracking

For the partial fulfillment of the course

BITS F463 - Cryptography

BITS Pilani, Pilani Campus

BITS Pilani

Department of Computer Science & Information Systems

Submitted by:

ANJAN DAS

[2015A7PS0150P]

ABSTRACT

In today's world when information flows through the web, it is crucial that we know, which content belongs to which organisation/person. With the boost in the volume of multimedia content, piracy as an issue plays a major role. And when it comes to giving credibility to the owner of a content, piracy harms the very aspect of it. In such a scenario, Digital Watermarking plays a major role, since it modifies every copy of the content which is theoretically inseparable from the embedded watermarks. This paper focusses on describing what it means by Digital Watermarking and how it is different from the other branches of its parent field, viz. Information Hiding. Digital Watermarking is compared with Steganography, which closely related with Watermarking. Transaction Tracking of multimedia is important in the digital world to identify any leak in the distribution of the content. The application of Digital Watermarking in the field of Transaction Tracking of Multimedia content is also described step by step. The work is an attempt to clarify the implementation of Digital Watermarking in the field of Transaction Tracking for a better understanding of the subject.

1. INTRODUCTION

Watermarking dates back to 13th century when people in Italy used this technique to identify papermakers. Later in the 13th century it became quite popular throughout Europe. In fact, Mona Lisa is quite an intelligent demonstration of watermarking by Leonardo da Vinci in the 15th century and was his state of the art painting of that time. However in the modern era, the term Digital Watermarking was first introduced in 1992 jointly by Charles Osborne and Andrew Tirkel. The first successful experimentation of Embedding and Watermarking happened in 1993.

The term **Digital Watermarking** describes the techniques and mechanisms to hide data, which may simple be a number or text, or in the digital world, may even be an image or a video.

On a more technical note, a **watermark** is a message that can be embedded into the digital data like video, audio, images, or text and the embedded data can be extracted later on the recipient side. Figure 1 describes the whole process.

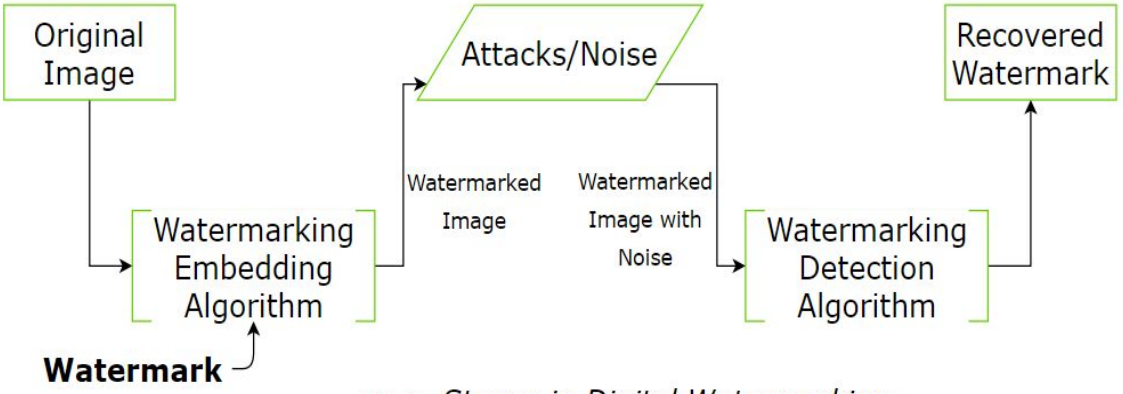

Fig 1. *Stages in Digital Watermarking*

Let us take the case of a Dollar bill. Under the daylight one can verify if a bill bears an image of the portrait of President Grant. It can be seen on all the bills after the 1996 series, except for $5

bills, on which it can be seen after the 1999 series. The watermark is embedded in the paper to the right of the portrait and is visible from both the sides of the bill. Figure 2 shows the unique portrait watermark present in a dollar bill.

