

Université de Bretagne Occidentale

--

Ecole Doctorale des Sciences de la Mer

Spécialité : Géomatique

Habilitation à diriger des Recherches

**Système d'information géographique temporelle maritime ;
Des distances linéaires à l'analyse temps réel des trajectoires**

M. Thomas Devogele

Soutenue le 1^{ier} décembre 2009 à L'Ecole navale

Jury :

Thérèse Libourel	Rapporteur	Professeur	Université Montpellier II
Hervé Martin	Rapporteur	Professeur	Université de Grenoble I
Anne Ruas	Rapporteur	Ingénieur HDR	IGN
Françoise Gourmelon	Examineur	Directeur de recherche	IUEM / UBO
Jacques Tisseau	Président du jury	Professeur	ENIB
Christophe Claramunt	Directeur d'habilitation	Professeur	Ecole navale

Remerciement :

Les travaux de recherche présentés dans ce mémoire ont été réalisés à l'Institut de Recherche de L'Ecole navale (IRENav). Ils n'ont jamais été accomplis *ex nihilo*. Ils sont le fruit de collaborations nombreuses et d'un terreau propice. Je tiens donc à remercier l'ensemble des personnes qui ont favorisé grandement ces travaux.

Je tiens à remercier l'ensemble du jury. Monsieur Christophe Claramunt, directeur de l'IRENav, a accepté de diriger cette habilitation. Il a toujours su, malgré son emploi du temps chargé, dégager du temps afin d'affiner ces travaux en leur apportant le recul et les ouvertures nécessaires. Merci de m'avoir initié aussi à des domaines tels que le raisonnement spatial.

Madame Thérèse Libourel, Professeur à l'Université de Montpellier, m'a fait le grand plaisir d'être rapporteur de ces travaux. C'est toujours une joie pour moi de croiser cette grande dame si attentive, lors des manifestations du GDR Cassini-SIGMA-MAGIS. Madame Anne Ruas, directrice du laboratoire COGIT de l'IGN, laboratoire qui est cher à mon cœur, m'a fait la joie d'être rapporteur de ces travaux. J'apprécie la force qu'elle déploie pour dynamiser la communauté SIG en France. Monsieur Hervé Martin, Professeur à l'Université de Grenoble I, qui m'a fait le plaisir de rapporter ces travaux. Je lui suis reconnaissant du souffle nouveau qu'il impulse dans le domaine des SIG en France. J'adresse mes vifs remerciements à madame Françoise Gourmelon directeur de recherche au CNRS (Géomer, IUEM) qui a accepté de faire partie de jury parmi des informaticiens et ainsi d'exprimer sa vision thématique de mes travaux. Monsieur Jacques Tisseau, professeur à l'ENIB, à qui j'exprime ma grande reconnaissance d'avoir accepté de faire partie de ce jury. J'espère que des collaborations futures pourront naître entre nos équipes. Je les remercie tous pour le temps qu'ils ont consacré à la lecture et l'analyse de ce manuscrit.

Un très grand merci à tous les doctorants anciens et actuels avec qui j'ai travaillé pour tous ces échanges fructueux : Sébastien Fournier, Valérie Noyon maintenant Mallé-Noyon, Ariane Mascaret, Thierry Le Pors et Laurent Etienne. Cette HDR leur doit énormément. Je leur souhaite à tous de s'épanouir dans ces métiers de chercheurs ou d'enseignants-chercheurs. J'en profite pour associer les co-encadrants et directeurs de thèse : Christophe Claramunt, Christine Chauvin et Alain Bouju.

Ces travaux se sont déroulés dans l'équipe SIG de l'IRENav, je tiens donc à remercier amicalement tous ses membres. Merci au canal historique de plus en plus actif que sont Éric Saux, Rémy Thibaud, Joseph Poupin, merci pour votre complicité. Merci aux nouveaux, Cyril Ray, Meriam Horchani et Mathieu Simonnet, qui sont déjà bien investis et qui apportent leur dynamisme. Je remercie aussi ceux qui se sont succédé comme responsable, enseignants-chercheurs ou doctorants dans ce groupe : Michel Farine, Thierry Huet, Éric Guilbert, Rodéric Béra, Myriem Sriti, Simona Niculescu, David Brosset, Mathieu Petit, Fabien Ramos, Thomas Le Bras, Jean-Marie Le Yaouanc, Géraldine Del Mondo. J'associe à ces remerciements tous les membres de l'IRENav, les enseignants-chercheurs, les administratifs et les techniciens pour leur aide dans les tâches de tous les jours. Je tiens à remercier particulièrement Abdel Boudraa pour ses conseils avisés et Laurent Guillon pour la relecture de ce manuscrit. Je ne veux pas oublier les autres professeurs de l'Ecole navale et plus particulièrement les enseignants de navigation et d'anglais pour leur collaboration. Beaucoup d'élèves ont été associés à ces travaux. Par leurs réflexions, ils ont contribué à la qualité des prototypes réalisés, qu'ils en soient remerciés.

Je tiens aussi à souligner le rôle important joué par le GDR Cassini-SIGMA-MAGIS dans ces travaux. C'est un facilitateur qui a été indispensable au développement de mes recherches. Je tiens donc à associer à ces remerciements mes collègues informaticiens, géomaticiens et thématiciens du GDR.

Ma gratitude va aussi à la Marine nationale et les cinq amiraux (VAE Pinon, VAE van Huffel, CA de Roquefeuil, CA Soudan, CA de Briançon) qui ont dirigé l'école pendant ces onze ans et qui ont eu un regard bienveillant sur la recherche de ces civils qui ne faisaient pas partie de leur culture d'origine. De même, je remercie les cinq directeurs de l'enseignement (CA Abadie, CV Deméocq, CV Rebour, CV Liot de Nortbécourt, CV Pouly) qui ont compris que l'association de l'enseignement et de la recherche malgré les contraintes est bénéfique à une grande école.

Je désire terminer, en remerciant ma famille : ma femme, Cristelle et mes trois filles Mathilde, Emma et Anne-Lou. Elles sont des sources de joie pour moi chaque jour. Elles savent à la fois m'encourager, me mettre la pression et me ramener à la réalité. Merci donc de supporter un chercheur à la maison et de l'équilibrer.

Résumé :

Les Systèmes d'Information Géographique (SIG) gèrent des informations complexes localisées dans un espace géographique. Durant la dernière décennie, ces systèmes ont vu l'émergence d'une part des SIG en trois dimensions et plus particulièrement des modèles numériques de terrain (MNT) et d'autre part de l'intégration de données temporelles. Les travaux de recherche en informatique présentés dans ce manuscrit s'attaquent à ces deux thématiques et sont appliqués au domaine maritime.

En termes de MNT, ces recherches ont essentiellement porté sur l'intégration de MNT terrestre et maritime. Une méthode de fusion générant un MNT continu a été définie. Elle est basée sur l'appariement d'éléments caractéristiques linéaires (crêtes, talwegs, lignes de rupture de pente). Elle emploie des distances linéaires discrètes introduites par Fréchet. Ces distances fournissent des vecteurs d'appariement autorisant la fusion de MNT à l'aide d'une technique de déformation élastique.

En ce qui concerne les SIG spatio-temporels, ces travaux concernent les déplacements d'objets mobiles et plus particulièrement les navires. Dans le cadre de la navigation maritime, les SIG ont une place prépondérante via les aides à la navigation, les systèmes de surveillance du trafic et les simulateurs de formation. Dans ce cadre, un nouveau modèle de représentation des déplacements, la vue relative, a été formalisé. Ce formalisme s'appuie sur la perception humaine de l'espace spatio-temporel et des objets mobiles. Cette représentation est complémentaire de la représentation absolue. Les déplacements relatifs vis-à-vis d'un objet référent sont ainsi plus facilement analysables. Les évolutions des distances et des vitesses relatives sont mieux perçues. Parallèlement, une simulation des activités de navires à l'aide de systèmes multi-agents a été conçue. Les déplacements simultanés d'un grand nombre de navires dans un environnement maritime sont ainsi simulés. Elle prend en compte les différentes activités concurrentes des objets mobiles et les contraintes géographiques. Qui plus est, elle intègre des raisonnements à base de patrons pour reproduire les raisonnements humains. Finalement, à partir des bases de données historiques des positions d'objets mobiles dans un environnement ouvert, des techniques d'extractions des trajectoires et de fouille de données spatio-temporelles ont été réalisées. Elles permettent pour chaque itinéraire et chaque type d'objets de définir des routes types spatio-temporelles et de détecter à la volée des comportements inhabituels.

Ces travaux, bien qu'appliqués au domaine maritime, sont génériques. Les méthodes de fusion de MNT peuvent être appliquées à tous types de MNT ayant des zones de recouvrement. De même, les travaux de modélisation, d'analyses et de simulation sur les objets mobiles sont adaptables à tous les déplacements d'objets mobiles dans des espaces ouverts.

Mots clés :

Système d'Information Géographique (SIG), Modèle Numérique de Terrain (MNT), objets mobiles, base de données spatio-temporelle, distance linéaire de Fréchet, simulation multi-agents cognitive, vues relatives spatio-temporelle de trajectoires, appariement et fusion de données, système de navigation.

Abstract:

Geographic Information Systems (GIS) manage information located in geographic space. These systems have included more and more complex information over the past decade. Indeed, they can now store temporal data as well as tree-dimensional data: for example, the digital elevation model (DEM) represents relief. My research in computing science deals with those two fields and is applied to maritime domain.

As far as the DEM is concerned, those works focus on DEM integration. A new approach based on landscape elements (ridges, talwegs, etc.) allows to merge land and maritime geographical datasets. This includes a process matching linear elements. It develops from several extensions of linear distance as defined by Fréchet. Those distances also provide data matching vectors that allow to merge DEMs thanks to rubber sheeting process.

In the spatio-temporal GIS area, my work studies moving object trajectories and more particularly vessel displacements. For maritime navigation, GIS are key systems. Navigation, monitoring and simulation learning systems are very close to Geographic Information System. First, a new conceptual model has been introduced: the relative view, oriented to the representation of trajectories in time and space. This model is based on the human perception of time and space for mobile objects. This representation complements the traditional absolute view. Displacements of moving objects relative to a referent object are qualified in terms of relative speed and distance. With this new representation, evolutions of displacements are better understood. In the same way, a new multiagent system (MAS) is designed to simulate maritime activities. Several vessel displacements can be simulated at the same time in a maritime space. This simulator takes into account the maritime space constraints and different activities of vessels with qualitative spatial reasoning and role paradigm. Moreover, to simulate human reasoning, it integrates a decision making system based on patterns. Finally, data mining on a huge quantity of positions stored in spatio-temporal databases is carried out in order to deduce the behavior of those mobile objects. We can define spatio-temporal patterns for each cluster of trajectories of mobile objects following the same itinerary.

Those works are applied to the maritime domain but can also be extended to other types of data. DEM integration processes can be useful for other DEMs with overlap area. In the same way, research on moving objects can be adapted to other kind of objects moving in an open area.

Keywords:

Geographic Information System (GIS), Digital Elevation Model (DEM), tracking systems, spatio-temporal database, Fréchet distance, Spatial and Cognitive Simulation with Multi-agent Systems, relative representation of trajectories, data matching, data merging, navigation systems

Table des matières

1	Introduction.....	12
2	La géométrie algorithmique	21
2.1	Les distances linéaires.....	22
2.2	Applications des distances de Fréchet discrètes maximale et moyenne	30
3	L'intelligence artificielle appliquée aux SIG.....	41
3.1	Les algorithmes génétiques et données spatiales	41
3.2	Les systèmes multi-agents agent-groupe-rôle avec un contexte spatial	45
3.3	Les systèmes multi-agents cognitifs intégrant des concepts de patrons	51
4	Les bases de données spatio-temporelles.....	59
4.1	Les vues relatives spatio-temporelles	60
4.2	Analyse de similarité de trajectoires d'objets mobiles.....	68
5	Conclusion	79
5.1	Résumé des recherches effectuées	79
5.2	Projets de recherche à 4 ans	81
6	Annexes.....	87
6.1	Curriculum vitæ	87
6.2	Liste détaillée des publications	91
7	Bibliographie.....	94

Figures

Figure 1: Reproduction du portulan de Moka (Yémen) de 1683 réalisé par Augustine Fitzhugh, conservé au département carte et plan la bibliothèque nationale de France (disponible sur http://gallica.bnf.fr/)	14
Figure 2: Exemple d'un navire suivant une route à l'aide du logiciel GECDIS	16
Figure 3 : Trafic Maritime au sud du Royaume-Uni (source : http://www.bmt-isis.com	17
Figure 4: Modèle altimétrique Litto3D ® avec un drapage de la partie terrestre à l'aide d'une image aérienne	18
Figure 5 : Thématiques de recherche traitées vis-à-vis des SIG classiques.....	19
Figure 6: Thématiques et domaines informatiques abordés.....	20
Figure 7 : Exemple de TIN (Triangular Irregular Network).....	21
Figure 8 : exemple de couple de lignes (L_1 , L_2).....	23
Figure 9 : Calcul de distance de Fréchet sur des lignes ou des parties de lignes homologues	26
Figure 10 : exemple de couple de lignes (L_1 , L_2) suréchantillonnées d'emprise différente	27
Figure 11 : Exemple de 5 traits de côtes numérisés avec des techniques et des sources différentes, image de fond SPOT 5.....	30
Figure 12 : exemples d'un appariement partiel 1-1 multiple et de deux appariements partiels simples	33
Figure 13 : fusion d'un couple de lignes après identifications des points homologues	34
Figure 14: coefficient d'amortissement en fonction de la distance entre le point et l'origine d'un vecteur $\vec{v}_i$ pour un RI de 4.....	35
Figure 15: vecteur de déformation en fonction des vecteurs locaux proches	36
Figure 16 : Déformations élastiques des parties propres à partir des couples de points homologues.....	36
Figure 17 : Chenal non bruité (a) partie basse bruitée (b) et parties fusionnées.....	37
Figure 18 : Schéma descriptif de la méthodologie de fusion de MNT	38
Figure 19: schéma des algorithmes génétiques.....	42
Figure 20: Codage sous la forme de voisinage de solution de segmentation et reconstruction des segments.....	43
Figure 21 : Comparaison de la sélection de sondes à l'aide de la méthode classique (algorithme d'Oraas) et à l'aide de l'algorithme génétique. La zone haute est plus rugueuse que la zone basse.	45
Figure 22 : perception de l'espace classique et maritime	47
Figure 23: perception de l'espace maritime selon les règles de navigation	48
Figure 24: Perception de l'espace déformé par le contexte spatial	49
Figure 25 : Métamodèle multi-agents Agent-Groupe-Rôle étendu	50
Figure 26 : Capture d'écran de la fenêtre principale de TRANS. Deux navires réalisant une manœuvre d'évitement sont représentés.....	51
Figure 27: Modèle cognitif de décision d'expert basé sur des patrons.....	53
Figure 28 : fonction d'appartenance en fonction de la vitesse en nœuds pour les sous-ensemble flous "inconnu", "petit cargo" et "grand cargo"	54
Figure 29: fonction d'appartenance en fonction de la distance minimum en mille marin pour les 5 sous-ensembles flous.....	55

Figure 30 : Situation courante détectée et décisions associées aux deux patrons proches P1 et P2	56
Figure 31 : Une situation issue de la plate-forme de gestion du suivi des déplacements des navires	60
Figure 32: positions relatives en fonction de la relation topologique	61
Figure 33: Evaluation temporelle des positions relatives entre deux objets mobiles surfaciques	62
Figure 34: Evolution temporelle vitesse relativisée entre deux objets mobiles surfaciques.....	63
Figure 35 : Combinaison des mesures de position relative et de vitesse relativisée.....	64
Figure 36 : Copie d'écran du prototype de suivi absolu et relatif des déplacements de navires non dangereuse.....	66
Figure 37 Copie d'écran du prototype de suivi absolu et relatif des déplacements de navires partiellement plus dangereux	67
Figure 38 : Deux des cas d'école visualisés à l'aide de la vue absolue et de la vue relative (situation 1 le navire noir ralentit, situation 2 : il accélère)	68
Figure 39 : Graphe de zones dans la région de Brest.....	69
Figure 40 : Schéma UML des trajectoires et itinéraires.....	70
Figure 41 : filtrage des parties des trajectoires dans la zone de départ (a) sans troncature, (b) avec troncature	71
Figure 42: Calcul de la trajectoire médiane et nuages de positions au même temps normalisé ...	72
Figure 43: Couloir spatial autour de la trajectoire médiane.....	73
Figure 44: Couloir spatio-temporel.....	74
Figure 45 : Partition de l'espace à un instant relatif. Les positions ayant servi à l'analyse spatio-temporelle sont cerclées de blanc.....	76
Figure 46 : Les différentes étapes de la détection de positions inhabituelles	77
Figure 47 : Résumé chronologique des travaux de recherches	81
Figure 48 : Classification des différentes phases de mes travaux de recherche (modélisation, analyse, visualisation, simulation) en fonction du type de données traitées (passées, présentes ou futures).	82
Figure 49 : Interactions futurs entre les travaux de recherche réalisés	83
Figure 50 : Position P du navire et couloir spatio-temporel (a). Trajectoire du navire (b), P' position lors de la sortie de la route, positions et directions d'autres navires (en rouge et en blanc) lors de la sortie.	84

Tableaux

Tableau 1 : Calcul de la d_{Fd} pour le couple de lignes de la figure 8	24
Tableau 2 : Tableau des couples $MF_{i,j}$ $MD_{i,j}$ pour les lignes de la figure 8	25
Tableau 3: distance de Fréchet discrète maximale et moyenne pour les deux lignes de la figure 8 en fonction du suréchantillonnage	26
Tableau 4: matrice de distance (MD) associée aux 2 lignes de la figure 10.....	28
Tableau 5: Distances de Fréchet maximales et moyennes en zone dunaire.....	31
Tableau 6 : Fonction d'appartenance des 3 sous-ensembles flous de type de cargos en fonction de la vitesse.....	54
Tableau 7 : ensemble des patrons des ferries croisant un cargo prioritaire en zone extérieur (4) sans être privilégiés (GW pour Give Way).....	56
Tableau 8 : valeur des deux patrons pour chaque critère	57
Tableau 9: Synthèse des activités de recherche par domaines informatiques et thématiques de recherche	80

1 Introduction

Les Systèmes d'Information (SI) regroupent l'ensemble des éléments participant à la gestion, au stockage, au traitement, au transport et à la diffusion de l'information au sein d'une organisation¹. Les systèmes d'information Géographique (SIG) sont donc par définition, des systèmes permettant de communiquer et de traiter l'information géographique (Denègre Salgé 2004). Les Systèmes d'Information Géographique (SIG) ont été conçus à l'origine afin d'acquérir, d'archiver, d'analyser et d'afficher des données géographiques à l'aide d'un modèle des phénomènes du monde réel. L'objectif des SIG est de fournir aux utilisateurs une meilleure compréhension de l'espace géographique et des phénomènes qui le composent. Pour décrire ces phénomènes dans le cadre des SIG, une description abstraite spécifique aux sujets d'étude, est donc dans un premier temps définie. Cette description a le plus souvent pour fondement les représentations cartographiques papier et est définie à l'aide d'un ensemble d'hypothèses simplificatrices (**abstraction**).

Les données localisées peuvent alors être **acquises** notamment à l'aide de méthodes très diverses :

- la photographie aérienne ou satellitaire,
- le relevé direct sur le terrain à l'aide de différents capteurs,
- la numérisation de documents papier déjà existants.

