# WATERMARKING OF DIGITAL GEOSPATIAL DATASETS: A REVIEW OF TECHNICAL, LEGAL AND COPYRIGHT ISSUES

CARLOS LÓPEZ
*Universitario Autónomo del Sur*
clv@ei.edu.uy
*Uruguay 1225, Montevideo, URUGUAY*

**Abstract.** It has been widely recognized that gathering data accounts for more than 80 per cent of the cost of any GIS project. Software and hardware costs are decreasing, but not datasets. In the analog era, unauthorized copy of the datasets was difficult to made, requiring not only access to them but also specialized equipment and procedures. Fast Internet connections and digital datasets threaten the data producer investments through piracy, which today is easier than ever.

Software providers have solved the piracy problem for their product supplying hardware keys, or through runtime license services. However, they do not offer similar protection to the dataset being used by their software. Reasons are both technical and legal. Outside the GIS community this problem has been known for a long time; we will consider current, possible solutions for digital imagery, formatted text, 3D meshes and so on, showing possible links to typical Geospatial datasets. The current proposed solution forces the data producer to rely in weaker protection means: the dataset delivered to customers is functionally valid, but some subtle changes identify uniquely the buyer, the supplier as well as other extra information. The process for embedding hidden information in a dataset without producing perceptible changes is denoted as watermarking. The producer can recover the embedded information on request in order to produce evidence of ownership in a court, so the overall strategy relies on legal basis rather than technical ones. The differences with cryptography will be made clear.

This paper analyzes the state-of-the-art for watermarking protection in digital geographic datasets, with emphasis to the technical aspects. We analyzed the situation of some typical formats: raster images, vector and point datasets, as well as text and databases. Some reference to the legal status of copyright digital geographic data in the US and the EU is also presented. The review makes evident that digital imagery has clearly a maturer situation, even with multiple commercial vendors offering watermarking protection. 2D vector and point datasets have received less attention from the research community; however, 3D meshes have been considered by the CAD community and a handful of techniques are available for that case, and they are reviewed here. Some references are provided for the case of text and databases.

## 1. Introduction

The cost of setting up a GIS system can be roughly attributed to hardware, software, salaries and source data. Some of them might decrease due to technology advances, while others do not. It is widely accepted that the cost of data acquisition and collection is the major one in setting up a GIS project. Some rule-of-thumb states that it might account for

80 per cent of the overall budget. Since many projects have similar requirements in terms of data, one possible way to drop costs is by a cost-sharing approach: the data is collected and maintained by someone (the *data provider*), and used many times by users that pay a fee for that. From the technical point of view, this approach is now practical. It has benefited in general from the communication infrastructure associated to the Internet era, and more specifically, by standardization efforts regarding uniform description of geographic datasets. From the policy side, the developing of concepts like Spatial Data Infrastructures (either at the local, national or global scale; see http://www.gsdi.org) created a framework suitable for coordinate data collection activities, virtual data catalogs, as well as means to buy, sell and distribute datasets. The catalog service is carried out in the so-called *Clearinghouse*, and after locating the dataset this is also the natural place where to offer the other services.

All of this is very new, and changes dramatically the way geographical data will be disseminated in the present and future. The data provider can market easily its product through the Clearinghouses, where the prominent client searches for the data. Through the catalog services, the product can be located without visiting all possible offices in the government, education and private sector, but just entering a few keywords in an electronic form. Once found the provider, and once decided that the dataset will fit the client needs, it is time for commerce. Whether the transaction is done electronically or not, the final situation is that the data provider will give the client a copy of their dataset. We will focus in this paper the case of digital datasets, because it has some peculiarities and differences mostly in the piracy risk. Traditional media (like paper, filmstrip, etc.) is certainly amenable to unauthorized copy; however, in order to produce a copy expensive equipment is usually required, which precludes the typical customer to do so. In addition, new copies are of lower quality than their originals, which certainly diminishes its market value. It is certainly possible to make high quality copies, but the associate cost is possibly higher than merely buy a new original one.

The situation is different for digital datasets. Perfect copies can be made easily with widely available equipment. This new fact threatens the data provider, because delivering the first copy of the dataset will open the door to unauthorized copies without any technical barrier. The cost-sharing basis assumption of many data collection projects (either at the government or private sector) can fail immediately, leading to highly restricted accessing to the data, or even worse, stopping completely its distribution. Limited access can be provided with the help of cryptography, as described by (Zhao *et al.* 1998). The dataset can be supplied only in encrypted form, and the client can use it only if using an appropriate software, hardware or combination of both with the right key. Today, and to the author knowledge, there is no GIS software capable of using encrypted datasets. It is certainly difficult to create one, because once decrypted, there are many valid manipulations within the GIS (like copying layers of information) that can be used by a malicious client to disassemble the dataset, saving to temporary files, and finally reassemble it creating a non-encrypted version of the dataset. A software capable for distinguish which layers can be copied, and which not, and which procedures can be performed, and which not, need to be highly sophisticated in its conception, because there will be the risk that the legitimate owner have some trouble using the dataset.

Other barriers can be set outside the technical field. The law can recognize copyright ownership for the dataset. A contract between the data producer and the client can be sign in a per-case basis, stating the limitations for use and further distribution as well (as done with satellite imagery). The legal and contract approaches have some drawbacks that will be analyzed later. However, the main one is that, if an unauthorized copy of the dataset is found, there should be some means to discover who was the original owner. If identical copies of the dataset were delivered to client A and B, how could we identify the source of the illegal copy found at Mr. C's office? The goal of this paper is to present the means to create functionally identical copies (but not binary identical!) carrying hidden information in order to do that. Considering not the source data but the output of the GIS analysis, we can even add value by producing maps, reports or certificates with hidden information that identify the customer, the operator, the date, etc. using basically the same technology to be described for the input data.

This paper is organized as follows: section 1 serves as an introductory background to the problem; Section 2 discusses the differences between watermarking and cryptography, and introduces the technique; Section 3 describes the situation for raster images while section 4 covers the situation of vector-like datasets; Section 5 analyzes briefly the situation for other GIS data types like text and databases; Section 6 is devoted to discuss the protocols involved in the trading of the data, while section 7 analyzes some legal aspects and in Section 8 we present our conclusions.

## 2. Cryptography vs. Steganography: brief introduction to digital watermarking

The goal of cryptography is to protect the content of the dataset from unauthorized users *during transmission*, modifying the original dataset to make it unreadable. Some secret numbers (named *keys*) are required to decrypt the files. See (Schneier 1995) for a good introduction to the subject. Once recovered the original dataset, there is no further protection: perfect copies of the original can be made without participation of the legitimate owner. Thus the usefulness of cryptography in the case under consideration is limited to the distribution stage. Steganography is a somewhat different technique, because it attempts to add extra information to the dataset. The information can be clearly visible or not. The former is used to clearly identify the source of the dataset without destroying its usefulness for the intended applications. An example, due to (Mintzer *et al.* 1996) is presented in figure 1. It shows a very clear stamp over an image from a Bible of the XV century belonging to the Vatican library. The user can still read the image, while preserving and providing information about the source.

