DIGITAL WATERMARKING

N. Srivadana [1]
[1]University of Petroleum and Energy Studies
Dehradun, India.

Abstract: ***This paper introduces an unfamiliar method of watermarking, the digital watermarking technique. Rights protection for the data which is confidential or belonging to a particular group which cannot be copied or used is watermarking process. This embeds a signal into data for copyright control. There are many algorithms and methods available which are ready and easy to implement in market. These have been successful to a level in multimedia applications, but these were unsuccessful in the issue of noise, secureness and robustness.In this paper we present the digital watermarking technique overview which approaches these problems in an unique way being highly effective and efficient against several forms of attacks that can destroy the originality of the content. It clearly describes how the process such as embedding, detection and response against the watermarks, which humans cannot identify and how triggering activity is done.***

I. INTRODUCTION

The technological development is both boon and curse for the same reasons as people can access the valuable data, easily copy and process the data and pirate it. Though several methods such as cryptography are used for reducing piracy by encryption and decryption process, it is not always possible as the data is not just in the form of words or values, but can be even in the form of images, video clips and audio tracks. Digitally represented copyrighted material such as files, images, movies, songs, and pictures offer many advantages. However, we even have an disadvantage of number of original files can be illegally produced which is a serious threat to the rights of content owners. The main way to protect content is encryption. Encryption is a process of converting the content during the data transmission from the sender to receiver.

A. Watermarking

Watermarking is an advancement of encryption and decryption. A digital watermark is a piece of information which is hidden in the content directly, which cannot be perceptible to human but can be read by the computer. The main advantage of this is that the content is inseparable from the watermark. It has several forms, such as

1. **Signatures**
   The watermark reflects the owner of the content. This information is used by an user to obtain the legal rights to publish the content from the owner.
2. **Fingerprinting Watermarks**
   This helps in identifying the buyers of the content. This helps in tracing the source of illegal copies. The only difference between fingerprinting and signatures are that, in signaturing the watermark identifies the content owner and in fingerprinting, the automated systems as computers and other visual channels identifies the content. It ensures the owner of the content that it is legally distributed.
3. **Copy control**
   The watermark contains information about the rules of usage and copying which the content owner wishes to enforce. It will have the sentences such as this content may not be copied.

Thus we approach it through other techniques such as algorithms but the noise and the robustness can spill over the efforts. Therefore the best method of protecting these content can be digital watermarking. By this technique we can ensure that the information that is encrypted into the digital format can be accessed only by the authorized people. It becomes difficult for the other parties to access this data without destroying the original data. This is almost similar to many well known approaches as watermarks which are visible , fragile, pattern matching(fingerprint based),steganography. Visible watermarking technique is quite a familiar one which does not encode the content thus making it easily read or copied by all people. The confidentiality cannot be maintained through this method. Mere visible watermarks are formed along with the content. Steganography is the other way of hidden watermarking which hides the encoded content which can be accessed only through a proper channel. These communication channels can also be retraced thus making it useless. Fragile watermarks are somehow the best when compared to the rest where the encoded data cannot be easily read, through the signatures the content can be accessed only by the authenticated parties. Other being fingerprinting which is based on the database of known content.

The most important classification that can be done is the kind of content that has to be watermarked. The data can be of any format which has normal files that contain text, images, audio content ,video content, voice clips, decoded scripts.The decoding and encoding process as well known end up with a certain key which may or may not distort the data. As well there is a possibility of errors along with limitations that some files cannot be encoded.