Les données collectées sont donc de qualité très variée. Elles servent de socle aux applications des SIG. Une fois acquises, elles sont **archivées** dans des Bases de Données Géographiques (BDG). Ces dernières sont des ensembles structurés d'informations. Comme toutes les BD, les informations sont placées dans des fichiers, et organisées de manière à pouvoir être facilement triées, classées et modifiées à l'aide d'un logiciel spécialisé appelé système de gestion de base de données (SGBD). Les SGBD Géographique (SGBDG) qui contiennent la plupart du temps de gros volumes de données sont optimisés pour extraire rapidement les données en fonction de critères spatiaux et de disposer de langage de requêtes spécifiques.

Ces données peuvent alors être **analysées** afin de répondre à des applications de domaines multiples (urbanisme, environnement, gestion de réseau, militaire, etc.). Dans ce cadre, les SIG sont donc des aides aux décideurs qui doivent aider à répondre à des questions variées, telles :

- Questions de localisation : Où se trouve tel objet ?
- Questions d'identification spatiale : Qui y a-t-il à tel endroit ?
- Questions d'analyse de l'espace : Comment sont répartis les objets de tel type sur le territoire ?
- Questions sur l'évolution de l'espace : À quel moment tel changement est-il intervenu ?
- Questions de simulation : que se passera-t-il si tel objet est modifié pour tel autre type d'objets ?

Les évolutions des outils d'analyses des SIG s'inscrivent pour la plupart dans la droite ligne de **l'analyse spatiale** (Bavoux 1998, Pumain et Saint-Julien 2004). L'analyse spatiale propose des modèles de l'espace géographique qui ont pour objectifs de comprendre les logiques, les causes et les conséquences de la localisation des humains et de leurs activités. Ces modèles sont capables de mettre en évidence des formes récurrentes d'organisation spatiales et des théories, notamment à travers diverses notions-clés : distance, réseaux, structure, etc. L'espace

¹ Norme ISO 5127-1-1983

géographique n'est donc pas simplement un environnement. Il est un des éléments fondamentaux du modèle.

Les données des SGBDG et issues des analyses doivent pouvoir être restituées de manière graphique (**affichage**) afin de communiquer avec l'utilisateur. La **carte** de type papier a servi (et continue de servir) de moyen de communication privilégié pour transmettre de tels messages ; c'est une vue, une représentation particulière d'un espace géographique adaptée à une utilisation précise, avec des informations différentes. Elle répond à des normes (échelle, légende, sémiologie graphique, etc.) qui doivent faciliter sa lisibilité. Elle demande cependant à l'utilisateur un effort d'abstraction pour associer les phénomènes aux symboles graphiques. Le passage du support papier au support numérique a permis à chaque utilisateur, à l'aide d'un SIG, de cartographier l'espace géographique selon ses propres besoins et son métier. De plus, la représentation n'est plus statique. Les objets de la représentation peuvent être mobiles, la légende être modifiée, l'échelle s'adapter à la taille de l'écran et de l'espace à cartographier.

Durant les 10 dernières années, de nombreuses évolutions ont été observées. Qui plus est, le déploiement d'un grand nombre de systèmes de réception GPS et le développement de la cartographie sur Internet ont stimulé l'utilisation des SIG auprès des utilisateurs novices. Cette révolution a été accompagnée par l'intégration dans les SIG de données de plus en plus précises et complexes à manipuler. La troisième dimension a ainsi fait son apparition. Les données planimétriques sont devenues volumiques. Les phénomènes du monde réel peuvent ainsi être modélisés en trois dimensions (3D). Le relief d'une zone terrestre naturel peut notamment être représenté à l'aide d'un modèle numérique de terrain (**MNT**). Ces modèles associent à chaque point de l'espace modélisé, une altitude à l'aide d'un modèle régulier ou irrégulier. De même, les SIG intègrent maintenant de plus en plus la dimension **temporelle**. La position des données n'est donc plus fixe. Elle peut évoluer dans le temps (déplacement d'un objet mobile, redécoupage administratif). Finalement, pour répondre aux besoins des applications de plus en plus variées et complexes, plusieurs sources sont souvent nécessaires. Or, les techniques d'acquisition sont hétérogènes et produisent des mesures qui peuvent s'avérer être incohérentes.

Mes travaux de recherche sont réalisés à l'Institut de Recherche de L'Ecole navale (IRENav). L'IRENav est un centre de recherche pluridisciplinaire orienté vers le domaine maritime. Il est le support de la formation scientifique au sein de l'Ecole navale. L'IRENav est membre du groupement d'intérêt scientifique " Europôle Mer ". Il est reconnu comme Equipe d'Accueil par la MSU (EA3634) dans le cadre de la contractualisation des laboratoires de l'Ecole des Arts et Métiers ParisTech. Le contexte d'application de mes recherches est donc naturellement le **domaine maritime**.

Le **groupe SIG** est une des trois équipes de cet institut. Les travaux de recherche de ce groupe sont divisés en deux axes depuis 2008 :

- Axe 1 : Modélisation et analyse spatio-temporelle d'environnements naturels 2D et 3D,
- Axe 2 : La modélisation et la visualisation d'informations spatio-temporelles mobiles.

Je suis responsable de l'axe 2. Cependant, mes travaux s'insèrent dans les deux axes. Ils traitent plus particulièrement de deux thématiques :

- Les systèmes d'aide à la navigation liés à la sécurité maritime (axe 2),
- Les modèles numériques de terrain (MNT) maritimes et côtiers (axe1).

Ses travaux bénéficient également de la dynamique du groupe de recherche sur les SIG (GDR SIGMA puis GDR MAGIS).

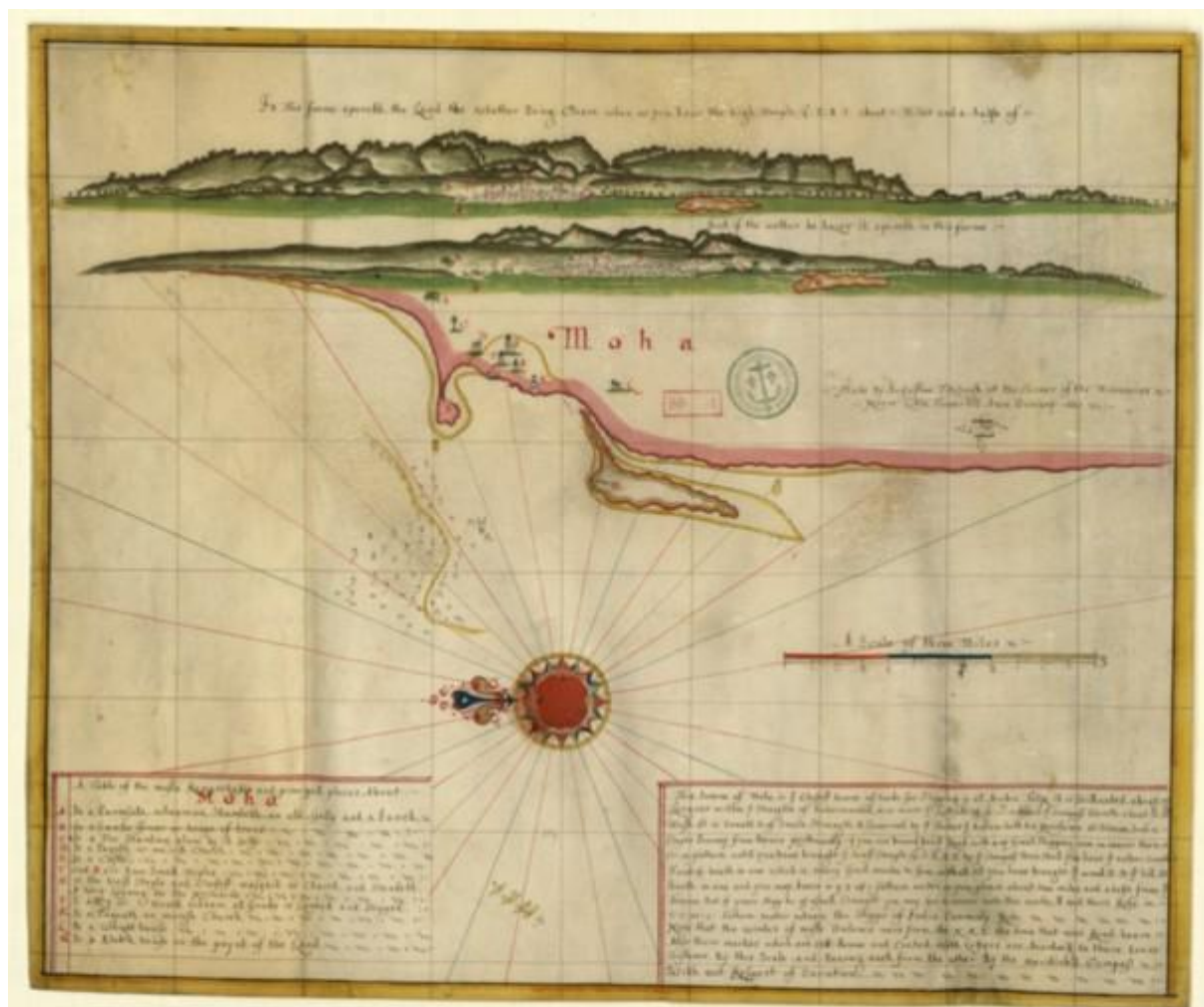


Figure 1: Reproduction du portulan de Moka (Yémen) de 1683 réalisé par Augustine Fitzhugh, conservé au département carte et plan la bibliothèque nationale de France (disponible sur <http://gallica.bnf.fr/>)

Les navigateurs font partie des premiers utilisateurs de cartes. Ces cartes ont toujours eu pour objectifs de se situer, de se diriger et d'identifier les dangers (récifs, hauts-fonds, épaves, munitions immergées). Les cartographes de marine ont souvent fait part d'originalité pour représenter l'espace marin. Les anciennes cartes nautiques servant à repérer les ports, les portulans, sont très innovants pour l'époque (XIV – XVIII siècles). Par exemple la figure 1 représente l'accès au port de Moka au Yémen et date de 1683. Son dessinateur Augustine Fitzhugh, a la volonté de regrouper l'ensemble des informations utiles aux navigateurs. Il fournit trois représentations complémentaires. Une représentation planimétrique terrestre et maritime, qui représente les dangers de la zone (les deux bancs de sable), et les points remarquables (amers) pour que les navigateurs se repèrent. Parallèlement, il inclut deux vues cavalières. La première vue (en haut) par temps clair est dessinée depuis l'endroit où les navires doivent mouiller. Elle représente une petite ville avec des montagnes en arrière-plan. Pour la deuxième, dessinée au même endroit, mais par temps couvert, les montagnes sont cachées et la ville sous l'effet de la brume (sorte de mirage), apparaît plus grande. Qui plus est, l'orientation (Nord à gauche) est telle que le port apparaisse de face afin de faciliter l'arrivée à Moka. Finalement, des informations sur les conditions de vent viennent parfaire cette représentation. Ce portulan qui s'affranchit des contraintes cartographiques classiques,

permet de fournir des représentations complémentaires, innovantes pour l'époque et fort utiles.

Aujourd'hui les passerelles des navires modernes embarquent de plus en plus de systèmes d'aide à la navigation. Ces systèmes ont pour objectif de faciliter le positionnement du navire vis-à-vis de son environnement tels les autres navires, les hauts fonds et la côte. Parmi ces systèmes, il faut citer :

- Les RADAR ARPA : radars qui facilitent le suivi des échos et donc le positionnement relatif d'autres navires afin d'aider le navigateur dans le choix d'une manœuvre pour éviter la collision.
- Les transpondeurs AIS () qui gèrent l'envoi et la réception via VHF des positions GPS, vitesse, cap, type, lieu et heure d'arrivée des navires, vers et depuis les navires environnants.
- Les systèmes d'information et de cartographie ECDIS (Electronic Chart and Information Display). Ces systèmes permettent de superposer sur une carte électronique les positions des navires et de gérer des alarmes (risque de collision, risque d'échouement). La figure 2 présente l'interface du logiciel GECDIS (ECDIS de la société Sodena). Un navire (en noir) suit une route préalablement définie (ligne rouge en pointillée) à l'intérieur d'un chenal de navigation. Une carte électronique de navigation (ENC) suivant la norme S57 est employée. Ces cartes vectorielles fournissent des descriptions numériques individuelles de chacun des objets qui figurent sur la carte papier (isobathes², sondes³, balisages, trait de côte, etc.) et des liens entre certains de ces objets.

² Lignes joignant des points d'égale profondeur

³ Profondeur du fond par rapport au zéro hydrographique (le niveau théoriquement atteint par les plus basses mers) associée à une position géographique (latitude, longitude)

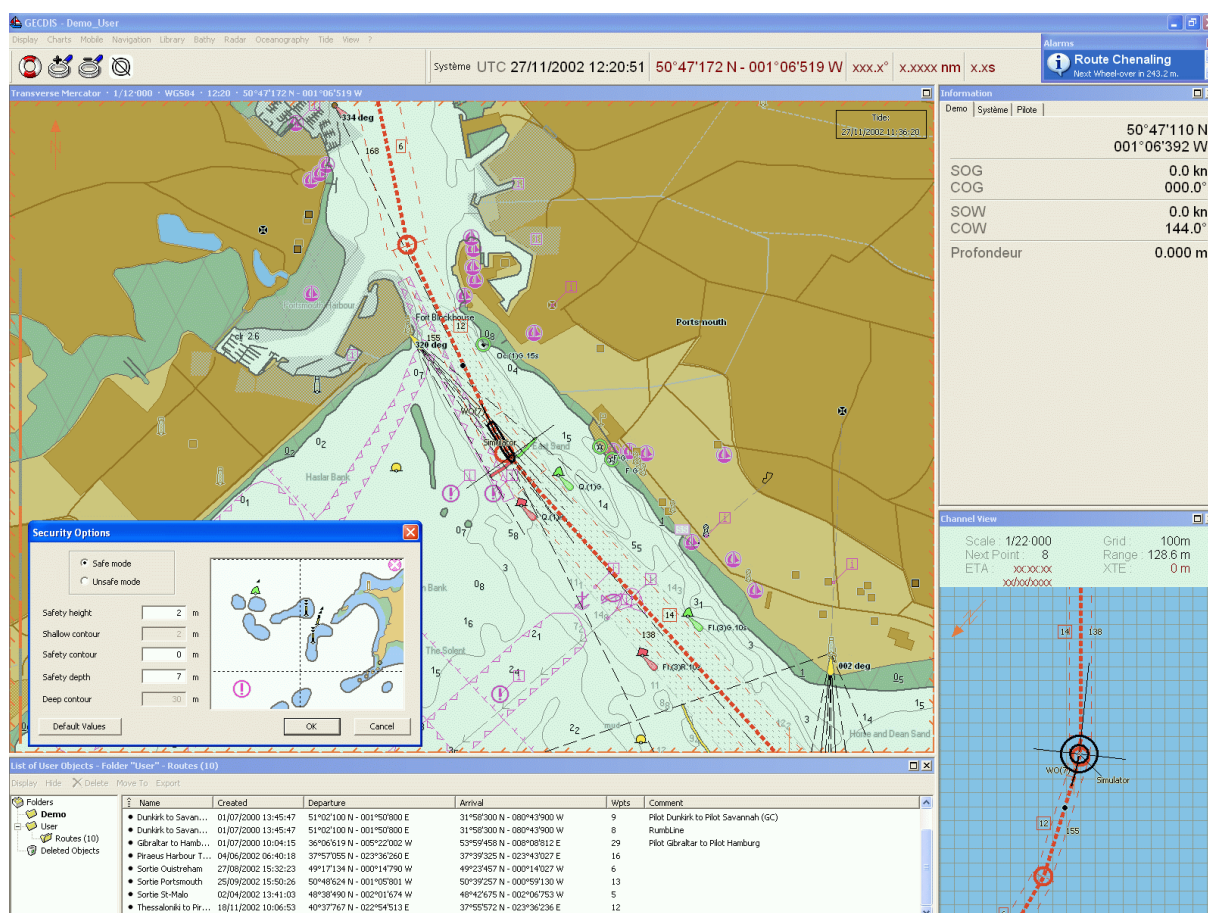


Figure 2: Exemple d'un navire suivant une route à l'aide du logiciel GECDIS

Parallèlement, les centres de surveillance à terre : les CROSS (centres régionaux opérationnels de surveillance et de sauvetage), sémaphore et capitainerie, sont équipés de systèmes favorisant le suivi des déplacements des navires sur de larges zones. La formation des marins a suivi la même direction. Elle intègre ces systèmes d'aide à la navigation et fait de plus en plus appel à des simulateurs de navigation qui se doivent d'être les plus réalistes possible. L'OMI (organisme maritime international) accorde une grande importance au développement de ces systèmes à travers de nombreux groupes de travail sur l'e-navigation. Néanmoins, l'ensemble des données produites par ces différents systèmes est encore sous-exploité en termes d'amélioration de la sécurité maritime. Le plus souvent, seule la dernière position de chaque navire est prise en compte. Or, l'évolution de la position en ce qui concerne la sécurité est primordiale. De même, l'analyse de l'historique des déplacements permet de détecter des comportements inhabituels.

Ces systèmes embarqués ou à terre sont d'autant plus importants que le trafic maritime est dense et augmente. La figure 3 fournit une image de la densité des positions de navires issues de transpondeurs AIS du sud du Royaume-Uni. En Manche, chaque jour, plus de 70 000 passagers traversent entre le Royaume-Uni et la France et 700 à 800 bateaux passent par jour dans le détroit du Pas de Calais⁴.

⁴ Source Atlas Transmanche de l'université de Caen

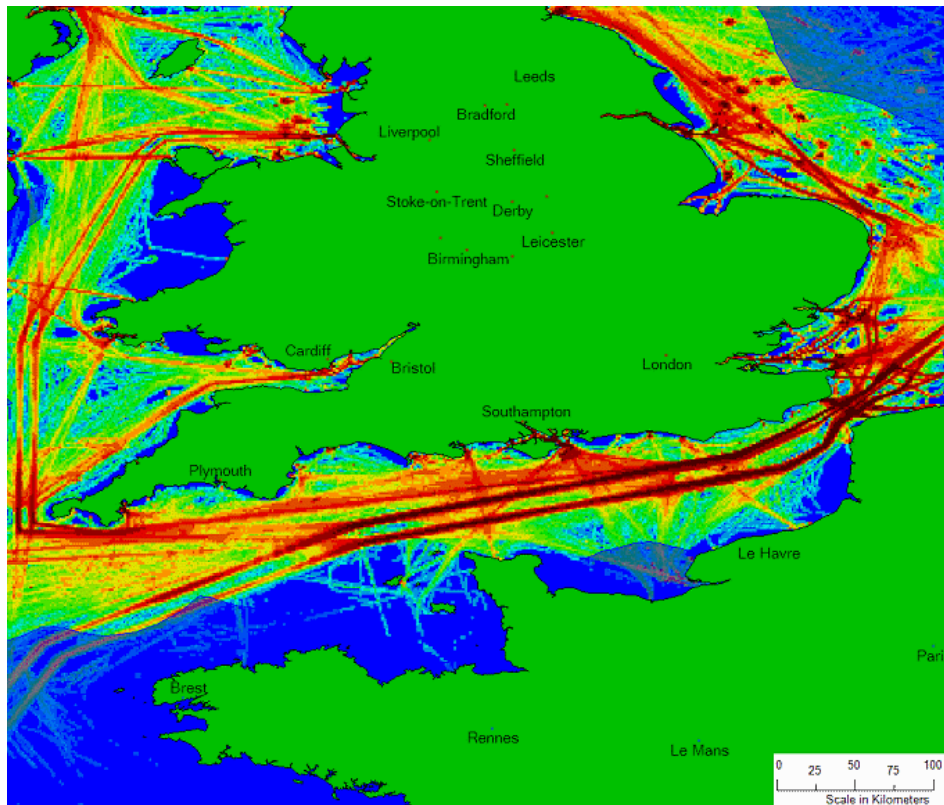


Figure 3 : Trafic maritime au sud du Royaume-Uni (source : <http://www.bmt-isis.com>)

En ce qui concerne les modèles numériques de terrain (MNT) maritimes et côtiers, les systèmes d'acquisition actuels (sondeurs multifaisceaux, LIDAR⁵ bathymétriques et topographiques) produisent des volumes de données ponctuelles de plus en plus importants. Toutefois, les utilisateurs vont manipuler des représentations du terrain simplifiées sous la forme notamment :

- de points remarquables,
- de courbes de niveaux ou d'isobathes,
- des partitionnements de l'espace à l'aide de grilles régulières,
- de Réseau de Triangles Irréguliers (TIN, Triangular Irregular Network).