Fig 2. *Watermarking in a Dollar bill*

Lately, due to rapid developments in the field of Computer Science and Technology, multimedia data i.e. audio, images and video has been associated with widespread applications. Digital watermarking is one of the best solutions to prevent Illegal Copying, Modifying and Redistributing multimedia data. To be able to achieve an absolute Copyright Protection, our Watermarking method should meet the following basic requirements:

1. **Robustness**: The watermarked data should be robust enough to resist common signal processing manipulations such as filtering, compression, filtering with compression, making it impossible for the unauthorized users to separate the watermark.
2. **Imperceptibility**: The quality of the signal content should not be affected while watermarking making it imperceptible for the human senses.
3. **Capacity**: The number of bits that can be embedded in one second of the host signal.
4. **Security**: Only an authorised person should be able to detect the watermark.
5. The original signals should not be referenced to while detecting the watermark..
6. Zero prior knowledge of the embedded watermark pattern should result in absolute non-detectability.
7. The watermark should not be embedded in the header of a signal but directly in the signal.

A trade-off is often needed amongst these requirements to obtain the best results, since they contradict with each other.

2. STEGANOGRAPHY and WATERMARKING

Both Steganography and Watermarking are the branches of the same tree, viz. *Information Hiding.*

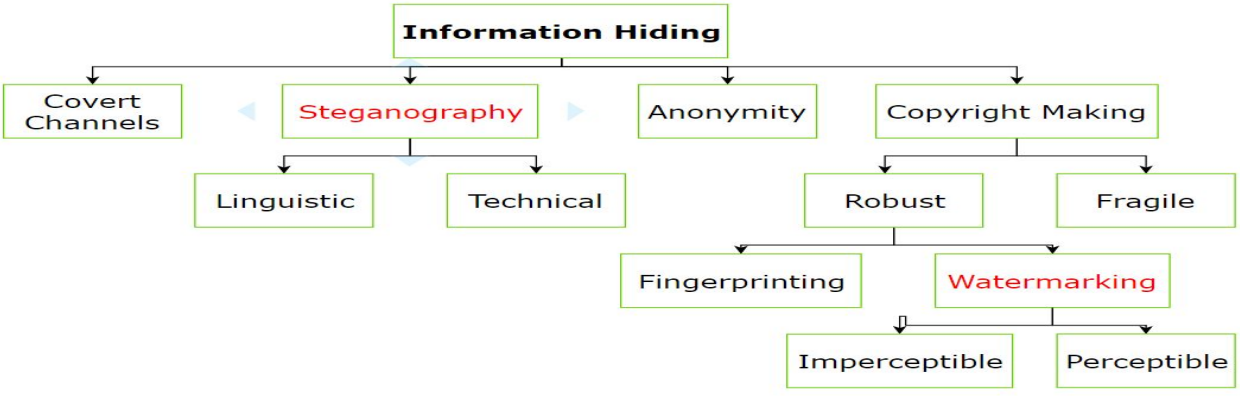

Fig 3. *Information Hiding Hierarchy*

Information hiding provides techniques to encrypt classified information so that, for any unauthorised user, it becomes undetectable. Although there are many branches of Information Hiding, there are two branches, viz. Steganography and Watermarking, which closely resemble with each other and understanding the difference between them is crucial in understanding what makes Digital Watermarking so special in the field of Digital Security. The following are the major differences between Steganography and Watermarking:

1. In steganography, we take a message **m** to hide it in a cover data **d**, to generate a new data **d'** which should be indistinguishable from **d** in practical sense by humans, so that an eavesdropper can never detect **m** in **d'**. In Watermarking, we take a message **m** to fuse it in a cover data **d**, to generate a new data **d'** which should be indistinguishable from **d** in practical sense by humans, so that an eavesdropper can never separate **m** from **d'**
2. Also, steganography is used generally in one-to-one communication systems, but watermarking is used in one-to-many communication systems.
3. In Steganography, an intruder cannot detect a message hidden in the cover data, whereas, in Watermarking, an intruder may see/detect a message but cannot remove/replace the watermark in any way, i.e., the cover data cannot be separated from the embedded watermark.
4. In watermarking, the carrier or the cover data is generally digital files such as text, images, audio or video. On the other hand, in Steganography, the carrier can be any file, protocol, service, environment employing digital representation of data.
5. Since watermarks are non-separable from the cover data in which they are embedded, so in addition to protecting content they provide many other applications also, like Copy Protection, Copyright Protection, Identity Card Security, Piracy Prevention and a lot others.

3. WATERMARKING CLASSIFICATION

Depending on a wide range of parameters, Digital Watermarking techniques can be classified into the following categories as shown in Figure 3.

