Part III

Towards 3D Shape Watermarking

Chapter 5

Introduction to Digital Watermarking

5.1 Introduction

Over the last decennium, the way that audiovisual content is being produced, distributed and consumed has dramatically changed. This transition is mainly due to the shift from the physical to the digital. Bits sequences transported on the Internet have almost completely replaced paper, vinyl and tape in the past ten years. This evolution to the digital allows an extremely flexible way of producing, transporting and consuming audiovisual content.

Unfortunately the ease related to digital media can be misused for piracy purposes at a much larger scale than previously. In the time of physical media carriers, piracy had to face inherent obstacles to be a large-scale phenomenon. Indeed, unauthorized copying on physical carriers (such as audio cassette tapes) was often characterized by degradations of content quality. The distribution of unauthorized content, before the evolution to the digital, was also often expensive, cumbersome, privacy sensitive and potentially dangerous. As unauthorized transport and consumption of digital content is now much easier, the importance of intellectual property rights (IPR) protection on these contents has become critical.

Moreover, a much finer control over how content is consumed is expected by content owners. For example, digital media should allow for one or a limited number of copies instead of choosing between «free copy»

and «copy never». Fighting piracy while enabling more flexibility to IPR protection is consequently a very challenging and complex task. The collective set of methods and techniques used to impose rules on how digital content is being produced, distributed and consumed is referred to as digital rights management (DRM). More precisely, DRM handles the description, layering, analysis, valuation, trading and monitoring of the rights held over a digital work [15, 48, 70]. This multidisciplinary technology has promoted innovative research and development in such diverse fields as authentication, biometrics, secure collaboration and data sharing systems, digital signatures and data hiding, smart-card technology, time stamping, watermarking, steganography, forgery detection etc.

Among these fields, this chapter is related to digital watermarking. Watermarking [118] refers to the *imperceptible*, *robust* and *secure* embedding of a given *amount of information* in a digital multimedia. The applications of watermarking range from copyright protection, labelling, monitoring, tamper proofing to conditional access. It is very close to steganography [71] which only intends to convey imperceptible and undetectable data in a media without any robustness requirements.

5.2 Terminology

This section introduces important vocabulary related to watermarking schemes. The media to be watermarked is often called *cover* or *host* and noted C. The hidden information or message to be embedded in the cover is named *watermark* and noted w. The exact way of embedding or decoding a watermark in a cover media is often ruled by a personal key which can be kept secret k_s or public k_p.

Watermarking schemes are usually characterized by four constraints:

- capacity
- robustness
- imperceptibility
- security

Capacity (of the hidden channel), also referred to as payload, is expressed in bits. The capacity needs of a watermarking scheme depend

on the targeted application. It may vary between 1 bit (in authentication), to 64 bits (for author serial numbers in copyright protection or fingerprinting), to 256 bits (for a URL redirecting to more information about the content and its use policy in labelling or conditional access) [19].

Robustness refers to the ability of the watermark to be retrieved even after modifications of the content. These modifications, also known as *watermarking attacks*, can be either malicious (performed with the intention of erasing the watermark) either not. There is no international consensus on which attacks should be taken into account for a given watermarking application and a given media type. This point will be developed in the case of 3D watermarking in the following sections.

The *imperceptibility* of the watermark must be understood as perceived imperceptibility (the human eye shoud not be able to detect differences between the watermarked and original versions of a media) and numerical or objective imperceptibility (non authorized people should not be able to numerically detect the presence of the watermark).

Security refers to the fact that non authorized people (those who do not know the necessary secret key) should not detect nor read the watermark.

These four constraints are usually antagonistic and an application-dependent trade-off must be found. Intuitively, hiding more information leads to less robustness and conversely. Generally, the imperceptibility constraint is the most important (the watermark should not degrade the audiovisual content so that it becomes unusable). Unfortunately, improving robustness often implies perceptible degradations and improving imperceptibility also leads to less robustness.

Another important characteristic of a watermarking scheme is the *minimum watermark segment*, which corresponds to the protected minimal content amount.

A very general scheme of watermak embedding is illustrated on Fig. 5.1. The original content c_o is the input of an extraction function e and possibly a transform t which generate a vector $\boldsymbol{x}$. This vector enters a watermarking function w which has the message m and a key k as input. A mixing

function p generates the watermarked data $\boldsymbol{y}$ from the output of the watermarking function and the data $\boldsymbol{x}$. The inverse transform t^{-1} and the inverse of the extracting function e^{-1} enable to obtain the watermarked content c_w.