In the second case, the extra information (*the watermark*) is included in a file without been noticed; a watermarked image is expected to be indistinguishable from the unwatermarked, original one. This is a significant difference to the encrypted message, which is unreadable without the right key. In contrast to cryptography, steganography does not immediately arouse suspicion of something secret or valuable. Instead, it hides an important message within an unimportant one. See (Bender *et al.*, 1996) or (Anderson and Petitcolas 1998) for a more detailed presentation.

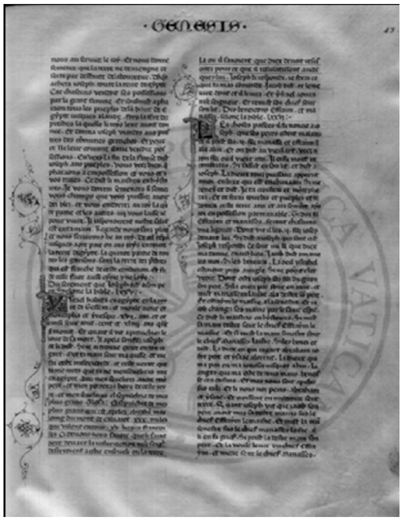

Figure 1 Visible watermark superimposed to an image taken from the Vatican library

The most interesting case for GIS applications is the case where the watermark is not evident, and we will discuss later how to accomplish that. Through the watermark, the dataset might include some extra information which might identify the distributor (if different from the data producer), the buyer, the date of the transaction, etc. making possible to trace back the dataset to its source.

Other application for watermarking is for use control: (Cox and Linnartz 1998) described the example of the Digital Video Disc (DVD). The data is in digital form, but encrypted. In addition, the (compliant) disk player produces an analog output suitable for displaying in a TV set, but unsuitable to be recorded in another system. And the third protection relies on watermarking. The absence of the watermark indicates that free unlimited copy is allowed. A second possibility includes “copy not allowed” and the third “copy once” (creating a copy that is itself marked as “copy not allowed”). Despite that the DVD experiment is not a success (because its encryption has already been broken) it illustrates a technological mean to deal with piracy while using the media. However, for GIS applications this example is of limited interest today, because current software does not expect to process watermarked datasets. This might change in the future, however.

The watermark can be used to prove data integrity. If the dataset is edited or modified *after* the watermark has been inserted, the watermark might reveal that, and in some cases also which part has been affected. If the watermark is very sensitive to changes in the dataset, it is denoted as a *fragile watermark.* Conversely, if the watermark is able to survive in the dataset even if it has been (inadvertently or deliberately) edited, it is denoted as a *robust watermark.* This is the most interesting case, because the malicious user might want to acquire a legitimate copy, edit it and later distribute the modified one arguing that he was the author. If the watermark is robust, it can be recovered and used to prove ownership.

The case of robust watermarking is the main subject of this paper. A short, good introduction is (Voyatzis and Pitas 1999). The term *robust* should be interpreted in a framework of possible *attacks*. An attack is loosely defined as any action that a user can perform on the given dataset that can affect the watermark, or its usefulness. An unintentional attack happens when a legitimate transformation (like format changes, rotation, JPEG compression, etc.) is applied without the removal of the watermarking in mind. Extreme cases imply the destruction of the dataset. The underlying principle below intentional attacks is that the resulting product should be of good quality, without destroying its commercial value. In addition, it is assumed that the commercial value of the dataset is not so extremely high to induce the attacker to use but a limited amount of computer resources. As many cryptographic algorithms, most watermarking schemes can be defeated with enough time and resources. We will see later that intentional attacks are illegal, and can be prosecuted under international copyright law.

(Kutter and Petitcolas 1999), (Mintzer *et al.* 1998), etc. describe and define a set of feasible attacks that an ideal robust watermarking system must survive in order to be useful. Some of them are unique to images, but others are applicable to other datasets as well. Among others, they define the following:

- *Noise addition or removal:* is to add (or multiply) every pixel by a random, uncorrelated value small enough (or close to 1.0) in order to make the changes unnoticeable. For geographic datasets this might be achieved by random perturbation of coordinates and/or attribute values.
- *Geometric transformations:* The possibilities include small angle rotation, mirroring, reflection, cropping, affine transformations, deletion + duplication of lines or columns, etc.
- *Loosy Compression:* The dataset is simplified by retaining its *most significant* features. A good example is JPEG compression for still images.
- *Enhancement techniques:* The possibilities include low pass filtering, sharpening, histogram modification, Gamma correction, color quantisation, dithering and restoration.
- *Format transformations:* like convert from shapefile to DXF and back to shapefile.
- *Overmarking*: embedding a second watermark with the same watermarking method damages the original watermark. The extreme case of many watermarks is also denoted as *watermark bombing.*
- *Inversion:* If the embedding process is invertible, the watermark can be removed by the Inversion attack. The attacker tries to first detect the watermark, then to isolate it, and finally through the inverse operation to remove it. For example, if the embedding process uses addition of images, the watermark is removed by subtraction. (Craver *et*

*al.* 1998) and (Ramkumar 1999) considered an extreme case denoted as *Counterfeit attack*, where the pirate can use the inadequacies of the protocol to “demonstrate” the presence of *his* watermark in the actual, original content.

- *Copy Attack:* The watermark is estimated from a watermarked image, extracted from there and then inserted in another one. Kutter *et al.* (2000) stated that this attack makes sense if the watermark is used for image authentication.
- *Averaging:* The attacker averages three or four valid images marked with legal and different watermarks, producing a new image which might be too close to the original source. All the watermarks are in fact removed because their individual strength is diminished below a threshold.
- *Collusion* or *conspiracy:* The attacker creates a new dataset using small parts taken from three or four valid dataset instances marked with legal and different watermarks. Another approach uses all the instances to locate the differences (due to the watermark) so they can be altered and removed. Deliberate errors can be removed this way.
- *Brute Force Key Search:* If the watermark algorithm is known, an exhaustive search is performed over all possible keys.
- *Oracle attack:* When a public decoder is available, an attacker can remove a watermark by applying small changes to the image until the decoder cannot find it anymore.
- *Mosaic attack:* if the watermarked images are expected to be published in the Internet, a pirate can download it, split into tiles and show the image as a mosaic. The end user does not notice any difference. The interest on this attack is because software like MarcSpider (Digimarc Inc.) look for watermarks in all images available in the WWW but only if they are larger than some preset size: otherwise, they are not processed. The extreme case is to split the original image in others of just one pixel; no watermark can be embedded in such individual tiles, but they can be jointly displayed looking as the original source.
- *Printing + scanning:* a self explained process, which introduces both geometrical and noise-like distortions.

For the case of digital images, (Petitcolas and Kuhn 1997) reported that some of these attacks have been included in special software tools intended to be used as benchmark for new algorithms. According to (Petitcolas *et al.* 1998), it is still an open question whether there is any watermarking scheme for which a successful chosen *distortion* attack cannot be found. For Geospatial datasets, some attacks are not of interest. For example, the print + scanning attack is valid for raster datasets, but not for vector ones, because the real value resides in the vector format itself.

