DESIDOC Journal of Library & Information Technology, Vol. 29, No. 3, May 2009, pp. 29-35

Safeguarding the Digital Contents: Digital Watermarking

M. Natarajan and Gayas Makhdumi[1]

NISCAIR, 14, Satsang Vihar Marg, New Delhi-110 067
E-mail: m_natarajan@hotmail.com

[1]Department of Library & Information Science, Jamia Millia Islamia University, New Delhi-110 025.
E-mail: gayas_makhdumi@yahoo.co.in

ABSTRACT

Digital watermarks are one of the tools which helps to make the distribution of digital material more secure. It is a kind of digital signal or pattern hidden directly in digital content. The paper deals with the multi-faceted aspects of digital watermarking (DWM) technology. It also discusses the need for DWM with the properties like robustness, security, invertibility, transparency, complexity, capacity, and verification. The key aspects are given with the solutions for DWM. Few companies involved in DWM activities are discussed with the technique of embedding with different types of watermarks. It discusses tool and techniques for images, text, and other applications of DWM. The possible attacks on DWM and the organisations involved in developing standards for it are described. It has been concluded that the protection of individual rights is a must, and deeper understanding of the DWM will lead to the design of more reliable systems for safe-guarding digital contents.

Keywords: Digital water marking, DWM, safeguarding, digital library, watermarking

1. INTRODUCTION

The popularity of World Wide Web demonstrated the commercial potential of offering multimedia resources through the digital networks. Since commercial interests seek to use the digital network to offer digital media for profit, they have a strong interest in protecting their ownership rights. The digital data can be processed, accessed, and it can be transmitted very quickly using networks. There are numerous technical, legal, and organisational problems which arise when there is wide-scale use of digital documents. Digital information can be copied any number of times from one medium to another; they can be transmitted through networks, etc., all without compromising the quality of the data. There is no way to distinguish between an original electronic documents and its copy. It is easy to change any part of an unprotected electronic document. One possibility here is to replace original signatures with cryptographic methods.

Digital signature is data items formed by the signatory and created from the document that is to be signed. It relates the documents to the signatory in a secure and reliable way. The signature of one document cannot be used to sign another, even if the two documents in question differ by just a single character. Digital watermarking has been proposed as one way to accomplish this. Also advanced Internet services enabled the users to create copy and distribute multimedia products such as audio, video, and still images with much ease and less effort, minimum or no cost, and in less time. Though it encouraged trading on the Internet, but on the other hand it has created the problem of illegal copying or copyright infringement. E-commerce has become a significant business with well-established online shopping services, and online delivery of digital media such as audio and video. Thus, protection of digital rights assumed a primary importance in the digital age.

2. PROTECTION TECHNIQUES

Significant research has been conducted since 1996 on copyright protection and authentication. Steganography, cryptography, digital watermarking, etc., have become the most relevant and significant topics for research scientists and vendor consortiums. Because of the massive illegal copying of media files, content owners

Received on 25 November 2008

utilise various protection technologies like digital rights management (DRM) and/or digital watermarking. DRM is one of the way to protect and secure the rights of the author (owner) on the contents of the product. DRM links specific user rights to media in order to provide persistent governance of user activities such as viewing, accessing, and duplication. The management technique is aimed at balancing information protection, usability and cost to provide beneficial environment for all parties concerned. It is achieved through the interaction of effective and economic models, social values, legal policy, and technology.

Copyright protection inserts authentication data such as owner registration, buyer and seller information and logo in the digital media without sacrificing quality. In case of any dispute, the data of authentication is extracted from the product and can be used as an authoritative proof to prove the ownership or copyrights. Copyright infringement includes the user activities beyond the fair use of digital media. DRM is regarded as an extension of copy protection measures[1].

3. DEFINITIONS

3.1 Digital Libraries

- A Digital Library (DL) is a collection of services and "information objects" that are available digitally. "Information objects" can be defined as anything in a digital format such as books, journal articles and sounds, since DLs organise and present information objects to users, and support them in dealing with these objects.

