Ministère de l'Enseignement Supérieur et de
la Recherche Scientifique
Université Abdelhamid Ibn Badis - Mostaganem

Faculté des Sciences Exactes et de l'Informatique
Département de Mathématiques et d'Informatique
Filière : Informatique

Mémoire de Fin d'Etudes
Pour l'Obtention du Diplôme de Master en Informatique
Option : **Systèmes d'Information Géographique**

Thème :

Protection des cartes géographiques par le tatouage numérique

Etudiant(e)s : « **OULD MILOUD Mohamed** »

« **REFES Nassim** »

Année Universitaire 2016/2017

Résumé

Le tatouage permet l'insertion robuste et discrète d'information dans un document, comme par exemple l'identité de son propriétaire. Les buts recherchés sont multiples, mais le plus important est la protection du droit d'auteur. On dissimule dans le document des informations relatives à son propriétaire pour que personne ne puisse se l'approprier. Le tatouage est également utilisé pour identifier les utilisateurs malhonnêtes et éviter la piraterie. En effet un propriétaire d'un document qui veut commercialiser son œuvre (par exemple un logiciel) peut marquer différemment tous les exemplaires autorisés (les différentes copies qu'il a vendues). S'il rencontre une copie illégale, le marquage lui permet de remonter à la personne responsable de la fraude. Ce travail considère la protection des cartes géographiques par le tatouage numérique. Le travail présente tout d'abord un état de l'art sur le tatouage numérique et ainsi qu'une étude bibliographique sur le tatouage des cartes géographiques et les techniques présents dans la littérature puis il présente en détail le schéma de tatouage implémenté dans l'application. Une carte géographique est une représentation conventionnelle, généralement plane, de phénomènes concrets ou même abstraits, mais toujours localisables dans l'espace géographique. Le tatouage numérique est une technique permettant d'ajouter des informations de droit d'auteur ou d'autres messages de vérification à un document. Actuellement se développent des techniques permettant de tatouer des données structurées où l'insertion d'information doit préserver la qualité d'un certain nombre de requêtes déclarées au préalable. L'objectif de cette étude est d'étendre ces résultats aux données géographiques comportant des informations textuelles habituelles et l'information géométrique.

Mots clés: Tatouage, Droit d'auteur, Piraterie, Protection, Donnés géographiques.

Abstract

The watermarking allows insertion and information in a document, such as the identity of its owner. The goals sought are multiple, but the most important is the protection of copyright. The information contained in the document is concealed from the owner so that no one can take ownership of it. The tattoo is also used to identify dishonest users and avoid piracy. Indeed, an owner of a document who wants to market his work (for example software) may mark all the authorized copies differently (the different copies he has sold). If it encounters an illegal copy, the marking allows it to trace back to the person responsible for the fraud. This work considers the protection of geographical maps by digital watermarking. The work presents first a state of the art on digital tattooing and a bibliographic study on the tattooing of maps and techniques present in the literature then he presents in detail the watermarking scheme implemented in the application. A geographical map is a conventional, usually plane representation of concrete or even abstract phenomena, but always locateable in geographical space. Watermarking is a technique for adding copyright information or other verification messages to a document. Currently, techniques for tattooing structured data are developed, where the insertion of information must preserve the quality of a certain number of previously declared requests. The objective of this study is to extend these results to geographic data with usual textual information and geometric information.

Keywords: Watermarking, Copyright, Piracy, Protection, Geographical data.

Sommaire

Remerciements……i

Résumé ...……ii

Liste des abréviations...…… iv

Liste des figures …….....v

Liste des tableaux …….vii

Sommaire ……...…...……..viii

Introduction générale..… …….....1

Chapitre I Problématique générale du tatouage numérique

I.1 Introduction……3

I.1.1 Cryptographie et Stéganographie ……..……...3

I.1.2 Historique……..……...4

I.2 Principe général d'un système de tatouage numérique ……5

I.2.1 Définition ……...5

I.2.2. Points clés d'un algorithme de tatouage ……....5

A) La capacité (payload) ou ratio ……...6

B) L'invisibilité …….....6

C) La robustesse …….7

I.3. Techniques du tatouage numérique …….....8

I.3.1. Classification selon le type d'algorithme …….8

I.3.1.1. Modes d'extraction du tatouage …….....8

I.3.1.2. La clé appliquée : Asymétrique et Symétrique 9

I.3.3.1. Authentification ……..……...11
I.3.3.2. Protection des droits d'auteur ……..11
I.3.3.3. Transmission secrète …….....12
I.4. Les méthodes d'attaques ……..……...12
I.4.1. Méthodes d'attaques courantes en traitement d'image ……..……12
I.4.1.1. Attaques géométriques……..……12
I.4.1.2. Attaques d'effacement ……...……...13
I.4.2. Méthodes d'attaques de nature cryptographique……..……...14
I.4.3. Méthodes d'attaques de protocoles…….....14
I.5. Conclusion…….....15

Chapitre II Tatouage des cartes géographiques

II.1 Introduction……...…….16
II.2 Données géographiques……..16
II.3 État de l'art du tatouage de données géographiques……...17
II.3.1 Méthodes basées sur des modifications géométriques……..17
II.3.1.1 Méthodes basées sur le déplacement de sommets……18
II.3.1.2 Méthodes basés sur l'ajout de sommets……18
II.3.2 Système de tatouage à grande capacité pour les cartes numériques…...18
II.3.3 Tatouage des cartes vectorielles 2D dans le domaine spectre-maillage .18
II.3.3.1 Algorithme de tatouage ……19
II.3.3.2 Algorithme de détection …….....20
II.3.2 Schémas basés sur le tatouage d'objets 3D……...21

II.4.4 L'ajout de bruit gaussien…….27

II.4.5 Tatouage aveugle ou non…….27

II.4.6 Tatouage par transformée ou par modification géométrique…….27

II.5 Conclusion……..28

Chapitre III Algorithme de tatouage

III.1 Introduction……...29

III.2 Données géographiques considérée …….29

III.3 Précision du document …….29

III.4 Sommets et arêtes du document…….....29

III.5 Préservation de la qualité …….29

III.6 Préservation de la triangulation ……30

III.7 Transformations appliquée sur le document…….31

III.7.1 Réorganisation des données ……..32

III.7.2 Découpage de la carte…….32

III.7.3 Transformations géométriques …….....33

III.8 Schéma aveugle ……33

III.9 Schéma 0-bit ……34

III.10 Idées directrices……...34

III.11 Préservation locale de la qualité……..34

III.12 Présentation du schéma de tatouage……35

III.13 Définition des sites……..35

III.14 Préservation de la qualité des sites…….....36

III.15 Détails de l'algorithme de tatouage ……36

III.17 Algorithme de détection…….....41
III.18 Conclusion……...42

Chapitre IV : Démonstration de l'application

IV.1- Introduction ……43
IV.2 Qu'est ce que GDW (Geographical Data Watermarker) …….....43
IV.3 Choix de l'algorithme …….....43
IV.4 Choix de l'environnement ……43
IV.5 L'interface graphique ……..44
IV.5.1 Menus ……45
A) Menus Fichier ……..45
B) Menus Affichage ……46
C) Menus Aide……..46
IV.5.2 Barre d'outils…….....46
IV.5.3 Barre de statuts……...47
IV.5.4 Les panneaux ……47
IV.6 Déroulement d'un test……...48
IV.7 Conditions expérimentales …….....53
IV.7.1 Dispositif expérimentale……53
IV.7.2 Détection de la marque……...53
IV.8 Conclusion……58

Introduction Générale

L'arrivée de l'ère numérique à la fin du 20^{ieme} siècle, a permis un accès à l'information beaucoup plus facile que par le passé. Les documents numériques sont devenus de plus en plus utilisés puisque leur diffusion est peu coûteuse et extrêmement rapide, même ceux de grand volume, grâce aux réseaux informatiques et aux supports numériques. Cette extraordinaire révolution technique de l'analogique vers le numérique ne s'est pas faite sans engendrer des inquiétudes puisque n'importe qui peut facilement copier, modifier et distribuer les documents numériques sans risque de les détériorer. Il est très difficile de trouver un compromis entre le libre accès à l'information et le respect des droits d'auteurs. Des personnes peuvent s'approprier un document numérique pour faire des profits aux dépends des personnes légitimes ayant les droits initiaux. Elles peuvent aussi changer facilement son contenu.

À cause de l'utilisation illicite des documents numériques, les propriétaires des documents numériques, ayant les droits initiaux, ont été touchés par une forte baisse de leur chiffre d'affaires. Pour cela, ils sont motivés plus que jamais à les protéger par de nouvelles techniques qui empêchent la duplication, la modification et la distribution des documents numériques. Dans cette optique, la cryptographie est utilisée comme une technique pour sécuriser les documents numériques. Une clé secrète permet de crypter l'information de document en clair et une clé publique est employée pour décrypter (schéma asymétrique) le message chiffré, permettant ainsi de signer et donc d'authentifier un document. Cependant, une fois décryptés les documents numériques ne possèdent aucune protection.

Pour pallier ce problème, une nouvelle technique a été introduite, s'inspirant principalement de la cryptographie et la stéganographie. Cette technique, nommée tatouage numérique, en anglais digital watermarking, a fortement émergée depuis le début des années 1990. Elle consiste à inscrire dans un document numérique une marque invisible ou, dans certain cas, une marque visible. Dans le cadre de la protection des droits d'auteurs, la marque insérée doit être la plus robuste possible contre les attaques malveillantes ou même bienveillantes. En général, la présence de la marque doit être imperceptible.

Depuis quelques années, ce domaine connaît, un développement phénoménal et le besoin d'une solution efficace et optimale est devenu primordial puisque les documents

Le 1er chapitre présente les notions générales du tatouage et ses domaines d'applications et une étude bibliographique sur ses différentes techniques.

Cependant le 2eme chapitre présente un état de l'art sur le tatouage des cartes géographiques avec l'étude des différentes techniques existant dans la littérature.

Le 3eme chapitre est consacré a présenter un schéma de tatouage des donnés géographiques vectorielle, aveugle et robuste tout en détaillons le schéma et précisant chacune de ces étapes.

Le 4eme chapitre a pour but de présenter l'application implémentée et les résultats expérimentaux obtenus.

Liste des abréviations

XML	**(**Extensible Markup Language) Langage de balisage extensible
GPS	**(**Global Positioning System) Système mondial de positionnement
IGN	Institut géographique national
SIG	Systèmes d'Information Géographique
OGC	Open Geospatial Consortium
DFT	Transformées discrètes de Fourier
DCT	Transformées en cosinus discrètes
DWT	Transformées en ondelette discrètes
ESRI	Environmental Systems Research Institute
ACI	Action Concertée Incitative
GREYC	Groupe de recherche en informatique, image, automatique et instrumentation de Caen
CNAM	Conservatoire national des arts et métiers
3D	Trois dimensions
2D	Deux dimensions
LSB	(*Least significant bit*) Bit de poids faible
JPEG	(Joint Photographic Experts Group) Une norme qui définit le format d'enregistrement et l'algorithme de décodage pour une représentation numérique compressée d'une image fixe

Liste des figures

Figure I.1 – Dispositif générique d'un système de tatouage d'image……..6

Figure I.2 – Compromis à réaliser en tatouage d'image…….7

Figure I.3 – Les techniques de tatouage numérique……8

Figure I.4 – Illustration des déformations géométriques aléatoires engendrées par Stirmark..13

Figure I.5 – Illustration d'un lissage appliqué sur une image……...14

Figure II.1 Carte originale a tatouée…….19

Figure II.2 Les sommets sont triangulés pour créer un maillage…….19

Figure II.3 Subdivision de la carte avec la méthode k-d-tree……...20

Figure II.4 BD orthophotographie numérique……..23

Figure II.5 Image cartographique numérique géoréférencées……...23

Figure II.6 Représentation d'une carte scannée……24

Figure II.7 Représentation d'une carte d'altitude…….24

Figure II.8 Types de données géométriques…….25

Figure II.9 Représentation de deux couches qui contienne des informations de type géométriques et attributaires…….26