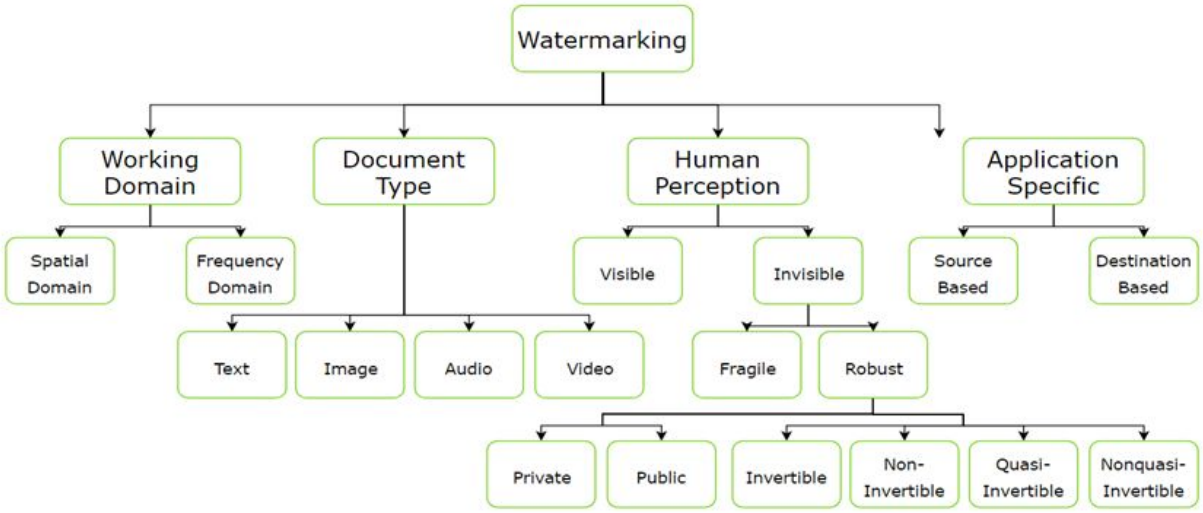

Fig 3. *Classification of Watermarking on the basis of various domains.*

4. TRANSACTION TRACKING

4.1 Why Transaction Tracking?

Previously, when videotapes were used in the entertainment industry, there was very little scope of piracy, since the quality goes down in creating copies of videotapes, each time! This would make piracy of no use as the content quality is protected. But, with the advancement of technology, media transferring tools such as Digital Versatile Discs(DVDs), Pen Drives, Blu-Ray Discs, etc came into existence. Today, even online transfer of multimedia content is so easy. To summarize, creating copies of multimedia content once you have access to it, is right at our fingertips. The same issue of piracy which was not important decades back, has become a major issue with the advancement of technology, as if it is a backdrop of technology.

Whenever a multimedia content is created, it is made available to the audience to enjoy it and understand the perspective of the creator. Today, we deal with multimedia in bulk, millions of photos, videos, gifs and what not, in the internet. Now, if we think of the studios and the people involved in the making of such content, which we enjoy and get entertained by, one thing is for sure- they should get the appraisal of the hard work they have done. When piracy comes into the picture, it causes huge loss to such studios and the people involved! Even the audience, who can get the pirated content of comparable quality right at home, why would they go for the original content and pay higher prices? This is why piracy is growing in the entertainment industry and the multimedia industry, hence.

Now, most of the piracy is done by pre-release organisations such as advertising agencies, media processing companies, review committees, television networks, etc. Recently, an episode of the famous blockbuster fantasy series- Game of Thrones was leaked by a Spanish Broadcaster, causing huge loss to HBO.

To prevent piracy, one needs to track not the end-users but the pre-market release organisations, before the multimedia content is released for mass-sellout and distribution. For this, Digital Watermarking can be used when combined with Cryptographic methods.

Here, the basis of this approach is to consider that the pre-market releasing organisations are completely unreliable and a state of mutual distrust is assumed in the beginning amongst all the organisational parties who will be involved in the transactions.