- The DL is the widely accepted term describing the use of digital technologies to acquire, store, preserve and provide access to information and material originally published in digital form or digitised from existing print, audio-visual and other forms.

- The above definitions reveal that a DL is a source of information, in different formats, e.g., text, video or audio, and that such information is stored digitally. Thus, the definition of DL is "a DL is a networked repository of digital content".

3.2 Digital Watermarking

- A Digital Watermarking (DWM) is a form of steganography in which copyright and other source information is hidden inside a document, image or sound file without the user's knowledge, but copies will retain the information. It can prove authorship and track copies to the original owner.

- A DWM is a digital signal or a pattern hidden directly in digital content. The digital content known as host data or cover could be any multimedia product such as still images, audio data, video clip, or text document. Watermark can be any useful information to prove the authenticity of the owner. The host and watermark are never separated so that the embedding watermark is the key point in this DWM. Once embedded, the watermark is invisible in the host. The original content is called the cover or the host and embedded message is known as watermark and the resulting cover after watermarking is called watermarked cover[18].

- The watermarking of the document involves the transformation of the original into another form. It is contrasted with public-key encryption, which also transform original files into another form. It is a common practice now-a-days to encrypt digital documents so that they become un-viewable without the decryption. Copyright infringements of digital audio, images, videos, and 3-D models can be detected by means of DWMs. They are invisible signatures that are embedded in a document in such a way that they are resistant to unintentional or malicious modifications of document: only the copyright holder owns the key that permits the decoding and/or removal of the watermark [11].

Therefore the watermarking is a process that embeds data into an object.

4. DIGITAL WATERMARKING

4.1 Need for Digital Watermarking

Digital watermarks are one of the various strategies which should help to make the distribution of digital material more secure. A distinction can be made between active and passive strategies:

(i) Active methods, such as cryptography, directly prevent unauthorized distribution of data; and

(ii) Passive methods, such as digital watermarks, serve more as a method to provide proof of ownership rights.

Unauthorised taping, reading, manipulating, or removing of data might lead to financial loss or legal problems for producers and creators. Thus, designers, producers, and publishers of digital data such as images, videos, audio sources, or multimedia material (for example, games or virtual environments) need technical solutions to deal with the problems associated with copyright protection of their data. They require systems in which digital data can be easily signed by authors/ producers to ensure and prove ownership rights on the produced audio-visual material after it has been publicly released[9]. DWM embeds secure invisible or inaudible

labels in multimedia data (such as images, audio, text, video, 3-D graphics) for identifying copyright-related information such as origin, ownership, use-control, integrity, or destinations. The DWM is integrated with the multimedia and tightly bound with the quality of the content. The benefits of DWM for content protection is twofold, viz., it provides evidence of illicit copying after the event and discourages such misuse in advance[10].

4.2 Properties of Digital Watermarking

- Robustness describes whether the watermark can be reliably detected after media operations, format conversion, rotation, scaling or cropping. It means resistance to build, non-targeted modifications or common media operations.
- Security describes whether the embedded watermarking information can be removed beyond reliable detection by targeted attacks. All watermarks except annotation should provide high security.
- Invertibility describes the possibility of extracting the watermark after embedding it to restore the origin. For example, integrity watermarks often need to verify a file's integrity and restore the original for medical images.
- Transparency relates to the properties of the human sensory system. A transparent watermark causes no artifacts or quality loss.
- Complexity describes the effort and time needed to embed and retrieve a watermark. This parameter is essential if we have real-time applications. Another aspect addresses whether the original data is required in the retrieval process.
- Capacity describes how many information bits can be embedded. It also addresses the possibility of embedding multiple watermarks in one document in parallel.
- Verification describes whether it has a private verification, like private key functions, or a public verification, like the public-key algorithms in cryptography[16].

4.3 Key Aapects of Digital Watermarking

(i) Images, videos (raw or compressed) and 3-D models can be marked.

(ii) The original document is not necessary to decode and read the watermark. Only the copyright holder's key and the watermarked document are required.