Figure II.10 Représentation d'une BD Corine Land Cover……..26

Figure III.1 Document original…….30

Figure III.2 Nuage de points associé…….31

Figure III.3 Triangulation de Delaunay……31

Figure III.4 Document original…….32

Liste des figures

Figure IV.4 Carte chargé……...49
Figure IV.5 Morceau sélectionné……..49
Figure IV.6 Les paramètres configurés. ……...50
Figure IV.7 Résultat du tatouage……..50
Figure IV.8 Carte tatoué……51
Figure IV.9 Zoom sur les sites bougés……..51
Figure IV.10 Choix du seuil de détection…….52
Figure IV.11 Résultat de la détection……52
Figure IV.12 Sélection d'au moins 100 sommets…….....54
Figure IV.13 Sélection d'au moins 200 sommets…….....55
Figure IV.14 Sélection d'au moins 300 sommets…….....56
Figure IV.15 Sélection d'au moins 400 sommets (carte zoomé)……57

Liste des tableaux

Tableau IV.1 Résultat de la détection sur le corpus de test (88 documents)……..4

Tableau IV.2 Résultats sur au moins 100 sommets……..48

Tableau IV.3 Résultats sur au moins 200 sommets……..49

Tableau IV.4 Résultats sur au moins 300 sommets……..50

Tableau IV.5 Résultats sur au moins 400 sommets……..52

I.1. Introduction

I.1.1. Cryptographie et Stéganographie

La cryptologie existe depuis des siècles. Depuis l'invention de l'écriture, le besoin de sécurité est motivé par les problèmes de confidentialité et d'intégrité : on souhaite éventuellement que l'information écrite ne soit accessible qu'à certaines personnes et qu'elle ne soit pas modifiée volontairement dans un but de mystification. La cryptologie regroupe à la fois la cryptographie, qui désigne l'art de chiffrer le contenu d'un message susceptible d'être intercepté lors de sa transmission, et la cryptanalyse, qui consiste à casser le code protégeant un message chiffré. Depuis son origine, où elle était principalement réservée à un usage militaire et diplomatique, la cryptologie a considérablement évolué, notamment avec l'apparition de l'ordinateur, et s'étend aujourd'hui au domaine civil pour la protection des données circulant sur les réseaux informatiques. Ainsi, la cryptologie moderne est maintenant une discipline de recherche publique de l'informatique théorique utilisant des outils mathématiques sophistiqués.

Le « watermarking » (littéralement filigrane) ou tatouage numérique peut être perçu comme une branche de la stéganographie. Le mot stéganographie vient du grec « steganos » (caché ou secret) et « graphy » (écriture ou dessin), et signifie littéralement « écriture cachée ». La stéganographie consiste à cacher, un message secondaire dans un message primaire. Le message primaire reste lisible de tous, tandis que le message secondaire n'est lisible que par une ou plusieurs personnes propriétaires d'une information secrète. Pour la petite histoire, les premières traces de stéganographie remontent à l'Antiquité. Un légataire romain voulant envoyer un message à César, le camoufla dans une amphore qu'il lui envoya en guise de cadeau. Une autre forme de stéganographie, elle aussi très rudimentaire, consistait à raser le crâne d'un esclave. On y tatouait alors le message, et l'esclave était envoyé lorsque ses cheveux avaient repoussés. Le destinataire n'avait plus qu'à le faire raser de nouveau pour faire apparaître le message. Une autre forme de stéganographie très connue est le principe de l'encre invisible. Cette technique était très utilisée au moyen âge pour envoyer des messages secrets. A l'époque, l'encre était fabriquée simplement à base de jus d'oignons et de chlorure d'ammoniac. L'écriture était alors rendue visible en approchant le papier d'une flamme de bougie.

La stéganographie se distingue de la cryptographie dans la mesure où l'objectif

vérification ou une extraction efficace et automatique de certaines informations liées à l'origine, au contenu ou même à la diffusion du document.

I.1.2. Historique

Les tatouages du papier sont apparus dans l'art de la fabrication du papier il y a presque 700 ans. Le plus ancien document tatoué trouvé dans les archives remonte à 1292 et a son origine dans la ville de Fabriano en Italie qui a joué un rôle important dans l'évolution de l'industrie papetière. A la fin du troisième siècle, environ 40 fabricants du papier partageaient le marché du papier.

La concurrence entre ces fabricants était très élevée et il était difficile que n'importe quelle partie maintienne une trace de la provenance du papier et ainsi que son format et sa qualité. L'introduction des tatouages était la méthode parfaite pour éviter n'importe quelle possibilité de confusion.

Après leur invention, les tatouages se sont rapidement étendus en Italie et puis en Europe et bien qu'au commencement utilisé pour indiquer la marque ou le fabricant du papier, ils ont servi plus tard pour indiquer le format, la qualité, et la force du papier, et ont été également employés comme une base pour dater et authentifier le papier [1, 2].

L'analogie entre le tatouage du papier et le tatouage numérique est évidente : les tatouages du papier des billets de banque et de timbres ont inspiré la première utilisation dans le contexte de données numériques. Les premières publications portant sur le tatouage de documents numériques ont été publiés par Tanaka et al. [3] en 1990 et par Tirkel et al. [4] en 1993.

En 1995, le temps est évidemment bien de prendre ce sujet, et il a commencé à stimuler l'augmentation des activités de recherche. Depuis 1995, le tatouage numérique a gagné beaucoup d'attention et a évolué très rapidement et alors qu'il y a beaucoup de sujets ouverts pour davantage de recherches, des méthodes de travail et des systèmes pratiques ont été développés et depuis, le nombre de publications et de brevets a fait du tatouage un domaine majeur en traitement de document numérique (voir tableau I.1).

I.2. Principe général d'un système de tatouage numérique

I.2.1. Définition

Le tatouage numérique, consiste à introduire, généralement de manière invisible, une information dans un document numérique, puis à tenter de la récupérer après que le document ait éventuellement subi des manipulations de natures variées.

Définition Kundur et Hatzinakos 1998

Le processus du tatouage numérique implique la modification des données multimédia originales pour insérer un watermark contenant des informations clés telles que le code d'authentification ou de droit d'auteur. La méthode d'insertion doit conserver les données originales visuellement inchangés, mais d'imposer des modifications qui peuvent être détectés à l'aide d'un algorithme d'extraction. Les types de signaux à tatouer sont des images, le son, vidéo et le texte [5].

Définition Petitcolas, Anderson et Kuhn 1999

Le tatouage numérique signifie l'intégration d'une information dans un document numérique de façon à ce que cette information soit imperceptible pour un observateur humain, mais facilement détectée par l'ordinateur. Le watermark est une information transparente, invisible qui est inséré dans un document source en utilisant un algorithme informatique [2].

Définition Christian REY et Jean-Luc DUGELAY 2001

Le tatouage numérique est une technique qui consiste à cacher dans un document numérique une information subliminale (i. e, invisible ou inaudible suivant la nature du

I.2.2. Points clés d'un algorithme de tatouage

Les principales contraintes techniques à prendre en compte pour concevoir un algorithme de tatouage performant sont les suivantes :

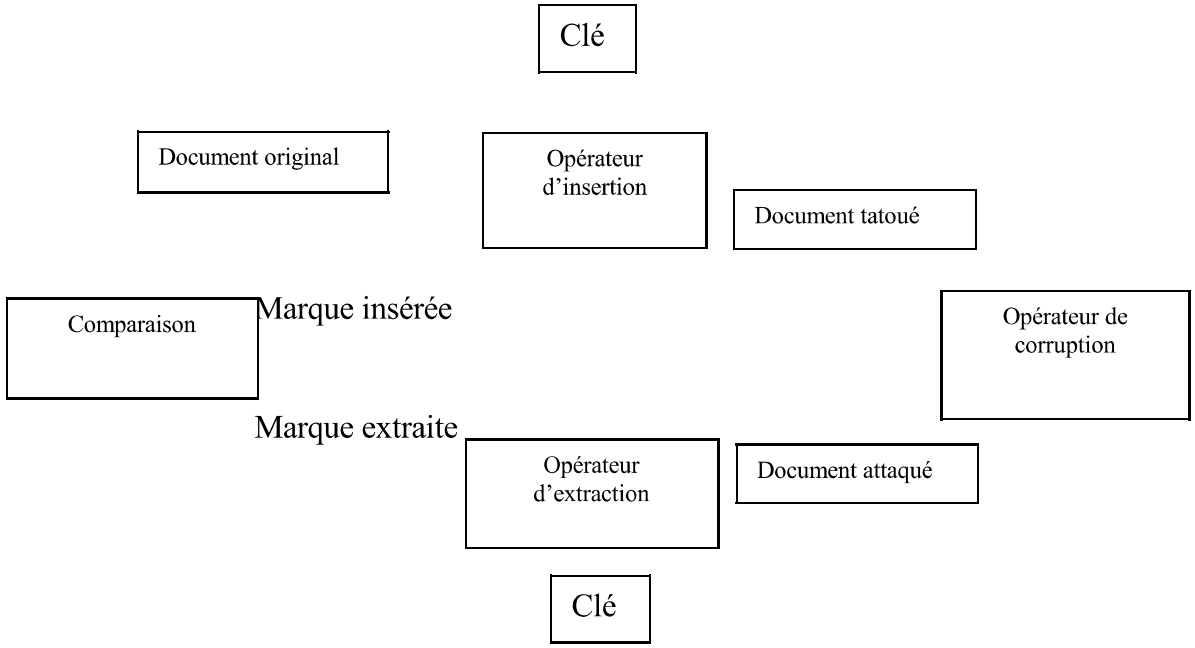


Figure I.1 – Dispositif générique d'un système de tatouage d'image.

A) La capacité (payload) ou ratio

C'est la quantité d'informations que l'on espère cacher par rapport à la quantité d'informations associée au support audio, image ou vidéo utilisé. Ordinairement de 16 à 64 bits sont suffisants pour assurer un service de droit d'auteurs à l'aide d'un identifiant, mais pas pour cacher des informations explicites comme un logo de société, assurer des services

Le but est de faire en sorte que l'impact visuel du marquage (i.e. distorsion) soit la plus faible possible afin que le document marqué reste fidèle à l'original (i.e. modifications imperceptibles). De nombreux algorithmes prennent d'ailleurs en compte un modèle psycho visuel [16].

C) La robustesse

Il s'agit ici de pouvoir récupérer la marque même si le document marqué a été manipulée. Il est nécessaire de distinguer plusieurs types d'attaques selon qu'elles sont considérées comme étant biens ou malveillantes, destructives ou non (en termes de dégradations visibles inacceptables et/ou d'utilisation commerciale rendue impossible). Les attaques bienveillantes regroupent les manipulations effectuées par un utilisateur de bonne foi. On trouve dans cette catégorie la compression, les conversions de format en général, les changements de résolution (zoom), etc. Il n'est pas possible d'énumérer l'ensemble des attaques bien ou malveillantes ne dégradant pas le document de façon significative mais qui néanmoins sont capables de « Lessiver » le document marquée afin de retirer la marque, ou plus simplement d'empêcher de l'extraire correctement. Il existe des logiciels libres spécialisés dans le lessivage : Stirmark [23], officiellement présentés comme logiciel d'assistance pour la mise au point et l'évaluation d'algorithmes de tatouage.