4.2 Usual Procedure of Content Distribution

In the usual Procedure, if lets say in a scenario, where A wants to pass some original content to B, but wants the content to be protected, then to avoid any kind of leak from B's side, A would just have to create a copy, watermark the copied content before sending it to B. Now, the content might also get leaked while it is en route from A to B. For that purpose, an encryption-decryption procedure is added to the process where A encrypts the watermarked content before sending it to B and B decrypts the same on the receiving end. Finally, B receives the watermarked copy of the content on the receiving end after the decryption process. But, the caveat here is that, the phase of the content after A's watermarking, but before A's Encryption and after B's decryption, are the same. This leads to a weakness in the whole process. If A has wrong intentions and A sells the copy of the content after just Watermarking it, A can just claim that the leak was caused from B's side and blame B. This would just lead to an ambiguous situation. No one can find out who caused the leak. Is it A or is it B? Hence we need to remove the weakness from the scenario. Figure 4 shows the usual procedure in the form of a flow chart.

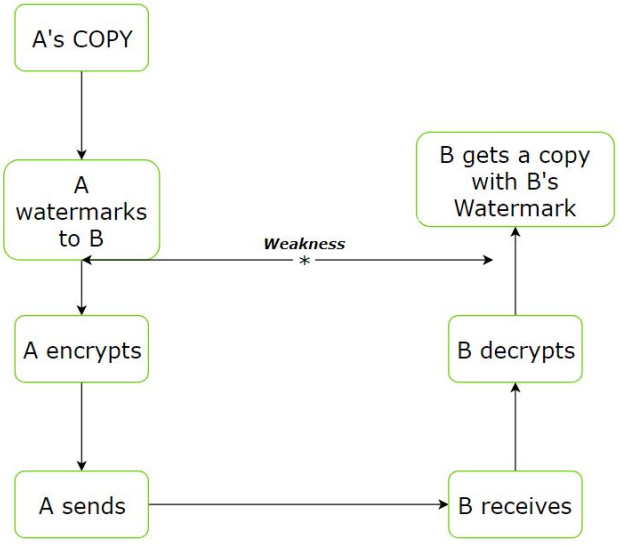

Fig 4. *The usual procedure of sending a copy*

4.3 Modified Approach

To remove the weakness, consider the scenario, where A first encrypts the copy of the content and then watermarks it, before sending it to B and on the receiving side, B finally decrypts the received content to get a copy of the content. Now, let's make the decryption procedure on B's side unique so that after the decryption, B's copy would be unique, i.e., B's copy would be distorted due to the cryptographic process. Now, the weakness has been taken care of and neither A nor B can blame the other for leaking the content. This process is called Staining. In this procedure, the sender has information from both the receiving and sending sides. In the above case, A would have it's own unique watermark and also B's public decryption key which, B sends to A for the encryption process. Figure 5 shows the modified procedure in the form of a flow chart.

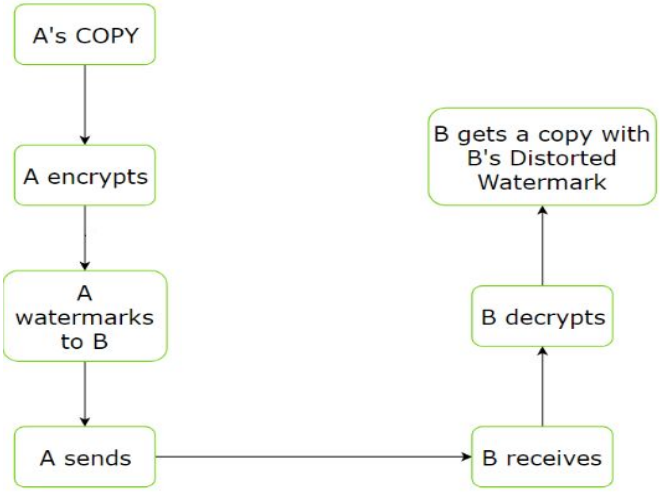

Fig 5. *The modified procedure after the removal of weakness*

4.4 Analysis

Now, we have to understand the implementation with some sample content and also test the feasibility of the procedure. FIgure 6 describes all the possessions which belong to A and B.