(iii) The watermarks are resistant to multiple watermarking, as well as to a variety of image manipulations, such as JPEG and MPEG compression, printing-rescanning, rotation, scaling, cropping, filtering, color changes, etc.. in case of videos, the watermarks are resistant to format conversion, frame-rate changes, etc.

(iv) A time-dependent identification is supported.

(v) The process is content adaptive to prevent any visible artifacts.

(vi) The embedding/detection is based on a cryptographic public key scheme. A third party may, therefore verify in a legal dispute who has embedded the watermark[12].

4.4 Solutions for Digital Watermarking

DWM technology builds on the idea of hiding meta-information-like an owner's signature-in physical material in which content is represented, such as the pixels of an image. This concept resembles watermarks in paper. In the digital form, one can extract the embedded information when necessary to show proof of ownership. It is a multidisciplinary field that combines media and signal processing with cryptography, communication theory, coding theory, signal compression and the theory of human perception.

Interest in this field has recently increased because of the wide spectrum of applications it addresses, such as in the Digital Versatile Disc (DVD) scenario or in the Secure Digital Music Initiative (SDMI) for copy control. An application-based methodology, can generally distinguish among the following classes:

Authentication or copyright watermark: watermarks the data with the owner's or producer's identification. This is the enabling technology to prove ownership on copyrighted material, to monitor the use of the copyrighted multimedia data and to analyze where the data is in use over networks and servers.

Fingerprint watermark: watermarks the data with customer identifications to track legal or illegal copies and detect the originator of illegally made copies.

Copy control or broadcast watermark: ensures copyrights with customer rights protocols such as copy or receipt control.

Annotation watermark: annotates the media data. This kind of watermark can be used to embed descriptions of the value, content of the data or dates.

Integrity watermark: ensures the data's integrity and recognizes manipulation. The advantage of all these methods is that relevant copyright information becomes part of the audio-visual object[2].

5. COMPANIES INVOLVED IN DIGITAL WATERMARKING

- *Digital Copyright Technologies (DCT) Ltd,* a start-up company based in Zurich Switzerland, is one of the technological leading companies in the copyright protection and secure distribution of digital multimedia data. Their technology is for multimedia service providers (image or video archives) and for small multimedia companies (photographers, graphic designers)[13].

- *Nextamp* was founded in 2002 as a spin-off from the Thales group (formerly Thomson – CSF), which started developing a video watermarking technology in the mid 90's. Nextamp developers and markets cutting edge video watermarking solutions for the broadcast industry. Nextamp's open products are designed to ensure graceful integration into existing systems. Nextamp solutions are designed for applications in production, post-production, asset management, distribution and content tracking[15].

- *Signum Technologies* solutions range from simple-to-use plug-ins for use with graphics programs such as Adobe Photoshop – for those users who work with small numbers of images, through to high-productivity batch processors – for those with many thousands. Signum Technologies provides a range of advanced digital watermarking solutions for copyright protection (SureSign) and communication applications embracing a wide range of business, professional and institutional sectors. Choose an application closest to your area of specialisation to learn more about how Signum can make your business more secure. SureSign provides protection for many organisations who market high value visual content via e-commerce solutions. The permanent nature of embedded watermarks facilitates the monitoring of published images—both in print and in digital media—by linking traded images to stored rights information. Users include museums, stock photo libraries, broadcasters, news agencies, photographers, illustrators, catalogue retailers and publishers. For partner companies providing SureSign watermarking functions within e-commerce and digital rights management (DRM) solutions please see Signum's partner section below[17].

- *Microsoft Corporation* has been awarded a patent for DWM that could be deployed even when digital music is distributed with DRM protection. The technology called "stealthy audio watermarking", inserts and detects watermarks in audio signals that can identify the content producer, "providing a signature that is embedded in the audio signal and cannot be removed", according to a filing with US Patent and Trade Organisation (US Patent No. 7,266,697).

There are many more companies working on DWM. The above four are only as a sample.