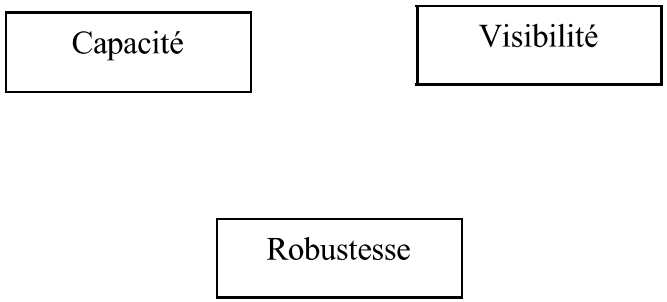


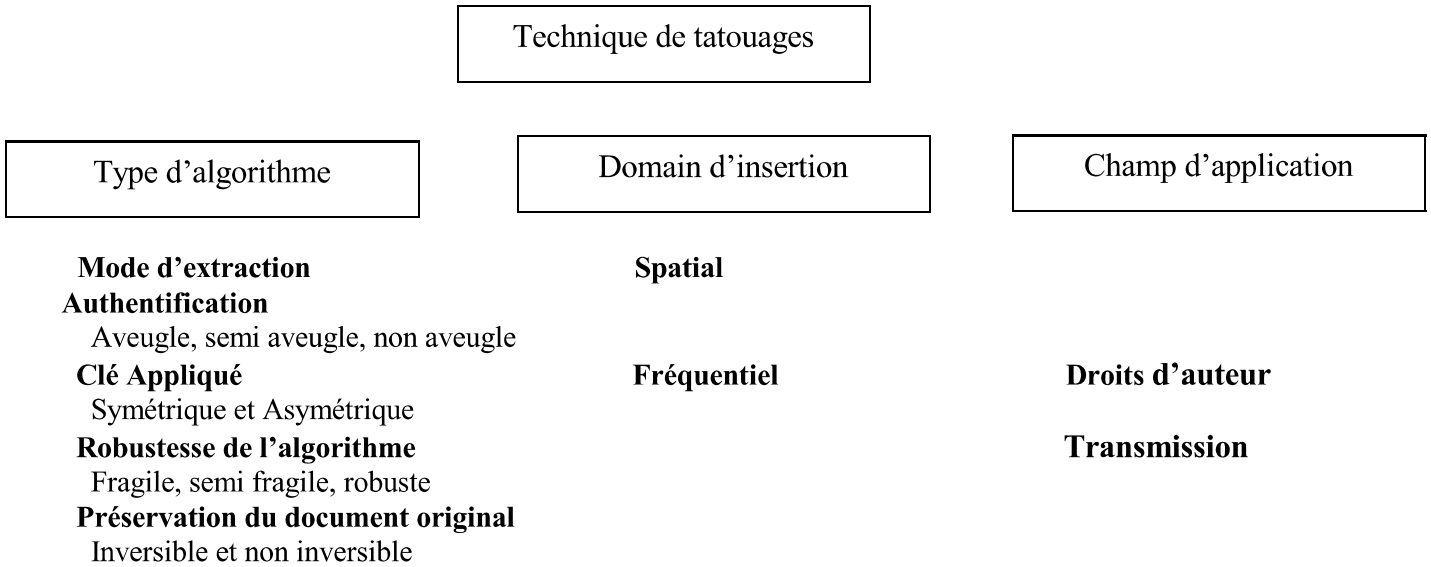
Figure I.2 – Compromis à réaliser en tatouage d'image.

l'intégrité des documents. De ce fait, si la marque est altérée ou lessivée, le document n'est plus considérée comme intègre.

I.3. Techniques du tatouage numérique

Cette section n'a pas pour objectif de dresser une revue exhaustive de toutes les techniques disponibles dans la littérature. Néanmoins, nous décrivons les grandes lignes de certaines catégories du tatouage numérique dans le but de montrer à quel point le sujet est vaste.

La technique de tatouage est présentée sur la base de plusieurs publications récentes. La raison de cet arrangement est de fournir une vue générale de plusieurs principaux domaines du tatouage. Nous avons choisi trois critères pour regrouper les techniques du tatouage : le type d'algorithme, le domaine d'insertion et le champ d'application (voir la figure I.3).



module d'extraction pour la vérification du tatouage. L'utilisation de tel ou tel mode dépendra de l'application visée et des protocoles utilisés.

- **Mode non-aveugle (ou tatouage privé):** Le récepteur dispose du document ainsi que du tatouage original. Ce contexte est bien évidement incompatible avec des applications visant à vérifier l'intégrité du document, ou à assurer la vérification en temps réel du copyright (problème de temps d'accès à la base de données contenant les informations originales). Il y a une branche de chercheurs qui disent que si le watremark est non-aveugle, alors elle peut être facilement attaquée. Néanmoins, nous trouvons des applications qui demandent que le watermark soit visible, c'est le cas du logo des sociétés dans les programmes télévisuels. Dans la catégorie du tatouage visible, nous distinguons les travaux [8,9, 10].

- **Mode semi-aveugle (ou tatouage semi-privé):** Le tatouage original est supposé connu lors de l'extraction et utilisé le plus souvent via un score de corrélation.
- **Mode aveugle (ou tatouage public):** C'est un concept beaucoup plus complexe. Le tatouage aveugle modifie le signal d'une manière imperceptible par l'utilisateur final. Pour reprendre l'exemple de l'agence de photo, les photos hautes résolutions vendues par l'agence possèdent elles au contraire un watermark invisible, qui ne dégrade donc pas le contenu visuel, mais qui permet de détecter l'éventuelle source d'un vol. Le message caché par le tatouage peut être un identifiant de l'acheteur par exemple. En cas d'utilisation non-autorisée, l'agence peut alors se retourner contre l'acheteur [9].

Le tatouage invisible est l'approche la plus développée qui attire la plupart des chercheurs [12, 13, 14]. Il s'agit du seul mode où l'on peut réellement parler d'extraction du tatouage (par opposition à la vérification intervenant dans les deux précédents modes) puisque l'on ne présuppose ni la connaissance du tatouage, ni la connaissance du document originale. C'est le mode d'extraction le plus intéressant, mais également le plus difficile à mettre en

l'extraire. Le rôle de clé-privée et clé-publique n'existe pas, l'insertion et l'extraction sont faites à l'aide de la même clé et procédure [15].

I.3.1.3. La robustesse de l'algorithme : Fragile, semi fragile et robuste

Dans le tatouage fragile, le watermark est fortement sensible aux modifications du document tatouée. Cette approche sert à prouver l'authenticité et l'intégrité d'un fichier tatoué [5]. Le tatouage semi-fragile a pour objectif de reconnaitre les perturbations mal intentionnées et de rester robuste à certaines classes de dégradations légères du document, comme la compression avec pertes par exemple. Diverses méthodes d'authentification de document de type image par tatouage semi-fragile ont été proposées [17]. Le tatouage robuste dispose d'un large champ de théories et de résultats. Celui-ci cherche à préserver les données cachées face aux attaques. Le watermark doit donc être suffisamment résistant aux attaques afin de rester identifiable [15].

I.3.1.4. Préserver la qualité du document

Le tatouage inversible permet de récupérer toutes les propriétés originales de document hôte après l'extraction du watermark [18]. La distorsion liée à l'insertion de la marque dans le document doit être la plus faible possible, de manière à ce que la qualité du document tatouée soit quasi identique à l'originale. Malheureusement cette notion d'invisibilité est très subjective et difficile à modéliser. D'autant plus qu'elle dépend de nombreux facteurs, tels que : la nature du document à tatouer (peinture, image médicale, photo satellite, etc.), la qualité du document original (ex: plus une image est de bonne qualité, plus il est difficile de garantir l'invisibilité de la marque) et des conditions de visualisation (problème bien connu en codage de source avec pertes). Dans le tatouage non-inversible, le document original est définitivement altérée par le mécanisme d'insertion du watermark.

I.3.1.5. Technique d'insertion

Dans le tatouage additif, le message à ajouter n'est pas corrélé au document hôte. La

Dans les techniques spatiales, le watermark est insère en modifiant directement les valeurs du document hôte. Ce sont des méthodes simples et peu couteuses en temps de calcul. Elles sont consacrées aux tatouages en temps réel demandés dans des environnements de faible puissance. Certaines techniques dans le domaine spatial peuvent être robustes aux attaques de type transformations géométriques [15].

La plus part des techniques spatiales sont basées sur l'addition d'une séquence pseudo-bruit d'amplitude fixe. Dans le cas des fonctionnalités utilisées, sont simplement des opérations d'addition et de soustraction, respectivement. Plusieurs méthodes, proposées dans la littérature, modifient les bits de poids faible LSB du document hôte de type image. L'invisibilité du watermark est obtenue par l'hypothèse que les données contenues dans les bits LSB sont visuellement insignifiantes. Le watermark est généralement inséré en utilisant la connaissance de la séquence pseudo-bruit (et peut être la connaissance d'une clé secrète, comme la location du watermark).

I.3.2.2. Domaine fréquentiel

Les méthodes présentées précédemment permettent en général de retrouver le watermark en faisant la différence entre l'image originale et l'image tatouée. Cela leur confère un sérieux désavantage : une personne qui voudrait attaquer ces documents et qui se serait procurée le document original, ou bien plusieurs personnes mettant en commun leurs documents tatoués peuvent détruire le watermark. Des algorithmes incluant le watermark non pas directement dans le document numérique (image, vidéo, son), mais dans une transformée du document seront à cet égard plus robustes, et permettent en plus de choisir les coefficients qui seront plus résistants a certains types d'attaques. Des schémas du tatouage peuvent effectuer l'insertion du watermark dans des espaces transformés. Un espace transformé est obtenu après l'emploi d'une transformée telle que : DCT, DFT, DWT, etc. Cette stratégie rend le watermark plus robuste à la compression, puisqu'elle utilise le même espace qui sert au codage du document. Contrairement au domaine spatial, le watermark inséré dans le domaine fréquentiel est très sensible aux transformations géométriques parce que ce genre de transformations modifie considérablement les valeurs des coefficients transformés [5].

et certains nous permettent de calculer une approximation du document originale dans les régions modifiées [15].

I.3.3.2. Protection des droits d'auteur

La protection des droits d'auteur a été une des premières applications étudiée en tatouage numérique. Ce service reste cependant toujours d'actualité et concerne encore la majorité des publications. L'objectif est d'offrir, en cas de litige, la possibilité à l'auteur ou au propriétaire d'un document d'apporter la preuve qu'il est effectivement ce qu'il prétend être, et ce même si le document concernée a subi des dégradations par rapport à l'original. La mise en place d'un tel service doit respecter les deux contraintes suivantes :

- **Garantir la non-ambiguïté de la preuve:** Le tatouage doit constituer une preuve irréfutable. Pour cela il convient d'assurer l'unicité (éviter les problèmes de collision) et de l'authenticité de l'identifiant, mais également de dater le dépôt (au cas où le document aurait été tatoué plusieurs fois avec des marques différentes). Pour cela, il convient de définir des protocoles stricts excluant toute ambiguïté.
- **Assurer la robustesse des éléments de preuve (tatouage):** L'algorithme utilisé doit être capable d'extraire correctement la marque cachée, même si le document a été manipulé.

I.3.3.3. Transmission secrète

On peut aussi tatouer un document, de façon aussi discrète que possible, dans le but d'échanger des messages secrets. Ceux-ci sont cachés dans le document, et nécessitent une clé secrète pour être décodés. Cette application a été formulée par Simmons [1] avec le problème du prisonnier, dans laquelle on suppose deux détenus dans des cellules séparées tentent de passer les messages secrets. Leur problème est qu'ils ne peuvent pas passer directement ces messages, mais plutôt, ces derniers doivent passer à travers le gardien de prison. Le gardien est prêt à transporter des messages inoffensifs entre eux, mais les punir s'il constate, par exemple, que leurs messages contiennent un plan pour s'enfuir. La solution consiste à déguiser le plan d'évacuation dans les messages inoffensifs. Il existe plusieurs programmes

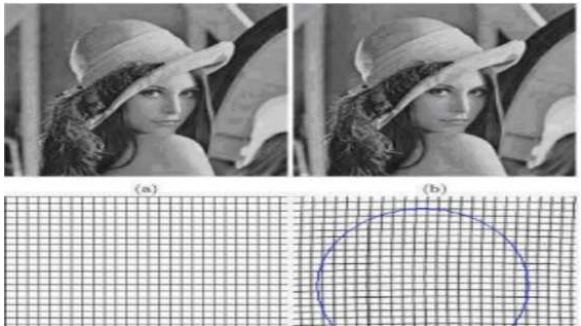
- Attaques d'effacement : visant à supprimer le watermark ;
- Attaques de nature cryptographique : visant à décrypter la clé secrète ;
- Attaques de protocoles : visant à trouver une faille dans le protocole de gestion des droits d'auteurs.

Les deux premières catégories d'attaques peuvent être considérées comme des attaques sur la robustesse, alors que les suivants sont des attaques sur la sécurité. De plus ces manipulations peuvent être aisément combinées entre elles de manière à créer des attaques plus complexes. Nous nous contenterons simplement de présenter brièvement les plus couramment utilisées.

I.4.1. Méthodes d'attaques courantes en traitement d'image

I.4.1.1. Attaques géométriques

Les manipulations géométriques, mêmes très simples, sont des attaques particulièrement sévères, face auxquelles beaucoup d'algorithmes de tatouage d'image se révèlent inefficaces, en particulier lorsque l'on impose une extraction en mode aveugle. En effet, pour la majorité des méthodes proposées, l'opérateur d'extraction a besoin de connaître la position exacte de la marque dans l'image. Or, les distorsions géométriques ont pour effet d'introduire une désynchronisation entre le tatouage contenu dans l'image et le détecteur. De ce fait, bien que la marque soit encore présente, les bits extraits ne correspondent plus à ceux qui ont été cachés.



- **Les symétries axiales (horizontale et verticale)** : elles ne sont pas forcément décelables si l'image présente naturellement un axe de symétrie ou aucune information textuelle, mais peuvent suffire à piéger l'algorithme d'extraction, si celui-ci ne les prend pas en compte.
- **Suppression de lignes et de colonnes** : ces manipulations sont généralement invisibles, mais suffisent à créer un décalage significatif pouvant rendre difficile l'extraction du tatouage.