## 3. The case of digital images (raster)

This is a typical dataset form used in many GIS operations. Its main characteristic is that there is a clear order (row, columns) in the data, in opposition to vector and point datasets, which will be considered later. Satellite imagery (LANDSAT, SPOT, etc.) as well as aerial photography falls within this classification. Watermarking of still images have received significant attention from the research community. The driving force is the copyright protection of artistic imagery, which has a rather different legal status than other images like satellite ones (see section 8).

The robust watermarking schemes can be applied in the spatial or spectral domain. The former applies the watermark keeping the (column, row) structure of the image. It is the option required for visible watermarks. For the more common invisible ones, the choice of the spatial domain produces weaker watermarks, because the changes need to be performed in the least significant bits (LSB) of the image in order to assure low perceptual changes. An immediate consequence is that few bits can be inserted; see the pioneer work of (van Schyndel *et al.* 1994) for details. For common GIS images, the meaning of "low perceptual" and "LSB" might be different. For the case of aerial photography, which will be processed by humans, the limit is related with the human visual system. For satellite data human limitations are (usually) unimportant because the image will be processed within a computer. In such case, the LSB limit might be more precisely defined: it is related with the properties of the sensor (an example is the remote sensing data) or the inherent uncertainty of the measured parameter. Equally valid values might be altered to amounts just below such threshold; if properly chosen, such alterations might serve as a watermark.

As another example of spatial domain watermarking, (Nikolaidis and Pitas 1996) suggested to divide the pixels in the image in two sets, A and B, by applying a pseudorandom partition using a secret key. The luminance of the pixels of class A is increased by a fixed integer *k*, small enough to produce an imperceptible change. Given the secret key and *k*, the watermark is detected by comparing the difference of the average luminance in sets A and B, which will be near k if the watermark is present, and nearly zero in other cases. The original image *is not* required for the test.

(Kutter *et al.* 1998) proposed a similar system that exploits the low sensitivity of the human visual system toward changes of high frequencies and blue color. The pixel modifications are proportional to the luminance and the signature bits determine the sign of the modifications. (Nikolaidis and Pitas 1998) recognized that a significant problem of all the spatial domain techniques is that the watermark might not survive JPEG loosy compression, which is a typical image transformation. This is due to the fact that the watermark is essentially low power, white noise. They modified their original method by varying the integer *k* added to each pixel, but keeping its total sum as before. The set A is formed in a different way, because the pixels are now grouped in small blocks of size 2x2 or 2x4. An optimum $k_{mn}$ is separately calculated for each block$_{mn}$ minimizing the contribution to the higher frequency components of the Discrete Cosine Transform (DCT) of the whole image.

The other possibility is to store the watermark in the spectral domain. The image can be transformed through well-defined procedures (Discrete Fourier Transform, DCT, Wavelets, etc.). The coefficients can be analyzed and modified according to some strategy, and the inverse transformation will produce a very similar image, but now with some extra information embedded. We denote as $\alpha$ as the vector holding the watermark, and we will assume that its elements are drawn at random from a gaussian pdf with zero mean and unit variance. The method by (Cox *et al.* 1997) suggested modifying just the largest coefficients of the DCT with the transformation:

$$c_i' = c_i + \varepsilon.\alpha_i \quad i = 1..n; n < N^2$$

being $c_i$' the new coefficient, $c_i$ the original one, ε a small scaling factor, $\alpha_i$ the i-th term of the watermark and n the length of the watermark. The alternative to modify the least significant terms of the DCT does not survive the JPEG compression, so it has not been further considered in the literature. The idea is that, if the watermark should not be evident, the changes should be small. However, small changes are badly affected by noise, except if they are concentrated in the most perceptual significant terms of the spectrum. Their method belongs to the class of *spread spectrum techniques.* Operationally, the image is decomposed in tiles of size NxN, and the watermark is applied independently to either all or selected tiles. The watermarked image is recreated through the inverse DCT transform. To recover the watermark, the DCT transform of the original image is usually required, in order to verify the relationship

$$\frac{c_i^{'} - c_i}{\varepsilon} = a_i'$$

This poses a serious threat to the data owner, because he needs to expose his original either in court or before to verify ownership. Notice that for some methods, once the original or at least the key is available, the attacker can create new instances without the watermark. On the other hand, the watermark is more robust for those methods that rely on the availability of the original at the detection stage. To solve this, (Zeng and Liu 1999) proposed an alternative method that does not require the original image, but has less robustness. If the correlation between α and α' is larger than a given threshold, the watermark is claimed to be detected. This approach is robust against some typical valid transformations, like JPEG compression, and also printing+scanning attacks. However, in its straightforward implementation version, it is amenable to protocol attacks (to be described later).

Two other important issues are the maximum lengths of the watermark, and how many watermarks can be reliably stored in a given image. It is customary to measure the strength of a cryptographic key in bits; larger keys imply stronger security. If the length of the key is small enough, it can be discovered by brute force with an exhaustive trial of keys, and once found, the removal is easy. Smaller watermarks can be easily removed, provided the watermarking algorithm is known. Long watermarks are also required in order to *uniquely* identify the owner, customer, etc. They might be difficult to produce and insert, because there are limits to be honored. (Servetto *et al.* 1998) assumed that the attacks can be modeled as additive noise, and derived upper bound formulas for the length in bits. The estimation of the number of different watermarks that can be stored in the same image is also a difficult problem, and as before relies heavily on assumptions about the noise.

Taking advantage of the intended use, there exist some algorithms that rely in the limitations of the human vision system, like the ones reported by (Podilchuk and Zeng 1998) or (Delaigle *et al.* 1998). For many applications, an image can be transformed and still be useful until the changes become noticeable for humans. One example is loosy image compression (algorithms which degrade the original image quality in order to achieve higher compression ratios than would be otherwise possible). Visual models provide a set of thresholds that describe the *Just Noticeable Differences (JND)* that can be perceptually detected. If the modifications are below such thresholds, the watermark can be strong but still unnoticeable. As mentioned before, the usefulness of JND based models for Geospatial datasets is however limited to manual photo-interpretation applications; they are useless where automatic processing is involved.

The author is not aware about current procedures from major data producers of (for example) satellite data. They sell the image under a contract that precludes the buyer for further redistribution, use, etc. of the material, but no information is given about any further protection apart from the contract. The reasons might be in the legal side, which will be treated in Section 8. The watermarking of digital still images is an active market; there are a number of commercial providers (Digimarc Inc., Blue Spike Inc., Signum Technologies, SysCoP, etc.). It is difficult to draw any conclusive statement about which is better than the other, because some of the companies do not provide detailed information about the algorithm embedded in their applications. A functional comparison can be made, however, analyzing the resistance of the watermark to different attacks. (Kutter and Petitcolas 1999) proposed a benchmark for comparison purposes, connected to the characteristics of the human visual system using the StirMark software, described by (Petitcolas and Kuhn 1997), to perform the attacks.