6. WATERMARK EMBEDDING

The concept of DWM consists of inserting information into the host signal under the condition that the modifications are not perceptible. In addition, it is desirable to put maximum energy into the watermark to achieve high robustness. This is a well known concept from communication theory: To decrease the error rate, the signal energy must be maximized. In mathematical formulation, the watermark embedding process can be considered as a constrained maximisation problem: Maximise the watermark energy under the visibility constraint. Although the problem is straightforward to formulate, it is extremely difficult to implement because of the visibility constraint, which is usually based on a highly nonlinear model of the human visual system.

Watermark embedding can be performed in a variety of ways. There are two main groups of watermark embedding technologies: coefficient-based and system-based. Coefficient-based approaches are the most obvious approaches since the embedding process is performed by a direct modification of pixel values or transform coefficient values. Examples of this group are approaches based on pixel modifications in the spatial domain, such as least significant bit watermarking where the least significant bit of the pixel values are replaced by the binary watermark values. The second group is less obvious to understand because the watermark embedding process is performed by slightly changing an existing processing system. Example of this group is fractal image watermarking[5].

7. TECHNIQUES FOR WATERMARKING

Watermarks were first used in Europe to identify the guild that manufactured paper. They were like trademarks or signatures and invisible. A watermark image becomes visible as darker and lighter areas when the paper is held up to the light. Digital watermarks for photographs work differently than those used for paper. With images widely available on the Internet, it is desirable to use watermarks. A watermark is a secondary image which is overlaid on the primary images and provides a means of protecting the image. There are two types of watermarks, viz., visible and invisible watermarks.

A visible watermark is a visible transcription image which is overlaid on the primary image. Perhaps consisting of the logo or seal of the organisation which holds the right to the primary image, it allows the primary image to be viewed, but still marks it clearly, as a property of the owning organization. It is important to overlay the watermark in a way which makes it difficult to remove, if the goal of indicating property rights is to be achieved. An

invisible watermark is an overlaid image which cannot be seen, but which can be deducted algorithmically. They are hidden in the image and can survive image cropping and file format changes. They are almost indestructible. However, with the reader, one can display them and who created the photograph and how to get in touch with them. This is like free advertising. If someone sees your image and wants to contact you about reusing it, they can easily do so. To be effective in the protection of the ownership of intellectual property, the invisibly watermarked document should satisfy several criteria such as follows:

- The watermark must be difficult or impossible to remove, at least without visibly degrading the original image.
- It must survive image modifications that are common to typical image-processing applications (e.g. scaling, color requantisation, dithering, cropping, compression).
- It should be imperceptible so as not to affect the experience of viewing the image and
- For some applications, it should be readily detectable by the proper authorities, even if imperceptible to the average observer [3].

Watermarking techniques tend to divide into two categories, text and image, according to the type of document to be watermarked. First, watermarking allows an image file to be"tagged" by embedding a unique and permanent message into the object data. Once applied, the embedded message cannot be modified. Such a message can be used for various purposes eg. to indicate date/time of creation, ownership, origin, identity of content or revision status. Second, by applying a sophisticated checksum 'pattern' into digital image data, any subsequent change to the structure can be detected and the site of the alteration within the document pinpointed. This is done by calculating a 'signature' or 'digest' of the values of the object data, then embedding this back into the object data using algorithms. Similarly, confirmation of data integrity will be displayed or flagged where the object file remains intact and unchanged. This form of watermark is sometimes called 'fragile'[1].

7.1 Techniques for Images

The simplest (too simple for many applications) is to just flip the lowest–order bit of chosen pixels in a grey scale noisy modification. A more robust watermark can be embedded in an image in the same way that a watermark is added to paper. Such techniques may superimpose a watermark symbol over an area of the picture and then add some fixed intensity value for the watermark to the varied pixel values of the image. The resulting watermark may be visible or invisible depending upon the value (large or small, respectively) of the watermark intensity. One disadvantage of spatial domain watermarks is that picture cropping can be used to eliminate the watermark. Spatial watermarking can also be applied using color separation.