I.4.1.2. Attaques d'effacement

Les attaques d'effacement les plus évolués sont :

- **Compression JPEG :** La compression JPEG est une technique de compression avec pertes qui supprime les informations redondantes des images dont le but de diminuer la taille du fichier image. Comme le watermark est invisible, il peut donc être considéré comme non significatif et donc aussi être supprimé.
- **Filtrage :** les filtre les plus utilisés sont : filtre médian, filtre gaussien, filtre laplacien et filtre de moyenne.
- **Rehaussement et lissage :** Le rehaussement correspond à l'augmentation des composantes hautes fréquences de l'image. L'image devient alors plus contrastée. Le lissage est l'opération contraire du rehaussement, il attenue les composantes hautes fréquences de l'image qui devient alors plus floue. Ces opérations peuvent modifier également les composantes hautes fréquences du message et leur faire perdre leurs particularités.

I.4.2. Méthodes d'attaques de nature cryptographique

Moins courantes, ces attaques suivent le modèle des attaques classiques en cryptographie. Certaines de ces attaques, telles que les « Brute Force Attacks », ont pour objectif de découvrir la clé secrète utilisée pour insérer la marque, en essayant de manière exhaustive toutes les clés possibles. Bien évidemment, ce genre d'attaque est très coûteux en temps de calcul, et n'est réellement efficace que sur des algorithmes utilisant des clés de petite taille. L'attaque oracle, quant à elle, est plus spécifique aux algorithmes de tatouage asymétriques. Si un pirate dispose du décodeur public, il peut appliquer de petites modifications successives à l'image jusqu'à ce que le décodeur ne décèle plus de marque. De cette façon, il a l'assurance de ne pas avoir dégradé l'image plus que nécessaire. Cette attaque, suggérée pour la première fois par Perrig [21], a depuis fait l'objet de nombreuses études, et des analyses théoriques ainsi que des contre-mesures possibles ont été récemment publiées.

I.4.3. Méthodes d'attaques de protocoles

Les attaques de protocoles se distinguent des autres familles d'attaques, dans la mesure où leur but n'est pas de détruire ou d'empêcher la détection de la marque par des manipulations du document. Ces attaques s'en prennent directement aux protocoles de l'application elle-même. Une des premières attaques de ce type a été proposée par Craver *et al.* [22]. Les auteurs introduisent la notion de tatouage inversible, et montrent que pour assurer certains services de sécurité, il est impératif d'utiliser des tatouages non inversibles. Ce qui signifie en terme de tatouage de document qu'il ne doit pas être possible d'extraire une signature depuis un document qui n'a pas été tatouée.

I.5. Conclusion

Bien que le tatouage de document numérique soit un domaine relativement jeune, les acquis disponibles en codage de source, codage canal, cryptographie et théorie de

II.1 Introduction

Comme on a vu dans le chapitre précédent qui a introduit le vocabulaire des différentes classes de schémas de tatouage numérique.

Dans ce chapitre, nous abordons la problématique du tatouage des cartes géographiques pour le but de la protection. Nous présentons les données géographiques considérées ainsi qu'une sélection de travaux de ce domaine.

II.2 Données géographiques

On appelle donnée géo référencée un ensemble de données géométriques et de données descriptives utilisées dans une application en géomatique. Les données géométriques renseignent sur la position et la forme d'une entité. Les données descriptives sont relatives aux attributs des entités. Si l'on traite une base de données routière, chaque tronçon de route est une entité. La géométrie de cette entité représente le tracé de la route. Les attributs peuvent être, par exemple, le nom de la route, le nombre de voies et sa vitesse maximale autorisée [41].

De nos jours, ce type d'information est de plus en plus utilisé. Les outils tels que les navigateurs GPS basent leurs calculs sur les cartes vectorielles. Les sites internet tels que Mappy, Googlemap et le site de l'IGN Géoportail fondent leur valeur ajoutée sur la pertinence de l'information géographique sur laquelle ils basent leurs calculs.

Les SIG (Systèmes d'Information Géographique) facilitent énormément l'accès et les traitements sur les données géographiques. Ces exemples sont très loin d'être exhaustifs et l'on voit de plus en plus de sites internet et de programmes qui utilisent la donnée géographique, afin de tirer parti des GPS sur les portables par exemple [41].

Ces données sont rassemblées par des organismes spécialisés tels que l'IGN. Dans ce cas, des équipes de professionnels recueillent ces données à partir de photos aériennes et de relevés terrain. Ce processus demande du temps et énormément de ressources humaines. C'est ce qui donne de la valeur à la base de données construite et qui explique que cette donnée est

organismes qui les ont collectées. Par conséquent il est pertinent de poser le problème de la protection de la propriété intellectuelle sur les données géographiques. Nous pensons que le tatouage de données est une façon efficace de lutter contre ce problème. En effet, il devient plus facile pour un propriétaire de faire preuve de ses droits devant un tribunal en utilisant le tatouage. Nous pensons que cela peut contribuer à décourager la revente et la dissémination illicite des données géographiques.

II.3 État de l'art du tatouage de données géographiques

Les données géographiques peuvent prendre plusieurs formes. Il peut s'agir, entre autres, d'images bitmap provenant de satellites (mode raster) ou de données vectorielles (mode vecteur) provenant de relevés sur le terrain. Cet état de l'art ne présente que des méthodes de tatouages appliquées aux données vectorielles. Ce type de données, que l'on retrouve par exemple dans les navigateurs GPS, est spécifié par le consortium OGC (Open Geospatial Consortium) qui regroupe un grand nombre des acteurs majeurs du domaine.

Une carte géographique est constituée d'un ensemble d'objets géographiques où chacun peut représenter par exemple une route ou une ville. Chaque objet est constitué d'un ensemble de propriétés que l'on nomme aussi données descriptives et d'un ensemble de géométries. Les données descriptives peuvent être, pour une ville, sa population, le nombre d'actifs, le nombre d'usines dans la ville, etc. Les géométries représentent l'emprise de l'objet sur la terre et sont composées de trois types géométriques primitifs : le point, la ligne et le polygone. Un document est donné avec un système de projection et une précision qui représente l'écart maximal entre un point dans le monde réel et son image dans la base de données [41].

Notons que beaucoup de travaux de tatouage de la littérature ne prennent pas en compte la notion d'objet géographique. Dans ces schémas, les documents géographiques vectoriels sont vus en tant que cartes vectorielles. Cela supprime toute notion d'objet géographique et de géométries et les auteurs travaillent sur un graphe constitué d'un ensemble de sommets et d'arêtes entre ces sommets.

Pour présenter cet état de l'art du tatouage de données géographiques nous avons choisi dans un premier temps de présenter les méthodes basées sur des transformées. Nous

II.3.1.1 Méthodes basées sur le déplacement de sommets

H. Kang [26] découpent le document en zones de surfaces égales. Dans chaque zone, on utilise un masque pour sélectionner un certain nombre de sommets. Chaque zone est ensuite découpée en deux en suivant la droite qui passe par le point le plus au sud-ouest et celui le plus au nord-est parmi ceux qui sont contenus dans le masque. On obtient ainsi deux ensembles de sommets (les sommets du nord-ouest et ceux du sud-est) qui permettent de coder un bit d'un message chiffré. Pour coder un bit à 0, on déplace les sommets appartenant au premier ensemble vers leur position symétrique par rapport à la droite. Pour coder un bit à 1, on effectue la même opération mais avec les sommets appartenant au second ensemble. Lors de la phase de décodage, on retrouve les deux ensembles de sommets. On lit un bit à 0 ou à 1 suivant l'ensemble majoritaire. On obtient ainsi un message que l'on peut comparer avec le message chiffré original. Le schéma est aveugle, et permet d'introduire un message dans le document. Il est robuste à l'ajout de bruit et à la suppression de sommets.

II.3.1.2 Méthodes basés sur l'ajout de sommets

Certains schémas préservent les positions des sommets. Ils se contentent d'ajouter des sommets le long des arêtes. Utilisant ce principe et une mise en œuvre pour ArcView, le logiciel de visualisation et de traitement de données géographiques d'ESRI. L'algorithme de tatouage sélectionne l'arête la plus longue du polygone.

Des sommets sont ajoutés par interpolation le long de l'arête. La distance entre deux sommets consécutifs d'une polyligne permet de coder un bit. La distance qui permet de coder un bit à 1 est paramétrée par l'utilisateur. Pour coder un bit à 0, on prend la moitié de cette distance.

L'algorithme de détection est trivial. On peut noter que l'algorithme assure la confidentialité du message grâce à un chiffrement par une clé secrète. Des codes correcteurs d'erreurs permettent de retrouver le message si celui-ci a été un peu altéré. Ce schéma est aveugle et permet d'inclure un message dans le document. Il est robuste à la translation et la rotation.

II.3.3 Tatouage des cartes vectorielles 2D dans le domaine spectre-maillage [35]

Ces algorithmes altèrent les coordonnées ou la connectivité des sommets pour le tatouage. Dans l'article de **(R.Ohbuchi, H. Ueda, S. Endoh)** proposent un algorithme de tatouage de cartes numériques vectorielles. L'algorithme insère le tatouage dans la représentation du domaine fréquentiel des cartes.

II.3.3.1 Algorithme de tatouage : L'algorithme consiste à insérer un message binaire dans la carte en modifiant les coefficients du domaine fréquentiel de la carte comme décrit ci-dessous.

Figure II.1 Carte originale a tatoué[35].

- On extrait les points des polygones ;

Pour le calcul de la représentation du domaine fréquentiel, l'algorithme établit une

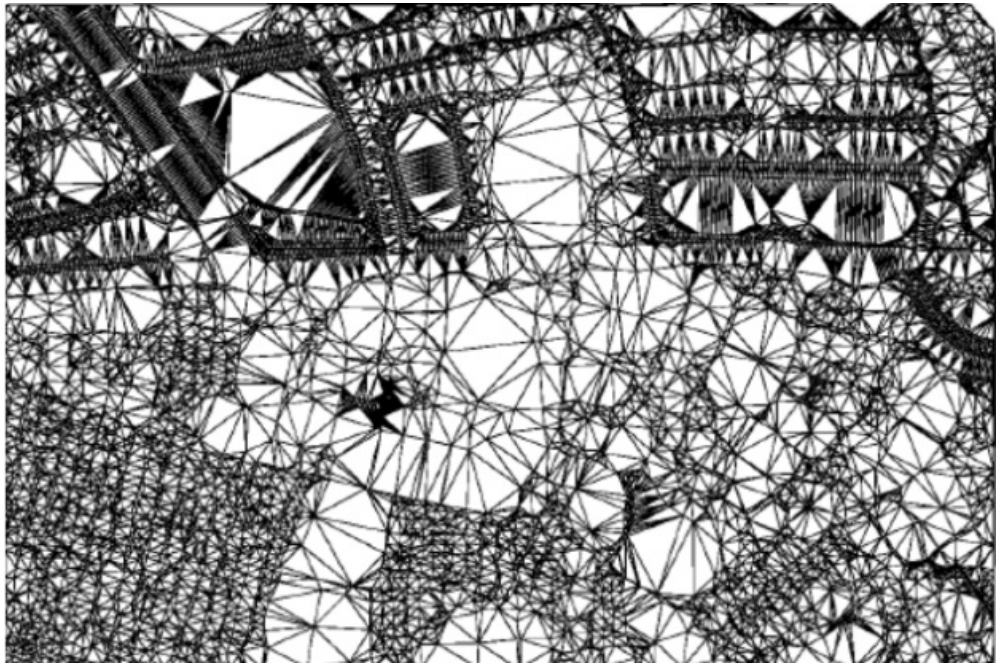


Figure II.2 Les sommets sont triangulés pour créer un maillage[35].

- Pour des raisons d'efficacité et de robustesse, la carte est divisée en plusieurs surfaces rectangulaires équilibrées du point de vue nombre de sommets contenus dans chaque zone, la méthode utilisée est la k-d-tree subdivision [26].

Figure II.3 Subdivision de la carte avec la méthode k-d-tree [35].

- La maille est ensuite transformée en domaine fréquentiel en utilisant l'analyse spectrale de maille proposée par Karni et al[27].

- Un message est inséré dans chacune de ces surfaces, en modulant (modifiant) les fréquences ainsi obtenues dans chaque surface.

- Les tatouages sont extraits en comparant la carte (éventuellement attaquée) à la carte originale, ce qui conclut que c'est un algorithme non aveugle.