Ces représentations nécessitent l'emploi d'outils de généralisation et de partitionnement. Par ailleurs, les MNT terrestres et maritimes sont à l'heure actuelle dissociés. Or, de plus en plus d'applications côtières nécessitent de disposer de MNT terre-mer continu. En France, il a été demandé au SHOM et à l'IGN de définir un référentiel géographique du littoral (RGL). Dans ce cadre, le projet Litto3D (Le Roux et Martinoty 2009) a pour objectif de définir un modèle "terre-mer" en trois dimensions. Les premiers modèles Litto3D ont été acquis dans le golfe du Morbihan. La figure 4 est une image d'une partie du modèle, les différences de granularité du modèle sont dues à l'acquisition à l'aide de différentes techniques (SONAR, LIDAR bathymétrique, LIDAR).

⁵ LIDAR est l'acronyme de "Light Detection and Ranging". Le LIDAR est dans ce cas, un système aéroporté de mesure du terrain par balayage laser.

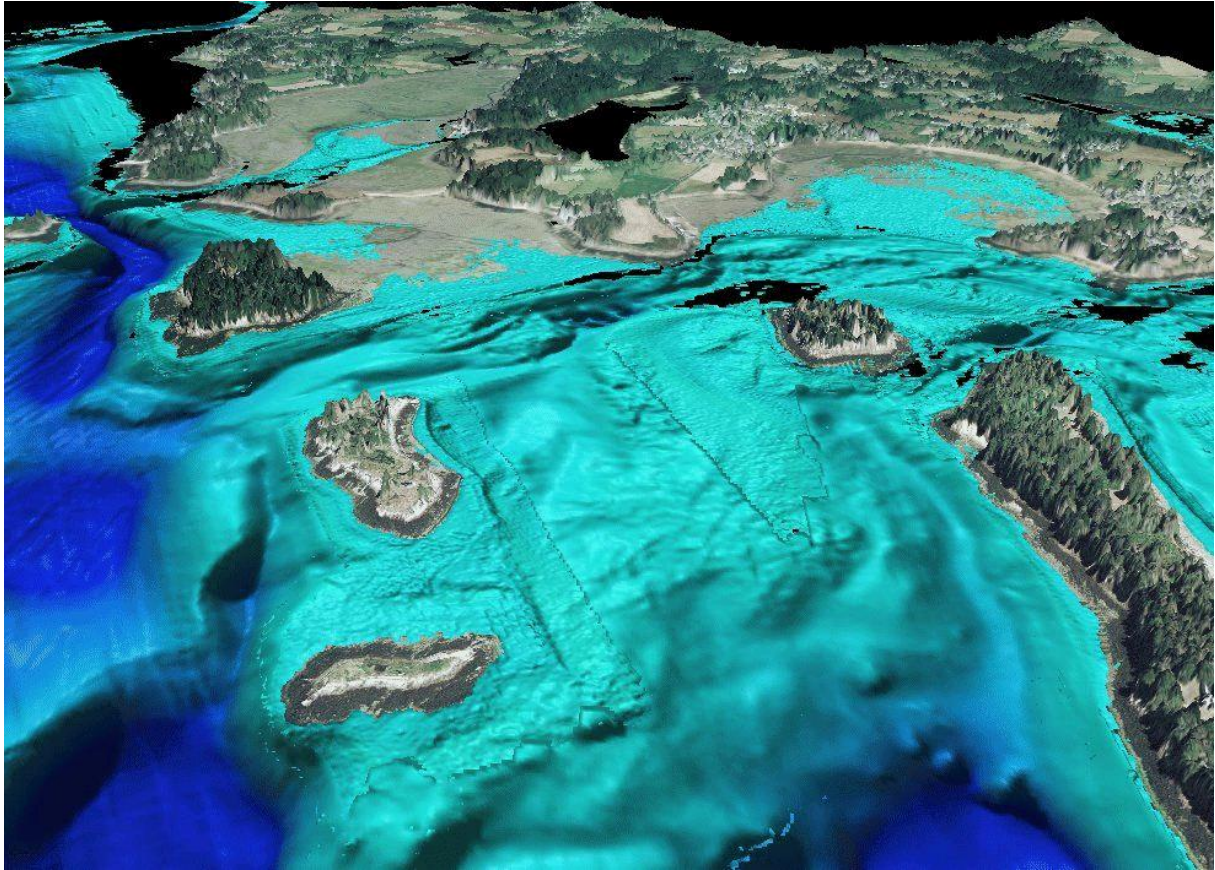


Figure 4: Modèle altimétrique Litto3D ® avec un drapage de la partie terrestre à l'aide d'une image aérienne

La figure 5 présente les thématiques abordées par mes travaux de recherches vis-à-vis des SIG classiques. La branche du haut s'attache à résoudre certains problèmes rencontrés par les SIG pour gérer les MNT. Les problèmes traités sont :

- La **généralisation** de MNT maritime qui consiste à obtenir une représentation plus simple et plus abstraite. Le but de la généralisation est donc de réduire le nombre de données transmises (sondes⁶) en conservant autant que possible l'essentiel de l'information qu'elles véhiculent. La carte marine devant assurer juridiquement la sécurité de l'utilisateur, sa généralisation est caractérisée par l'existence de contraintes imposées par des normes de l'Office Hydrographique International (OHI).
- La **segmentation** de MNT qui consiste à définir une partition d'un MNT. Chaque partie définie doit être homogène en termes de pente et de rugosité et la plus étendue possible.
- L'**appariement** d'éléments caractéristiques de MNT qui consiste à détecter dans la zone de recouvrement des deux MNT, les éléments morphologiques linéaires homologues (crêtes, talwegs, rupture de pente).
- La **fusion** qui a pour objectif de fusionner les éléments appariés et à déformer à l'aide de technique de déformation élastique les MNT afin de tenir compte des déplacements liés à la fusion des éléments caractéristiques.

⁶ Un point de sonde est un emplacement où la profondeur a été mesurée.

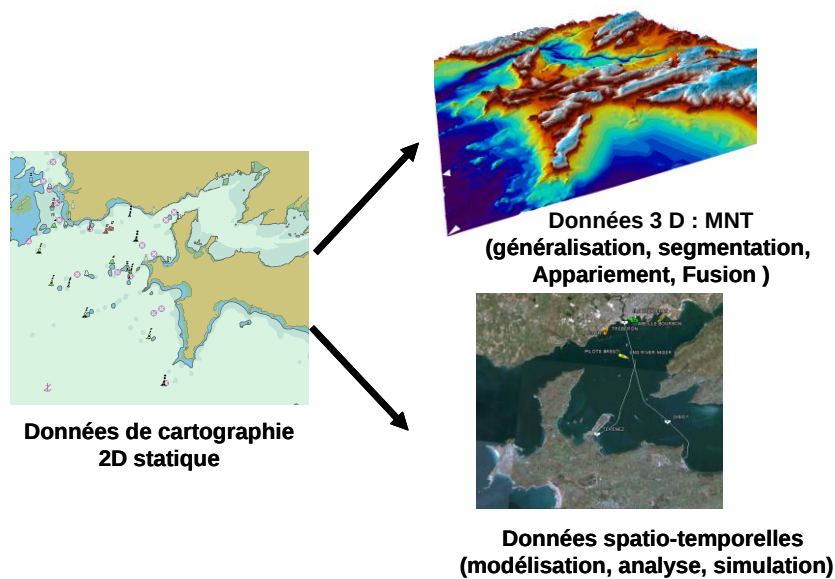


Figure 5 : Thématiques de recherche traitées vis-à-vis des SIG classiques

La partie du bas de la figure 5 a elle pour objectif de prendre en compte la dimension temporelle. Les problématiques abordées concernent les déplacements d'objets mobiles. Elles portent sur :

- La **modélisation** des déplacements relatifs pour mieux faire ressortir l'évolution des distances et des différentiels de vitesses entre les objets mobiles,
- L'**analyse** des trajectoires des objets mobiles suivant un même itinéraire afin de déduire des comportements génériques des objets de même type et ayant un même itinéraire. Cette analyse devant permettre dans un deuxième temps, d'identifier des trajectoires inhabituelles lors du suivi,
- La **simulation** des déplacements des objets mobiles, en tenant compte des différentes actions des navires, de l'environnement géographique et des comportements hétérogènes des pilotes de ces objets.

J'ai abordé ces deux thématiques maritimes, d'un point de vue informatique, selon trois axes :

- Les **bases de données spatio-temporelles** : Les vues spatio-temporelles et la fouille de données spatio-temporelles,
- L'**intelligence artificielle appliquée aux SIG** : Les algorithmes génétiques, la logique floue et les systèmes multi-agents,
- La **géométrie algorithmique** : les distances linéaires, la fusion de données et les déformations élastiques, les distances spatio-temporelles sur des trajectoires.

La figure 6 résume les domaines et les thématiques informatiques abordées dans ce manuscrit. La valeur de gris traduit mon investissement à l'intersection de ce domaine et de cette thématique.

	Géométrie Algorithmique	Intelligence Artificielle appliquée aux SIG	Bases de Données Spatio-temporelles
Données 3D : MNT			
Données spatio-temporelles : objets mobiles			

Figure 6: Thématiques et domaines informatiques abordés

La suite de ce manuscrit se décompose de la manière suivante. La section 2 décrit les travaux réalisés en termes de géométrie algorithmique. Ces derniers portent entre autres sur l'intégration de MNT terrestre et maritime. La thèse de M^{elle} Mascret traite de cette problématique. Récemment, ils ont été étendus aux trajectoires maritimes dans le cadre de la thèse de M. Etienne. La section 3, synthétise les résultats obtenus dans le domaine de l'intelligence artificielle appliquée aux SIG. Ces études ont été réalisées en partie dans le cadre du stage de DEA de M. Brosset sur la généralisation et la partition de MNT à l'aide d'algorithme génétique. Parallèlement, des travaux ont été réalisés sur la simulation des déplacements des navires s'appuyant sur des systèmes multi-agents. Dans ce domaine, la thèse de M. Fournier a montré l'intérêt de définir plusieurs rôles pour chaque objet mobile afin de mieux modéliser leurs déplacements. La thèse de M. Le Pors a prolongé ces travaux en définissant des rôles cognitifs permettant de modéliser les raisonnements d'humains. La section 4, se focalise sur mes recherches en base de données spatio-temporelles. Dans ce contexte, les travaux concernent la définition de vues se focalisant sur les déplacements des navires vis-à-vis d'un navire-référent. Ces vues sont complémentaires des représentations absolues classiques et ont été modélisées lors de la thèse de M^{elle} Noyon. Finalement, des processus de fouille de données ont été initiés en 2008. Ils portent sur l'extraction de routes types à partir des bases de données de trajectoires et la qualification en temps réel de trajectoires vis-à-vis de ces routes types (thèse de M. Etienne). La 5^{ème} et dernière section résume les apports principaux de ces travaux de recherche et expose les principales pistes de recherche que j'envisage dans les prochaines années.

2 La géométrie algorithmique

Pour décrire l'environnement géographique, deux modes de représentations ont été développés :

- Le mode **raster** (ou mode maillé) où l'espace est découpé en une grille composée de pixels. Pour chaque pixel, un ensemble de valeur permet de décrire les caractéristiques de ce pixel (hauteur pour les MNT, occupation du sol, nombre d'habitants, etc.).
- Le mode **vecteur** où l'espace est composé d'objets ayant chacun une géométrie spécifique. Un objet peut donc être décrit par un point. Les coordonnées du point permettent de rattacher cet objet à un lieu sur terre. Ils peuvent être géographiques (latitude, longitude) ou cartographiques (X, Y) et être associés à un système de projection cartographique. Une troisième dimension : l'altitude, est associée si la modélisation n'est pas seulement planimétrique. Une représentation linéaire sous la forme d'une suite ordonnée de points reliés par des segments de droite (polyligne) ou par toute autre ligne définie mathématiquement, est aussi autorisée. La représentation surfacique est le troisième type possible. La géométrie de l'objet est alors décrite par une ligne fermée décrivant sa frontière. Des trous à l'intérieur de la géométrie peuvent être associés à la représentation. Ces trous sont des surfaces. Finalement, la géométrie d'un objet peut être composée d'un ensemble de géométries de ces trois types.

Les travaux présentés dans ce document concernent quasi uniquement des représentations objet de l'environnement géographique à l'aide du mode vecteur. Pour les modèles numériques de terrain (MNT) des représentations sous la forme de **TIN** (Triangular Irregular Network) sont employées. Dans cette représentation vecteur, un ensemble de points en trois dimensions (sommets) répartis irrégulièrement sur l'espace, sont reliés deux à deux par des arcs. Ces triangles forment une partition de l'espace. La figure 7 fournit un exemple de TIN. Les MNT définis par les SIG, ne sont pas des objets 3D au sens strict. A chaque coordonnée planimétrique, une seule altitude est généralement associée. Le terme de 2.5D est souvent utilisé pour les décrire.

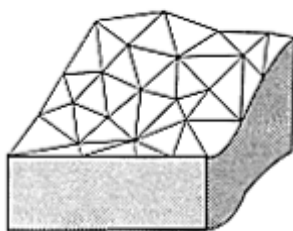


Figure 7 : Exemple de TIN (Triangular Irregular Network)

De même pour étendre le mode vecteur, aux objets mobiles, les notions de position et de trajectoire sont définies. La **position** d'un objet mobile est un point auquel une estampille temporelle (une date) est associée. Cette estampille peut être absolue (15 mai 1988 à 10h30 et 15 secondes) ou relative (10 min et 30 secondes depuis le départ de l'objet mobile). Des attributs qui décrivent la cinématique de l'objet mobile peuvent être associés à la position. Leur présence dépend du type de capteur employé pour recueillir les données. La vitesse, la direction, le taux de giration et l'accélération sont des attributs qui améliorent la connaissance sur le déplacement de l'objet mobile à cet instant.

Une **trajectoire** est une suite ordonnée temporellement de positions. A chaque instant, une seule position est possible pour chaque objet. La représentation étant discrète, une fonction

d'interpolation est indispensable pour connaître l'emplacement d'un objet entre deux positions connues. Cette fonction est plus ou moins précise en fonction de l'intervalle temporel entre deux positions et des attributs sur la cinématique qui sont gérés.

Afin d'extraire et de modifier les objets géométriques, des algorithmes sont définis. La **géométrie algorithmique** est la partie de l'algorithmique qui traite des algorithmes qui s'attachent à la manipulation des objets géométriques. L'infographie est le domaine qui a le plus contribué à cette discipline. Dans le domaine des SIG, de nombreux travaux ont aussi été menés. Les premiers travaux concernaient essentiellement la cartographie. Les algorithmes définis permettent de changer de représentation géométrique, de simplifier des représentations cartographiques. Un des algorithmes le plus connu est celui dit de Douglas et Peucker (Douglas et Peucker 1973) qui permet de compresser une géométrie linéaire en supprimant les points intermédiaires ayant les positions les moins significatives. Mustière (Mustière 2001) décrit un très grand nombre d'algorithmes lié à la cartographie numérique.

Dans le domaine de la géométrie algorithmique, mes travaux concernent la définition de distances linéaires et leur utilisation pour fusionner des représentations. Les objets étudiés sont les polygones et les trajectoires.

2.1 Les distances linéaires

Les distances linéaires servent à calculer l'écart entre deux lignes. Cet écart peut être minimal, moyen ou maximum. En termes d'étude de similitude, l'écart maximum est le plus intéressant. Il fournit une indication de l'espacement maximal entre deux lignes. L'écart moyen est aussi utile. Il fournit un indicateur de ressemblance global. Par contre l'écart minimal, est une mesure de proximité entre deux objets linéaires. Cette mesure a peu de sens dans ce contexte.

Les deux distances linéaires maximales les plus usuelles sont :

- la distance de Hausdorff (Hausdorff 1919),
- la distance de Fréchet : d_F (Fréchet 1906).

La distance de **Hausdorff** recherche pour chaque point de chaque ligne, le point le plus proche de l'autre ligne. Cette distance souffre de ne pas tenir compte de l'ordre des points. Deux polygones semblables ont nécessairement une distance de Hausdorff faible. Par contre, deux polygones dissemblables peuvent avoir une distance de Hausdorff très faible (Devogele 2000).

La deuxième, plus complexe, définie par **Fréchet**, s'appuie sur la propriété suivante : toute polygone orientée est équivalente à une application continue $f : [a, b] \rightarrow V$ ou $a, b \in \mathcal{R}$, $a < b$ et V est l'espace vectoriel. La distance de (d_F) est la suivante :

Soit $f : [a, a'] \rightarrow V$ et $g : [b, b'] \rightarrow V$ deux polygones et $\| \cdot \|$ une norme,

$$d_F(f, g) = \inf_{\alpha: [0,1] \rightarrow [a,a']} \max_{\beta: [0,1] \rightarrow [b,b']} \|f(\alpha(t)) - g(\beta(t))\|$$

Une illustration intuitive de la distance de Fréchet est la suivante : un maître et son chien suivent deux chemins. Ils avancent ou s'arrêtent à volonté, indépendamment l'un de l'autre, mais ils ne peuvent pas revenir sur leurs pas. La distance de Fréchet entre ces deux chemins est la longueur minimale de la laisse qui permet de réaliser un cheminement de concert satisfaisant ces conditions.

La distance de Fréchet d_F a l'avantage d'être calculée uniquement à partir de couples de points qui peuvent être mis en correspondance visuellement (points homologues). La distance de Fréchet est donc très proche d'une distance maximale "visuelle" entre deux lignes. Hélas, cette distance a l'inconvénient d'être complexe à programmer. Alt et Godau (Alt et Godau 1992) fournissent une comparaison très exhaustive de ces deux distances linéaires maximales.

2.1.1 Distance de Fréchet discrète maximale et moyenne

Une approximation discrète, d_{dF} (Eiter et Mannila 1994) a donc été retenue. Elle fournit une mesure dont l'erreur peut être bornée à l'aide d'un suréchantillonnage des lignes :

$$d_F(L_1, L_2) \leq d_{dF}(L_1, L_2) \leq d_F(L_1, L_2) + LongMax$$

$LongMax$ est la longueur maximale des Segments de L_1 et L_2 .