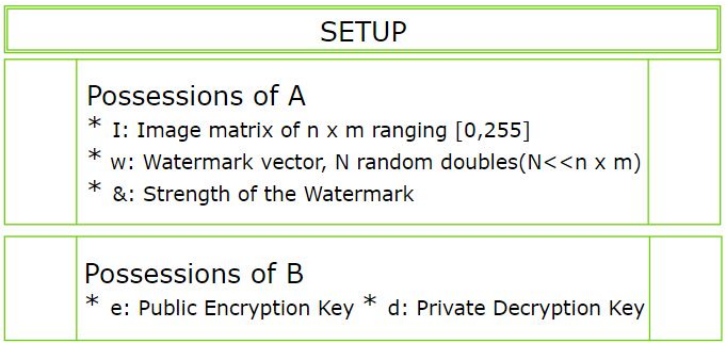

Fig 6. *Prerequisite setup of possessions of both parties*

Here **I** is the content **A** wants to send to **B**. and **w**, **N** are used as part of the algorithm, **&** signifies the Strength of the watermark. These are all the possessions of **A**. The possessions of **B** are **d**, which is the Private Decryption key and **e** which is the public Encryption key, which it sends to A before A's encryption process starts. After we have this setup, we can proceed to apply the algorithm.

The algorithm starts with B sending the Public Encryption key to A. A then uses this public key **e** to encrypt the image using modified Cox's Algorithm in Discrete Wavelet Transform domain(Discrete Wavelet Transform has been selected over Discrete Cosine Transform since it is the basis for basic image compression standards and is more reliable). After the encryption we obtain **E(I)**. For the watermarking procedure, A first applies a 3 level wavelet decomposition on E(I) and selects largest **N** values of **second level** (It is observed that the second level values gives the best results). After that, it is again altered by the watermarking procedure. And finally A applies 3 level wavelet reconstruction on the altered content to produce a watermarked-encrypted content ready to be sent to B. When B receives the watermarked-encrypted content, B then applies the standard decryption procedure using the private key **d** to produce a unique distorted watermarked content(The decryption function is so made that B always gets a unique copy of the content which is different from A's copy of the content). Figure 7 describes the whole Algorithm in the form of a flow chart.