II.3.3.2 Algorithme de détection

- D'abord les deux cartes (originale et attaquée), sont alignées en utilisant un processus d'optimisation itératif pour minimiser la distance entre un ensemble de points repères. De cette manière, on supprime les transformations affines appliquées à la carte tatouée.

- Ensuite, la même subdivision identique à celle utilisée lors de l'insertion est recréée sur la carte originale, et cette subdivision est transférée sur la carte tatouée.

- Pour chaque sous-surface correspondante sur la carte tatouée, on applique une analyse spectrale.

- La comparaison des coefficients spectraux extrait le tatouage inséré.

- **Apport de l'algorithme:** Les résultats des expérimentations effectuées ont montré une nette amélioration par rapport au précédent article proposé par Ohbuchi[28], notamment la résistance aux déformations locales de la carte et l'ajout aléatoire de bruit, et aussi le cropping (enlever une partie de la carte). L'analyse spectrale est d'une grande utilité, car elle permet d'obtenir des coefficients qu'on peut perturber pour l'insertion du tatouage, mais aussi ces coefficients sont obtenus en utilisant une clé privée, ce qui renforce la fiabilité et la sécurité de l'algorithme. Aussi, on trouve dans l'algorithme le processus de normalisation qui consiste à éliminer les transformations

- L'amélioration de la robustesse aux attaques, en calculant une meilleure représentation du domaine fréquentiel ;
- Enumérer l'ensemble des attaques possibles sur les cartes, ainsi que la perception humaine de la déformation des cartes.

II.3.4 Schémas basés sur le tatouage d'objets 3D

R. Ohbuchi et al. ont commencé par proposer des schémas de tatouage d'objets 3D [25] qu'ils ont ensuite adaptés au tatouage de documents géographiques. Le schéma est robuste, non-aveugle et permet d'introduire un message au sein du document en utilisant une analyse spectrale.

Le schéma ne travaille pas sur le document lui-même, mais sur la triangulation de Delaunay des sommets du document. Cela se justifie pour plusieurs raisons. Tout d'abord, la triangulation de Delaunay permet aux auteurs de se rapprocher du tatouage d'objets 3D. Ainsi, pour les deux types de données il s'agit de tatouer un maillage de triangles. De plus, elle fournit une notion de voisinage des sommets qui est corrélée à la topologie du document. Enfin, les auteurs vont préserver cette triangulation lors du tatouage. Comme elle est très représentative de la topologie du document, ils vont ainsi limiter la dégradation du document.

II.4 Terminologies

Dans cette section nous discutons la notion de document géographique et de qualité. Ensuite, nous expliquons la différence entre les données de type raster et les données de type vecteur, puis nous abordons le sujet des transformations qui sont susceptibles d'être appliquées à un document géographique. Enfin, nous montrerons les avantages des schémas aveugles pour les documents géographiques.

Tant qu'il s'agit de tatouer des polygones, par exemple des bâtiments, il est tout à fait possible de travailler avec des ensembles de géométries ou avec un graphe. Cependant, travailler au niveau des géométries semble mieux adapté. Dans le cas contraire, on perd par exemple la notion d'intérieur et d'extérieur. D'autre part, on voit difficilement dans quel cas un utilisateur a intérêt à transformer un ensemble de polygones en graphe.

En revanche, lorsque l'on veut tatouer des polylignes, comme des fleuves ou des routes par exemple, on a intérêt à travailler sur un graphe. En effet, les polylignes peuvent être découpées ou recollées afin de produire un document dont les géométries sont complétement différentes, sans perdre d'information. Par coupage et collage des polylignes, on peut obtenir une multitude de documents. Cependant, on peut noter que le graphe issu du document transformé demeure identique à celui issu du document de départ.

On peut donc conclure que si l'on souhaite concevoir une méthode de tatouage applicable aux polylignes ou à la fois aux polygones et aux polylignes, on a tout intérêt à considérer les données comme un graphe.

II.4.2 Le mode raster et le mode vecteur

Il existe deux modes fondamentaux de représentation numérique des données géographiques :

- Le mode matriciel : **RASTERS**
- Le mode vectoriel : **VECTEURS**

II.4.2.1 Les Rasters

Les Raster sont des images (plans scannés, photographies aériennes, images satellitaires) repérées dans l'espace. Une donnée au format raster est une grille régulière de cellules (les pixels = pictures elements) qui formes une matrice de lignes et de colonnes. A chaque cellule on associe une valeur chiffrée. Le traitement des images est plus complexe et nécessite des outils très spécialisés.

A) Exemples de données Rasters

Figure II.4 BD orthophotographie numérique [29].

- Scan25® : il s'agit d'un document papier qui a été numérisé. La valeur de chaque pixel représente la couleur de la carte [39].

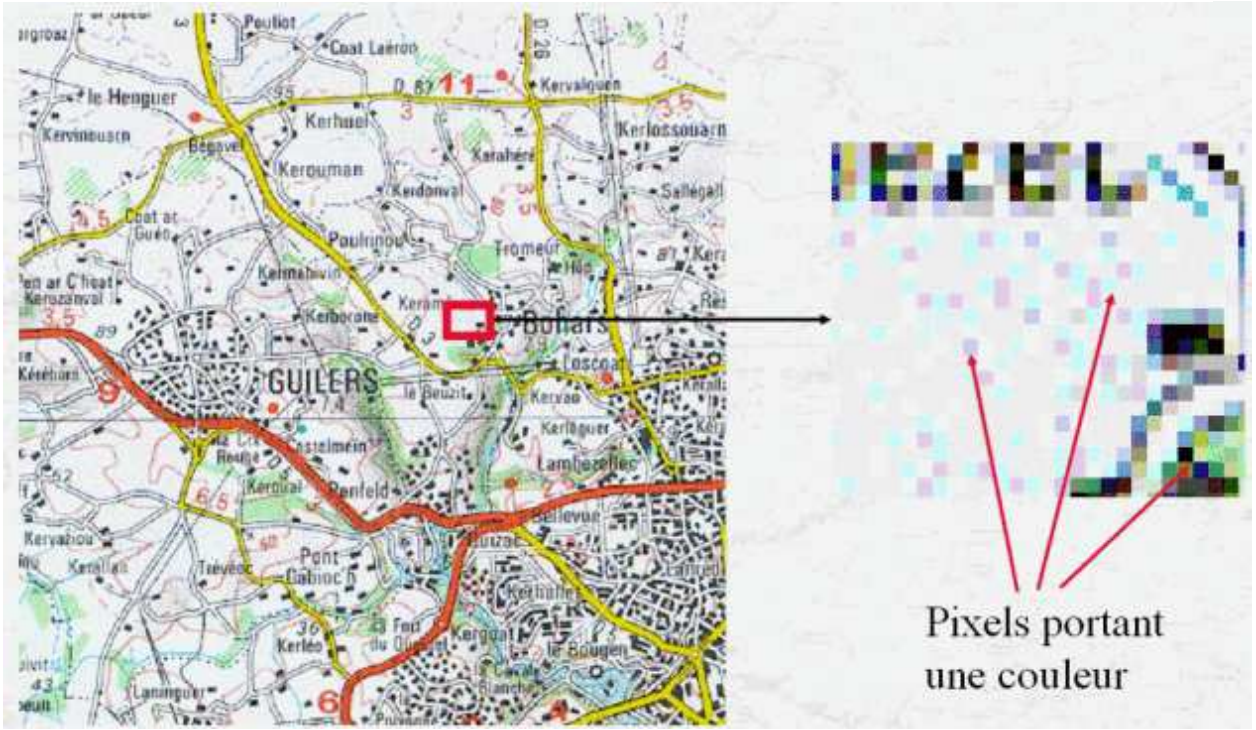


Figure II.6 Représentation d'une carte scannée [30].

- Carte d'altitude : La valeur des pixels d'une image Raster peut représenter autre chose qu'une vraie couleur. Ici par exemple la valeur de chaque pixel représente l'altitude au lieu considéré. Sa visualisation permet d'avoir une représentation du relief [39].

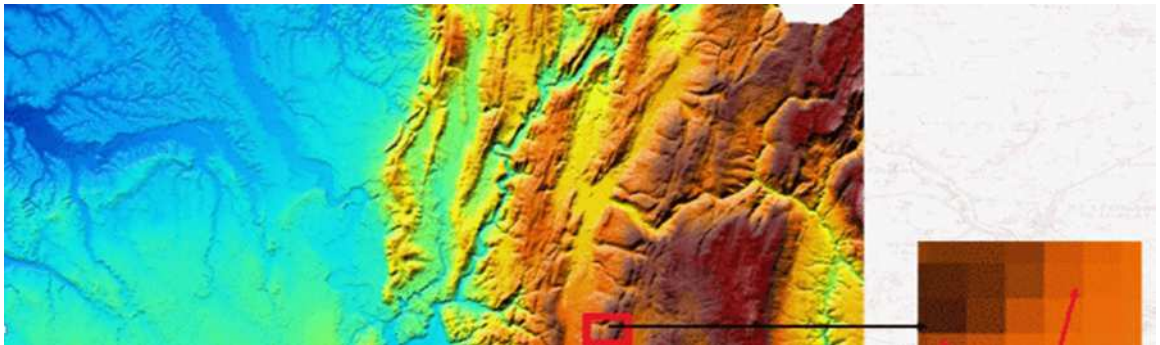


Figure II.7 Représentation d'une carte d'altitude[31].

II.4.2.2 Les Vecteurs

Le format vectoriel utilise le concept d'objets géométriques (points, lignes, polygones) pour représenter les entités géographiques. Ces objets géométriques sont définis par leurs coordonnées dans un système de projection :

- Localisation précise des objets,
- Possibilités de modélisation plus poussée,
- Facilité de manipulation des objets.

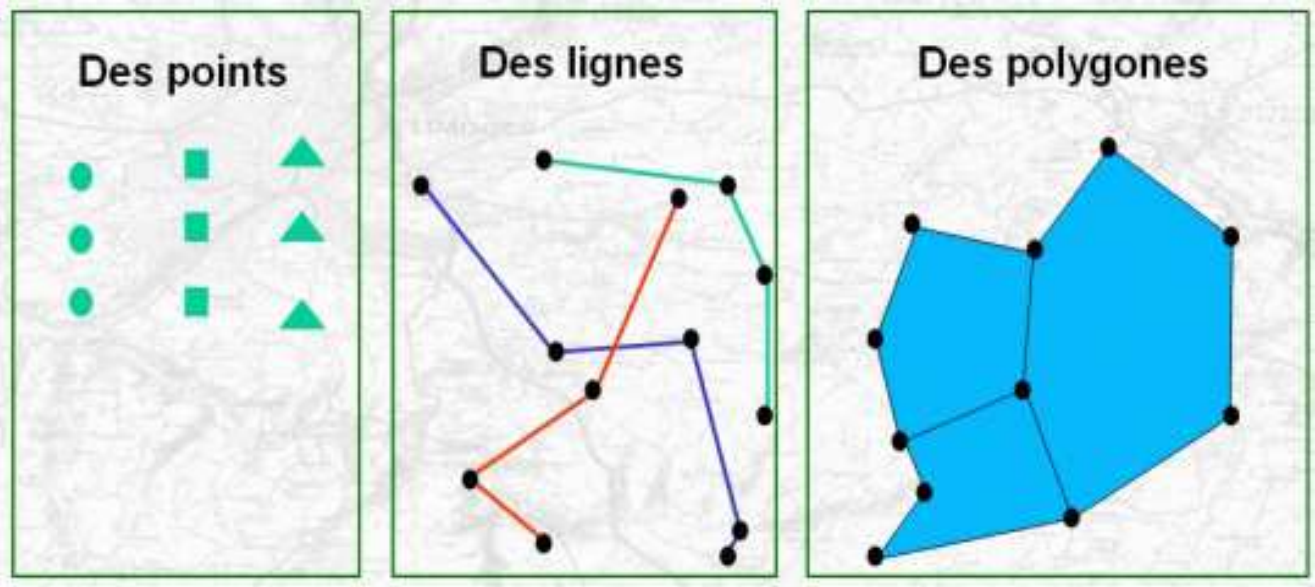


Figure II.8 Types de données géométriques [32].

(symbolisés ici par une locomotive). A chaque surface ou point sont attachées les données attributaires descriptives de la commune ou du point considéré.