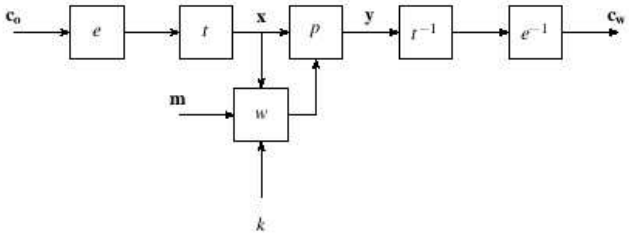

Figure 5.1: *Watermark embedding scheme.*

Finally, the *detection* and the *decoding* of the watermark do not have the same meaning. The detection of a watermark only intends to check whether the media is watermarked or not. The decoding of the watermark stands for reading the watermark bits embedded in the media with a given secret or public key. Some applications related to authentication or steganography only need a simple detection.

The detection (or decoding) step is also classified between:

- informed or non-blind detection : the original version of the media is needed to detect whether the suspect media is watermarked or not. The knowledge of a key and of the embedding algorithm are not sufficient at the detector side.
- blind detection : in this case, only the key and the knowledge of the algorithm enable to read or detect the watermark without any original not watermarked data.
- side-informed detection : close to the blind detection, the original version is not needed but only some information related to it.

A general scheme of blind watermak decoding is illustrated on Fig. 5.2. The possibly attacked content c_a is the input of the extraction function and the transform used at the embedding in order to generate a data vector $\boldsymbol{y'}$. This vector and a key k are the input of the detection function which generates the estimated message $\hat{m}$. If the decoding is successful, we have $m = \hat{m}$.

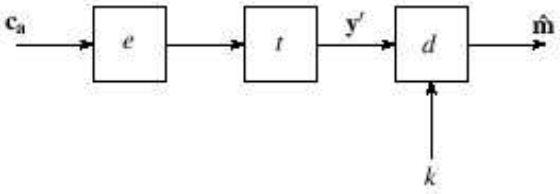

Figure 5.2: *Watermark blind decoding scheme.*

5.3 Applications of watermarking

This section presents several applications that make use of watermarking. Each context of application presents different requirements in capacity, security, robustness and imperceptibility. It is consequently obvious that there exists no watermarking scheme which would be optimal for every application.

5.3.1 Copy protection

Copy protection [52] intends to embed in the cover content the information needed to specify its use. For example, the information describes whether the content may be copied and, if so, how many times it can be copied. The most well-known example is the DVD [93]. Each traded DVD is encrypted and watermarked. However, everybody is allowed to burn personal videos on a DVD. This one will not be encrypted nor watermarked. DVD players are configured to read only encrypted and watermarked (commercial) DVDs and not encrypted and not watermarked (personal) DVDs. A watermarked but not encrypted DVD is detected as a pirate DVD and cannot be read. In this case, the *detection* of the watermark is preferred to the decoding of the watermark.

For this application, imperceptibility is the most important criterion. An effective robustness is also required while security is not dealt and capacity can be limited to a few bits.

5.3.2 Copyright protection

Copyright protection [3, 133] aims at ensuring that the identifier of the content buyer or licenser will always be present and detected in the cover media. As mentioned before, copying on physical carriers implies perceptual degradations and limits the number of acceptable copies. For a digital carrier, the reproduction is perfect no matter the number of copies. The notion of *original* content does not exist anymore. According to the applicable legislation, one first has to declare himself before a trusted third party (TTP) before being allowed to lodge contents before it. The TTP then delivers an identifier (ID) for each lodged content and watermarks the ID inside it. It also records the date of deposit, so as to keep evidence of precedence of rights.
Robustness and imperceptibility are the most important constraints in this application scenario. Robustness means that the watermark should resist to any combination of attacks until the cover media becomes too much degradated to be used. The capacity is constraint to 64 or 128 bits which corresponds to memory needed to encode the serial number identifying the owner or the licenser. As for copy protection, security is not dealt.

5.3.3 Integrity protection

The goal of integrity protection is to ensure that the received content is conform to the original content [46, 136]. A well-known example is given by press pictures. These sometimes undergo malicious modifications difficult to perceive but which change their semantic content. The idea is to embed a fragile watermark that disappears after any modification. The *absence* of watermark proofs that the integrity of the content is questionable. Security and imperceptibility are the most important criteria of a good fragile watermark. Capacity can be limited to a hundred of bits.