Dans un premier temps, la rapidité de calcul de cette approche discrète a été améliorée (Devogele 2000). Si nous notons L_1 et L_2 les deux polygones chacune composée d'une suite ordonnée des extrémités des segments $\langle L_{1,1} \dots L_{1,n} \rangle$ pour L_1 et $\langle L_{2,1} \dots L_{2,m} \rangle$ pour L_2 . La distance de Fréchet discrète entre L_1 et L_2 ($d_{Fd}(L_1, L_2)$) peut se calculer récursivement à partir de la formule suivante :

$$d_{Fd}(L_1, L_2) = \max \left(\begin{array}{c} d_E(L_{1,n}, L_{2,m}) \\ \min \left(\begin{array}{c} d_{Fd}(\langle L_{1,1} \dots L_{1,n-1} \rangle, \langle L_{2,1} \dots L_{2,m} \rangle) \text{ si } n \neq 1 \\ d_{Fd}(\langle L_{1,1} \dots L_{1,n} \rangle, \langle L_{2,1} \dots L_{2,m-1} \rangle) \text{ si } m \neq 1 \\ d_{Fd}(\langle L_{1,1} \dots L_{1,n-1} \rangle, \langle L_{2,1} \dots L_{2,m-1} \rangle) \text{ si } n \neq 1, m \neq 1 \end{array} \right) \end{array} \right)$$

A l'aide de technique de programmation dynamique, le temps de calcul a été optimisé. Pour cela deux matrices sont calculées :

- la matrice de distance (MD) dont la cellule MD_{ij} a pour valeur la distance euclidienne (d_E) entre $L_{1,i}$ et $L_{2,i}$,
- la matrice de Fréchet (MF) qui permet de calculer la distance de Fréchet. La formule générale pour calculer MF_{ij} est :

$$MF_{ij} = \max (D_E(L_{1,i}, L_{2,j}), \min(MF_{i-1,j}, MF_{i,j-1}, MF_{i-1,j-1})).$$

La distance de Fréchet discrète (d_{Fd}) a pour valeur $MF_{n,m}$. Les valeurs des autres cellules $MF_{i,j}$ correspondent à la $d_{Fd}(\langle L_{1,1} \dots L_{1,i} \rangle, \langle L_{2,1} \dots L_{2,j} \rangle)$. Ces valeurs sont donc des distances partielles.

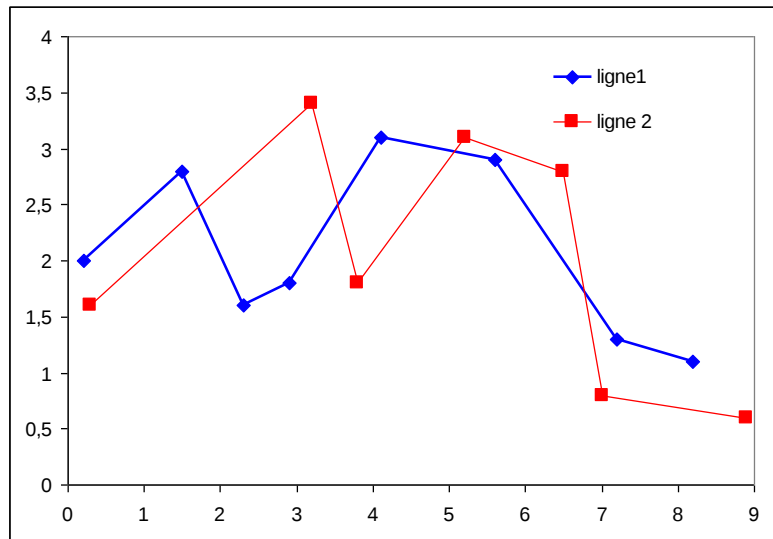


Figure 8 : exemple de couple de lignes (L_1, L_2)

Pour illustrer ce calcul, le couple de lignes de la figure 8 va être employé. Le tableau 1 fournit les valeurs des cellules de MD et MF pour ce couple de lignes. La distance de Fréchet discrète (d_{Fd}) est la valeur de $MF_{8,7}$, soit 1,8. Elle correspond à la distance euclidienne (d_E) entre les

points $L_{1,2}$ et $L_{2,2}$ ($MD_{2,2}$). Pour améliorer la mesure d'écart, un suréchantillonnage peut être réalisé. Avec un pas de 0,1, deux matrices (107×129) sont calculées et la d_{Fd} est égale à 1,2.

$L_{2,j,x}$	0,3	3,2	3,8	5,2	6,5	7	8,9
$L_{2,j,y}$	1,6	3,4	1,8	3,1	2,8	0,8	0,6

		Matrice de distance (MD)							
$L_{1,i,x}$	$L_{1,i,y}$	1	2	3	4	5	6	7	
0,2	2	1	0,41	3,31	3,61	5,12	6,35	6,91	8,81
1,5	2,8	2	1,70	1,80	2,51	3,71	5,00	5,85	7,72
2,3	1,6	3	2,00	2,01	1,51	3,26	4,37	4,77	6,68
2,9	1,8	4	2,61	1,63	0,90	2,64	3,74	4,22	6,12
4,1	3,1	5	4,09	0,95	1,33	1,10	2,42	3,70	5,41
5,6	2,9	6	5,46	2,45	2,11	0,45	0,91	2,52	4,02
7,2	1,3	7	6,91	4,52	3,44	2,69	1,66	0,54	1,84
8,2	1,1	8	7,92	5,50	4,46	3,61	2,40	1,24	0,86

		Matrice de Fréchet (MF)						
$L_{1,i,x}$	$L_{1,i,y}$	1	2	3	4	5	6	7
1	0,41	3,31	3,61	5,12	6,35	6,91	8,81	
2	1,70	1,80	2,51	3,71	5,00	5,85	7,72	
3	2,00	2,01	1,80	3,26	4,37	4,77	6,68	
4	2,61	2,00	1,80	2,64	3,74	4,22	6,12	
5	4,09	2,00	1,80	1,80	2,42	3,70	5,41	
6	5,46	2,45	2,11	1,80	1,80	2,52	4,02	
7	6,91	4,52	3,44	2,69	1,80	1,80	1,84	
8	7,92	5,50	4,46	3,61	2,40	1,80	1,80	

Tableau 1 : Calcul de la d_{Fd} pour le couple de lignes de la figure 8

Une fois cet algorithme optimisé, il a été amélioré afin d'en déduire :

- les couples de points homologues,
- une distance moyenne dérivée des distances entre ces couples.

A partir des valeurs des deux matrices, un algorithme (Mascret et *al.* 2006) a donc été défini afin d'identifier les **couples de points homologues**. L'algorithme de recherche des points homologue est défini par analogie avec le chemin réalisé par le maître et son chien. Le chemin entre les couples de points ($L_{1,1}$, $L_{2,1}$) et ($L_{1,n}$, $L_{2,m}$) doit donc être compatible avec une laisse de longueur d_{Fd} . De plus, pour que le chemin soit optimal, la laisse doit être en moyenne, la moins tendue. Pour l'exemple des lignes la figure 8, 32 chemins compatibles sont envisageables. Par exemple, après le couple ($L_{1,1}$, $L_{2,1}$) deux couples sont possibles ($L_{1,2}$, $L_{2,1}$) ou ($L_{1,2}$, $L_{2,2}$). Le chemin optimal (CO) est donc celui où le chien est le plus proche possible du maître.

Dans le but de construire le CO, L'opérateur $\leq$ entre deux couples de réel, est défini de la manière suivante :

Soit (a, b) et $(c, d) \in \mathbb{R}^2$

$(a, b) \leq (c, d)$ si $a < c$ ou
si $a = c$ et $b \leq d$

Le CO est construit par backtracking (retour sur trace) à travers les matrices. Le CO est initialisé avec le couple des points finaux : $(L_{1,n}, L_{2,m})$. Le couple précédant est choisi parmi les trois couples de points suivant : $(L_{1,n-1}, L_{2,m})$, $(L_{1,n}, L_{2,m-1})$, ou $(L_{1,n-1}, L_{2,m-1})$ à l'aide de l'opérateur $\leq$. Cet opérateur est appliqué au couple de réels MD_{ij} et MF_{ij} associé au couple de points. Le couple de points choisi est ajouté à la tête du CO et le couple de points précédant est à nouveau sélectionné. Cette opération est répétée tant que les indices i et j ne sont pas égaux à 1. Cet opérateur permet de s'assurer que :

- Le passage par le couple de points sélectionné, ne pas construit pas un chemin qui passe au préalable par des couples de points éloignés. Cette propriété est vérifiée en comparant des distances partielles à l'aide de la première inéquation ($a < c$) appliquée à deux cellules de la matrice MF .
- Entre deux distances partielles identiques de MF , le couple ayant les points les plus proches est choisi. Cette propriété est vérifiée à l'aide de la deuxième inéquation. Elle est appliquée à deux cellules de la matrice MD .

Cette méthode de backtracking est rapide et de complexité $O(n+m)$.

	L ₂													
	1	2	3	4	5	6	7							
L ₁	1	0,41 0,41	3,31 3,31	3,61 3,61	5,12 5,12	6,35 6,35	6,91 6,91	8,81 8,81						
	2	1,70 1,70	1,80 1,80	2,51 2,51	3,71 3,71	5,00 5,00	5,85 5,85	7,72 7,72						
	3	2,00 2,00	2,01 2,01	1,80 1,51	3,26 3,26	4,37 4,37	4,77 4,77	6,68 6,68						
	4	2,61 2,61	2,00 1,63	1,80 0,90	2,64 2,64	3,74 3,74	4,22 4,22	6,12 6,12						
	5	4,09 4,09	2,00 0,95	1,80 1,33	1,80 1,10	2,42 2,42	3,70 3,70	5,41 5,41						
	6	5,46 5,46	2,45 2,45	2,11 2,11	1,80 0,45	1,80 0,91	2,52 2,52	4,02 4,02						
	7	6,91 6,91	4,52 4,52	3,44 3,44	2,69 2,69	1,80 1,66	1,80 0,54	1,84 1,84						
	8	7,92 7,92	5,50 5,50	4,46 4,46	3,61 3,61	2,40 2,40	1,80 1,24	1,80 0,86						

Tableau 2 : Tableau des couples MF_{ij} MD_{ij} pour les lignes de la figure 8

Pour l'exemple, les lignes de la figure 8, le chemin optimal est initialisé avec le couple $(L_{1,7}, L_{2,8})$. Puis le couple $(L_{1,7}, L_{2,6})$ est ajouté. En effet, $(1,8, 0,54) \leq (1,8, 1,24)$ et $(1,8, 0,54) \leq (1,84, 1,84)$. Le tableau 2 fournit l'ensemble des couples de réels : MF_{ij} et MD_{ij} . Les cellules grisées correspondent au chemin optimal. Ainsi pour ces deux lignes, le CO est composé de 9 couples de points : $(L_{1,1}, L_{2,1})$, $(L_{1,2}, L_{2,2})$, $(L_{1,3}, L_{2,3})$, $(L_{1,4}, L_{2,3})$, $(L_{1,5}, L_{2,4})$, $(L_{1,6}, L_{2,4})$, $(L_{1,6}, L_{2,5})$, $(L_{1,7}, L_{2,6})$, $(L_{1,7}, L_{2,8})$.

Si les deux lignes correspondent au même phénomène du monde réel, les points des couples de ce CO sont les points homologues. Un point de la ligne 1 peut avoir plusieurs points homologues dans la ligne 2. Pour l'exemple, $L_{1,6}$ est homologue avec les points $L_{2,4}$ et $L_{2,5}$. De même, un point de la ligne 2 peut avoir plusieurs points homologues dans la ligne 1. Pour l'exemple, $L_{2,3}$ est homologue avec les points $L_{1,3}$ et $L_{1,4}$. Par contre, par construction, les relations sont toujours de type 1–1, 1–N ou N–1 mais jamais de type N–M.

Le CO étant connu, il est possible de connaître l'écart moyen entre les points homologues. Cet écart est appelé distance de Fréchet discrète moyenne (d_{Fdm}). La d_{Fdm} pour les deux lignes de la figure 8, est de 0,95. Cependant, comme la d_{Fd} , cette mesure est sensible à l'échantillonnage. Le tableau 3 fournit les valeurs des d_{Fd} et des d_{Fdm} en fonction du suréchantillonnage des lignes.

	sans	1	0.1	0.01	0.001
d_{Fd}	1.8028	1.2260	1.2015	1.2012	1.2012
d_{Fdm}	0.9524	0.5843	0.5116	0.5030	0.4997

Tableau 3: distance de Fréchet discrète maximale et moyenne pour les deux lignes de la figure 8 en fonction du suréchantillonnage

2.1.2 Adaptation des distances de Fréchet discrète sur d'autres types de lignes

En termes de calcul de similarité, la distance de Fréchet ne peut être employée que sur des lignes ouvertes d'emprises similaires (figure 9 (a)). Pour des lignes fermées (lignes dont le premier et le dernier point sont confondus) ou des lignes d'emprise différentes, le calcul de similarité à l'aide de distance de Fréchet doit donc être adapté. Cette adaptation s'appuie sur la recherche de parties homologues. Deux parties de lignes sont dites homologues si visuellement elles se correspondent.

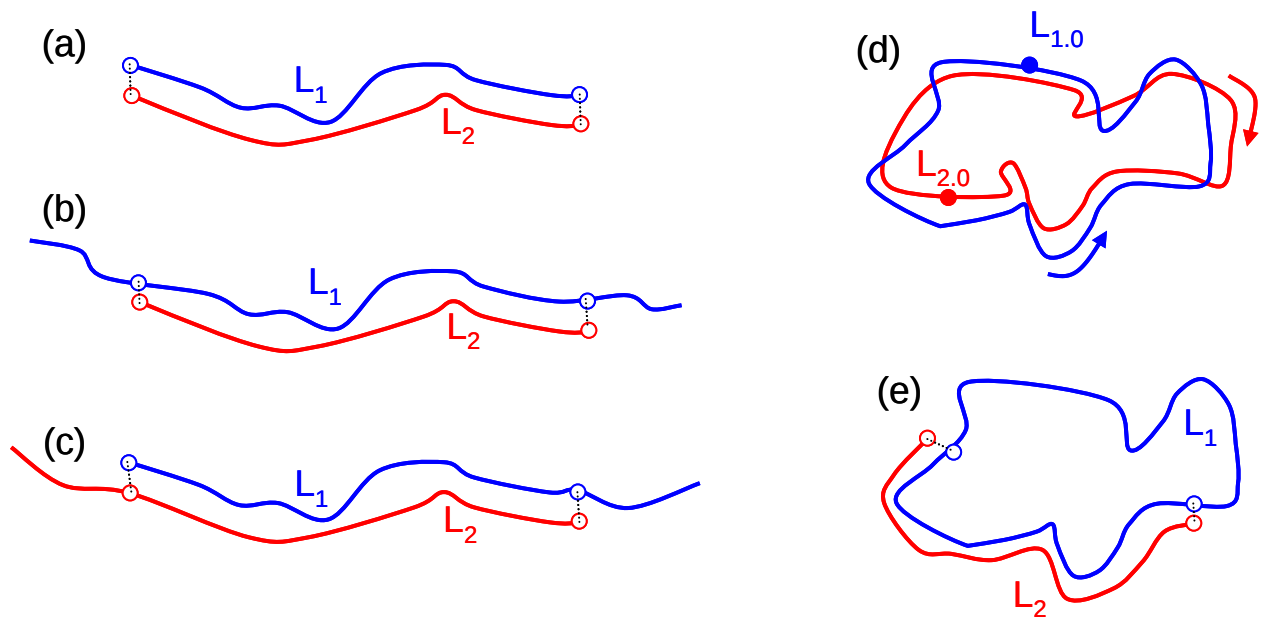


Figure 9 : Calcul de distance de Fréchet sur des lignes ou des parties de lignes homologues

Afin de pouvoir généraliser le calcul de similarité à l'aide de distance de Fréchet, au plus grand nombre de lignes, 4 adaptations ont été proposées :

1. Le calcul de distance de similarité entre une ligne L_1 et une ligne L_2 d'emprise plus petite (voir figure 9 (b)).
2. Le calcul de distance de similarité entre deux lignes L_1 et L_2 ayant seulement une partie de leur emprise commune (voir figure 9 (c)).
3. Le calcul de distance de similarité entre deux lignes fermées (voir figure 9 (d)).
4. Finalement, le calcul de distance de similarité entre une ligne ouverte L_2 et une partie homologue d'une ligne fermée L_1 de plus grande emprise (voir figure 9 (e)).

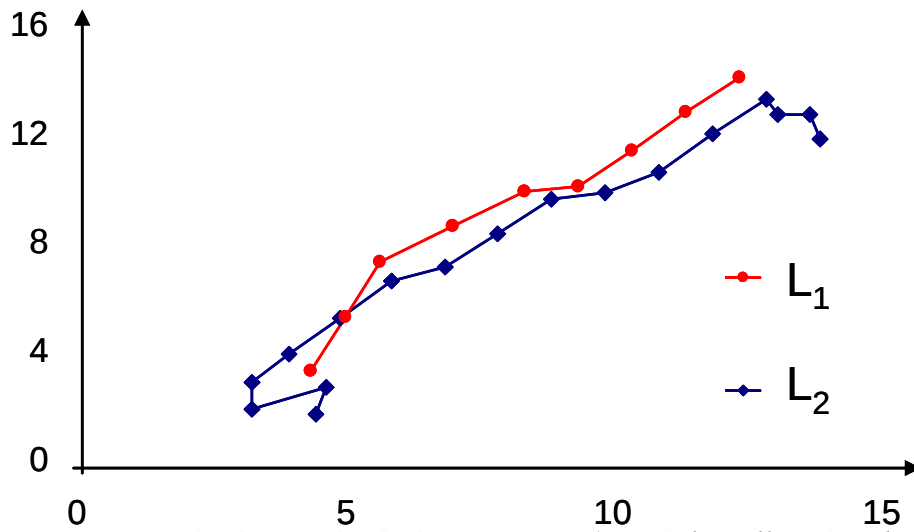


Figure 10 : exemple de couple de lignes (L_1 , L_2) suréchantillonnées d'emprise différente

Pour des lignes d'emprise différentes, la distance de Fréchet discrète est égale le plus souvent à la distance euclidienne (d_E) entre deux extrémités. Pour l'exemple de la figure 10, la d_{Fd} est égale à la distance entre $L_{1,9}$ et $L_{2,17}$. Cette mesure ne correspond pas au résultat d'un écart visuel entre les parties homologues. Il faut donc définir une variante de cette distance permettant de mesurer l'écart entre ces parties. Pour cela la distance de Fréchet discrète partielle (d_{Fdp}) est définie. Elle va permettre de :

- Détecter la **ligne partielle homologue** à l'intérieur de la ligne d'emprise supérieure,
- Calculer la distance maximale entre la ligne d'emprise inférieure et la ligne partielle homologue.