Finally, an interesting application is watermarking the output from the GIS system. If we agree that a typical output is in raster format (a digital or printed map, a processed image, etc.) it might be possible to insert information about the date, producer, customer, license information of the software, etc. which can be recognized by appropriate (usually digital) means. The implications are huge. We can trace back to the watermark of the sources used, tracking theft through their products. If the map has complex texture and color in the background, it might be treated with the same procedures already presented for general images. If no background color exist, only methods valid for vector data could be applied, and they will presented below. An interesting intermediate example is a categorical map: it holds just a few different categories organized in large contiguous areas. Most methods that apply the watermark in the spatial domain badly affects either texture or color rendering it trivially noticeable. In addition, they can be easily removed by lossy image compression. This situation is not unique to cartography: the same problem is observed with cartoon images. The alternative is to store the watermark in the spectral domain, as proposed by (Su *et al.* 1999). The author proposed a procedure based upon independent wavelet decomposition in each YUV channels. Detection does not require the original, unwatermarked image. Other weak alternatives exist, however, which rely in the selection of the color palette, and are applied typically to GIF images.

## 4. The case of the vector datasets

It is surprising that, despite the large costs associated to the collection and assembling of vector datasets, the “copy protection means” has not caught interest from the GIS research community. In fact, the author is not aware for any procedure specifically designed/applicable to 2D vector data. The closest area is the creation of 3D mesh models for Virtual Reality and CAD applications. In some cases, the dataset elements have been *designed* (from building blocks) and not *observed* (through 3D scanners, for example). We emphasize here the dataset format and not the source itself. Since some ideas can be borrowed from there, we will summarize the most relevant references now.

Virtual Reality Modeling Language (VRML) scenes are becoming increasingly popular in the Internet. They are composed of audio samples, textures and background images, and 3D geometry (model) based data. The most expensive part to develop is the last one, which should be also the target for an aspiring forger. (Ohbuchi *et al.* 1997); (Benedens 1999) and others claim that fortunately it is also the one more likely to hold the watermark. Notice that the VRML standard allows inserting information in the file, through comments and annotations. Format converters however, easily strip them out, so they are useless for watermarking purposes.

One important characteristic of the 3D model is that it lacks from an implicit ordering. Audio, video and still images are sequences of time series. Of course that vertices, edges and faces in a 3D model can be ordered, but maybe requiring an orientation frame and an origin defined in advance. A second characteristic is that no visually unique representation of the model exists. It can be modified, for example, by moving vertices to large amounts without significant change in the overall visual quality. To be rendered at reasonable speeds, 3D models are usually compressed through simplification as presented by (Garland 1999). In such process, they might loose even 86% of their faces without noticeable changes. This explain why is customary to store the same watermark more than once in the 3D model, allowing for recovery even under splitting of the model.

(Ohbuchi *et al.* 1997) discusses alternatives for watermarking 3D meshes. For example, coordinates of points and vertices can be modified to embed data, modifying scalar or vector quantities (like the area of a triangle, or the normal to a surface). However, some simple transformations can destroy it, so it is interesting to consider just the quantities that are invariant under a class of geometrical transformations. A hierarchy of transformations is established, as presented in Table 1. A second possibility is to embed the watermark in the topology, taking advantage of its lack of uniqueness. For example, given four vertices forming a square, they can be converted to two triangles in two different ways. Thus, one bit of information can be stored depending on the position of the diagonal. This approach can survive a geometrical transformation, but not a topological modification or remeshing.

1) Altered by almost any transformation
    a) Coordinates of points
2) Invariant to rotation and translation
    a) Length of a line
    b) Area of a polygon
    c) Volume of a polyhedron
3) Invariant to rotation, translation and uniform scaling
    a) Two quantities that define a set of similar triangles (e.g., two angles)
    b) Ratio of areas of two polygons
4) Invariant to affine transformation
    a) Ratio of lengths of two segments of a straight line
    b) Ratio of the volumes of two polyhedrons
5) Invariant to projection transformation
    a) Cross-ratio of four points on a straight line

Table 1 Alternatives for embedding information in the geometry of a 3D model (from Ohbuchi *et al.* 1997)

The class of expected valid geometrical transformations for the case of Geospatial datasets is more restricted. In some cases, Geospatial datasets have either absolute coordinates, or local coordinates linked with some reference system. In any case, substantially changing the coordinates might render a useless dataset, because it will not fit with others using the same original system. Thus, an appropriate watermarking system for GIS dataset might lack robustness against transformations 4 and 5 without penalty. However, coordinates might be known with uncertainty (which should not be confused with limited machine precision). This imply that changing randomly its given values by an amount below the uncertainty will produce a semantically equivalent dataset, and thus giving room to store a watermark. Despite simple, this procedure is amenable to many attacks.

To be useful, the watermark information should have enough bits, requiring a number of primitives to store it, and an order (explicit or implicit) among them. (Ohbuchi *et al.* 1998) considers three possibilities: a) there is a global arrangement in the primitives, b) there is a local arrangement or c) there is no arrangement, but subscript information is also encoded with the primitive. Examples of local or global arrangements are 1D sequences generated by sorting triangles according to their areas, and 2D arrangements of embedding primitives based on the connectivity of triangles in an irregularly tessellated triangular mesh. Global arrangements tend to have higher information density than the other methods. Local arrangements and subscript arrangements have the advantage that the watermark might be robust to a resection of a model, because the same part of the watermark can be repeatedly embedded in the mesh.

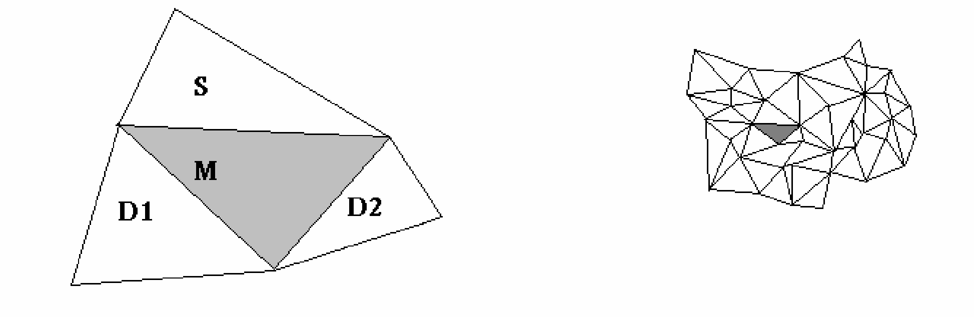

Figure 2 Basic unit for Triangle Similarity Quadruple Embedding

To illustrate this, we will show how might work the subscript arrangement. First, a set of four triangles sharing one side (as presented in Fig. 2 left) is identified in the mesh (Fig. 2 right). The gray one will be used as the reference and its shape will be modified in order to denote that it is part of the watermark by forcing its two smaller inner angles to be (for example) 33 and 57 degrees. Then we modify briefly the one denoted with S to have 20 and 60 degrees, encoding the number 3 according to a previously agreed lookup table. This number is the subscript, and indicates that the information corresponds with the 3rd element of the watermark. Triangles D1 and D2 store the information itself with the same lookup table, holding D1 the first element because its area is larger than the one of D2. To recover the whole watermark, we look for all triangles with internal angles of "exactly" 33 and 57 degrees, and that share each side with just one triangle. Using the lookup table, we identify the subscript, the information from D1 and D2, recovering one element of the watermark at a time. The same is repeated until all possible subscripts are found.