5.3.4 Self-indexed contents

Indexing aims at giving a semantical description of the content [90]. These metadata can be encoded in the header of a dedicated file format. The problem arises when a format conversion is needed, the metadata

disappear if the targeted format does not take them into account. If these metadata are embedded by a watermarking scheme able to resist to file format conversions, then no file format standard would be discarded for applications using such metadata.
This kind of watermarking scheme is of course not intended to be secure.

5.3.5 Augmented contents

Similarly to self-indexed contents, augmented contents also contain metadata enabling different levels of use. For example transmitting hidden depth maps in images could allow everyone to watch these images and allow those who buy the dedicated decoder to watch the images with the augmented 3D.

5.3.6 Steganography

Steganography finds its roots in the Antiquity [71] and saved many Empires when used to transmit *undetectable* secret information. Let us imagine that Alice and Bob exchange information with George looking at them. George can read all the information transmitted between Alice and Bob and has the power to stop the transmissions if he does not like their content. If Alice or Bob sends encrypted data, George knows they do not want him to read them and he can stop the transmission. A possible way to transmit secret data to Bob is to hide a secret in a day-to-day message. George is unable to know whether there is a secret because he cannot even detect whether the message contains more information. Steganography is the complement of cryptography: encrypted information should also be transmitted with discretion. Imperceptibility and numerical undectability are the most important constraints. Security is ensured by cryptography and robustness is useless in this case.

5.3.7 Fingerprinting

Fingerprinting aims at ensuring that a user has not pirated the rights attribution protocol [86]. The corresponding licenser identifier of the users asking for a media content is embedded in each version of the content they receive. If the content is retrieved in a pirat channel, owners are able to determine who is the seed of this piracy. Here again, imperceptibility is very important while security and robustness are relatively important.

5.4 Watermarking of 3D Data

Now that the general context of application and the technical vocabulary related to watermarking have been presented in the previous sections, the following sections focus on watermarking schemes dedicated to 3D models[1]. Indeed, the digital rights management problem of protecting data from theft and misuse has been addressed for many information types, including software code, digital images, videos and audio files. Demand for ways to protect 3D graphical models is significant and growing [84, 85].

5.4.1 Why should 3D graphics content be protected?

Contemporary 3D digitization technologies allow the efficient creation of accurate 3D models of many physical objects. For example, the Stanford Digital Michelangelo Project [87] has developed a high-resolution digital archive of ten of Michelangelo's large statues, including *David*. These statues represent the artistic patrimony of Italy's cultural institutions. Italian authorities permit distribution of the 3D models only to established scholars for noncommercial use [84]. The digital 3D model of the David would quickly be pirated if it were distributed without any protection; simulated marble replicas would be manufactured outside the provisions of the parties authorizing creation of the model.

Digital archives of archaeological artifacts are another example of cultural heritage 3D models that could require piracy protection. Curators of such artifact collections increasingly turn to 3D digitization as a way to preserve and widen scholarly use of their holdings, but they often want strict control over the manner of that use of the 3D data and to guard against theft. An example of such a collection is the Stanford Digital Forma Urbis Project[2] undertaken by Stanford with Italian archaeological officials to digitize more than a thousand marble fragments of an ancient Roman map and make them publically available through a Web-based database provided the 3D models have adequate protection.

[1] «3D model», «3D graphics», «3D graphical model» and «3D object» are synonyms, while «3D mesh» is a particular representation of a 3D model

[2] http://formaurbis.stanford.edu

Other application areas (such as entertainment, collaborative design of CAD mechanical pieces and online commerce) also require protection for 3D graphics content. Valuable 3D animated character models developed for use in motion pictures and 3D body scans of high-profile actors may be repurposed for widespread use in video games and promotional materials. Content developers might be reluctant to distribute the 3D models in interactive applications without control over piracy and reuse. A number of Internet application developers have reported that their clients are unwilling to pursue online 3D graphics projects due to the inability to prevent theft of 3D content [84].

5.4.2 Are there alternatives to 3D watermarking?

Technical research in intellectual property protection for 3D data has concentrated on 3D digital watermarking techniques which are addressed in this thesis. These techniques embed hidden information into 3D graphical models, with varying degrees of robustness to attacks aimed (or not) at disabling the watermarks by altering 3D shape or data representation. Complementary technologies search collections of 3D models, examining them for the presence of digital watermarks in an effort to detect piracy.

For the digital representations of valuable 3D objects (such as cultural heritage artifacts), it may not sufficient to detect piracy after the fact; piracy should even be prevented. The computing industry has experimented with a number of techniques for preventing unauthorized use of digital data: software access keys, node-locked licensing schemes, copy-prevention software, obfuscation, and encryption with embedded keys [84]. Most are either broken or bypassed by determined attackers, causing undue inconvenience and expense for nonmalicious users. High-profile data and software are particularly susceptible to attackers.