Pour calculer la d_{Fdp} et déterminer la partie homologue, il est nécessaire de détecter automatiquement les points de départ ($L_{2,begin}$) et d'arrivée ($L_{2,end}$) de la partie homologue. Ces points sont des points de L_2 tels que la distance de Fréchet discrète entre L_1 et la partie $\langle L_{2,begin} \dots L_{2,end} \rangle$ soit minimale. Pour déterminer $L_{2,begin}$, $L_{2,end}$ et la d_{Fdp} , les couples de points ($L_{2,j}$, $L_{2,jj}$) de $L_2 \times L_2$ avec $jj \geq j$ sont donc testés par ordre des distances à $L_{1,1}$ pour $L_{2,j}$, des distances à $L_{1,n}$ pour $L_{2,jj}$. Cet ordre est utilisé pour optimiser la recherche de $L_{2,begin}$ et $L_{2,end}$. Ainsi, le nombre de couples testés est inférieur à un parcours séquentiel des points de L_1 et de L_2 . La liste des couples de points à tester, a pu être réduite afin de ne tester que les couples potentiellement candidats (Devogele 2000).

L'algorithme de recherche de $L_{2,begin}$, $L_{2,end}$ et d_{Fdp} se décompose en 2 temps :

1. Dans un premier temps, les listes ordonnées de candidats (LCB et LCE) sont construites.
 LCB : liste ordonnée des points $L_{2,j}$ candidats pour $L_{2,begin}$ par ordre de distance à $L_{1,1}$,
 LCE : liste ordonnée des points $L_{2,j}$ candidats pour $L_{2,end}$ par ordre de distance à $L_{1,n}$.
2. Dans un deuxième temps, $L_{2,begin}$ et $L_{2,end}$ sont recherchés respectivement dans LCB et LCE et la d_{Fdp} est calculée à l'aide de l'algorithme suivant :

$d_{Fdp} = +\infty$;
 For $L_{2,j}$ in LCB
 If $d_E(L_{1,1}, L_{2,j}) < d_{Fdp}$ then
 For $L_{2,jj}$ in LCE
 If $j \leq jj$ and $d_E(L_{1,m}, L_{2,jj}) < d_{Fdp}$ then
 If $(d_{Fd}(<L_{1,1} \dots L_{1,n}>, <L_{2,j} \dots L_{2,jj}>) < d_{Fdp})$
 then
 { $d_{Fdp} = d_{Fd}(<L_{1,1} \dots L_{1,n}>, <L_{2,j} \dots L_{2,jj}>);$
 $L_{2.begin} = L_{2,j};$
 $L_{2.end} = L_{2,jj};$

Pour les deux lignes de la figure 10, $L_{2.begin}$, $L_{2.end}$ sont déterminés et d_{Fdp} est calculée. Premièrement, LCB et LCE sont instanciées. La première et la dernière ligne de la matrice de distance (voir tableau 4 pour l'exemple) permettent de définir l'ordre :

- **LCB** = $\{L_{2,2}, L_{2,5}, L_{2,4}, L_{2,1}, L_{2,6}, \dots\}$, $L_{2,3}$ n'est pas candidat, car $L_{2,2}$ est un meilleur candidat : $d_E(L_{1,1}, L_{2,3}) > d_E(L_{1,1}, L_{2,2})$, cette optimisation est justifiée dans (Devogele 2000).
- **LCE** = $\{L_{2,14}, L_{2,15}, L_{2,16}, L_{2,17}, \dots\}$, $L_{2,13}$ n'est pas candidat, car $L_{2,14}$ est un meilleur candidat : $d_E(L_{1,9}, L_{2,13}) > d_E(L_{1,9}, L_{2,14})$ cette optimisation est la symétrique de la précédente.



L1.i.x	L1.i.y
4,5	3,2
5,15	5,2
5,8	7,2
7,15	8,5
8,5	9,8
9,5	10
10,5	11,3
11,5	12,7
12,5	14

L2.j.x	4,6	4,8	3,4	3,4	4,1	5,05	6	7	8	9	10	11	12	13	13,2	13,8	14
L2.j.y	1,6	2,6	1,8	2,8	3,8	5,15	6,5	7	8,25	9,5	9,75	10,5	11,9	13,2	12,6	12,6	11,7
Matrice de Distance																	
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	1,60	0,67	1,78	1,17	0,72	2,03	3,62	4,55	6,14	7,74	8,55	9,77	11,49	13,12	12,81	13,22	12,75
2	3,64	2,62	3,82	2,97	1,75	0,11	1,55	2,58	4,17	5,77	6,65	7,89	9,58	11,21	10,93	11,38	10,98
3	5,73	4,71	5,91	5,01	3,80	2,18	0,73	1,22	2,44	3,94	4,91	6,16	7,78	9,37	9,16	9,65	9,35
4	7,36	6,35	7,68	6,82	5,60	3,95	2,31	1,51	0,89	2,10	3,11	4,34	5,92	7,50	7,31	7,81	7,56
5	9,08	8,10	9,49	8,66	7,44	5,79	4,14	3,18	1,63	0,58	1,50	2,60	4,08	5,64	5,47	5,99	5,82
6	9,72	8,77	10,22	9,44	8,22	6,58	4,95	3,91	2,30	0,71	0,56	1,58	3,14	4,74	4,52	5,02	4,81
7	11,35	10,40	11,86	11,08	9,86	8,22	6,58	5,54	3,94	2,34	1,63	0,94	1,62	3,14	3,00	3,55	3,52
8	13,07	12,12	13,58	12,79	11,57	9,93	8,29	7,26	5,66	4,06	3,31	2,26	0,94	1,58	1,70	2,30	2,69
9	14,70	13,76	15,22	14,43	13,21	11,57	9,92	8,90	7,30	5,70	4,93	3,81	2,16	0,94	1,57	1,91	2,75

Tableau 4: matrice de distance (MD) associée aux 2 lignes de la figure 10

Par rapport à la matrice de distance (MD), le calcul de la distance de Fréchet discrète consiste à trouver le cheminement entre une cellule $(1,j)$ et une cellule (n,jj) tel que la valeur maximale des cellules de ce cheminement soit minimale.

Deuxièmement, les d_{Fd} entre les lignes L_1 et $<L_{2,j} \dots L_{2,jj}>$ vont pouvoir être calculées :

1. La première d_{Fd} calculée est $d_{Fd}(L_1, <L_{2,2} \dots L_{2,14}>)$. Elle renvoie comme d_{Fd} : 1,78 ($d_E(L_{1,1}, L_{2,3})$). Cette valeur ainsi que $L_{2,2}$ et $L_{2,14}$ sont retenues comme résultat intermédiaire. La seconde distance calculée : $d_{Fd}(L_1, <L_{2,2} \dots L_{2,15}>)$, renvoie le même résultat. Les d_{Fd} avec les $L_{2,jj}$ suivants ne sont pas calculées, car $d_E(L_{1,9}, L_{2,jj}) > 1,78$ quelque soit $L_{2,jj}$, la d_{Fd} est donc forcément supérieur au résultat intermédiaire.
2. Nous pouvons passer au deuxième candidat pour $L_{2.begin}$, $L_{2,5}$. $d_{Fd}(L_1, <L_{2,5} \dots L_{2,14}>)$ est la première distance calculée. Elle renvoie 1,22 ($d_E(L_{1,3}, L_{2,8})$) qui est inférieur à 1,78. Cette valeur ainsi que $L_{2,5}$ et $L_{2,14}$ sont retenues comme résultat intermédiaire. Un cheminement permettant d'obtenir ce résultat est fourni par les cellules en caractères gras italiques dans le tableau 3. Les d_{Fd} avec les $L_{2,jj}$ suivants ne sont pas calculées, car $d_E(L_{1,9}, L_{2,jj}) > 1,22$ quelque soit $L_{2,jj}$.

3. Nous pouvons passer au troisième candidat pour $L_{2,begin}$: $L_{2,4}$. Les distances calculées quelque soit $L_{2,jj}$, ne changent pas le résultat intermédiaire.
4. Les candidats suivants pour $L_{2,begin}$: $L_{2,1}$, $L_{2,6}$, ... ne sont pas testés, car les distances euclidiennes entre ces points et $L_{2,1}$ sont supérieures à 1,22.

La partie homologue est donc $\langle L_{2,5}...L_{2,14} \rangle$ et la distance de Fréchet discrète partielle associée est de 1,22.

Cet algorithme est d'ordre $O(n^2m^2)$. Mais, sa complexité en moyenne pour des lignes ayant une partie homologue, est de l'ordre de $O(nm)$ car peu de couples de points candidats sont testés.

Une distance de Fréchet discrète partielle moyenne (d_{Fdpm}) peut être aussi obtenue entre une ligne et la partie homologue en construisant au préalable le chemin optimal et en calculant ensuite l'écart moyen entre les points homologues.

Pour le calcul de distance de similarité entre deux lignes L_1 et L_2 ayant seulement une partie de leur emprise commune (voir figure 9 (c)), il n'a pas été possible de définir un algorithme générique. Une limitation importante a été posée. L'algorithme se focalise sur la recherche de deux parties homologues du type $\langle L_{1,1}... L_{1,end} \rangle \langle L_{2,begin}... L_{2,m} \rangle$. En effet, ce type d'approche se limite à la détermination de deux points extrémités. En conséquence, une des deux extrémités de chaque partie homologue doit être une des deux extrémités de la ligne associée. Cette limite est aussi valable pour le cas précédant où seul $L_{2,begin}$ et $L_{2,end}$ pouvaient varier. Par rapport à l'algorithme précédant deux modifications sont à noter :

- Les listes LCB et LCE sont remplacées par les listes $LC1$: liste ordonnée des points $L_{1,i}$ candidats pour $L_{1,end}$ par ordre de distance à $L_{2,m}$ et $LC2$: liste ordonnée des points $L_{2,j}$ candidats pour $L_{2,begin}$ par ordre de distance à $L_{1,1}$.
- Vis-à-vis de la matrice MD , le chemin va partir de la première ligne correspondant à $L_{1,1}$ et aller vers la dernière colonne correspondant à $L_{2,m}$ (diagonale haut gauche vers bas droite).

L'algorithme est sinon similaire au précédent. Il définit parallèlement les deux parties homologues de L_1 et L_2 et calcule la distance maximale entre ces parties homologues. Une distance de Fréchet discrète partielle moyenne (d_{Fdpm}) peut alors être mesurée sur ces deux parties homologues.

En ce qui concerne, le **calcul de distance de similarité entre deux lignes fermées** (voir figure 9 (d)), les lignes L_1 et L_2 n'ayant ni des extrémités homologues ($L_{1,0}$ et $L_{2,0}$) ni un sens commun de rotation, deux opérateurs sont employés pour transformer L_2 :

- La permutation circulaire de l'ordre des points. La permutation de X points de L_2 transforme $\langle L_{2,1}... L_{2,m} \rangle$ en $\langle L_{2,X+1}... L_{2,m-1}, L_{2,1}, ... L_{2,X+1} \rangle$. Pour rappel, pour une ligne fermée, le premier point est identique au dernier.
- L'inversion du sens. L'inversion de L_2 transforme $\langle L_{2,1}... L_{2,m} \rangle$ en $\langle L_{2,m}, L_{2,m-1}, L_{2,m-2}... L_{2,1} \rangle$

L'objectif de l'algorithme est de trouver s'il est nécessaire d'inverser L_2 et de définir le paramètre X de la permutation circulaire. Suite à ces modifications, si L_1 et L_2 sont homologues, alors elles ont des extrémités homologues et un sens commun de rotation.

Finalement, en ce qui concerne le calcul de distance de similarité entre une ligne ouverte L_2 et une partie homologue d'une ligne fermée L_1 de plus grande emprise (voir figure 9 (e)). Ce calcul combine les deux méthodes définies précédemment:

- distance de similarité entre deux lignes L_1 et L_2 ayant seulement une partie de leur emprise commune,
- calcul de distance de similarité entre deux lignes fermées.

Les méthodes de calcul de l'ensemble de ces distances linéaires maximales et moyennes ont été codées en C++. Elles ont été testées dans le cadre de la thèse de M^{elle} Mascaret sur des données réelles maritimes (lignes de côtes, ligne de changement de pente, lignes de crête, lignes de talweg). Les lignes peuvent être en 2D ou en 3D. Quatre applications directes de ces méthodes sont maintenant présentées.

2.2 Applications des distances de Fréchet discrètes maximale et moyenne

Les distances de Fréchet discrètes maximales et moyennes sont des mesures fiables pour qualifier la similitude entre deux lignes. Elles peuvent donc être employées une fois les lignes suréchantillonnées en fonction de la précision requise, dans différents domaines :

- contrôle de la qualité des lignes en la comparant avec des lignes plus précises,
- appariement de données entre les lignes de deux jeux de données représentant les mêmes phénomènes du monde réel,
- fusion de données à l'aide de la détection des points homologues,
- déformation élastique.

2.2.1 Contrôle de la qualité

En termes de contrôle de la qualité, ces travaux ont été par exemple, appliqués sur des traits de côtes (Mascaret et *al.* 2006) (Le Berre et *al.* 2005) dans la zone des Abers (Nord de Brest). Par exemple, la figure 11 fournit un échantillon de plusieurs traits de côtes issus de digitalisations différentes à partir de sources différentes :

- image satellite SPOT5 avec une résolution de 2,5 m prise le 17 mai 2003,
- photo aérienne de la BD Ortho® (Orthophoto) de l'IGN avec une résolution de 50 cm de juin 2000,
- saisie terrain à l'aide d'un tachéomètre et d'un GPS différentiel avec une précision centimétrique, réalisée en mai 2003.

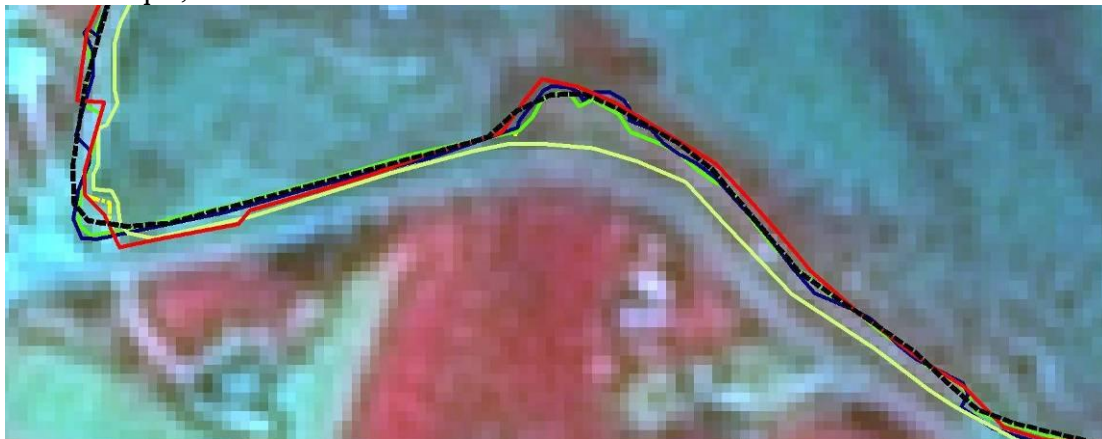


Figure 11 : Exemple de 5 traits de côtes numérisés avec des techniques et des sources différentes, image de fond SPOT 5.

Le calcul des distances maximales et moyennes (voir tableau 5) entre les deux traits (SPOT5 et orthophoto) et un trait de référence (tachéomètre), a montré l'importance de disposer de données récentes pour les zones ayant un relief changeant (zones dunaires). Pour ces zones des données récentes moins précises sont préférables à des données plus précises, mais obsolètes. A l'opposé pour des zones moins changeantes (falaises), les données plus précises bien que plus anciennes sont préférables.

Zones dunaires	d_{Fd} (X, tachéomètre)	d_{Fdm} (X, tachéomètre)
SPOT 5	3.79	1,17
Orthophoto	8.64	2,1

Tableau 5: Distances de Fréchet maximales et moyennes en zone dunaire

Un autre exemple d'utilisation possible de la distance de Fréchet partielle a été donné dans (Devogele et Riffault 2000). La d_{Fdp} a permis de simuler la vérification de la cohérence entre les lignes de côte d'une carte numérique et les lignes de côte issues d'un radar. Cette simulation a été réalisée, à l'aide d'un jeu de données du SHOM (Service Hydrographique et Océanographique de la Marine) de la région de la pointe du Raz (lignes de 1935 points, 34 km) et de lignes de côte radar simulées (≈ 100 points, 1 km). Le temps de calcul des distances de Fréchet discrètes partielles est compatible avec son utilisation en tant qu'alarme à bord d'un navire associé à un radar.

2.2.2 Appariement de données linéaire

Les phénomènes du monde réel sont actuellement représentés dans plusieurs Bases de Données géographiques (BDG), la réutilisation de telles BDG nécessite un processus d'intégration complète pour :

- unifier la sémantique : schémas et métadonnées,
- éliminer les objets redondants et regrouper les parties complémentaires.

Ce **processus d'intégration** (Spaccapietra et al. 1992) est nécessaire étant donné le coût d'acquisition des données géographiques (de cette manière des saisies d'information sont évitées) et permet de fédérer l'information provenant de différentes sources. Il consiste à :

- prendre en entrée :
 - un ensemble de bases de données (schémas et populations),
- produire en sortie :
 - une **description unifiée** des schémas initiaux (le schéma intégré),
 - les **règles de traduction** qui vont permettre la migration des données,
 - des **liens** entre les objets des différentes représentations représentant le même phénomène du monde réel.

Pour les BDG, le processus permettant de définir ces liens est appelé **appariement** ou encore conflation.

En ce qui concerne le littoral, dans ce cadre du projet Litto3D (Le Roux et Martinoty 2009) la définition d'un processus d'intégration est complémentaire des campagnes de saisie de données de grande précision. Dans ce contexte, pour les MNT, les saisies pourraient être limitées, dans un premier temps, aux zones avec des données :

- peu ou pas nombreuses,
- obsolètes,
- incohérentes,
- très changeantes.

Dans les zones où les données terrestres et maritimes en 3D sont récentes, un processus d'intégration doit permettre d'obtenir une représentation continue du littoral de qualité acceptable. Qui plus est, pour les nouvelles données acquises, les capteurs étant différents (sonar, lidar, etc.) un processus d'appariement est aussi nécessaire. Il permet avant d'intégrer les modèles numériques maritimes et terrestres de se recalculer en fonction des données homologues appariées.

Dans le cadre de la thèse de M^{elle} Mascaret, il a été décidé pour les zones ayant des reliefs significatifs de se baser sur les éléments caractéristiques linéaires (ligne de changement de

penne, lignes de crête, lignes de talweg) et les éléments ponctuels (sommets, cuvettes) pour intégrer les MNT sur leur zone de recouvrement.

Le processus d'appariement des MNT, dans les zones littorales, est décomposé en trois grandes étapes :

- enrichissement des MNT terrestre et maritime en détectant les éléments caractéristiques,
- mesures de distance de similarité entre les éléments caractéristiques de même type, des deux zones,
- analyse des résultats de similarité.

La phase **d'enrichissement** consiste à détecter les éléments caractéristiques. Les méthodes automatiques testées (méthode du module BathySIS de la société Geomod) ont permis de détecter ces éléments avec des résolutions faibles (grille avec des pas de 50 m et plus) par contre pour des données plus précises, les méthodes automatiques actuelles doivent être améliorées. Cette phase n'étant pas le cœur de notre recherche, les MNT testés ont donc été enrichis par des numérisations manuelles des éléments caractéristiques.

Une fois les éléments caractéristiques ajoutés, pour chaque type d'éléments, des **mesures de similitude** sont réalisées :

- Distances maximales de Fréchet discrètes pour les éléments linéaires,
- Distance euclidienne pour les éléments ponctuels.