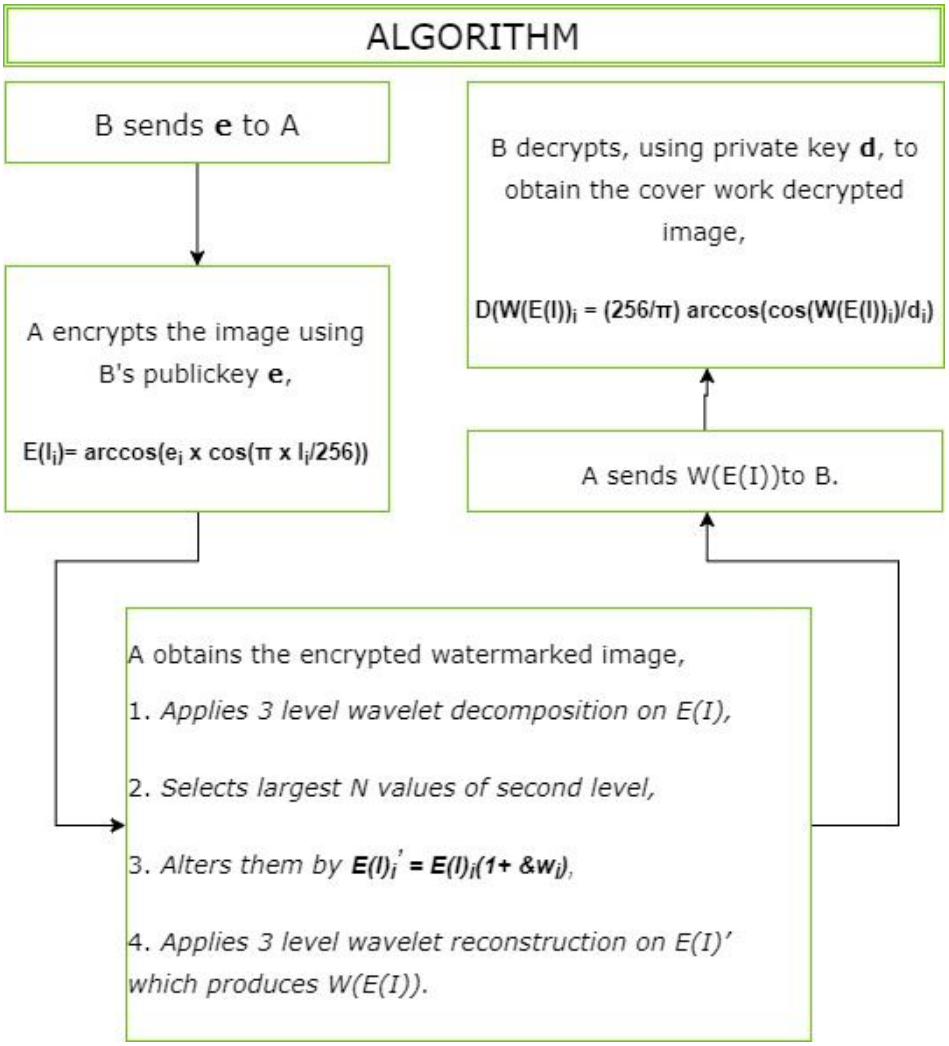

Fig 7. *Algorithm for the Watermarking-Encryption procedure*

If we take an image I, which is the image of **Lena** here, and apply the encryption-watermarking procedure we get an image which is considerably modified. Now, after the decryption process, the image D(W(E(I))) closely resembles to the original image, which signifies the success of the procedure, since the change is imperceptible to human senses. Figure 8 describes the transition of the image over the entire process.

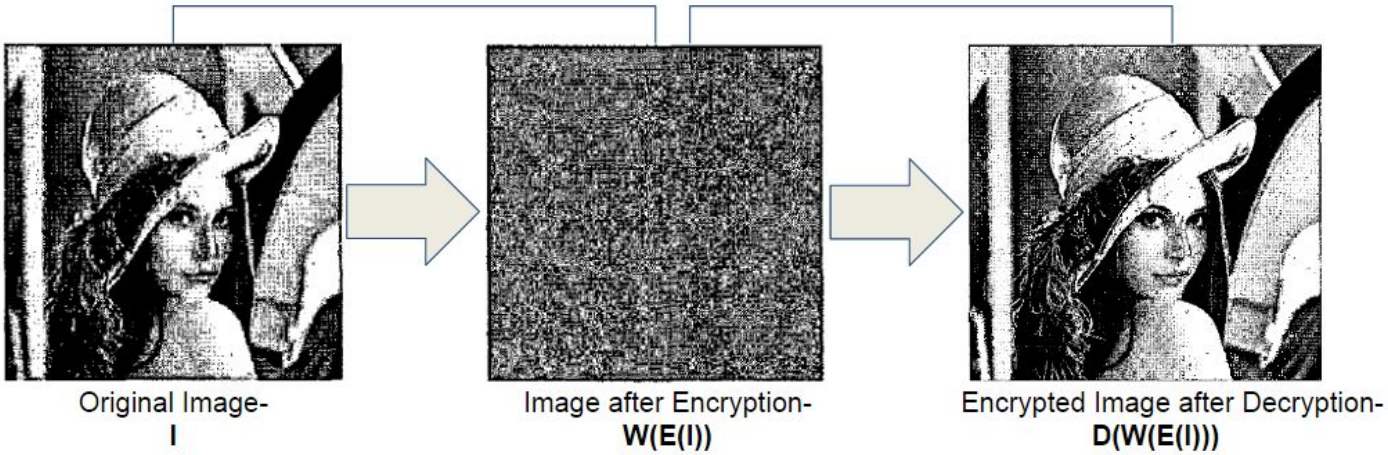

Fig 8. *Image stages during the entire procedure*

Now, if we overlay the original image with the decrypted one, and find out the residual of the difference, it is found out to be negligible and the pixel differences are scattered. This difference can further be improved by selecting more compatible Cryptosystems and watermark-embedding algorithms. Figure 9 shows the residual of the difference between the original image sent from A's side and the decrypted image from B's side.