According to the author, for 3D models the changes in the original geometry can be minimal, and unnoticeable for humans. The set of three triangles plus the gray one cannot share vertices with other similar sets. In addition, triangles with too small inner angles should be avoided, because they are very unstable under even simple geometrical transformations. To reduce the risk of missing parts of the watermark, the same information is stored many times, allowing that even splitted, the model still will hold most of the watermark. If multiple copies for the same subscript element are found, a simple majority voting is performing. The watermark can be destroyed by randomization of coordinates, by a more general class of geometrical transformations, or by extensive topological alterations like remeshing. One interesting feature is that the original model is not required to recover the watermark. Another important property is that, given the model and the watermark, the exact original model cannot be derived. This might have implications regarding the ownership protocol.

(Benedens 1999) presented also a spatial domain method that stores the watermark in the surface normals of the model. He argues that such elements are somewhat persistent in the model even under moderate modifications affecting the geometry. He maps surface normals onto the unit sphere, and then modified the location of certain vertices altering the surface normal's distribution. His procedure requires a large amount of extra information to recover the watermark; in addition, it requires a somewhat precise reorientation of the model.

Since previous methods modify the coordinates, they can be classified as a spatial domain ones. Like for images, there are also spectral domain methods, like the ones presented by (Date *et al.* 1999) or (Praun *et al.* 1999). In the first reference, the wavelet transform of the 3D model has been used to represent it at different resolutions, allowing easy compression and efficient rendering. The spread spectrum principle can be applied to the coefficients of the wavelet transform, as it has been presented before for images. As before, to recover the watermark the original unwatermarked model is required. In the second paper, the authors used a different but otherwise equivalent function basis decomposition. A registration process is necessary to recover the watermark, implying not only similar orientation and rescaling, but also producing a mesh with the same connectivity as the original.

For GIS application, there is also a need for integrity verification, a problem only considered by (Yeo and Yeung 1999) again for the case of 3D models. In short, the verification problem is how to prove to a legitimate owner that the file he has is in fact the one produced by the producer. For this application, the watermark is like a hash function: any modification of the 3D model has devastating effects on the watermark, which will no longer be present in the file.

We have discussed so far the state-of-the-art regarding robust watermarking for 3D models, and pointed out the similarities that can be used while implementing a 2D system. This is a yet unexplored research area, and specific procedures should be derived for Geospatial datasets. In the spatial domain a number of alternatives might be implemented by carefully perturbing the coordinates; some of them might be easily destroyed as well. The spectral domain concept is general, and on principle it can be implemented for 2D datasets. However, geometric accuracy is an issue in GIS data and not so much in 3D models, which will be rendered and analyzed by humans.

## 5. Other data types: text and databases

This section is concealed with two important data formats not considered before: text and databases. The output of many GIS tasks or queries can take the form of a written report, a certificate, a map, etc. all of which should be considered as part of the GIS system. The state of the art shows that the efforts are also at their early stages, and far to be considered mature. In addition, despite that the goals and assumptions of the researchers might not be readily compatible with GIS needs, we feel that some words are worth considering anyway.

Text poses some difficulties to be watermarked, because of the limited possibility to include unnoticeable changes. (Low *et al.* 1998), (O'Gorman 1993) as well as (Brassil et al. 1995) consider the problem of marking an image version of the formatted text (like a fax, a photocopy or a bitmap file) and not the ASCII sequence of characters. The large areas blank areas preclude the use of general image watermarking techniques. If the text should be protected once printed, the most popular solution is to slightly move up or down an entire row (lets say, 1/300th inch) encoding a "1" or a "0", or shifting certain words left or right. According to (Low and Maxemchuk 1998), such watermarks can be detected even in the tenth photocopy, a process that introduces the most noise in the image. Of course, the watermark will not survive to a retyping of the document, a not too heavy task that today can be automated with character recognition software. The technique can be applied, however, to digital versions of the document. A postscript file can be watermarked, and distributed in bitmap form. The original document is assumed to have equally spaced rows, so usually it is not required to detect the watermark. For word shifting the original document (either as a formatted file or an image bitmap), or more specifically, the spacing between words is required to detect the watermark. Word shifting is less noticeable for the reader than line shifting, but in turn, it is more affected by noise.

Yet another possibility is Feature Coding, which stands for changing slightly some characteristics of individual letters. For example, every occurrence of letter "t" can be modified (coding a 1) or not (coding a 0) by altering upward the length of the letter by one pixel. Clearly the technique require an enlarged set of fonts, which might be a nuisance for the producer. All three techniques are amenable to collusion attacks plus sophisticated digital image processing; the ultimate protection relies in the assumption that unwatermarking the document should be more expensive than obtaining a legitimate copy.

In all three approaches, digital image processing techniques are required in order to analyze the given document. The elimination of the skew is a crucial step, and (Brassil *et al.* 1995) analyzed various means to do that. Further steps including detection of line and word shift. (Low *et al.* 1998) analyzed and compared the available techniques.

Databases are a little trickier to protect, since the order, field names, etc. are trivially easy to change, leaving no obvious place to hide information. A possible, naive approach is to introduce formally valid but otherwise wrong records in the database, and trace back the source through them. This is particularly easy if the database has a field for names of individuals, because an artificial foreign one can be forged and introduced there. Also some misspelled ones can be used. One significant drawback is that this procedure deliberately introduces errors, lowering the accuracy of the dataset. The watermark can survive a format translation, but is completely vulnerable to collusion attacks. This technique was used in an important cased analyzed by the U. S. Supreme Court (*Feist Publications vs. Rural Telephone Service Co., 499 U.S. 340)* under a trial regarding copyright ownership of a telephone directory (an example of text database). The case itself is more important because of its consequences, and will be analyzed later.

Other databases hold numeric records of limited precision, which can be modified to some extent without making them wrong. For example, an online service providing GPS reference data might identify records requested by registered customers exploiting the Selective Availability (SA) feature. SA is a system used by the US government to degrade the signal coming from the satellite constellation by a pseudo random algorithm seeded with a secret key. The introduced variance can be estimated, and the database can in turn be also altered with another pseudo random algorithm (now know by the service provider) but keeping the mean value and standard deviation equal to the recorded levels. Each customer will receive a uniquely watermarked database.

# 6. Protocols

We have presented so far different techniques designed to hide information in digital datasets. This is just the technological side, and we need to consider also the legal side and the protocol side. All three parts interact, and some valid solution for one of them can be useless if not considering the whole picture. We here define protocol as a procedure that unambiguously produces evidence of facts. Following the common practice in cryptography, let's denote as Alice the originator of the image and Bob as an aspiring forger. We will use the case of images as an example, but the conclusions are valid for more general cases. We will consider just two problems and propose suitable protocols. The first one is how to prove that Alice is the owner of a given image. The second one is how to prove that the seller cannot create many identical copies of the same image, sell one to Bob and later accuse him to illegally redistribute his copy. Notice that one possible strategy for Bob is to attack the watermarked image obtained from Alice and remove the watermark from it. We will assume that the watermarking schema used is robust, and thus we will discard such option, but concentrate in other weakness associated with the rest of the process. Both cases will illustrate the crucial role of the protocol in the intended use of the watermarking as a tool, but in no way they are the single protocols to consider. Other protocols have been devised with special goals, and this is a very active subject for research.