Un seuil S est défini empiriquement en fonction de la qualité des MNT à intégrer. Seuls des éléments de même type, à une distance inférieure à ce seuil S peuvent être appariés. Pour les éléments linéaires plusieurs cas d'appariement sont à prendre en considération :

- **Appariement 1-1 simple** pour une ligne caractéristique L_1 du MNT_1 et une ligne caractéristique L_2 du MNT_2 si :
 - une seule distance de Fréchet discrète maximale (d_{Fd} , d_{Fdp}) est inférieure au seuil S ,
 - il n'existe pas d'autres lignes $L_{1'}$ du MNT_1 ou $L_{2'}$ du MNT_2 telle qu'une des distances maximales de Fréchet soient inférieures au seuil.

Ce type d'appariement où un objet du MNT_1 est apparié à un objet du MNT_2 est le plus courant. Cet appariement est total si une d_{Fd} est inférieure à S . Pour la figure 9 (page 26), les cas a et d sont des appariements totaux. Cet appariement est partiel si seulement une d_{Fdp} est inférieure à S . Pour la figure 9, les cas b, c et e sont des appariements partiels.

- **Appariement 1-1 partiel multiple.** S'il existe plusieurs d_{Fdp} inférieures au seuil S pour deux mêmes lignes, deux cas sont possibles. Pour le couple de lignes de la figure 12 (a) les deux lignes sont appariées deux fois partiellement. Cependant, l'appariement 1-1 peut aussi être simple ou faire partie d'un appariement plus complexe. Par exemple, pour les lignes de la figure 12 (b), la même partie de L_2 est candidate à un appariement partiel simple avec la partie haute ou la partie basse de L_1 . Pour ce genre de situation, le calcul de la d_{Fdp} n'est pas suffisant. Il faut aussi tenir compte de l'environnement, l'appariement possible avec d'autres lignes est un des éléments à considérer. Pour cet exemple, la partie du bas de L_1 étant candidate à l'appariement avec $L_{2'}$, L_2 est donc appariée avec la partie haute de L_1 .

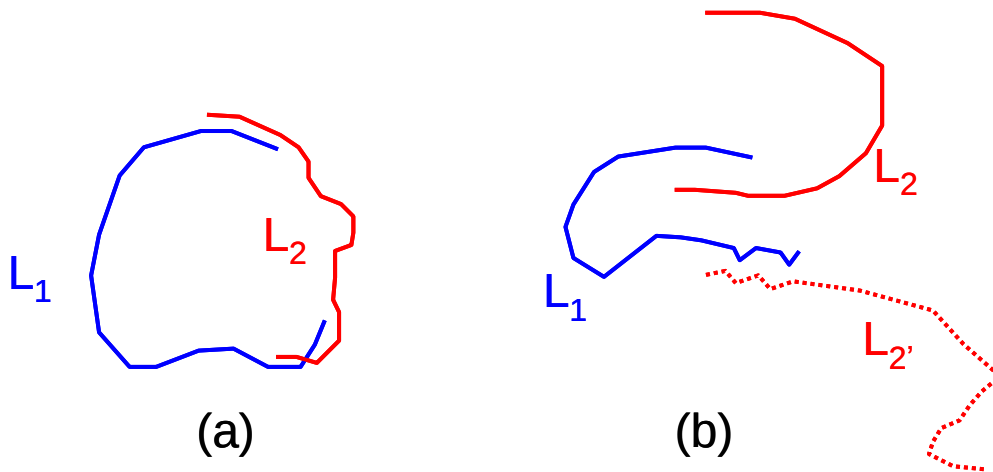


Figure 12 : exemples d'un appariement partiel 1-1 multiple et de deux appariements partiels simples

- **Appariements 1-N et N-M.** Ces cas plus complexes ont été détaillés dans le cadre de la thèse de M^{elle} Mascaret. La figure 12 (b) illustre un appariement de type 1-N, L_1 est appariée partiellement avec L_2 et L_2' .

Une **méthode générique d'appariement** de données linéaires a été détaillée dans (Mustière et Devogele 2008), elle reprend les étapes détaillées ci-dessus. D'autres étapes plus spécifiques à l'appariement de données formant des réseaux à des niveaux de détails différents sont aussi présentes. Cette méthode a été validée sur plusieurs jeux de données de l'IGN ayant des thématiques variées (routier, hydrographie, réseaux électriques).

2.2.3 La fusion de données

L'intégration de représentations géographiques a pour objectif de manipuler simultanément des données issues de plusieurs représentations. Il est alors nécessaire au niveau des données d'éliminer les parties redondantes et de regrouper les parties complémentaires. Une des solutions consiste à fusionner les données géographiques de deux représentations. Ainsi, une seule représentation basée sur les éléments les plus précis est définie pour chaque phénomène du monde réel. Les représentations à intégrer peuvent être deux vues différentes d'un même espace ou deux espaces limitrophes.

Pour fusionner les objets ayant une géométrie linéaire, les couples de points homologues des lignes appariées, jouent un rôle fondamental. Pour chaque couple, un nouveau point peut être défini par moyenne pondérée. La suite de points ordonnée obtenue forme la ligne fusionnée. Un point pouvant appartenir à plusieurs couples, le nombre de points de la ligne résultante est plus important. De même afin de favoriser les représentations les plus détaillées, pour chaque couple, les poids attribués aux points dépendent du nombre de couples dont fait partie ce point. Cette méthode est détaillée dans (Devogele 2002).

La figure 13 illustre cette technique de fusion à partir des couples de points homologues. Premièrement, deux lignes sont appariées (a). Les couples de points homologues sont définis, sur la figure ils sont reliés par des traits noirs fins (b). Pour chaque couple de points, un nouveau point est défini par moyenne pondérée (c). Le poids d'un point apparié est inversement proportionnel au nombre de couples de points homologues auquel il fait partie. Un point appartenant à N couples a pour poids $1/N$. D'autres pondérations sont envisageables. La suite de points obtenue forme la ligne fusionnée (d).

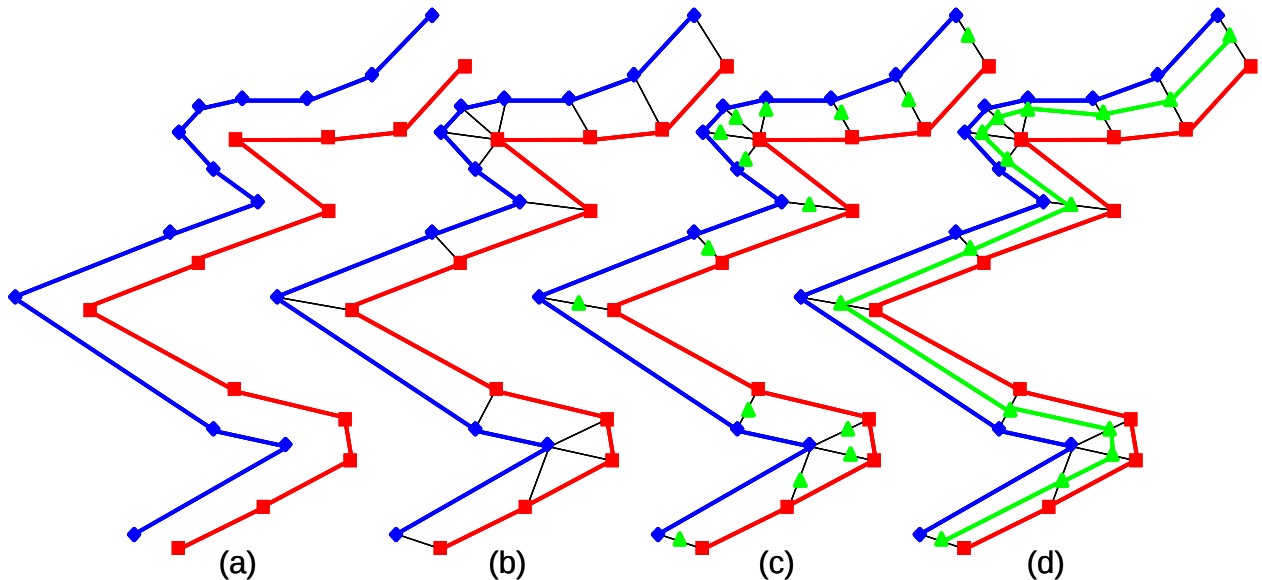


Figure 13 : fusion d'un couple de lignes après identifications des points homologues

Cette méthode proposée ci-dessus considère les deux lignes comme ayant la même importance. Si une ligne est plus précise, que l'autre ou si les données d'une des BDG doivent être recalées sur les données de l'autre BDG, il vaut mieux considérer la ligne la plus importante comme ligne fusionnée et recalculer les autres données.

2.2.4 La déformation élastique

Lors de fusion, les objets homologues sont déplacés, par contre l'emplacement des points sans homologues, est maintenu. Afin de garder une représentation cohérente au sein de la représentation finale, il est donc nécessaire d'effectuer des déplacements des objets (ou partie d'objet) présents dans une seule des représentations. Les processus de déformation élastique réalisent une distorsion non linéaire de manière à forcer l'ajustement d'éléments géographiques (points, lignes). Ces déformations s'appuient sur :

- des **vecteurs de déformation**. Deux types de vecteurs peuvent être employés :
 - Des **vecteurs locaux** ($\vec{v}_l$) qui relient les points homologues aux points fusionnés associés. Ces vecteurs représentent des déformations locales. Plus un objet sans homologues est proche d'un vecteur local, plus il subit la déformation liée à ce vecteur.
 - Des **vecteurs globaux** ($\vec{v}_g$) qui sont des moyennes des vecteurs locaux. Ces vecteurs permettent de définir des biais globaux des données. Si ces vecteurs sont employés, les vecteurs locaux sont alors corrigés des vecteurs globaux afin de ne pas déformer deux fois.
- Une **zone d'influence** aussi appelée bande de corrections élastiques (Langlois 1994, Laurini 1998). Dans cette zone et uniquement dans cette zone, les données sans homologues sont déformées. Deux zones peuvent être définies, une large pour le vecteur de déformation global et une plus étroite pour les vecteurs de déformation locaux. Ces zones

sont caractérisées par des rayons d'influence de la déformation : RI . Seuls les objets sans homologue dans la zone d'influence sont déformés.

- Un **coefficient d'amortissement** $coef_i$ qui permet de moduler les déformations des objets sans homologues en fonction de leur emplacement dans la zone d'influence. Si un point sans homologue est très proche du point d'origine p_i d'un vecteur $\vec{v}_{li}$ alors ce vecteur est appliqué de manière quasi totale. Par contre si ce point est à une distance légèrement inférieure à RI de p_i , alors ce point subit une influence quasi nulle. Afin d'avoir une fonction continue dérivable sur $\mathcal{R}^+$ le coefficient d'amortissement suivant a été choisi :

$$coef_i = \frac{\left(\cos \frac{d_E(p, p_i) \times \pi}{RI} + 1 \right)}{2} \text{ si } d_E(p, p_i) \leq RI, 0 \text{ sinon}$$

La figure 14 fournit la valeur du coefficient qui est appliquée au vecteur $\vec{v}_{li}$ en fonction de la distance entre le point à déformer et p_i avec un rayon d'influence RI de 4.

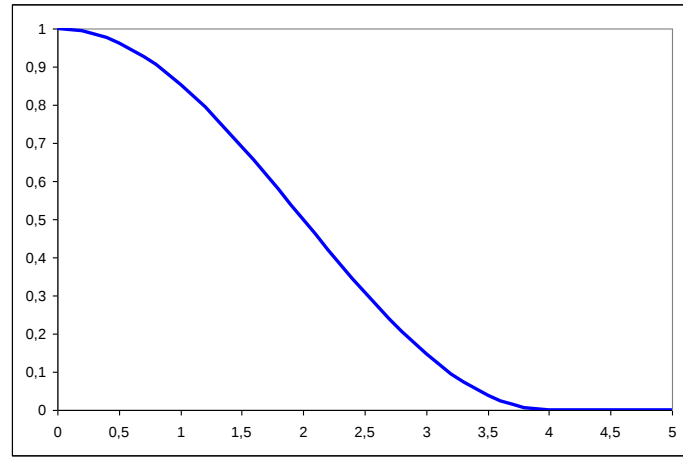


Figure 14: coefficient d'amortissement en fonction de la distance entre le point et l'origine d'un vecteur $\vec{v}_i$ pour un RI de 4

Le point p , est donc déplacé d'un vecteur $\vec{v}$ qui varie en fonction de la position de p . Ce vecteur $\vec{v}$ est calculé à l'aide d'une somme pondérée des $\vec{v}_{li}$. Le poids associé à chaque $\vec{v}_{li}$ est fonction du coefficient d'amortissement $coef_i$. La formule pour calculer $\vec{v}$ est donc :

$$\vec{v} = \sum_{i=1 \dots n} \vec{v}_{li} \times \frac{coef_i}{\sum_{j=1 \dots n} coef_j}$$

Cette fonction est plus fine que celle proposée dans la littérature (Langlois 1994, Laurini 1998). La figure 15 illustre le calcul de $\vec{v}$ pour un point p . Ce point est proche d'une ligne qui a été appariée lors de l'étape (a). Les points p_1 , p_2 et p_3 de cette ligne ont été déplacés selon des vecteurs locaux $\vec{v}_1$, $\vec{v}_2$ et $\vec{v}_3$. Pour la déformation élastique de p , le point p_3 n'est pas pris en considération, la distance à p étant supérieure à RI . Donc, $\vec{v}$ est calculé à partir de $\vec{v}_1$ et $\vec{v}_2$. $\vec{v}$ est plus court du fait des coefficients d'amortissement et ressemble plus à $\vec{v}_1$ qu'à $\vec{v}_2$ car p est plus proche de p_1 . Le point p' est donc issu de la déformation élastique de p selon cette technique.

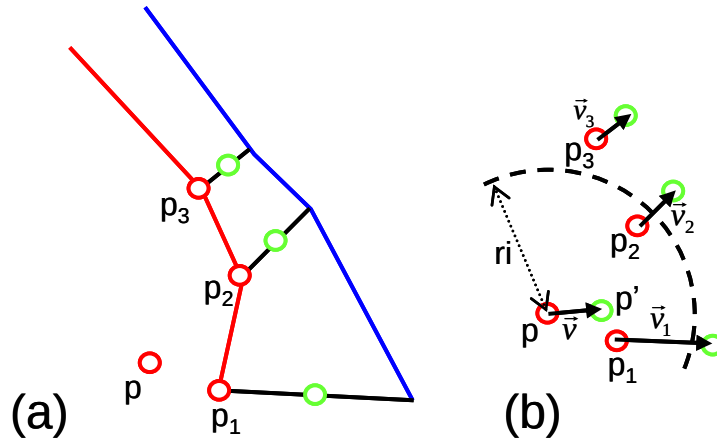


Figure 15: vecteur de déformation en fonction des vecteurs locaux proches

La figure 16 est un exemple qui reprend les différentes phases du processus d'intégration. Premièrement, les couples de points homologues sont détectés (a). Puis, pour chaque couple, un nouveau point est défini (b). Les deux parties de lignes homologues peuvent donc être fusionnées (c). Finalement, les points sans homologues proches sont déformés, à l'aide d'un processus de déformation élastique (d).

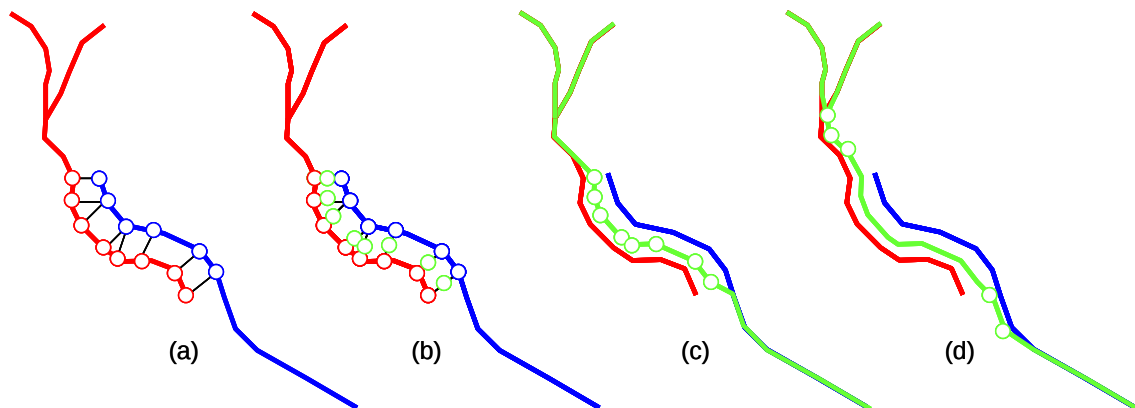


Figure 16 : Déformations élastiques des parties propres à partir des couples de points homologues

Cette technique de déformation élastique définie en 2000 (Devogele 2000) a été améliorée et est en cours de validation sur les points des MNT dans le cadre de la thèse de M^{elle} Mascaret. En fonction de différentes configurations, le processus de déformation élastique est optimisé en définissant:

- le rayon d'influence RI optimal en fonction de la longueur des vecteurs locaux et globaux,
- le rôle et l'ordre des déformations élastiques locales lié aux vecteurs locaux et des déformations élastiques globales liées au vecteur global.

Afin de montrer l'intérêt de cette méthode un chenal de forme très régulière a été défini (figure 17 (a)). Sa longueur est de 18 unités. Il est divisé en deux parties qui se recouvrent sur 2 unités (zone de recouvrement ZR). La partie basse est bruitée avec trois types de bruits planimétriques et altimétriques différents :

- bruit blanc,
- biais systématique,
- bruits locaux.

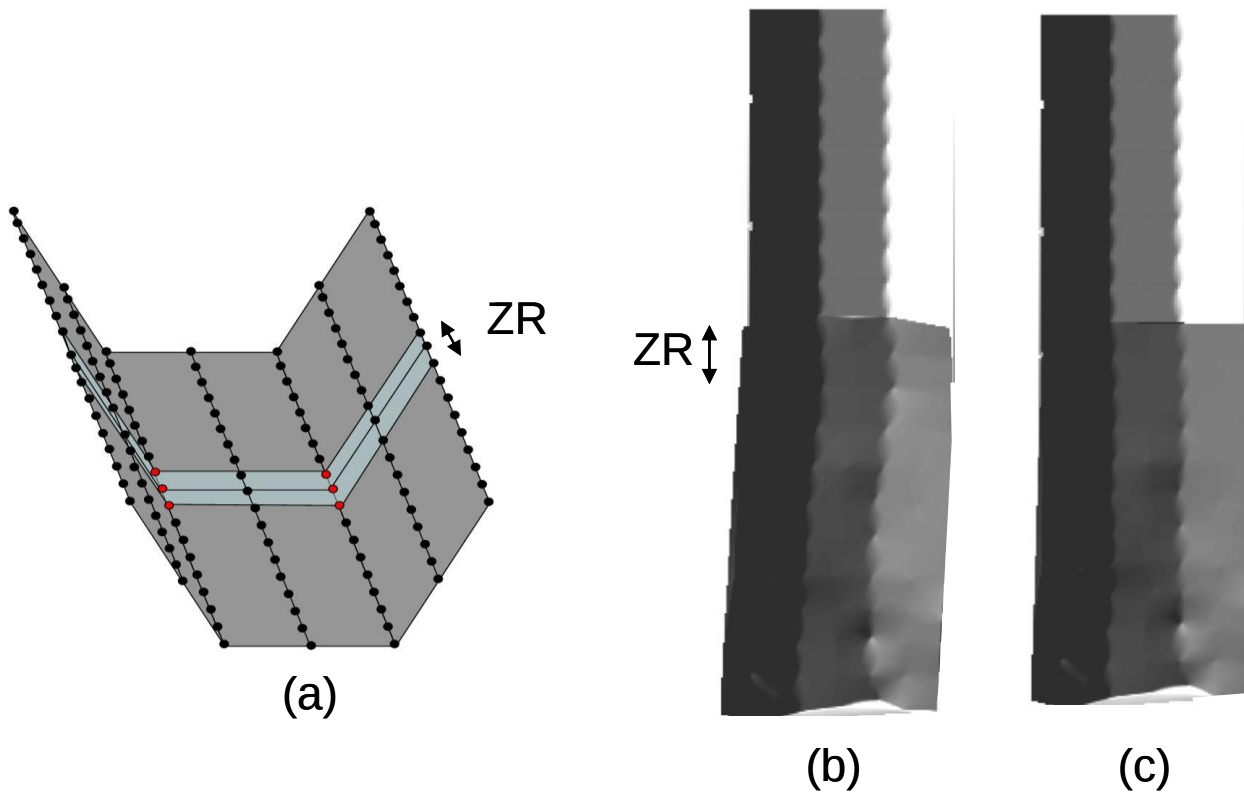


Figure 17 : Canal non bruité (a) partie basse bruitée (b) et parties fusionnées

Ces deux derniers types de bruits sont visibles sur la figure 17 (b). La partie basse dévie sur la gauche. Des bosses et des creux apparaissent. Deux ombrages différents sont employés afin de distinguer les deux parties.