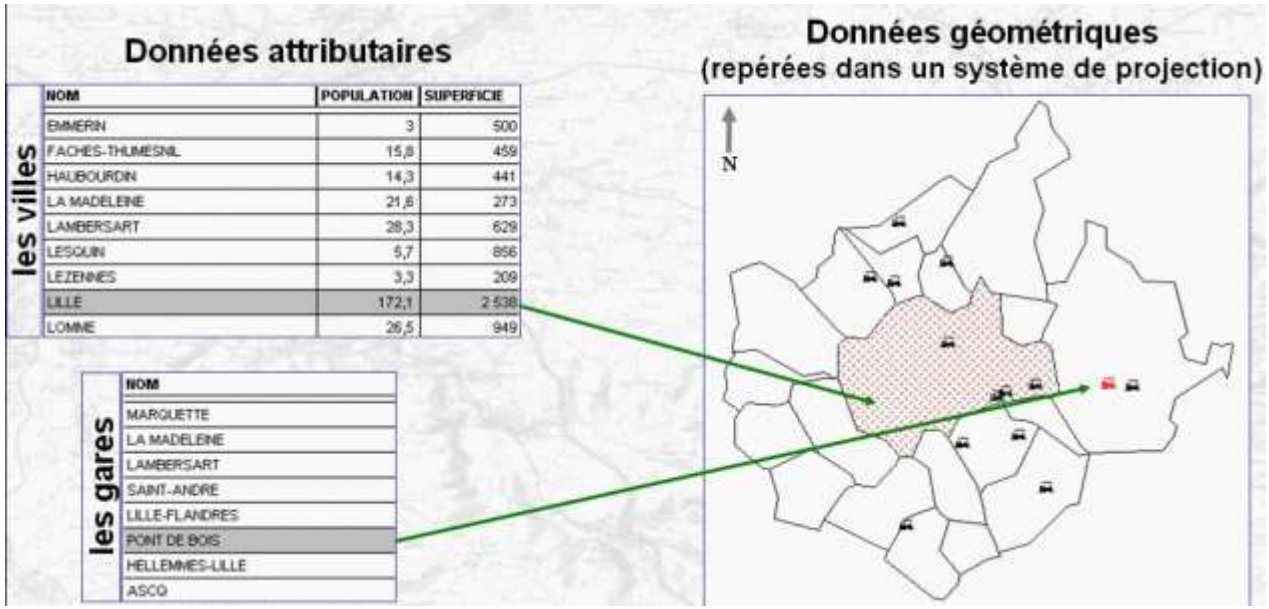


Figure II.9 Représentation de deux couches qui contienne des informations de type géométriques et attributaires [33].

- Base de données (Corine Land Cover): La base de données Corine Land Cover est une base de données vectorielles composée de surfaces. Le territoire est divisé en surface présentant un usage du sol homogène. Les données attributaires précisent l'usage en question selon une nomenclature à plusieurs niveaux.

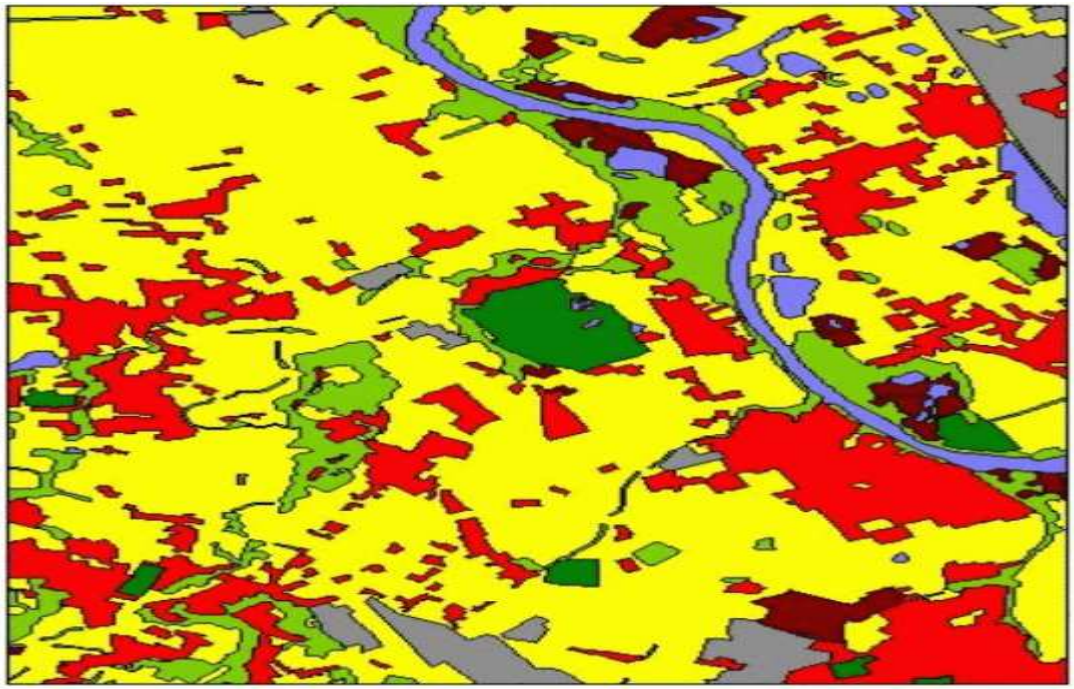


Figure II.10 Représentation d'une BD Corine Land Cover [34].

La remarque sur ces deux mode de cartes géographiques est que le traitement des données rasters et celui des données vectorielles sont très différents et ne font pas appel aux même outils, ni aux mêmes compétences.

Le traitement des rasters et des données qu'ils contiennent est complexe. Il réclame des outils spécialisés. C'est pourquoi dans une majorité de cas ils ne servent que de fond de plan en tant que support à des couches vectorielles. Par contre, les données vectorielles sont plus simples à manipuler, notamment à des fins d'analyse.

II.4.3 Réordonnancement des données

Le réordonnancement des données regroupe plusieurs transformations. Parfois, Il s'agit de modifier l'ordre dans lequel sont stockés les polygones. On peut aussi changer l'ordre dans

L'ajout d'un bruit gaussien va à la fois diminuer la précision du document et risquer d'ajouter des erreurs topologiques. Par exemple, deux routes parallèles risquent de devenir sécantes.

II.4.5 Tatouage aveugle ou non

Certains des schémas ne sont pas aveugles. Par définition, ils nécessitent de conserver le document original afin de pouvoir l'utiliser lors de la détection. Cela pose plusieurs problèmes. Tout d'abord, pour prouver que le document nous appartient, il faut pouvoir donner un couple composé du document original et de la clé. Or, dans les schémas non-aveugles que nous avons étudiés, un attaquant peut créer un tel couple à partir du document tatoué. La seule façon d'éviter ce problème est de faire appel à un tiers de confiance pour archiver le document original. L'archivage peut alors s'avérer coûteux car les documents géographiques sont souvent lourds et pour chaque document publié, il faut conserver l'original en lieu sûr. Les schémas aveugles ne présentent pas ces problèmes d'archivage [38].

II.4.6 Tatouage par transformée ou par modification géométrique

Les méthodes de tatouage par transformées garantissent que l'aspect des géométries sera préservé lors de la phase de tatouage. Par contre, elles ne permettent pas de contrôler le déplacement des sommets du document. On ne peut donc pas facilement borner la perte de précision due au tatouage. Les méthodes utilisant des déplacements de sommets paraissent donc mieux adaptées pour tatouer des données géographiques.

II.5 Conclusion

Dans ce chapitre, nous avons présenté différents schémas de tatouage de données géographiques et avons dégagés les principaux aspects de ces schémas. On constate que peu des schémas présentés tiennent réellement compte de la spécificité des documents géographiques vectoriels.

Généralement ils bornent le déplacement des sommets mais peu s'intéressent à la

III.1 Introduction

Dans ce chapitre, nous présentons le schéma de tatouage pour les documents géographiques vectoriels qui a été élaboré dans le cadre du projet Tadorne (Tatouage de Données Contraintes) de l'ACI Sécurité et Informatique auquel participaient les laboratoires du GREYC, du Cédric CNAM [27],du Lamsade (Paris-Dauphine), du Le2i (Université de Bourgogne) et enfin le laboratoire Cogit de l'IGN (Institut National Géographique).

Dans un premier temps, nous présentons les données géographiques considérées ainsi que les idées directrices du schéma. Nous détaillons enfin le schéma et précisant chacune de ces étapes.

III.2 Données géographiques considérées

L'algorithme propose un schéma de tatouage de données géographiques vectorielles bien qu'il puisse exister des informations descriptives associées aux objets géographiques, nous utilisons uniquement les données géométriques car ils estiment que ces données sont suffisamment riches pour y insérer une marque. De plus, cela permet d'être totalement indépendant de la présence, de l'absence ou de la modification des données alphanumériques [43].

III.3 Précision du document

L'algorithme borne le déplacement de chaque sommet du document. Cette borne est un paramètre de l'algorithme choisi par l'utilisateur. La donnée géographique originale est fournie avec une précision qui indique la distance maximum entre la position réelle et la position représentée des sommets du document. Le déplacement des sommets va engendrer une perte de précision de la donnée géographique. En bornant le déplacement des sommets, l'algorithme peut garantir, à priori, la valeur de la précision après tatouage [48].

III.4 Sommets et arêtes du document

Nous souhaitons que l'algorithme de tatouage préserve la qualité du document original. Cette notion de préservation de qualité est centrale pour le schéma de tatouage qu'ils ont construit car c'est la qualité de l'information qui lui confère sa valeur. Ils doivent donc garantir que la dégradation du document engendrée par l'algorithme de tatouage sera contrôlée et bornée. L'approche consiste à préserver cette qualité à tout moment pendant le processus de tatouage. Pour cela, ils définissent la notion de préservation de qualité pour les documents géographiques vectoriels. Rappelons que la préservation de qualité est représentée par une relation Qd et qu'un algorithme de tatouage préserve la qualité des documents si cette relation est vérifiée entre tout document et son tatoué. La relation Qd pour les données géographiques vectorielles fait intervenir des notions de topologie et de métrique [44].

III.6 Préservation de la triangulation :

Ils s'inspirent des travaux de R. Ohbuch [35], ils utilisent la triangulation de Delaunay de l'ensemble des sommets du document. Le schéma garantit que celle-ci reste inchangée avant et après tatouage. Plus formellement, toute paire de sommets qui sont reliées dans la triangulation originale le restera dans la triangulation du document tatoué. Ils posent l'hypothèse raisonnable qu'en préservant la triangulation de Delaunay, ils modifient très peu la topologie du document.

La triangulation de Delaunay sur un nuage de points permet d'obtenir une triangulation telle que tout triplet de sommets forme un triangle si et seulement si aucun sommet n'est à l'intérieur du cercle circonscrit à ce triangle. Par construction, cette triangulation maximise l'angle minimum des triangles et elle favorise l'équilatéralité des triangles [46].

Les figures III.3 illustrent une triangulation de Delaunay associée à un graphe issu de données routières réelles. Les figures III.1 et III.2 représentent respectivement un réseau routier et son nuage de points associé.

Figure III.1 Document original [46].

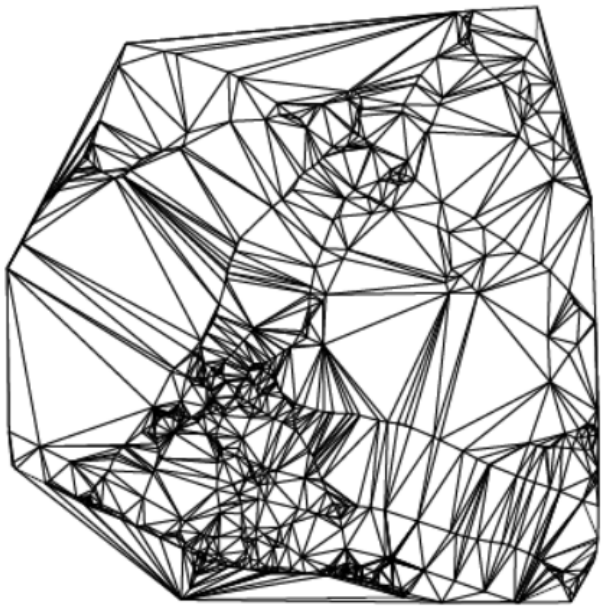


Figure III.3 Triangulation de Delaunay [46].

III.7 Transformations appliquée sur le document

Le schéma qu'ils présentent doit être résistant à un ensemble de transformations sur le document au moins dans une certaine mesure que les expériences permettront de quantifier. Parmi ces transformations on trouve :

III.7.1 Réorganisation des données

La réorganisation des données est une transformation potentielle que le schéma prend en compte. Il est possible de définir cette transformation de plusieurs façons. Une telle transformation peut impliquer une grande perte d'information, par exemple, d'un changement de format du document.