## 6.1. Protocol and issues involved in proving ownership

As pointed out by (Craver *et al.* 1998), the mere existence of a watermark is not enough to prove ownership in a court. Some characteristics of the watermarking scheme as well as the choice of the key should be taken into account. Let's consider one possible scenario, where both Alice and Bob are claiming ownership of a watermarked image $\hat{I}$, in which the watermarks of both Alice and Bob have been detected. If Alice keeps her original image and watermark locked away, she can ask Bob for his fake original, and test whether or not it has her watermark. Bob can do the same, of course. As (Craver *et al.* 1998) shows, the ownership decision will be taken based on the results of Test I (Alice and Bob both test an image for presence of their watermark) and Test II (Alice and Bob both test each other's "original" images for presence of a watermark). The possible results of the tests are summarized in Table 2, where $x_A$ and $x_B$ checks for the existence of Alice's and Bob's watermark respectively.

| Scenario | Test I | | Test II | | Derived Ownership |
|---|---|---|---|---|---|
| | $x_A$ | $x_B$ | $x_A$ | $x_B$ | |
| Case 1 | 1 | 0 | d | d | Alice |
| Case 2 | 0 | 1 | d | d | Bob |
| Case 3 | 1 | 1 | 1 | 0 | Alice |
| Case 4 | 1 | 1 | 0 | 1 | Bob |
| Case 5 | 1 | 1 | 1 | 1 | ??? |

Table 2 Determination of ownership from watermark presence tests. “1” indicates the presence of watermark, “0” indicates the absence, and “d” represents *don’t care*’s (from Craver *et al.* 1998)

The Case 5 is the most interesting, because it shows a potential ownership deadlock. If Bob is able to engineer a process in order to make appear his watermark in the original (and locked away!) image hold by Alice, this situation will arise. The watermark is still in the images, but its existence does not allow the legitimate owner (Alice) to conclusively prove that she is so. (Craver *et al.* 1998) showed that this inverse engineering is possible for some of the most popular robust watermarking methods. In the absence of a central watermarking authority, non-invertibility of the watermark schema is necessary to avoid this ownership deadlock. (Craver *et al.* 1998) proposed some enhancements that made more difficult for Bob to provoke the Case 5 by using signatures derived from hash functions. (Ramkumar and Akansu 1999) extended such approach by increasing the computational complexity of such attack by a factor over $10^{100}$ thus making it virtually attack proof.

### 6.2. A Buyer-Seller Protocol

(Memon and Wong 1998) considered the following problem. Assume that Alice is the owner or reseller of an image, and Bob is a legitimate buyer. If the watermark inserted by Alice in the copy delivered to Bob is found in unauthorized copies, Alice can claim that Bob is guilty. However, Bob can claim in turn that Alice has created such copies, or that they arise after a security breach in Alice’s system. The authors propose a protocol to allow Alice to insert a unique watermark in the image, but which preclude her to know which exact watermark is inserted. The protocol involved a watermark certification authority that generates random watermarks on request, and the existence of a Public Key Infrastructure (PKI). Some features of the public-key cryptosystem are also required. In short, Alice inserts an encrypted version of the watermark (encrypted using the public key of Bob), producing an image which in turn could only be decrypted by Bob himself using his private key. The protocol has some disadvantages despite it protocol solves the problem. One is the need of a watermark authority (supplier of valid watermarks) and a PKI. There are also some mathematical requirements on the PKI and the *watermarking insertion* operator which restrict somewhat the applicability of the schema in general.

## 7. Legal aspects

Once analyzed the technological means to include hidden information in datasets, and the protocols required to unambiguously prove ownership, what is required is proper legal protection to punish the pirate. We will consider briefly the situation in this area in particular to geospatial datasets, and comparing the current status in the US and the European Union (EU). A throughout review of this topic is outside from our goals, and the reader should go to specific literature

It has been customary to protect the product with either a specific contract between customer and supplier, a patent, some technological mean (like those available for DVD), or under the general umbrella of copyright protection. Geospatial datasets are, however, not clearly amenable to copyright protection. According to (Thorner 1997) copyright protection is allowed for any "original works of authorship fixed in any tangible medium of expression". The important thing is the need of *originality*. The mere compilation of entries (like the ones of a telephone directory) is not original, and thus it is not amenable to copyright protection. This concept is accepted by international agreements, like those arising from the Berne Convention of 1886. As (Karjala 1995) points out, copyright law has explicitly protected maps. However, the paradox is that maps apparently contain only unprotectable elements.

This principle of originality has been made evident in the US by the Supreme Court in 1991, in the case *Feist Publications vs. Rural Telephone Service Co.* (for an extended report see http://laws.findlaw.com/us/499/340.html). Feist wanted to publish a unified directory of an area operated by eleven telephone companies, based upon their white page telephone directories. Feist negotiated and agreed a payment with most of them, but Rural Telephone refuses to do so. Since for their project the directory should be complete, Feist simply used the data available (from Rural) and included it in his product. Rural noticed that, and thus litigated against Feist. Feist accepted that they have copied the contents, and despite that, the Supreme Court understood that Rural couldn't claim originality in a mere compilation of a list of phone numbers, addresses and names. Other cases after Feist followed the same trend, limiting the protection for compilations.

In addition, (Onsrud and López 1998) assert that if there are a single or few ways to describe a feature, it cannot claim for copyright protection in the US. For example a line of constant elevation above a standard and widely used datum collected and expressed at a standard and widely used mapping scale exhibiting a reasonable degree of accuracy for many practical purposes cannot be protected by copyright, because any other person or sensor attempting to represent these physical facts would have little choice but to do so in much the same way. Another typical situation is exemplified by the compilation of the daily value of the NASDAQ index, which is *unique*. Every compiler must produce *identical* databases. Under the US doctrine, such databases cannot be protected by copyright. An equivalent doctrine is not clearly established in the EU.

These arguments show that GIS database producers have a potentially unstable base to fill a suit based upon copyright arguments. The current alternative in the US is issuing a patent (a workaround only considered by Thorner, but with little support) or to rely in protection by contracts. The latter is considered also a very weak protection mean (at least in the US), because it provide little protection against third parties that gain access to the data but have no contractual relation with the supplier. This situation highlights the need for a special, different in concept but similar in its operation, protection umbrella for compilations of facts. The EU has lead that way by the adoption of the Database Directive in 1996 (CEC 1996). According to (Band 1999), under this regime, a second-generation publisher could not extract a substantial part of a first-generation database, even if the second publisher did not extract any protectable expression. The protection is also granted to non-EU producers, but only on a reciprocity basis. For example, a producer based in the US might receive similar protection in the EU only if the US grants an equivalent level of protection to EU producers.

Some authors suggest that the time the database is protected might be shorter than for traditional copyrighted products. For example, (Onsrud and López 1998) stated that protection for a mere compilation in the EU is for 10 years, while for original work is 70 years. Both (Karjala 1995) and (Thorner 1997) suggested that this period should be even shorter, because factual compilations tend to be obsolete in shorter times. Thorner suggested that a two-year period is enough.