Pour fusionner ces deux parties :

- les parties communes des lignes de changement de pente sont appariées sur la zone de recouvrement (ZR),
- les points homologues sont fusionnés (les points de la partie basse sont déplacés vers les points homologues de la partie haute),
- les vecteurs locaux et globaux de déformation sont calculés,
- les autres points de la partie basse sont déformés à l'aide de ces deux types de vecteurs :
 - déformations globales sur la totalité de cette partie,
 - déformation locale sur une zone plus étroite délimitée par un *RI* de 8.

La figure 17 (c) fournit le résultat de cette fusion. Environ 90% du biais systématique est corrigé. Les bruits locaux proches de la ZR sont bien corrigés. Par contre, il est impossible de corriger les bruits locaux éloignés, c'est dernier n'étant pas détectable.

2.2.5 Méthodologie d'intégration des MNT

Dans le cadre de l'intégration de MNT terrestre et maritime, une méthodologie d'intégration des MNT (Mascret Devogele 2006) est proposée (voir figure 18).

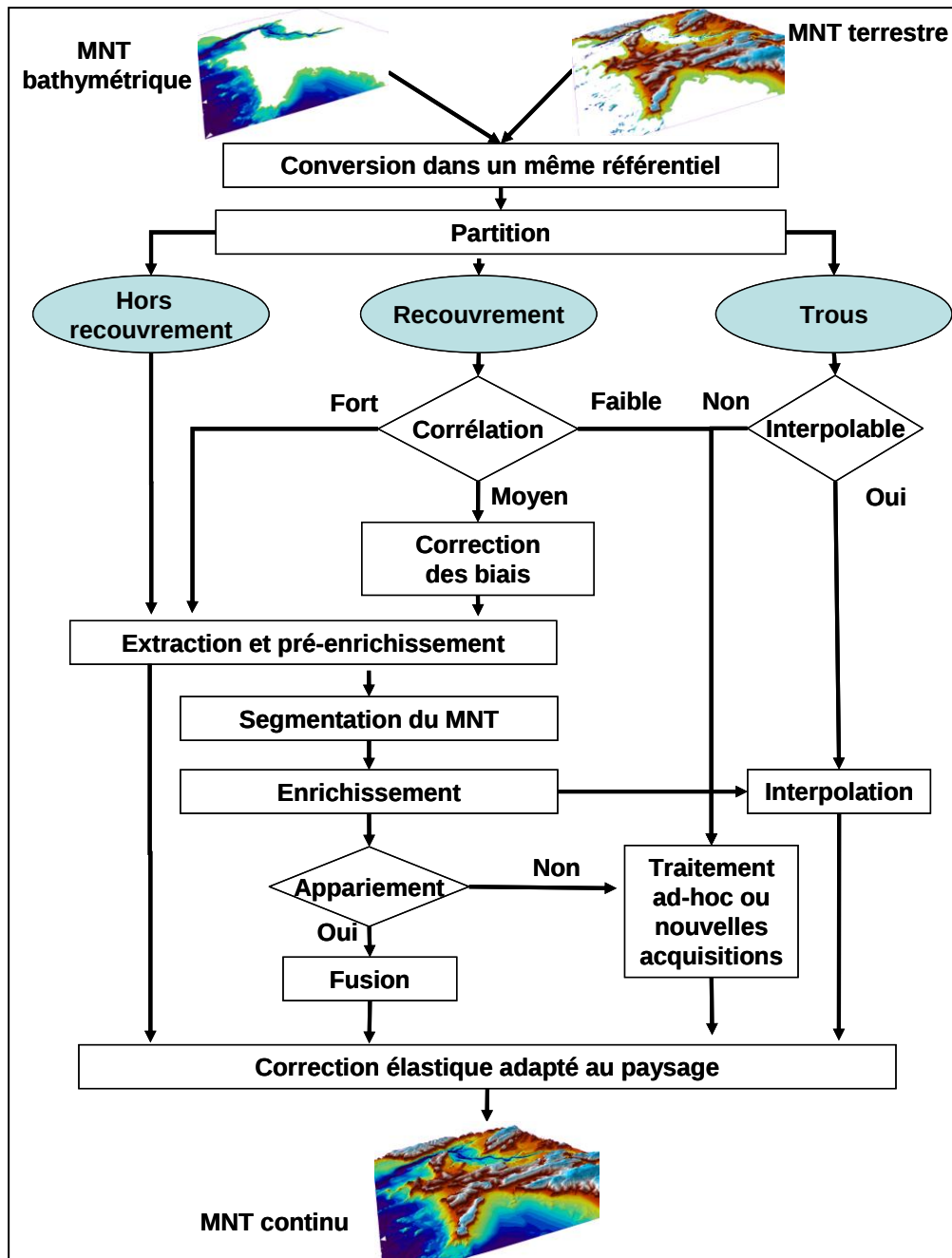


Figure 18 : Schéma descriptif de la méthodologie de fusion de MNT

Cette méthodologie est suffisamment générique pour être capable d'intégrer des données issues de MNT très divers et suffisamment spécifiques pour tenir compte des caractéristiques propres à différents types de relief (plages, dunes, rochers, falaises, etc.). Elle met en œuvre plusieurs outils dans un ordre précis pour répondre aux besoins posés lors de la fusion :

- conversion dans un même référentiel,
- correction des biais locaux,
- interpolation des zones sans informations,

- détection des objets homologues,
- fusion des données appariées.

Elle effectue donc un enchaînement de traitements répartis en trois phases principales :

- comparaison globale du MNT qui comprend : la conversion dans un même référentiel, une étape de partition et une corrélation et correction des biais.
- Analyse paysagère qui effectue : une extraction d'éléments caractéristiques, une segmentation (Saux et *al.* 2004), et un enrichissement.
- Génération du MNT avec tout d'abord l'interpolation des trous (Mascret Devogele 2006), ensuite l'appariement, la fusion puis finalement les déformations élastiques.

Quatre applications directes des distances de Fréchet discrète (contrôle qualité, appariement, fusion et déformation élastique) et une méthodologie d'intégration de MNT terrestre et maritime ont été présentées. Des travaux en géométrie algorithmique ont aussi été réalisés sur les trajectoires. Ils font partie d'un projet plus large sur la fouille de données spatio-temporelles dans le cadre de la thèse de M. Etienne. Ils seront présentés dans la section 4.2

.

3 L'intelligence artificielle appliquée aux SIG

L'intelligence artificielle a pour but de traiter des connaissances et des raisonnements afin de doter les systèmes informatiques de fonctions normalement associées à l'intelligence humaine : compréhension, raisonnement, dialogue, adaptation, apprentissage, etc.

À l'origine à la fin des années 50, cette discipline scientifique avait pour objectif de concevoir une "machine intelligente". Aujourd'hui, l'intelligence artificielle regroupe l'ensemble des cas où un traitement informatique ne peut être ramené à une méthode simple sous la forme d'algorithme.

Dans le domaine des Systèmes d'Information Géographique, les algorithmes classiques ne permettent pas de résoudre un grand nombre de problèmes. Effectivement, la compréhension d'un espace géographique est complexe. Elle nécessite de combiner des connaissances très variées à différent niveau de granularités. De même, la manipulation des données géographiques répond souvent à des objectifs contradictoires où un algorithme optimisant un seul des objectifs n'est donc pas approprié.

Dans le cadre de la simulation des déplacements dans un espace géographique, le système doit reproduire de manière réaliste les raisonnements humains. Or, chaque humain, quelque soit son mode de déplacement perçoit son environnement géographique différemment. Les notions de distances perçues, de proximité, de vitesse ressentie sont des concepts flous, qui dépendent de l'environnement. Qui plus est, les déplacements sont optimisés en fonction de critères contradictoires (gain de temps, sécurité, respect des règlements formel ou informel, etc.).

Mes travaux dans le domaine de l'intelligence artificielle appliquée aux SIG, concernent deux branches de l'intelligence artificielle :

- les algorithmes génétiques,
- les systèmes multi-agents.

Ces études sont à l'interface de l'intelligence artificielle et des SIG. Ils consistent à tenir compte des spécificités des données spatiales ou spatio-temporelles et à adapter des techniques récentes d'intelligence artificielle à ces données. Les algorithmes génétiques ont été appliqués à la généralisation et à la segmentation de MNT. En ce qui concerne les systèmes multi-agents, dans un premiers temps, des objets mobiles (navires) jouant chacun plusieurs rôles sont simulés de manière concomitante à l'aide d'un modèle agent-groupe-rôle (AGR). Dans cette simulation, l'espace géographique a une place importante. Dans un deuxième temps, des raisonnements de niveau expert sont intégrés pour simuler le rôle le plus complexe : la gestion de l'anticollision. Dans cet objectif, la logique floue et un raisonnement à base de patrons sont employés.

3.1 Les algorithmes génétiques et données spatiales

Dans le domaine des SIG, les problèmes d'optimisation multicritères sans heuristique connue, sont nombreux (van Dijk et *al.* 2002). Par exemple, il n'existe pas d'algorithme relativement simple pour partitionner l'espace ou l'aménager selon des critères corrélés négativement. Pour ces problèmes d'optimisation, les métaheuristiques sont employées. Elles procèdent par itération. A partir de solutions possibles, elles les améliorent en ajoutant des connaissances par analogie avec d'autres processus du monde réel itératifs (métallurgie, évolution des espèces, colonie de fourmis, etc.) et des mécanismes stochastiques. Les algorithmes génétiques (Holland 1975) (Coello et Romero 2002) sont des métaheuristiques qui proposent

de rechercher la meilleure solution à un problème en se laissant guider par les principes d'évolution biologique des espèces.

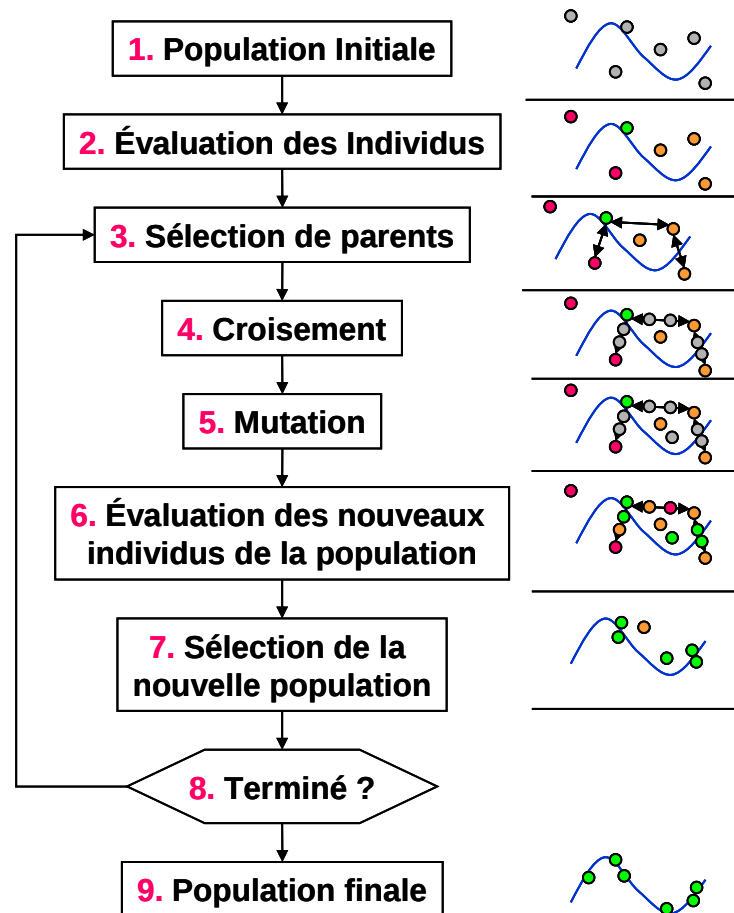


Figure 19: schéma des algorithmes génétiques

Leur principe est résumé et illustré par la figure 19. Pour l'exemple de droite de la figure 19, l'algorithme génétique recherche des bonnes solutions dans $\mathbb{R}^2$ tel que plus le point solution est proche de la ligne (sa position est inconnue) plus la solution est optimale. Initialement, un ensemble de solutions possibles au problème à traiter est défini aléatoirement (étape 1). Une solution est représentée sous la forme d'un codage reprenant l'ensemble de ses caractéristiques. Ce codage équivaut aux gènes d'un individu. Chaque solution est alors évaluée (étape 2). Les meilleurs ont plus de chance de survivre et de se reproduire. La reproduction consiste à sélectionner des couples de parents (étape 3) puis à les croiser pour obtenir deux solutions enfants (étape 4). Le processus de croisement doit alors assurer la transmission des bons gènes des parents à leurs enfants. Des mutants issus de la modification d'une solution existante sont aussi ajoutés à la population (étape 5). Les nouveaux individus (enfants et mutants) sont évalués (étape 6). La population des solutions est alors réduite (étape 7). Plus un individu est bon plus il a de chances d'être conservé. Les étapes 3 à 7 sont alors répétées pour la nouvelle population. Cette boucle s'arrête quand un critère d'arrêt est validé. Une solution doit alors être choisie parmi celles de la dernière génération.

Lors de mes travaux, J'ai proposé des adaptations des algorithmes génétiques aux données géographiques pour répondre aux problèmes suivants :

- sélection de sondes sur une carte marine,
- recherche du plus court chemin météorologique⁷,
- segmentation de MNT.

Premièrement, la notion de **voisinage** propre aux données géographiques, a été intégrée dans le codage des solutions et dans la phase de croisement (Brosset et Devogele 2004). Par exemple, dans le cadre de la segmentation de MNT, de nouvelles formes de codage ont permis de représenter non plus directement les notions de segments, mais les relations de voisins faisant partie d'un même segment. Pour chaque côté des triangles d'un MNT, un booléen est codé. Si ce côté relie deux triangles d'un même segment alors ce booléen vaut vrai. La figure 20 fournit deux solutions respectant ce codage. Les segments peuvent par la suite être reconstruits. Pour cet exemple, la solution de gauche définit 5 segments et celle de droite 4. La segmentation peut alors être évaluée. Elle est jugée bonne si les segments sont homogènes et leur nombre faible. L'homogénéité d'un segment se mesure à l'aide de l'écart type des vecteurs de pente⁸ de ses triangles. Le nombre de segments et l'homogénéité sont des critères corrélés négativement. De même, le processus de croisement prend en compte cette notion de voisinage en coupant le minimum possible de relation. Les deux solutions parents sont coupées en deux selon un axe $y = a \cdot x + b$. Les paramètres a et b sont défini aléatoirement. Les enfants sont alors composés de la partie haute d'un parent et de la partie basse de l'autre. Cette méthode a été testée lors du projet d'élèves (Brunod et Got 2006). Les méthodes classiques de segmentation ont le désavantage de devoir fixer au préalable soit le nombre de segments soit des critères d'homogénéité sur la direction et la longueur des vecteurs de pente. Cet algorithme génétique de segmentation bien que plus complexe ne souffre pas de cette contrainte.

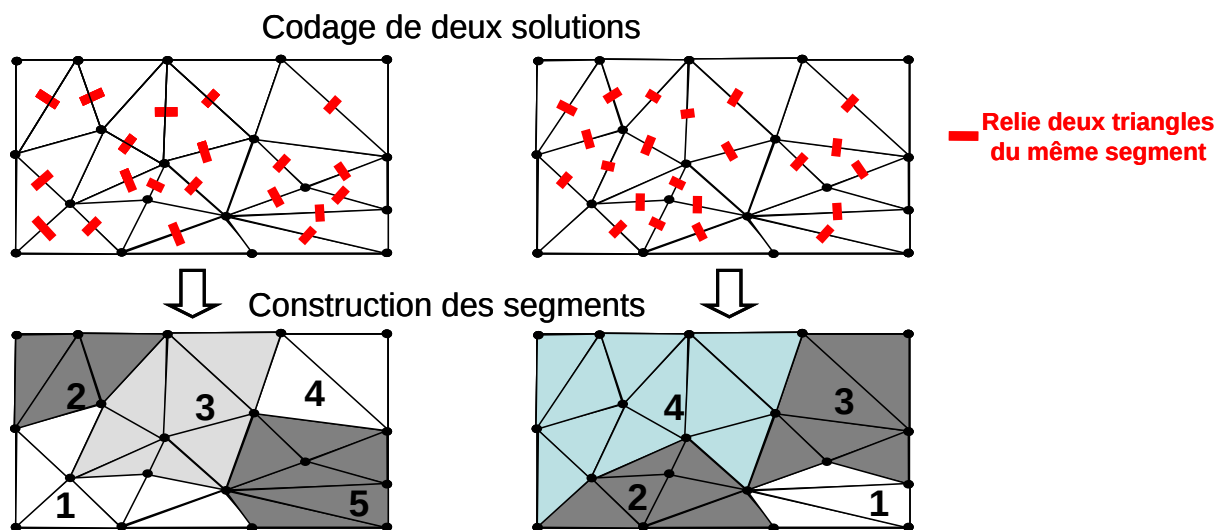


Figure 20: Codage sous la forme de voisinage de solution de segmentation et reconstruction des segments

Deuxièmement, pour croiser deux ensembles de gènes (codages), leur longueur doit être identique donc fixe. Le gène i du parent 1 correspond au gène i du parent 2. Pour des problèmes spatiaux ou spatio-temporels, cette correspondance élémentaire peut être

⁷ Le plus court chemin météorologique est une route maritime optimale qui prend en compte la résistance à l'avancement liée aux conditions météorologiques représentées sous la forme de grilles spatio-temporelles.

⁸ Un vecteur de pente (x, y) a une longueur proportionnelle au pourcentage de la pente et a pour direction celle de la pente.

remplacée par d'autres processus d'appariement plus complexes, mais plus souples. Des correspondances géométriques s'appuyant sur la distance de Fréchet ou des correspondances temporelles s'appuyant sur des estampilles temporelles ont été définies. Ces techniques d'appariement autorisent le croisement de solutions ayant des **codages de longueur variable**. Ces techniques ont été testées dans le cadre de la recherche du plus court chemin météorologique (Edus, Salis 2005). Deux chemins possibles ayant un nombre différent de positions, peuvent ainsi être codés comme des solutions de longueur différente.