7.2 Techniques for Text Images

Three types of coding techniques can be used text images such as: i) text line coding ii) word space coding and iii) character encoding. For text line coding, the text lines of a document page are shifted imperceptibly up or down. For a 40-line text page, for instance, this yields 2**40 possible code words. For word-shift coding the spacing between words in a line of justified text is altered. For character coding, a feature such as the end line at the top of a letter, "t" is imperceptibly extended. An advantage of these methods over those applied to picture images is that, by combining two or three of these to one document, two documents with different watermarks cannot be spatially registered to extract the watermark. Of course, the watermark can be defeated by retyping the text.

8. APPLICATIONS OF DIGITAL WATERMARKING

A number of application areas, with different requirements and limitations, have been envisioned for data hiding. These include (i) Copyright protection and finger printing; (ii) Authentication for data integrity check; (iii) Covert communication; and (iv) Labelling and annotations. In a practical application, the watermarking techniques can also trace the illegal users (forensic watermarks) so that the owner can approach the regulating authority. In fingerprint watermarking the buyer and seller information with date stamp can regulate copyright infringement. Hence, without an effective means for monitoring, tracking, and deterring the unlawful distribution of content once it reaches consumers, the bright expectation of content providers could not be achieved[8].

9. POSSIBLE ATTACKS ON DIGITAL WATERMARKING

An attack on a watermark can be defined as an operation, (coincidental or hostile) that may degrade a watermark and possibly make it unreliably detectable. Some of the practical attacks include the following.

9.1 Compression Methods

JPEG is currently one of the most widely used compression algorithms for images, and any watermark should be resilient to some degree of compression.

9.2 Geometric Transformations

Transformations such as flipping, rotation, scaling, cropping, and other transformations with compression.

9.3 Image Enhancement Techniques

Low-pass filtering including linear and non-linear filters such as median, Gaussian, and standard average filters, sharpening methods such as photo editing and denoising, and histogram modification generally associated with image smoothening, noise addition, statistical averaging, collusion and image fusion[6].

10. STANDARDS

The Copy Protection Technical Working Group (CPTWG) and Digital Audio-Visual Council (DAVIC) are working hard to develop digital watermarking standards that will allow wide deployment and acceptance of this technology. Other visible initiatives are underway at the International federation of Phonographic Industry (IFPI) and the Secure Digital Music Initiative (SDMI) for audio watermarking and at the DAVIC for watermarking in e-commerce. Standardization will likely happen in different sectors such as physical media (DVD, CDs), online media delivery, broadcasting, and document authentication in the imaging industry[14].

11. LIMITATIONS OF DIGITAL WATERMARKING

Watermarking relies on a secret key, which serves as a carrier modulated to transport a message and embed it in the original content. The drawback is that knowing the secret key implies the ability of modifying or removing the watermark—which makes a public watermarking scheme infeasible for the near future. This means that major developments in copyright technology will be oriented toward the rights of the individual, but exclude the information demands of the general public[7].

12. FUTURE OF DIGITAL WATERMARKING

Though publishers' have been clamoring for some means to protect their material on electronic networks, there has been no rush yet to embrace any of the current schemes. This could be just due to a period of inspection and appraisal, but the opinion is that publishers and scientists have yet to fully understand the practical satisfactions associated with the problem.

As scientists purpose solutions and publishers experiment with them and debate their merits, some methods of watermarking will emerge as useful and widely used. When that happens, there will also be the emergence of external agencies for monitoring electronic copyright infringement (much the same as there are agencies for music and print copyright management). In the meantime, the challenges is for the scientists to develop ever more invisible, decodable, and permanent watermarking methods and perhaps to meet even more specification as they demand.

16. CONCLUSION

Protecting intellectual property is currently a "hot" topic for scholars, for university administrators and for the media and publishing industries. This has drawn lawyer, legislators and computing service professionals. For fine arts and manuscript sources, quality of representation in the Internet and other digital images distributions is as important to the curators as aspects more commonly considered intellectual property rights. This can be seen as an extension of an author's right not to be misrepresented. DWM is an exciting for researchers because it is a new field and there is an opportunity to do pioneering work. It exists for entertainment companies, museums and libraries because it offers the promise of better protecting their multimedia content from piracy. The possibilities for digital watermarks in the fields of geo-imaging and e-commerce are only just being realized. Watermarking can support a broad range of imaging data types whilst supporting industry standards security implementations. It is transparent in use, does not increase files sizes, and yet is highly robust, secure and customisable. The technology has unique ability to safeguard, both digital and printed work, and secures the complete workflow with minimal impact on existing processes. A deeper understanding of the theory behind DWM will lead to the design of more robust and reliable systems for a variety of applications.