The above mentioned legal protection is useful once you are in a court. However, it is worth mention other complementary alternatives, related with technological measures. They have been illustrated with the DVD case. They cannot be used alone, because on principle there exist also technological workarounds, making it useless. In the US, the Digital Millennium Copyright Act (DMCA) explicitly protects this type of measures, by prohibiting the manufacture and sale of devices that can circumvent technological protection measures.

The watermarking scheme itself is also protected, even at the international level. The World Intellectual Property Organization (WIPO) since 1996 explicitly protects content owners against attempts *"...to remove or alter any electronic rights management information..."*. The limitations also include *"...to distribute, import for distribution, broadcast or communicate to the public, without authority, works or copies of works knowing that electronic rights management information have been removed or altered without authority..."*. Fredricsson (2000) pointed out that the *"rights management information"* needs not to be the author name, but also numbers or codes, covering appropriately the case of digital watermarks.

## 8. Conclusions

In this paper we have addressed the problem of protecting authorship and other commercial rights over GIS databases against unauthorized copy. In the past, Geospatial datasets were usually presented and stored in analog media, so they were difficult to copy without expensive equipment. Today the widespread availability of computers makes piracy a significant threat to digital Geospatial datasets, usually collected under a cost-sharing basis. Despite not complete, there exist good solutions to at least part of the business. Cryptographic tools have solved secure transfer between producer and consumer, because encrypted datasets can be copied, but they are unusable without the key. The problem arises once the dataset is properly decrypted, because perfect copies can be made with growing simplicity. Internet distribution makes even easier the dissemination of stolen copies.

Thus, the current trend is to embed some secret information within the dataset without noticeable disturbances. The ancient science that deals with this problem is Steganography, and the information is denoted as watermark. The watermark can be recovered by appropriate software, secret keys and extra data, and used in a court to prove ownership.

The protection is achieved by the combined use of watermarking techniques, legal protection and suitable protocols. For this purpose, watermarking techniques should survive some malicious attacks attempting to remove or damage the watermark. Legal protection is mandatory; we can show that the watermark is in the dataset, but no defense is possible without legal support. Protocols are standard procedures that both buyer and seller should follow in order to protect each one rights, and produce evidence of facts. If properly applied, and provided there exists legal protection, we have shown that there exist suitable technical procedures that on principle can protect our rights (either as a customer or a producer).

We analyzed the techniques in use for still images, 3D vector datasets, text and databases. None of them are completely immune to *attacks* devoted to either remove the watermark, or to degrade it significantly. The legal protection has also been analyzed; in the past, only the copyright could be argued. Only recently the European Union has specifically considered also databases. The main conclusion is that not all Geospatial datasets are amenable to be protected under the copyright umbrella, because Geospatial datasets typically hold a collection of facts, and facts themselves are not amenable to copyright protection.

In order to complete the scene, suitable protocols must assure the rights of the consumer and the supplier. For example, they should assure that: a) a reseller cannot produce many identical copies of the same watermarked dataset, sell one and distribute the others, and later accuse the recipient for make the extra copies. b) Formally valid but forged watermarks could not suddenly appear in the well-locked owner’s dataset, raising doubt regarding who is the true owner. Other protocols might be required for other circumstances; they differ if a Public Watermarking Authority is available or not, etc.

Geographic Information Systems uses and produces many datasets. We have suggested that both inputs and outputs can be properly modified in order to carry secret information, which might be used to prove ownership or to trace output to the sources. We have disguised some traditional methods, like introducing tiny errors in the dataset; see (Monmonier 1996) for a complete list of tricks. Such procedures degrade the geometric accuracy, the attribute accuracy or both and are unacceptable; in fact, no map publisher will admit that their maps have deliberate errors. Using watermarking, we can even create unique but very similar instances of the same dataset incorporating a serial number, customer identification, etc. In the foreseeable future, appropriate GIS software might recognize such elements and even check rights against a server for use the dataset.

## 9. Acknowledgments

This work has benefited from the very active exchange of ideas and material at the watermarking mailing list operated by Martin Kutter. The author also wishes to thank Mrs. Verónica Oliveira-Mattos for carefully translate to spanish part of the material originally in swedish, and Ana Inés Mora for carefully locating papers through the Internet.

## 10. References

Anderson, R. J. and Petitcolas, F. A. P., 1998, On the limits of Steganography. *IEEE Journal of Selected Areas in Communications*, 16, **4**, 474-481

Band, J., 1999, Making the World Safe for Databases. *IP Magazine*, April 1999 http://www.ipmag.com/monthly/99-apr/band.html

Bender, W.; Gruhl, D.; Morimoto, N and Lu, A., 1996, Techniques for data hiding. *IBM Systems Journal*, **35**, 3-4, 313-336

Benedens, O., 1999, Geometry-Based Watermarking of 3D models. *IEEE Computer Graphics and Applications.* 46-55

Brassil, J., Low, S., Maxemchuk, N. and O'Gorman, L., 1995, Electronic marking and identification techniques to discourage document copying. *IEEE J. Select. Areas Commun.*, **13**, 8, 1495-1504

Commission of the European Communities (CEC), 1996, Council Directive on the Legal Protection of Databases. 96/9/EC COM(95) 382 Final. Doc Num 31996L0009 *Official Journal of the European Communities of 27/3/1996* no. L 77 p. 20-28, Brussels http://europa.eu.int/smartapi/cgi/sga_doc?smartapi!celexapi!prod!CELEXnumdoc&lg=EN&numdoc=31996L0009&model=guichett

Cox, I. J. and Linnartz, J. P. M. G., 1998, Some General Methods for Tampering with Watermarks. *IEEE Journal on Selected Areas in Communications*, **16**, 4, 587-593.

Cox, I. J.; Kilian, J.; Leighton, T. and Shamoon, T., 1997, Secure Spread Spectrum Watermarking for Multimedia. *IEEE Trans. on Image Processing*, **6**, 12, 1673-1687

Craver, S.; Memon, N.; Yeo, B. L. And Yeung, M., 1998, Resolving rightful ownerships with invisible watermarking techniques: limitations, attacks and implications. *IEEE Journal on Selected Areas in Communications*, **16**, 4, 573-586

Date, H.; Kanai, S. And Kishinami, T., 1999, Digital Watermarking for 3D polygonal model based on wavelet transform. *Proceedings of DETC'99 – 1999 ASME Design Engineering Technical Conferences. September 12-15, 1999, Las Vegas, Nevada.* DETC99/CIE-9031, 10pp. http://minf.coin.eng.hokudai.ac.jp/members/kanai/asme-seq.pdf

Delaigle, J. F.; Vleeschouwer, C. D. and Macq, B., 1998, Watermarking algorithm based on a human visual model. *Signal Processing*, **66**, 3, 319-335.

Fredricsson, P., 2000, Vattenmärkning i digitala miljöer. (Watermarking in digital media). *Unpublished manuscript in swedish*.