II. DIGITAL WATERMARKING

Here comes the master technique, the digital watermarking which functions against the deployment of the content encoded

with the ability of identifying the infringed copyright content while ensuring that the copyright material can be easily accessed by the authenticated users. This can be called as the digitized data which cannot be easily converted and processed. They can be embedded into all formats such as text, images, audios and videos. These watermarks allow the data to be self defined,letting the information in between the content. This watermark is not perceptible to human, but can be read by the computers. Machine readable watermarks are more preferred than the human readable which is not generally necessary as it allows the active marking which helps in reading the actions that can be called as self defined as well for better encoding. The watermark cannot be removed or stripped off without any loss of content. It consists of many relevant information such as flags, trigger bits, copy control information, serial numbers, some code related to the content. These digital watermarks have several applications recently in copyright protection of movies and music, in FBI. These have serial numbers as requirements that help in making watermarks. They have bounds while preserving the data from several attacks such as alternations, subset selection. We have several improvements and alternative techniques validated with several experimental analysis that can withstand loss of data, alterations and subset selection kind of problems.

III. CLASSIFICATION OF DIGITAL WATERMARKING

The process of digital watermarking consists of three main activities which are interdependent. They are Embedding, Detection and Response. These all activities again depend on identification which is the main step. In Embedding process the watermark is embedded into the content in various methods and at different levels such as at the point of distribution or during the time of preparation of the content and so on. As the embedding process completes, detection of the watermark starts along with the processes of routing and caching of flow of work. The location of the watermarking can be detected in this process. As the position is detected the several responses such as triggering, data hash evaluation and identification happens.

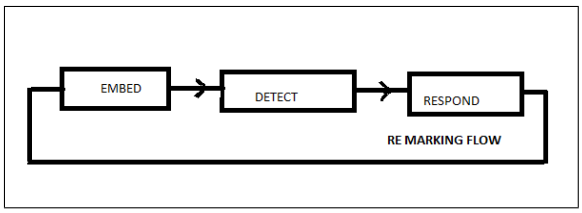

Figure 1: Activities of digital watermarking

The watermarking system has an algorithm which can be read as well written and its specific hardware components in the system. It even includes a database for storing of the information about the watermarks written and the watermarks that are read. The below is the block diagram representation of the process and its components. Classical cryptographic tools are used for checking data integrity of the data. In order to ensure the authenticity and integrity of the data, algorithms are designed which consider the special properties of such data types such as visual appearance or perceptual content. Thus an algorithm made should be robust against compression, format conversion and such other manipulations .The algorithm should be even able to handle the large number of data placed.

The digital watermarking can be classified into the below :

Case 1: Visible watermarking

Here the data is made available through the internet and the owner will be knowing that it is available to all the users commercially, without any payments. If the owner desires that the content can just be visually available but cannot be used in any purposes in any form. He can ensure that it can be done with the help of visible watermarking. There are two purposes even in this. This helps in collecting the revenue, other in watermarking the content so that it cannot be used even after paying.

Case 2: Invisible watermarking

Here the data is duplicated so that after inclusion the originality of the data can be verified with the duplicated copy. Invisible watermarks embedding is done during the preparation and distribution of the original content. It helps to detect the alteration of the data stored in a database. This will ensure that the data is neither altered nor replaced since it enters into an exposed networking as internet. It has several advantages over the visible watermarking as to detect the misappropriated files and even as an evidence of ownership. It can even determine the identity of misappropriator. These properties makes it robust. It is generally desired that the watermarks are sensitive.

IV. PROCESS OF DIGITAL WATERMARKING

The process of watermarking is clearly explained as the data is generally represented either in pixels, wavelets, transforms, fourier and other components which make the original data transformed. As the representation is done into transformed form, a subset of data must be marked. The choice of the subset data is made by checking the security. The data processing and compression techniques are also taken into regard. In this process of choosing components for a subset several parameters are considered, as quality, intensity, brightness for an image file. These components are randomly chosen and are scaled or cropped to different levels and combine them with the original values and form an encoded data. In general the watermark value is multiplied to the strength parameters which are accepted across the world. Thus by this we replace the original content values with the strengthened watermark value. There are several process such as multiplying factor, or by addition, by logarithmic transformations that we can obtain this strengthened watermark values. We should simultaneously see that the watermark reader must match with the writer. The

transformations represented by the above said forms should be decodable by the reader, and should be presentable by replacing the duplicated parameters with original values. There can be many ways to decode the duplicated values as hamming codes, error detection codes or with the help of correlation. There are even more effective methods such as signal detection test (which is limited to hypothetical basis),CLKS method which indirectly uses the correlation method for watermarking. It adds to the point of robustness along with securing the data along with the biggest advantage of reducing the noise which is generally caused by the original data when watermarking is being read. For obtaining this we use the noise reduction techniques applied along with the methods of subtracting the watermarked values from original and by standard correlation and signal detection techniques.