REFERENCES

1. Cappellini, V. *et al.* Information theoretic aspects in digital watermarking (Editorial). *Signal Processing*, 2001, **8**(1), 1117-19.

2. Dittmann, Jana & Nack, Frank. Copyright-Copywrong. *IEEE Multimedia*, 2000, **Oct-Dec**, 14-17.

3. Hernandez, Martin; Juan, R. & Kutter, Martin. Information retrieval in digital watermarking. *IEEE Commu. Mag.*, 2001, **August,** 110-16.

4. Jonker, W. & Linnartz, J.P. Digital rights management in consumer electronics products. *IEEE Signal Proc. Mag.*, 2004, **March,** 82-91.

5. Kundur, Deepa. Watermarking with diversity: Insights and implications. *IEEE Multimedia*, 2001, **8**(4), 46-52.

6. Page, Thomas. Rights management: Digital watermarking as a form of copyright protection. *Computer Law & Sec. Rep.*, 1998, **14**(6), 390-92.

7. Rao, N.V. & Pandit, S.N.N. Multimedia digital rights protection using watermarking techniques. *Inform. Sys. Sec.*, 2007, **September**, 93-99.

8. Rosenblatt, Bill. DRM, law and technology: An American perspective. *Online Inform. Rev.*, 2007, **31(**1), 73-84.

9. Wolf, Patrick, *et al.* Complementing DRM with digital watermarking: Mark, search, and retrieve. *Online Inform. Rev.*, 2007, **31**(1), 10-21.

10. Yeo, Boon-Lock & Yeung, Minerva M. Watermarking 3D objects for verification. *IEEE Comp. Graph. Appli.*, 1999, **Jan/Feb**, 36-45.

11. http://www.alpvision.com

12. http://www.computerworld.com/action/

13. http://www.dct-ch.ch

14. http://www.ifpi.org

15. http://www.nextamp.com

16. http://www.nikkeibp.com/nea/may97/mayspecial/index.html

17. http://www.signumtech.com

18. http://www.watermarkingworld.org

About the Authors

Dr M. Natarajan is a Senior Scientist at the National Institute of Science Communication And Information Resources (NISCAIR),New Delhi. He obtained Doctorate in LIS from University of Madras; Associateship in Documentation and Information Science from DRTC, Indian Statistical Institute, Bangalore; and MSc in Mathematics. He also passed Computer Systems and Design from Thiagarajar College of Engineering, Madurai, and PGDCA from SISI, Chennai. He joined INSDOC, Chennai, in 1981 and worked on Networking (SIRNET, MALIBNET), searching from national and international databases, conducted many training programmes of INSDOC and UGC Refresher courses for LIS Professionals and delivered lectures. He headed the position of Scientist-in-Charge of INSDOC, Chennai, and Bangalore, and delivered special lectures at various places. He acted as Programme-in-charge for IGNOU courses and still Councillor for IGNOU courses like MLIS and PGDLAN. He is presently working as a Faculty for AIS course in NISCAIR. He has published more than 100 papers in national and international journals / conferences / seminars / Festschrift volumes, etc. He is a Life member of many professional associations in India like ILA, IASLIC, SIS, SALIS, MALA, Academy of Information Science, Mysore, etc. and ordinary member of TLA, RRC, etc.

Dr Gayas Makhdumi has done his PhD from Universtiy of London. He is a Commonwealth scholar and currently the University Librarian at Jamia Millia Islamia, New Delhi. He has vast experience in the library management and conducting the BLIS courses at Jamia. He also conducted many UGC Refresher courses for Librarians with the help of Academic Staff College. He has published many papers in national and international journals / conferences / seminars. He is a life member of many professional associations in India.