Garland, M., 1999, Multiresolution Modeling: Survey & Future Opportunities. *EUROGRAPHICS'99*, ISSN 1017-4956, 111-131

Gopalakrishnan, K; Memon, N. and Vora, P., 1999, Protocols for Watermark Verification. *Proceedings of the Multimedia and Security Workshop (held as a part of the 7th annual ACM International Multimedia Conference at Orlando, Florida in Oct 1999)*, GMD Report No. 85, Dec. 1999, 91-94.

IEEE Multimedia, 8, 4, 66-70 http://csfac.csci.ecu.edu/faculty/gopal/galley.pdf

Karjala, D., 1995, Copyright in electronic maps. *Jurimetrics J.*, **35**, 395-415

Kutter, M. and Petitcolas, F. A. P., 1999, A fair benchmark for Image watermarking systems. *P. W. Wong and E. J. Delp, Eds. Security and Watermarking of Multimedia Contents, ISBN 0-8194-3128,* 226-239 proceedings of *Electronic Imaging '99*, *Security and Watermarking of Multimedia Contents*, vol. 3657, pp. 226–239, San Jose, California, U.S.A., 25–27 January 1999. The Society for Imaging Science and Technology (I.S.&T.) and the International Society for Optical Engineering (S.P.I.E.). ISSN 0277-786X. ISBN 0-8194-3128-1. http://www.cl.cam.ac.uk/~fapp2/publications/ei99-benchmark.pdf

Kutter, M.; Jordan, F. and Bossen, F., 1998, Digital signature of color images using amplitude modulation. *Journal of Electronic Imaging*, **7**, 2, 326-332

Kutter, M.; Voloshynovskiy, S. And Herrigel, A., 2000, The Watermark Copy Attack. *In Proceedings of SPIE: Security and Watermarking of Multimedia Content II*, Vol 3971, San José, CA, US, 10 pp.

Low, S. H. and Maxemchuk, N. F., 1998, Performance Comparison of Two Text Marking Methods. *IEEE Journal on Selected Areas in Communications*, **16**, 4, 561-572

Low, S. H., Maxemchuk, N. F. and Lapone, A. M., 1998, Document identification for copyright protection using Centroid detection. *IEEE Trans. Commun.*, **46**, 372-383

Memon, N. and Wong, P. W., 1998, A Buyer-Seller Watermarking Protocol. *IEEE Workshop on Multimedia Signal Processing (MMSP-98), Dec. 7-9, Los Angeles, California, US*, 278-283

Mintzer, F. C.; Boyle, L.E.; Cazes, A. N.; Christian, B.S.; Cox, S.C.; Giordano, F.P.; Gladney, H. M.; Lee, J. C.; Kelmanson, M. L.; Lirani, A. C.; Magerlein, K. A.; Pavani, A. M. B. and Schiattarella, F., 1996, Towards on-line, worldwide access to Vatican Library materials. *IBM Journal of Research and Development*, **4**, 2, http://www.research.ibm.com/journal/rd/mintz/mintzer.html

Monmonier, M., 1996. How to Lie with Maps. (2nd edition ISBN 0226534200. University of Chicago Press, Chicago). 207 pp.

Nikolaidis, N. and Pitas, I., 1996, Copyright protection of images using robust digital signatures. *IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP-96)*, **4**, 2168-2171

Nikolaidis, N. and Pitas, I., 1998, Robust image watermarking in the spatial domain. *Signal Processing*, **66**, 385-403

O'Gorman, L., 1993, The document spectrum for structural page layout analysis. *IEEE Trans. Pattern Analysis and Machine Intelligence*, **15**, 11, 1162-1173

Ohbuchi, R.; Masuda, H. and Aono, M., 1997, Watermarking Three-Dimensional Polygonal Models. *ACM Multimedia 97*, ACM Press, 261-272

Ohbuchi, R.; Masuda, H. and Aono, M., 1998, Watermarking Three-Dimensional Polygonal Models through Geometric and Topological Modifications. *IEEE J. on Selected Areas in Communications*, **16**, 4, 551-560

Onsrud, H. J and López, X., 1998, Intellectual property rights in disseminating digital geographic data, products and services: conflicts and commonalities among European Union and United States approaches. In *European Geographic Information Infrastructures: Opportunities and Pitfalls*. Masser, I. and Salgé, F., Eds. Taylor and Francis. 153-167

Petitcolas, F. and Kuhn, M., 1997, StirMark 2. http://www.cl.cam.ac.uk/~fapp2/watermarking/stirmark/

Podilchuk, C. and Zeng, W., 1998, Image Adaptive Watermarking Using Visual Models. *IEEE Journal on Selected Areas in Communications*, **16**, 4, 525-540

Praun, E.; Hoppe, H. and Finkelstein, A., 1999, Robust mesh watermarking. *Computer Graphics (SIGGRAPH 1999 Proceedings)*, 69-76. Also available at http://www.cs.princeton.edu/gfx/proj/meshwm.

Ramkumar, M. And Akansu, A. N., 1999, Image Watermarks and Counterfeit Attacks: Some problems and Solutions. *Content Security and Data Hiding in Digital Media, Newark, NJ*, 102-112

Ramkumar, M., 1999, Data hiding in Multimedia – Theory and Applications, *Ph. D. Dissertation, New Jersey Institute of Technology, Department of ECE, University Heights, Newark, NJ 07032*, 68 pp.

Schneier, B. 1995, Applied Cryptography: Protocols, Algorithms and Source Code in C. John Wiley & Sons, ISBN: 0471117099

Servetto, S. D.; Podilchuk, C. I. and Ramchandran, K., 1998, Capacity issues in Digital Image Watermarking. *In Proc. of the IEEE Int. Conf. on Image Processing. Chicago, IL*. 5pp. http://www-ima.enst.fr/~maitre/tatouage/icip98/ma11_05.pdf

Su, P. C.; Wang, H. J. and Kuo, C. C. J., 1999, Blind Digital Watermarking for Cartoon and Map Images, *In Proc. SPIE Vol. 3657, p. 296-306, Security and Watermarking of Multimedia Contents, Ping Wah Wong; Edward J. Delp; Eds.*

Thorner, B. B., 1997, Copyright Protection for Computer Databases: The Threat of *Feist* and a Proposed Solution. *Virginia Journal of Law and Technology*, **1**, 5. http://vjolt.student.virginia.edu/graphics/vol1/home.html

van Schyndel, R. G.; Tirkel, A. Z. and Osborne, C. F., 1994, A digital Watermark. *Proc. of the IEEE Int. Conf. on Image Processing (ICIP´94)* vol II, 86-90 http://goanna.cs.rmit.edu.au/~ronvs/papers/ICIP94.PDF

Voyatzis, G. and Pitas, I., 1999, Protecting Digital-Image Copyrights: A Framework. *IEEE Computer Graphics and Applications*, 19, **1**, 18-24

Yeo, B. L and Yeung, M. M., 1999, Watermarking 3D Objects for Verification. *IEEE Computer Graphics and Applications*, 36-45

Zeng, W. and Liu, B., 1999, A statistical watermark detection dechnique without using original images for resolving rightful ownerships of digital images. *IEEE Trans. Image Processing*, **8**, 11, 1534-1548

Zhao, J.; Kock, E. And Luo, C., 1998, In Business Today and Tomorrow. *Communications of the ACM*, **41**, 7, 67-72