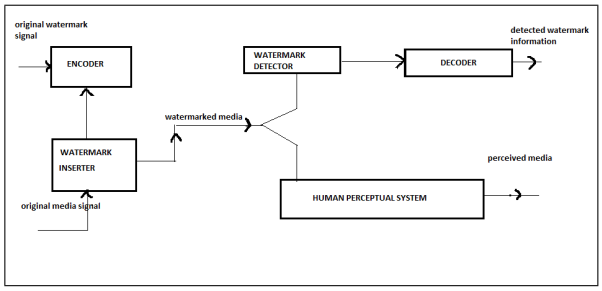

Figure 2: The process of digital watermarking and the signal flow diagram

Several other processes during the digital watermarking are explained briefly below :

A. NOISE ADDITION

Generally it is misunderstood that a watermark of small amplitude can be removed by addition of random noise of equal amplitude, but the noise appears to be robust after the addition of random noise. In being practical, this is not be considered as a serious issue until the noise is compared with the watermark.

B. SPATIAL FILTERING

Filtering of an image can influence the properties such as luminance, contrast and brightness. Thus the redundancy is controlled.

C. COLLUSION ATTACKS

It can be described that the signal versions have all different watermarks, each corresponding to a version. The intersection of two signals can be easily approximated as the region of signals is well known. The intersection point of two signals is equivalent to the watermark region. This gives rise to the collusion attacks. It can be described as the intersection of several watermarked signals combining to form an unwatermarked signal. The attacker can thus easily access the detection region, with the key watermark. The hacker can easily use any of the multiple watermarked versions of the signal to obtain closer approximations of the original region, which is originally watermarked .Resistance to the collusion attacks is also a varied function of the structure of the watermark.

V. BEHAVIOR OF DIGITAL WATERMARKING

The special features or traits can be mentioned as robustness, security ,fragility, false positive rate, cost computation. The Watermark is expected to resist the common distortions, malicious attacks, and can coexist with other watermarks and carries other information regarding the signal and computational details as inserting and detecting key. The robustness depends upon two key factors

a) whether the watermark is present in the data or not even after distortion.

b) can the watermark detector detect it.

Geometric distortions like cropping and scaling the signals inserted into images with the help of algorithms, the detection algorithms can detect the watermark if the distortion is also removed.

A. ROBUSTNESS

This can be achieved by two types of processing techniques which are most commonly implemented. They being alignment preserving transformation and alignment altering transformation. Data compression, the common conversion methods such as ADC and DAC can be observed in the preserving technique. cropping ,rescaling with a factor, rotation and zooming are some of the methods observed with the altering techniques. There are several algorithms in both the types, which depend upon the frequency for representation of the original data with the watermarked values. Then these data are compressed and protected. Several algorithms allow reading the watermarked values from compressed data with the online based systems. some watermarks does not even allow decompression which has to be checked empirically before the compression takes place. When compared among both the techniques, the altering transformation is regarded to be the difficult one as the data compression technique is difficult after cropping ,scaling and rescaling and other processes, but with an added advantage of making the detection process simpler by increasing the SNR ratio of the watermark extremely high. Several algorithms proved that the robustness can be attained if the watermark is placed in perceptually regions of signals because, in a distorted signal its fidelity is only preserved as the perpectual regions remain intact with no effect on fidelity.