La figure III.7 présente un exemple de découpage d'un document. Le document original est représenté par la figure III.7 (a). Dans cette figure, le rectangle foncé représente la zone découpée. La figure III.7 (b) illustre la zone découpée.

1) Exemple de découpage

Figure III.4 Document original.

Figure III.5 Document découpé.

III.7.3 Transformations géométriques

Ils considèrent aussi les transformations géométriques telles que la rotation des données ou leur translation. Bien que ces transformations ne soient pas les plus utilisées, la plupart des méthodes de tatouage présentées dans la littérature [47] les considèrent comme des transformations légitimes. Cela s'explique essentiellement par le fait qu'il est facile de laver une marque qui ne résisterait pas à ces modifications.

III.8 Schéma aveugle

non-tatoué à un tiers, par exemple à l'expert indépendant qui doit effectuer le test de détection pour un tribunal [21].

III.9 Schéma 0-bit

Au contraire d'un schéma n-bits, qui consiste à cacher un message au sein du document, un schéma zéro bit permet juste de répondre à la question : « Le document est-il tatoué avec cette clé ». Les schémas 0-bit sont donc moins spectaculaires car aucun message n'apparaît lors de la détection mais ils s'avèrent tout aussi efficaces. Pour les schémas aveugles n-bits, le principal problème est de resynchroniser l'algorithme de détection sur le message caché [25].

Le schéma présenté est de type zéro-bit, c'est à dire qu'aucun message n'est inséré dans le document. Il introduit un biais statistique en fonction d'une clé. Sachant que ce biais a une chance extrêmement faible d'être présent dans un document qui n'a pas été tatoué par la clé, lorsque nous observons le biais, nous concluons que la clé a été utilisée pour tatouer le document.

III.10 Idées directrices

Ils travaillent sur des petites parties significatives extraites du document original que nous nommons sites.

Pour chaque site, ils calculent un identifiant robuste, qu'ils l'appellent codage. Un sous ensemble des sites du document est sélectionné en fonction de ce codage et d'une clé secrète. Les sites sélectionnés sont ensuite forcés à satisfaire ou non une propriété Φ dont nous spécifierons le rôle. Pour forcer un site à satisfaire Φ nous modifions celui-ci. Pour garantir que ces modifications n'entraineront pas de perte de qualité du document, nous définissons une notion de qualité locale au site. Nous verrons que respecter cette qualité locale suffit à conserver la qualité globale du document.

III.11 Préservation locale de la qualité

III.12 Présentation du schéma de tatouage

Dans cette section, nous présentons le schéma de tatouage aveugle et robuste applicable aux documents géographiques vectoriels. Nous commencerons par expliquer la notion de site et les grandes lignes de l'algorithme qui composent le schéma de tatouage.

Rappelons que les algorithmes de tatouage et de détection prennent en entrée un graphe issu d'un document géographique dont les sommets sont étiquetés par leurs coordonnées.

En forçant certains sites à satisfaire la propriété, nous modifierons cette distribution. Rappelons que les sites modifiés sont choisis en fonction de leurs codages et de la clé secrète.

L'algorithme de tatouage retourne un graphe avec les mêmes arêtes mais dont les étiquettes sont modifiées.

L'algorithme de détection sélectionne les sites en fonction de leur codage et d'une clé secrète. Un comptage permet de déterminer le nombre de sites qui satisfont la propriété Φ. Pour décider si un document est tatoué, nous majorons la probabilité que la distribution mesurée s'écarte de la distribution modélisée. Lorsque cette probabilité franchit un seuil, nous dirons que le document est tatoué.

III.13 Définition des sites

L'algorithme construit autant de sites que le document contient de sommets. Chaque site extrait du document original est formé tout d'abord d'un sommet que nous nommons sommet central du site et de ses voisins dans la triangulation de Delaunay (ordonnés dans le sens trigonométrique).

Le sommet central et ses voisins sont tous issus du document original et sont donc étiquetés par leurs coordonnées dans le plan. Celles-ci sont utilisées pour vérifier si le site satisfait la propriété Φ que nous avons choisie. Les arêtes entre ces sommets dans le document

Les sommets n_i sont numérotés de façon croissante dans le sens trigonométrique. Le premier voisin, noté n_1, est choisi arbitrairement. Les sommets miroirs m_i sont numérotés de sorte que le triangle formé par les trois sommets n_i , n_{i+1} et m_i soit une face de la triangulation.

La figure III.6 donne un exemple de site. Dans cette figure, chaque triangle représente une face dans la triangulation de Delaunay.

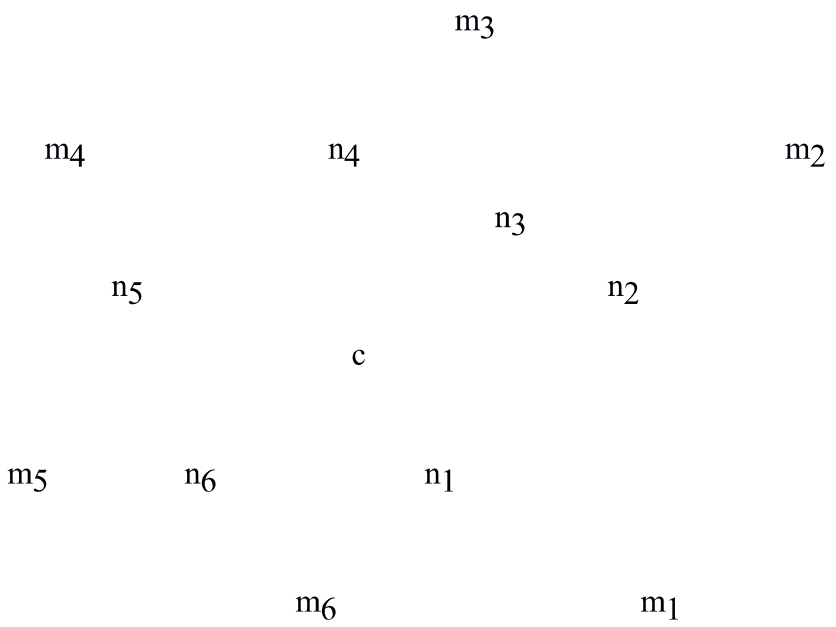


Figure III.6 Un exemple de site.

(c) est le sommet central du site, n_1 à n_6 sont ses voisins, m_1 à m_6 sont les sommets miroirs de c par rapport à chaque triangle adjacent à c.

III.15 Détails de l'algorithme de tatouage

Le calcul de la triangulation de Delaunay est une étape préliminaire au tatouage. Une fois cette étape achevée, nous nous servons à la fois du document original et de sa triangulation pour énumérer l'ensemble des sites du document. Chaque site extrait est ensuite traité individuellement et localement, sans référence au document global. Il est ensuite réintroduit dans le document original lorsque les modifications sur le site ne dégradent pas la qualité du document. Dans le cas contraire, on n'applique aucune modification dans le document pour ce site. Seules les fonctions qui permettent d'extraire un site du document et de le réintroduire interagissent avec le document. Le traitement séquentiel des sites est jalonné de plusieurs étapes.

Pour chaque site, ces étapes sont les suivantes :

1. La sélection ou non du site en fonction d'une clé et de son codage.

2. La modification du site si celui-ci est sélectionné, sinon on passe au site suivant.

3. Un test pour déterminer si la modification préserve la qualité du site. Si la qualité du site n'est pas préservée par la modification, alors celle-ci est abandonnée. Sinon on répercute la modification du site dans le document.

Il y a cinq étapes pour traités un site :

1. extraction du site;

2. sélection du site;

3. modification du site;

4. test de préservation de la qualité du site;

5. application des modifications du site dans le document.

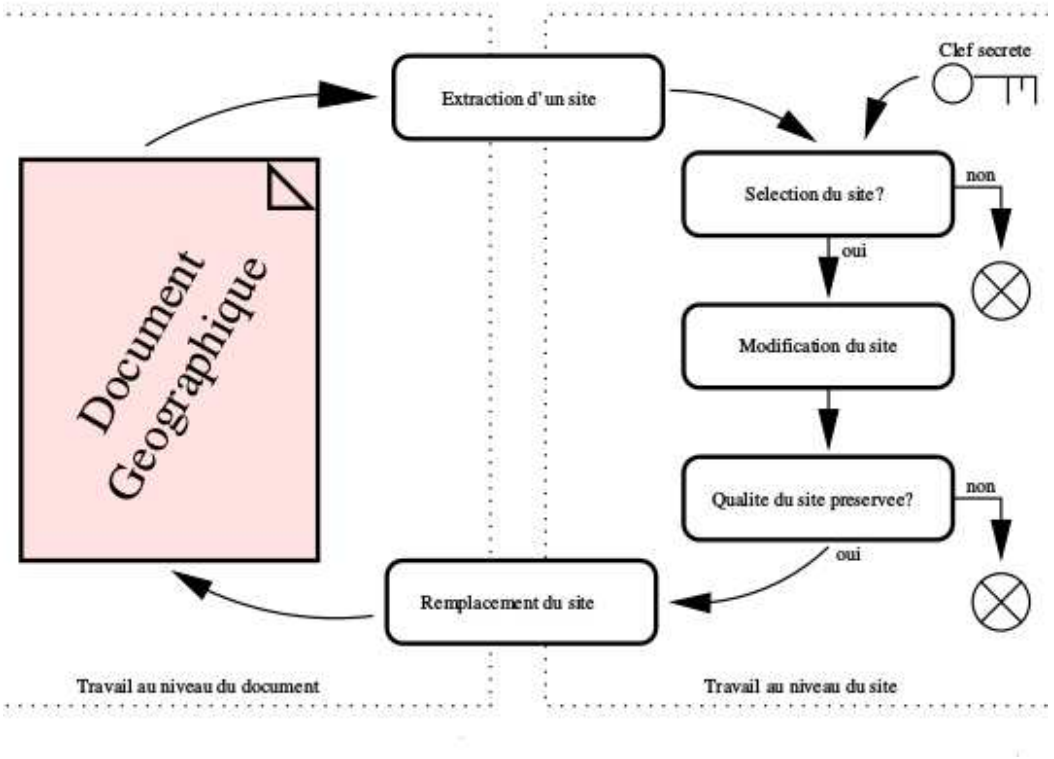


Figure III.7 Le processus de tatouage [27].

III.15.1 Extraction des sites

L'extraction des sites du document est effectuée séquentiellement. Ainsi, nous appliquons la fonction d'extraction sur chaque sommet du graphe pour traiter une seule fois chaque site du document. Pour obtenir une notion de localité au sein du document, nous calculons une triangulation de Delaunay à partir du nuage des points du document original [27].

Pour définir le codage d'un site, il associe à chaque site une matrice binaire M qui représente la connexité entre les sommets (c, n_1 , ..., n_N) où N représente le nombre de voisins de c. La matrice M est composée de N lignes et de N colonnes. La i-ème ligne représente la connexion de n_i, le i-ème voisin de c avec c ainsi que les connexions entre n_i et les sommets $\{n_1, ..., n_N\}$. Les coefficients de la matrice sont déterminés par les règles suivantes:

– $M_{i,1} = 1$ avec $1 \leq i \leq N$, lorsqu'il existe une arête entre n_i et c dans E_c , sinon $M_{i,1} = 0$.

– $M_{i,j} = 1$ avec $1 \leq i \leq N$ et $2 \leq j \leq N$, lorsqu'il existe une arête entre n_i et n_{i+j-1} modulo N dans E_c , sinon $M_{i,j} = 0$.

III.15.3 Sélection des sites

En fonction de son codage et de la clé secrète, nous décidons si un site doit être:

– forcé à satisfaire la propriété choisie;

– forcé à ne pas satisfaire la propriété choisie;

– laissé inchangé.

Cette décision dépend également du paramètre $p \geq 2$ qui sert à régler la proportion de sites du document qui doivent être modifiés. Soit la fonction P_p, pour une clé k, un paramètre p et un site s dont l'identifiant robuste est représenté par la matrice M, retourne une valeur comprise entre 0 et $p-1$.

P_p (s, k) = hash(id(s), k) modulo p

- P_p (s, k) = 0, on forcera le site à ne pas satisfaire la propriété choisie;
- P_p (s, k) = 1, on forcera le site à satisfaire cette propriété;
- dans les autres cas, le site est laissé inchangé.