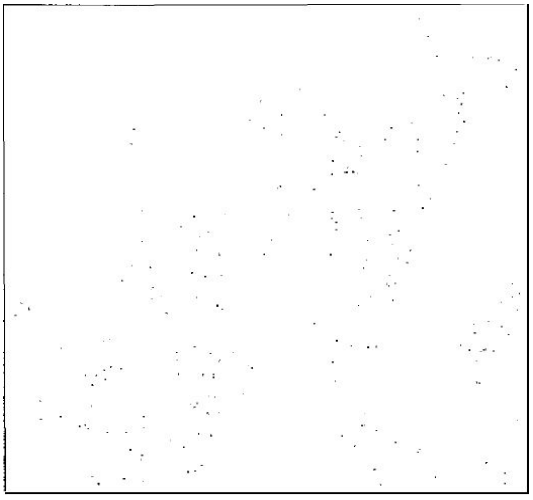

Fig 9. *Residual of the Difference between* ***I*** *and* ***D(W(E(I)))***

4.5 Requirements

The choice of Cryptosystem needs to satisfy some conditions.It needs to asymmetric to avoid any reversal of the encryption process using the public decryption key at the sender's side. It would also be required of the Cryptosystem to distort the watermark. Apart from the Cryptosystem, the algorithms also need to satiate the condition of it being linear for it might be subject to amplification if it was non-linear.

Also the watermarking embedding procedure needs to satisfy the basic requirements of watermarking mentioned in the Introductory section of this paper, viz., Robustness, Imperceptibility, Capacity, Security, Fidelity, etc.

5. CONCLUSION

Hence, we observed how Watermarking differs from Steganography and the other branches of Information Hiding. We saw how it can help publishers, content distributors, etc. protect their work from being pirated. It can also be used to trace where a copy of a film or an album has been leaked so that appropriate measures can be implemented. In the future, the described procedure would need more robust algorithms and sophisticated detection systems for better transaction tracking of multimedia.

6. QUESTIONNAIRE

1.What was the objective of your study?
Answer: The objective of my study was to understand Watermarking and its types, and how is it different from other branches of Information Hiding and also implementing the watermarking procedure in the field of Transaction Tracking.

2.What are the outcomes of your study, in terms of the significance, strength, limitations, future work, potential research problems in the area of your study?
Answer: The study was important to understand Digital Watermarking thoroughly and it provides a better understanding on how to implement it in Transaction Tracking. The compatibility issue of the Cryptosystem and the Watermarking embedding algorithms used is a challenge, since any alteration to the encrypted data tends to destroy it. Hence finding such compatible Algorithms is crucial and can be subject to research in the future.

3.How would you differentiate your work from the works already present in the literature?
Answer: The report provides a deeper insight to Digital Watermarking and how it can be applied to the field of Transaction Tracking step by step, making the reader understand both- the subject and its implementation.

4.Justify your study being comprehensive and in-depth.
Answer: The study was focussed on understanding Digital Watermarking and how it is different from the other branches of Information Hiding. I have tried to understand the subject enough to implement it in the field of Transaction Tracking. I believe that my study was thorough.

5.Highlight any taxonomy that you have proposed in your report?
Answer: [1] The Information Hiding Hierarchy
[2] The Watermarking Classification

6.Give the citations of top 5 references from the literature (it could be a book or paper) which you have consulted extensively to understand the topic.
Answer:
[1] Cox's Modified Algorithm - **A variant of the Cox algorithm for the imputation of non-response of qualitative data, Anne-Catherine Favre, Alina Matei, Yves Tillé**

[2] Discrete Wavelet Transform Algorithm - **Digital Image Watermarking Using 3 level Discrete Wavelet Transform, Pratibha Sharma, Shanti Swami, RTMNU**

[3] **Digital Watermarking with a new algorithm, Afroja Akter, Muhammad Ahsan Ullah, Chittagong University of Engineering and Technology**

[4] **Multiresolution Watermark Based on Wavelet Transform for Digital images, Vallabha VH, Cranes Software International Limited**

[5] **Digital Image Watermarking: A Formal Model, Fundamental Properties, and Possible Attacks, Hussain Nyeem, Wageeh Boles, Colin Boyd**

7.Are there any observations from your side which you think were not present in the literature at the time when you performed the study?
Answer: Yes. If we use alpha blending technique, better results can be obtained. I could not accommodate it in the study, since it requires a higher level of understanding.

8.Did you try to correlate the topic with Indian context?
Answer: Yes, the topic is applicable on any industry which involves mass distribution of content and needs to prevent piracy from happening. So it can easily be applied to the Indian entertainment industry, for example.
9.Are you OK if we upload your video presentation over YouTube so that others can be benefitted by watching it?
Answer: Yes