B. SECURITY

This is the main feature for which several algorithms have been developed. The main concern while seeking for security would be the cryptography. There are several cryptographic

methods with which an assurance can be made that the watermarks cannot be read by the third parties. But still there is a risk of an individual claiming rights creating his or her own watermarks over the CKLS algorithm by subtracting the original watermarks. The forgery attacks can be in different manners as above said, then comes into play the method of cryptographic techniques which safeguard against such. By randomizing the LSB (least significant bit) which contains the watermark or by making all the bits to zero, the watermarks can be made unreadable by human. The watermarks which are scaled at a higher extent are difficult to erase. By copying a set of rows and columns, by zooming or rescaling when the factor is known the CKLS marks can be readable. In such cases we can register to the original images or by using a watermark that can withstand against such transformations. These attacks can be commonly observed in forgery data.

The watermarks readers and writers which are available online publically which can be accessed by everyone creates the issue. This is a concern of risk though the algorithms which hide the process, cannot easily withstand the forgery. There was a proposed method of watermarking which uses key dependent basis functions which allows the construction of a secure public reader. The collusion type of attack, in which several different marked copies of same data or different data which can be marked in same way. This is a potent kind of data attack. There are correlative reader watermarks that can use component wise distribution or uniform way of distribution, which are at risk. The resistance to such collusive attacks is focusing on the several models of security, which is theoretical having several assumptions for modeling. There is a proposed model which use randomly selected watermarks which are resistant to collusive attacks. There are several famous models resistance to collusive attacks such as chor,et al, Boneh- shaw which is based on multiple key protocol and defense against k-way collision conveying some bits in the watermark size respectively. The shaw model is a combinational work that reduces collusion resistance, but ignores the issues of security. The watermark consists of n positions with each position of different sizes of alphabets. The watermarker identifies the information contained in the positions in which all the watermarks are same, though it does not answer to the question how to provide individual marking positions as robust nor can consider the possible ways of marked positions, which is difficult to detect and to spoil that makes the method more suitable for representing the watermarking perceptual content more better. There exists another model by Leighton, which assumes a model that the original data is of an n-dimensional vector. It has components which are independently distributed where the mean is considered as zero and variance as 1.It works on the Euclidian distance basis and the perceptual distance is a correlative original based watermark reader. The k-way collusive attacks can be resisted with high probability with the help of this model. Further there were many models as of Ergun, Kilian, Kumar, Mitchell, Tarjan and Zane with different algorithms which were highly resistant to collusive attacks and protect the erasing of watermarks with higher probabilities. All the theories thereafter where related to each other but independently developed algorithms. They came with a strategy of using pseudo random sequence bits as watermarks rather than the Gaussian noise. There were even models which attempted with random watermark directional property. The direction is hidden where the strength of the watermark was determined by the perceptual model. Small signaling patterns in combinational or distributed patterns as the components of the subsets was an efficient manner .Thus several models came into existence with different strategies to increase the efficiency of the reader.

Attacks can be prevented by employing key dependency but the parameters which attain a key role in data compression robustness which is compensated by decrease of compression which ofcourse reduces security.

C. FRAGILITY

Fragility, the opposite of robustness is one of the most important feature observed. For instance the watermark on bank notes, they do not survive any kind of copying or forging, and thus indicates authenticity. this property of watermarks is known to be fragility. Fragile watermarks are difficult to prepare when compared to preparation of the robust ones. Integrity cannot be achieved with the ones that are tampered and modified in other forms due to any sort of distortion. Thereby we prefer tamper resistant. The successful attack on watermark could be called for when watermark is removed from the signal. Without any change in the perceptual region or quality of the signal.

D. KEY RESTRICTIONS

Key restrictions are considered as another distinguishing characteristic. It is the level of restriction placed on the ability to read a watermark. Unrestricted key watermarks are those in which the key is available to a large number of detectors. Restricted, as name implies those watermarks, in which key is kept as secret to detectors. The difference being the usage and their respective algorithms.