Interestingly, 3D graphics data differ from most other forms of digital media in the fact that the presentation format (2D images) is fundamentally different from the underlying representation 3D geometry. 3D graphics data are usually displayed as a projection onto a 2D display device, resulting in a large information loss for single views. This property supports an optimistic view that protected 3D graphics systems can still be

useful to users, without making the 3D data as vulnerable to piracy as other types of digital content.

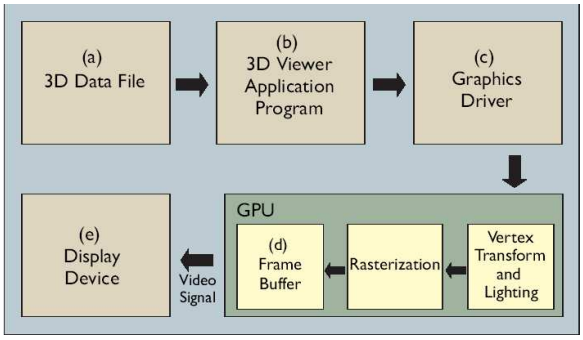

Figure 5.3: *Rendering pipeline protection issue [84]. Possible attack locations (a-e) for recovering 3D graphics data from an interactive viewing program. (a) 3D model file reverse engineering. Users with full access to 3D model data files can reverse engineer obfuscated of encrypted formats; (b) 3D viewer application tampering. Hackers use program tracing, memory dumping and other techniques to obtain access to data being used by application programs; (c) Graphics driver tampering. 3D data passes through graphics driver software on its way to the graphics hardware where the drivers are vulnerable to tampering or replacement by attackers trying to capture streams of 3D data; (d) Reconstruction from the frame buffer. By accessing the rendered images from the graphics memory, using 3D computer vision techniques to reconstruct the original model; and (e) Reconstruction from the final image display. The final video images output from a graphics system are vulnerable to capture and reconstruction.*

Protecting the 3D graphics content through its rendering is complementary to watermarking and involves the security of the visualization pipeline (see Figure 5.3). Several strategies to resist against attacks on the various stages of this pipeline are possible:

- Software rendering : it involves bypassing the graphics processing unit (GPU) driver and hardware, and using software-only graphics rendering for at least a portion of the data. Software rendering keeps control of the rendering process in the hands of the viewing application programmer, allowing for specialized data encryption or obfuscation techniques to be used to protect the data in the early stages of

the pipeline, trading off display performance.

- Deformations before visualization: introduction by the viewing application of subtle deformations in the geometry of the model before passing the 3D vertex data to the graphics driver; attackers would have difficulty for reconstructing the full 3D model due to the distortions.

The drawback of such protection techniques is that they all eventually rely on security through obfuscation, which is unsound from a computer-security point of view. Any 3D graphics protection technique that makes the actual 3D data available to potential attackers in software can be broken [27], as any attacker with enough time and resources will be able to reverse engineer the protections on the data. It is possible that future trusted computing platforms for general-purpose computers will make software tampering difficult or impossible, but few such systems are deployed today.

Another approach for protecting 3D graphics is hardware GPU decryption. If the 3D models were encrypted using public-key encryption when they are created, then custom GPUs could accept encrypted data and perform on-chip decryption and rendering. This technology would provide robust protection for the 3D data, but such GPUs do not exist today, and it will take years before it is widely available in personal computers.

Koller et al. [85] propose a scheme for securing 3D graphics by retaining the 3D model data on a secure server controlled by the content owner and pass only 2D-rendered images of the models back to user-client requests. The 3D geometry is safe from all types of graphics pipeline attacks (except reconstruction from images), though the server itself is still vulnerable to direct attack. To prevent recontruction from images, the 3D data is strongly simplified in order to erase high detailed features which are believed to be the content to be protected since they are the key to a valuable reconstruction.

These 3D data protection techniques are complementary to 3D watermarking and make part of a more complex system which would better tackle the digital rights management issue of protecting intellectual property of 3D content. The remainder of this section presents in a more detailed manner the issues related to 3D watermarking.

5.5 Conclusion

This Chapter has introduced the motivations of digital watermarking and its application contexts. The specificities of 3D data security have also been presented as well as other technologies that can complement digital watermarking. The next Chapter presents a detailed survey of 3D digital watermarking. It also introduces the contributions of this thesis which are the subject of the remaining Chapters of this thesis.