Input: δ′: la perte de précision autorisée lors du tatouage

Output: s′: une copie du site s dont le sommet central est déplacé de façon à satisfaire Φ

begin

b ← Barycentre des(n_1 , ... , n_N) ;

d ← Distance entre c et b ;

v ← ⌊ d/δ' ⌋ mod 2 ;

if v = 1 then

return s

c′ ← Bouger c vers b d'une distance de δ′ ;

return (c ′ , (n_1 , ..., n_N), (m_1 , ..., m_N), E_c)

end

B) Algorithme pour forcer un site à ne pas satisfaire Φ

Input: s = (c, (n_1 , ... , n_N), (m_1 , ... , m_N), E_c) : le site qui ne doit pas satisfaire Φ

Input: δ′: la perte de précision autorisée

Output: s′: une copie du site s dont le sommet central est déplacé de façon à satisfaire Φ

begin

b ← Barycentre des(n_1 , ... , n_N) ;

d ← Distance entre c et b ;

III.15.5 Test de préservation de la qualité des sites

Input: $s = (c, n_1, \ldots, n_N, m_1, \ldots, m_N, E_c)$: le site original

Input: $s' = (c', n_1, \ldots, n_N, m_1, \ldots, m_N, E_c)$: le site modifié

Input: δ': la perte de précision autorisée

Output: vrai : si s et s ′ ont la même qualité, faux sinon

```
begin
    if distance(c, c′ ) > δ′ then return faux ;
    if c′ n'est pas dans le polygone formé par n1 , ... , nN then return faux ;
    for i de 1 à N do
        j ← (i + 1) mod N ;
        k ← (i + 2) mod N ;
        if c′ est dans le cercle circonscrit à (ni , nj , mi ) then
            return faux ;
        if c est dans le cercle circonscrit à (ni , nj , nk ) then
            if c′ n'est pas dans le cercle circonscrit à (ni , nj , nk ) then
                return faux ;
    return vrai ;
end
```

III.16 Algorithme de tatouage

Output: w ∈ D : le document tatoué

```
begin
    w ← copie de d ;
    foreach site s de w do
        j ← P_p (s, k) ;
        if j = 0 then
        if s satisfait Φ then
            s′ ← Modification de s forcée pour ne pas satisfaire Φ ;
            if qualité préservée entre s′ et s then
                Réintroduction de s′ à la place de s dans w ;

    else if j = 1 then
        if s ne satisfait pas Φ then
            s′ ← Modification de s forcée pour satisfaire Φ ;
            if qualité préservée entre s′ et s then
                Réintroduction de s′ à la place de s dans w ;
    return w ;
end
```

III.17 Algorithme de détection

Input: $\lambda \in [0, 1]$: le seuil de détection

Output: [Vrai, Faux] : le document est-il tatoué par la clé k.

begin

$n_0 \leftarrow 0$; $m_0 \leftarrow 0$;

$n_1 \leftarrow 0$; $m_1 \leftarrow 0$;

foreach site s de w do

$j \leftarrow P_p(s,k)$;

if j = 0 then

$n_0 \leftarrow n_0 + 1$;

if s ne satisfait pas Φ then $m_0 \leftarrow m_0 + 1$;

else if j = 1 then

$n_1 \leftarrow n_1 + 1$;

if s satisfait Φ then $m_1 \leftarrow m_1 + 1$;

return λ

end

L'algorithme reprend les différentes étapes qui permettent d'effectuer la détection. En fixant un seuil λ, il est possible de décider si oui ou non un document est tatoué. La valeur calculée par la formule est la probabilité de présence du tatouage dans le document.

III.18 Conclusion

Dans le chapitre qui suit, nous présentons les résultats expérimentaux obtenus et le manuel d'utilisation de l'application qu'on a nommé GDW (Geographical Data Watermarker) qui implémente l'algorithme vu, qui permet de tatouer et détecter la marque dans des documents géographique de taille réelle.

Conclusion Générale

Le tatouage numérique est une technique permettant d'ajouter des informations de vérification à un fichier digital de format numérique, tel qu'un signal audio, une vidéo ou une image afin de préserver sont originalité et d'évité un éventuel vol intellectuel et protégé les ayants droits.

Ce travail s'inscrit dans le domaine du tatouage de données géographiques. Nous avons commencé par introduire le vocabulaire qui caractérise les différentes classes de schémas de tatouage numérique. Puis, nous avons présenté les données géographiques considérées ainsi qu'une sélection de travaux de ce domaine.

Après nous avons décrit en détail un algorithme de tatouage de données géographiques, robuste, 0-bit, aveugle et rapide.

L'aspect le plus important du schéma et qui le différencie des autres schémas présentés dans l'état de l'art est de donner des garanties sur la préservation de la topologie et la métrique du document lors du tatouage. Nous pensons que la préservation de la qualité du document lors du marquage est un aspect fondamental pour les données géographiques dont les schémas existants n'ont pas assez tenu compte.

Nous avons vu que le schéma pour les données géographiques vectoriels présente de multiples avantages. Il a cependant été suffisamment détaillé dans ce travail pour que l'on ait pu donner une implémentation logicielle avec des perspectives mentionnées dans le chapitre IV.

Bibliographies

[1] J. Seitz. Digital Watermarking for Digital Media. Information Science Publishing, 2004.

[2] F. Petiscolas, R. Anderson, and M. Kuhn. Information Hiding Terminology : A Survey. IEEE Signal Processing, 78(7) :1062–1078, 1999.

[3] K. Tanaka, Y. Nakamura, and K. Matsui. Embedding Secret Information into a Dithered Multilevel Image. In 1990 IEEE Military Communications Conference, pages 216–220, 1990.

[4] A. Tirkel, G. Rankin, R. Schyndel,W. Ho, N. Mee, and C. Osborne. ElectronicWatermark. In DICTA 1993, pages 666–672, 1993.

[5] D. Kundur and D. Hatzinakos. Digital Watermarking Using Multiresolution Wavelet Decomposition. In IEEE International Conference on Acoustics, Speech and Signal Processing, Seattle, Washington, volume 5, pages 2969–2972, 1998.

[6] C. REY and J. DUGELAY. Un panorama des m´ethodes de tatouage permettant d'assurer un service d'int´egrit´e pour les images. Traitement du Signal, 18(4) :283–295, 2001.

[7] I. Cox, M. Miller, and J. Bloom. Digital Watermarking : Principles & Practices. Morgan Kaufmann Publisher, San Francisco, CA, USA, 2002.

[8] Y. Hu, J. Huang, S. Kwong, and Y. Chan. Image Fusion Based Visible Watermarking Using Dual-Tree Complex Wavelet Transform. In IWDW'2003, pages 86–100, 2003.

[11] S.Mohanty, N. Ranganathan, and K. Namballa. VLSI Implementation of Visible Watermarking for a Secure Digital Still Camera Design. In 17th International Conference on VLSI Design, pages 1063–1068, 2004.

[12] M. Yeung and F.Mintzer. On Resolving Rightful Ownership's of Digital Images by Invisible Watermarks. 1997.

[13] M. Yeung and F. Mintzer. An InvisibleWatermarking Technique for Image Verification. In International Conference on Image Processing (ICIP'97), pages 680–683, 2007.

[14] F. Melgani, R. Benzid, and F. De Natale. Near-Lossless Spread Spectrum Watermarking for Multispectral Remote Sensing Images. Journal of Applied Remote Sensing, 1(013501), 2007.

[15] P. Wong and A. Memon. Secret and Public Key Image Watermarking Schemes for Image Authentication and Ownership Verification. IEEE Transactions on Image Processing, 10(10) :1593–1601, 2001.

[16] F. Bartolini, M. Barni, V. Cappellini and A. Piva. Mask building for perceptually hiding frequency embedded watermarks. In Proceedings of the International Conference on Image Processing (ICIP'98), vol. 1, pp. 450-454, Chicago, Illinois, US, Oct. 1998.

[17] K. Maeno, Q. Sun, S. Chang, and M. Suto. New Semi-fragile Image AuthenticationWatermarking Techniques Using Random Bias ND Non Uniform Quantization. IEEE Transactions on Multimedia, 8(1) :32–45, 2006.

[18] Z. Zhang, Q . Sun, and W.C. Wong. A Novel Lossy-to-lossless Watermarking Scheme

[21] A. Perrig. A copyright protection environment for digital images. Diploma dissertation, EPFL, Lausanne, Switzerland, Feb. 1997.

[22] S. Craver, N. Memon, B.-L. Yeo and M. Yeung. Resolving rightful ownerships with invisible watermarking techniques: limitations, attacks, and implementations. IEEE Journal on Selected Areas in Communications (Special issue on Copyright and Privacy Protection), 16(4):473-586, May 1998.

[23] F.A. Petitcolas and M. Kuhn. Stirmark – Image Watermarking Robustness Test. http://www.cl.cam.ac.uk/~fapp2/watermarking/stirmark/

[24] sejal visawadia digital watermarking Published in: Technology, Art & Photos on Nov 19, 2013

[25] Ohbuchi, R., Mukaiyama, A. et Takahashi, S. A frequency-domain approach to watermarking 3d shapes. In EUROGRAPHICS. (2002).

[26] Kang, L. H., Kim, K. L. et Choi, J. U. A vector watermarking using the generalized square mask. In International Conference on Information Technology : Coding and Computing, pages 234–236. (2001).

[27] Lafaye, J. Tatouage des bases de données avec préservation de contraintes. Thèse en sciences du cnam - spécialité informatique, Centre d'études et de recherches en informatique du CNAM. (2007).

[30] école nationale des techniciens de l'équipement (ENTE), http://www.ente-aix.fr/documents/118-demoGeo/demo/4_BasesIG/res/i20_DonneesRaster2_1.gif, (22/04/2017)

[31] école nationale des techniciens de l'équipement (ENTE), http://www.ente-aix.fr/documents/118-demoGeo/demo/4_BasesIG/res/i20_DonneesRaster3_1.gif, (22/04/2017)

[32] école nationale des techniciens de l'équipement (ENTE), http://www.ente-aix.fr/documents/118-demoGeo/demo/4_BasesIG/res/i20_DonneesVecteur2.gif, (23/04/2017)

[33] école nationale des techniciens de l'équipement (ENTE), http://www.ente-aix.fr/documents/118-demoGeo/demo/4_BasesIG/res/i20_DonneesVecteur3.gif, (23/04/2017)

[34] école nationale des techniciens de l'équipement (ENTE), http://www.ente-aix.fr/documents/118-demoGeo/demo/4_BasesIG/res/i20_DonneesVecteur.gif, (24/04/2017)

[35] Ryutarou Ohbuchi, Hiroo Ueda, Shuh Endoh Watermarking 2D Vector Maps in the Mesh-Spectral Domain - Computer Science Department, University of Yamanashi (2003)

[36] Rakesh Agrawal, Jerry Kienan.,Watermaking Relational Data Bases, Procedings of the 28th VLDB Conference, (Hong Kong, China :2002).

[37] M. Bader, X. Barillot, S. Mustière, M. Barrault, C. Duchêne., Line gaussian smoothing

Bibliographies

[40] Gerrit Schultz, Michael Voigt.,I High Capacity Watermarking System for Digital Maps, MM&Sec'04, (Magdeburg, Germany : September 20-21 2004).

[41] Les bases de l'Information Géographique, école nationale des techniciens de l'équipement (ENTE), version 6, 15 mai 2013

[42] OpenStreetMap France, http://openstreetmap.fr

[43] Lafaye, J., Beguec, J., Gross-Amblard, D. et Ruas, A. Geographical database watermarking by polygon elngation (technicl report). Rapport technique, Cedric, CNAM. (2007)

[44] Lopez, C. Watermarking of digital geospatial datasets : A review of technical, legal and copyright issues. International Journal of Geographical Information Science, 16:589–607 (2002)

[44] Lafaye, J., Gross-Amblard, D., Guerrouani, M. et Constantin, C. Watermill : an optimized fingerprinting system for databases under constraints. In IEEE Transactions on Knowledge and Data Engineering (2007).

[45] Raynal, F., Petitcolas, F. et Fontaine, C. Evaluation automatique des méthodes de tatouage. Traitement du signal. (2001)

[46] Yvinec, M. (2007). 2d triangulations. In CGAL User and Reference Manual. 3.3

[47] Niu, X., Shao, C. et Wand, X. (2006). A survey of digital vector map watermarking. ICIC International.

Bibliographies

[51] vim, http://www.vim.org/

[52] geany, https://www.geany.org/

[53] wxPython, https://wxpython.org

[54] Linux, linuxfr.org/