E. FALSE POSITIVE RATE

False positive rate is the most important consideration to be noted as in many applications, it is necessary to examine if the data contains watermark or not. It is the probability of the detection system that will identify a watermarked data from an unwatermarked one. This factor does not allow to copy or play the content if once it is detected that the data is watermarked. False positive causes errors, which is one in a million frames. This is the most important criterion that the companies designing the watermarks would look for.

F. MULTIPLE WATERMARKS

Modification and multiple watermarks can be although considered as an important feature, it does not have so much of usage. If after the watermark any insertion has to be done, this method comes into play. Changing an watermark is done after removing the first watermark which is made to restrict the

copying of the content and then adding a new watermark or by the other process in which inserting another watermark such that both the first and the second are readable, but making sure that both override each other. The second method is widely accepted and implemented.

G. DATA PAYLOAD

It is the amount of information. The data stored (expressed usually in bits is considered as same number of bits that can be inserted into the signal).

H. COMPUTATIONAL COST

Computational cost are the costs involved in inserting and detecting the watermarks. The speed and the cost requirements may vary each and every application. The hardware required can be expensive for few applications as well can be simple chips making it inexpensive in other case.

VI. RESULTS

This paper has given an overview of how digital watermarks have the capability of leading over the existing watermark technologies. This paper describes principles of watermarking with several characteristics which are desirable for all of its applications. Though it is a new entry into the field, it combines several disciplines as cryptography, encoding-decoding, probability, signal processing and so on. This makes the process secured and robust. The digital watermark technology is not just useful in protecting the content, but also helps in including the mechanisms to detect and to respond for the rights holder efficiently besides the several existing approaches such as fingerprinting. The use of digital watermarks embarks an monitoring and broadcasting content rights holder and distributors for unique identifier. An example of digital watermarks used now-a-days can be a detector infrastructure, a monitoring station. The embedded watermark can be detected online. Thus when inappropriate use of the information is detected a report will be sent to the content owner, which can be helpful in properly licensing the information. This action will lead to beneficiary in providing the rights as well increase the revenue for distribution thereby providing a better and efficient technology. The signal processing operations, resistance to tampering are difficult to achieve. With the help of all such operations the content is made available to open source networks as watermarking acts as a potential solution. Though misappropriation of the data can be reduced with set of other techniques such as multiple watermarking , cryptography, this can serve the diverse problems protecting the content in a better way.

VII. REFERENCES

[1] R. Anderson, Ed., Information Hiding, First International Workshop Proceedings, Lecture Notes in Computer Science 1174, Springer-Verlag, Berlin,1996.
[2] Preliminary Proceedings of the Second International Information Hiding Workshop, 1998
[3] D. Boneh and J. Shaw, "Collusion-secure Fingerprinting for digital data," Crypto '95, Lecture Notes in Computer Science 963, Springer-Verlag, Berlin, 1995, pp. 452465.
[4] G. W. Braudaway, "Protecting publically-available images with an invisible image watermark," Proc.IEEE Int. Conf. on Image Procesing, ICIP-97(1997), Vol. I, pp. 524-51.
[5]I. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multi- media," IEEE Trans. on Image Processing 6 (1997), 1673-1687.
[6] B. Chor, A. Fiat, and M. Naor, "Tracing traitors," Crypto '94, Lecture Notes in Computer Science,963, Springer-Verlag, Berlin, 1995, pp.452-465.
[7] J. Fridrich, A.C. Baldoza, and R.J. Simard, "Robust digital watermarking based on key- dependent basis functions," Preliminary Proc. Second International Information Hiding Workshop (1998).
[8] C. Podilchuck and W. Zeng, "Digital image watermarking using visual models," IS&T/SPIE Electronic Imaging 3016 (1997).
[9] R. van Schyndel, A. Tirkel, and C. Osborne, "A digital watermark," Proceedings 1994 IEEE Inter- national Conference on Image Processing, (1994), pp.86-90.