Finalement, des approches multicritères spatiales ont été définies en adaptant l'algorithme classique NSGA (Non-dominated Sorting Genetic Algorithm) de (Deb et *al.* 2000), aux SIG. Cette approche a été validée dans le cadre de la problématique de généralisation d'un MNT maritime défini à base d'un ensemble de sondes (Brosset et Devogele 2004). La généralisation de représentation géographique consiste à modifier les données, afin d'obtenir une représentation plus simple et plus abstraite. Le but de la généralisation est donc de réduire le nombre de données transmises en conservant autant que possible l'essentiel de l'information qu'elles véhiculent. La carte marine devant assurer juridiquement la sécurité de l'utilisateur, sa généralisation est caractérisée par l'existence de contraintes imposées par des normes de l'Office Hydrographique International (OHI). Les sondes ne pouvant ni être déplacées ni leur profondeur modifiée (contraintes définies par l'OHI), la généralisation du MNT marin consiste uniquement à sélectionner un ensemble de sondes, en respectant au mieux les contraintes. Trois contraintes principales sont identifiées :

- les contraintes applicatives de **sécurité** : la représentation issue du processus de généralisation doit fournir une "enveloppe haute" de la représentation initiale. En conséquence, les profondeurs interpolées sur la carte doivent être inférieures ou égales aux profondeurs réelles.
- Les contraintes graphiques de **lisibilité** : la représentation généralisée doit respecter des règles cartographiques de lisibilité. Pour toute sonde, la sonde la plus proche doit être à une distance comprise entre 0,4 cm et 4 cm carte.
- Les contraintes structurales de maintien de la **géomorphologie** : l'utilisation de la carte marine impose que le caractère géomorphologique des fonds (pente, rugosité) soit au mieux maintenu. Dans le même temps, les éléments caractéristiques du relief doivent être conservés et mis en évidence. On préserve entre autres les sommets et les cuvettes.

L'heuristique la plus employée pour sélectionner les sondes a été définie par Oraas (Oraas, 1975). Elle consiste à trier les sondes, de la moins profonde à la plus profonde, puis à sélectionner la moins profonde non encore traitée et supprimer celles qui sont à une distance R de cette sonde. Ce processus est répété jusqu'à ce que toutes les sondes soient traitées.

La figure 21, illustre les améliorations obtenues en définissant un algorithme génétique basé sur NSGA. La recherche de la solution combinant des critères corrélés négativement, permet d'obtenir une meilleure représentation que la sélection à l'aide de l'algorithme d'Oraas. L'algorithme génétique est moins lisse, les cuvettes sont mieux conservées et la distribution des sondes reflète mieux la granularité du fond marin. La partie la plus profonde, plus lisse a moins de sondes sélectionnées que la partie haute, plus rugueuse.

Parallèlement à ces travaux portant sur l'emploi d'algorithme génétique dans le cadre des SIG, des recherches ont été menées sur l'utilisation des Systèmes Multi-Agents pour modéliser les déplacements réalistes d'objets mobiles : les navires.

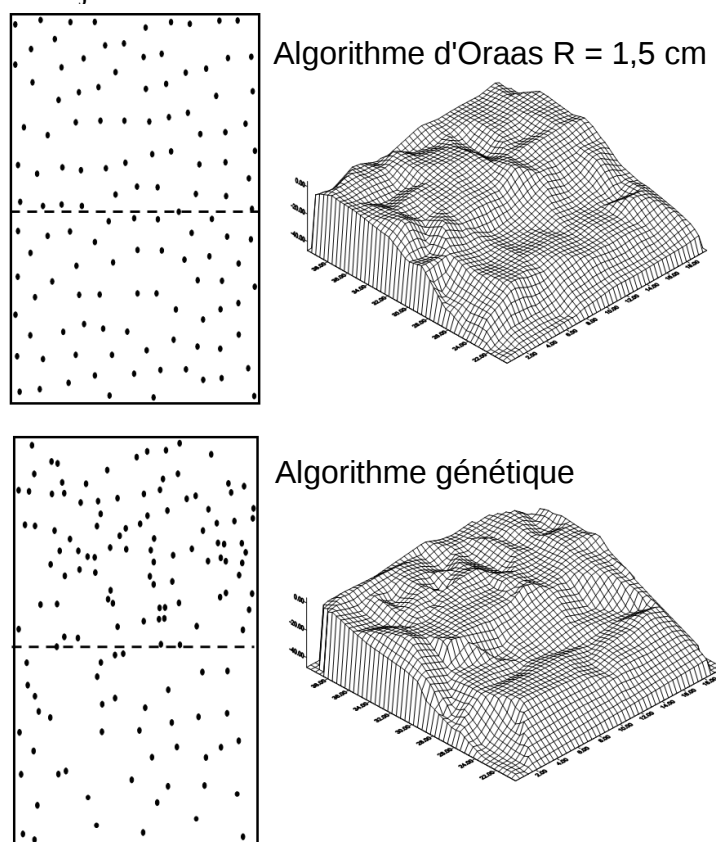


Figure 21 : Comparaison de la sélection de sondes à l'aide de la méthode classique (algorithme d'Oraas) et à l'aide de l'algorithme génétique. La zone haute est plus rugueuse que la zone basse.

3.2 Les systèmes multi-agents agent-groupe-rôle avec un contexte spatial

L'application des Systèmes Multi-Agents (SMA) à des contextes géographiques est améliorable. Les travaux réalisés lors de la thèse de M. Fournier ont permis d'améliorer l'intégration de la dimension spatiale au sein des systèmes multi-agents pour la simulation. Ils sont appliqués à un environnement géographique réglementaire : l'espace maritime. Cet espace est complexe et dynamique. Un nombre élevé d'entités autonomes avec des comportements diversifiés et changeants s'y meuvent. Les interactions entre les entités et l'environnement y sont importantes. Il est donc intéressant de disposer de simulateur pour le contrôle et la prévision de déplacements de navires intégrant le raisonnement spatial.

L'objectif de ce système est :

- de simuler des trajectoires de navires,
- de détecter et éviter les collisions basées sur l'application de règles de navigation et la prise en compte des relations d'orientation et de proximité.
- De détecter et éviter les échouements,
- d'analyser des trajectoires des navires.

Trois futures déclinaisons de cette plate-forme seront proposées afin de répondre aux besoins :

1. des élèves marins pour se former à l'aide de simulateur sur des situations réalistes complexes pour les préparer à la gestion de situations à risque,
2. des centres de surveillance de la navigation maritime (CROSS (Centres Régionaux Opérationnels de Surveillance et de Sauvetage), sémaphores, capitainerie) pour anticiper les trajectoires probables des navires afin de prévoir les risques de collision,
3. des chefs de quart à bord des navires, pour proposer des routes d'évitement de collision dans le cadre d'un processus d'aide aux décideurs.

Les travaux portant sur l'intégration de la perception de l'espace géographique et du raisonnement spatial associé vont maintenant être présentés.

3.2.1 Perception de l'espace et raisonnement spatial

Le cadre de cette étude est donc celui de la simulation de navigations maritimes. La représentation des relations d'orientation et de proximité est déterminante en navigation (Röffer 1999). En conséquence, il s'avère intéressant d'intégrer les travaux de **raisonnement spatial**. Ce dernier concerne l'acquisition, la gestion, la déduction de connaissance sur l'environnement spatial. Ces capacités sont propres à l'humain, qui est capable dans ses tâches cognitives de percevoir des informations spatiales, de les analyser et d'effectuer des actions appropriées en fonction de son environnement.

Ce type de raisonnement est basé essentiellement sur des relations entre des objets. Ces informations sont principalement les relations :

- topologiques (Egenhofer et Franzosa 1991),
- d'inclusion,
- d'orientation,
- de distance.

Elles peuvent être **quantitatives** comme :

- la distance à laquelle se trouve une entité par rapport à un point de référence (en général l'observateur).
- Le cap de cette entité par rapport à une direction de référence (la direction du déplacement si l'observateur est mobile ou du regard sinon).

Néanmoins, les informations quantitatives dans le cadre du raisonnement spatial posent des problèmes en termes de discrétisation de l'espace et de gestion des informations partielles. Les humains préfèrent le plus souvent pour raisonner à partir de son environnement, des informations **qualitatives** telles que :

- la proximité (proche, lointain),
- la direction (droite, gauche, nord, sud).

La figure 22 (a) fournit une perception de l'espace classique défini par Zimmerman et Freksa (Zimmerman et Freksa, 1996). Plusieurs thématiques scientifiques traitent du raisonnement spatial (Spatial cognition) notamment :

- Les SIG,
- Le Traitement du langage naturel,
- La vision artificielle.

Les contextes applicatifs forts telle la navigation maritime modifient la perception de l'espace et des relations spatiales. Cette affirmation était déjà vraie au XVII^{ème} siècle. Elle se matérialisait par des cartes spécifiques à la navigation tel le Portulan de la figure 1. La modélisation du raisonnement spatial doit donc prendre en compte cette perception qualitative spécifique. La figure 22 (b) illustre la perception qualitative spécifique au domaine maritime (Bâbord, Tribord, Arrière).

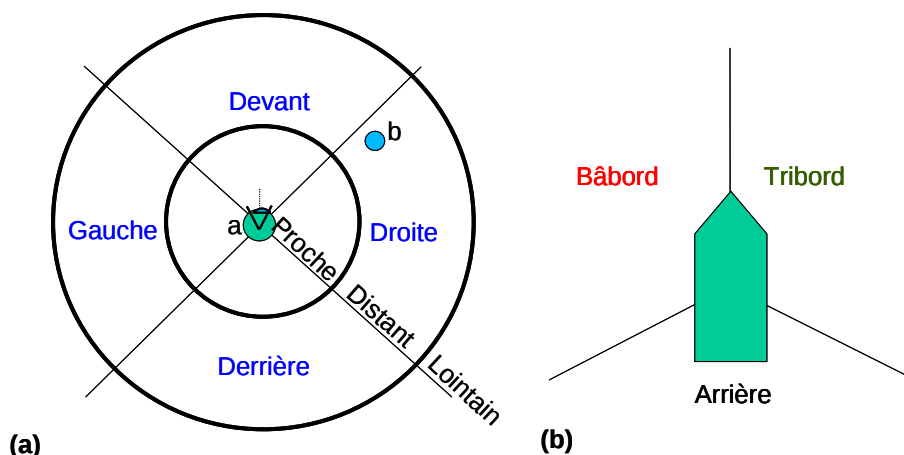


Figure 22 : perception de l'espace classique et maritime

De même, les notions de proximité sont fonction d'une part de la vitesse et de la manoeuvrabilité de l'objet mobile et d'autre part du règlement : Règlement international pour prévenir les abordages en mer COLREG (OMI 1972). Ce règlement est décrit de manière qualitative et est sujet à interprétation. Par exemple la règle 17 dit :

Manoeuvre du navire privilégié

- a) i) Lorsqu'un navire est tenu de s'écarter de la route d'un autre navire, cet autre navire doit maintenir son cap et sa vitesse.
- ii) *Néanmoins, ce dernier peut manoeuvrer, afin d'éviter l'abordage par sa seule manoeuvre, aussitôt qu'il lui paraît évident que le navire qui est dans l'obligation de s'écarter de sa route n'effectue pas la manoeuvre appropriée prescrite par les présentes Règles.*
- b) Quand, pour une cause quelconque, le navire qui est tenu de maintenir son cap et sa vitesse se trouve tellement près de l'autre que l'abordage ne peut être évité par la seule manoeuvre du navire qui doit laisser la route libre,

Ce type de règles peut être traduit en termes de raisonnement spatial qualitatif pour un navire qui est en route de collision. Le risque de collision est évalué à l'aide de la distance minimale à laquelle les deux navires vont se croiser. Cette distance est appelée DCPA (Distance of Closest Point of Approach). Plus le **DCPA** est proche de zéro, plus le risque de collision est important. Ce paramètre est complété par le **TCPA** (Time to Closest Point of Approach) qui mesure le temps qui nous sépare du croisement au plus proche. Plusieurs zones floues sont ainsi définies (voir figure 23 (a)) :

- la zone où si l'objet mobile (l'agent) n'est pas privilégié, il doit manoeuvrer pour éviter la collision. Cette zone est appelée Zone de Distance Minimale de Manoeuvre (ZDMM),
- la zone où si l'autre navire n'a pas manoeuvré, l'agent doit manoeuvrer malgré qu'il soit privilégié. Cette zone est appelée Zone de Distance Minimale d'Urgence (ZDMU),

- la zone où aucun navire ne doit passer. Cette zone est appelée Zone de Distance Minimale de Passage (ZDMP).

Les navires venant de droite étant privilégiés, cette zone est plus étendue.

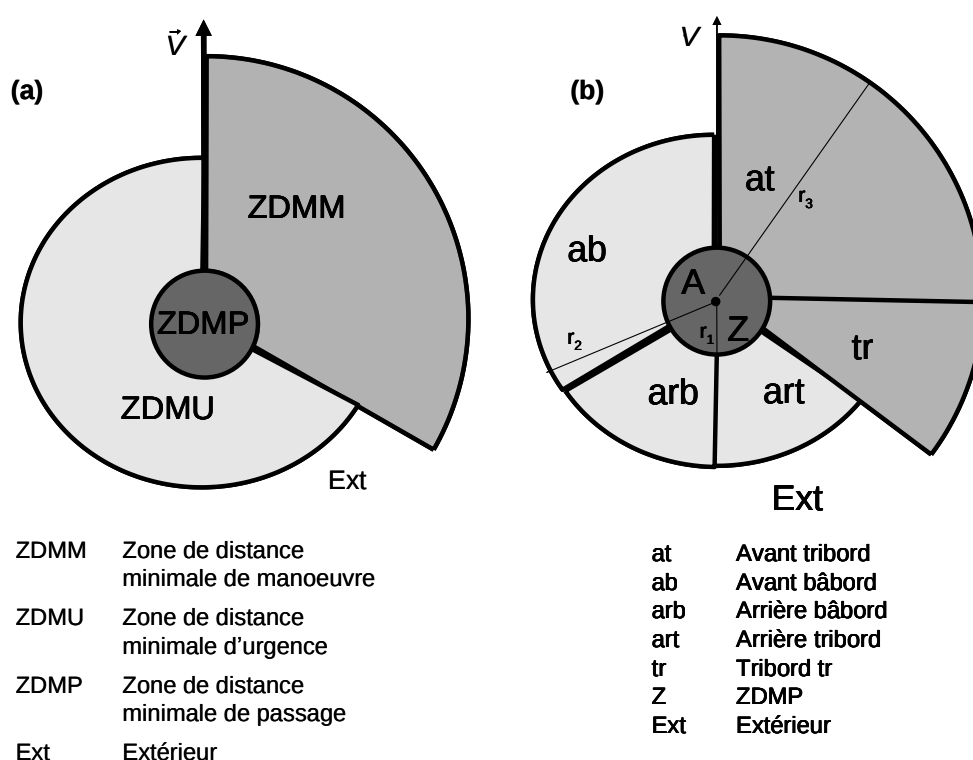


Figure 23: perception de l'espace maritime selon les règles de navigation

Ces trois zones peuvent être découpées en fonction du type de manoeuvres à effectuer (voir figure 23 (b)). Par exemple, pour la région avant bâbord qui fait partie de la ZDMU. Le navire qui est en collision potentielle avec le navire de référence n'est donc pas privilégié, il doit manoeuvrer. S'il ne le fait pas, le navire de référence doit réagir. Le navire de référence modifie alors son cap de la manière significative de manière à ce que l'autre navire constate qu'il a manoeuvré. Par exemple le nouveau cap NC peut être défini de la manière suivante :

$$NC = AZ + 180 - 30^9$$

L'azimut (AZ) entre les deux navires est l'angle que fait le segment de droite entre leurs centres par rapport au nord en prenant comme origine le centre du navire de référence dans le sens trigonométrique inverse.

Pour chaque navire, les rayons (r_1 , r_2 et r_3) ainsi que les manoeuvres à effectuer, sont précisés par le commandant du navire conformément au règlement.

Le contexte spatial peut aussi modifier le raisonnement. La présence de **barrière** va modifier l'emprise des zones. Par exemple, la proximité d'une digue (voir figure 24 (a)) entraîne la déformation des zones. Un navire de l'autre côté de la digue (navire c) n'est pas dans la zone "at". Par contre, un navire du même côté de la digue à une distance plus importante (navire d) est inclus dans la zone "at" déformée par la présence de cet obstacle. Les barrières sont

⁹ Ce nouveau cap est celui qui est préconisé lors des cours de navigation de la marine nationale.

nombreuses dans l'espace maritime (côtes, hauts fonds, digues, etc.) et doivent être prises en compte afin de simuler de manière cohérente la navigation maritime.

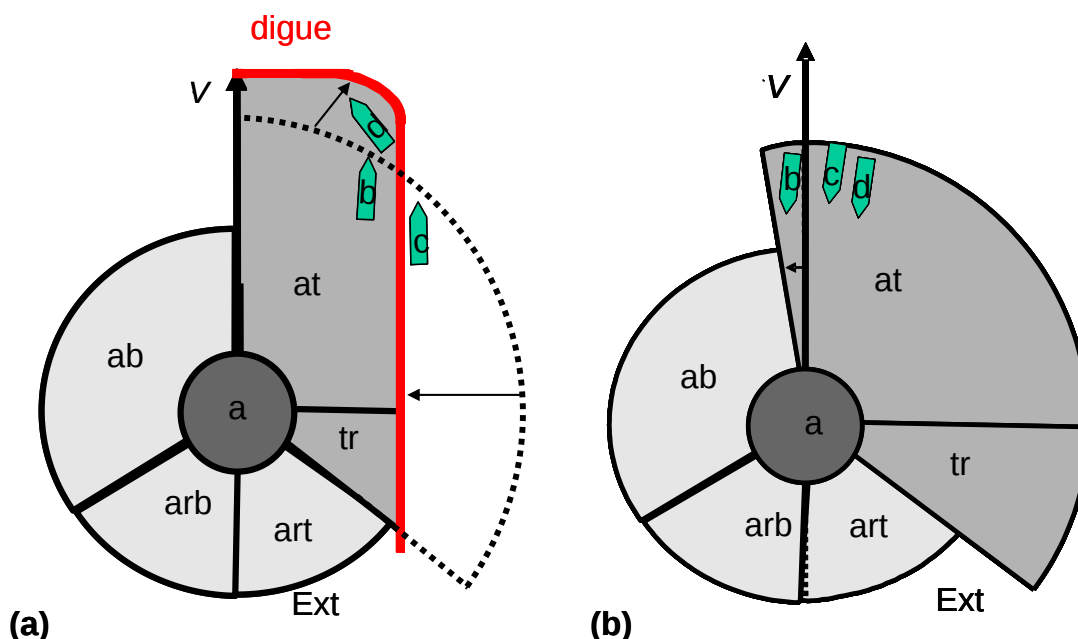


Figure 24: Perception de l'espace déformé par le contexte spatial

De même, la notion de **groupe spatial** doit être prise en considération. Pour chaque entité, des régions d'influence doivent être définies. Un groupe spatial se forme si au moins une entité se trouve dans la région d'influence d'une seconde entité. Ces groupes doivent être intégrés dans un raisonnement spatial mésoscopique¹⁰. Par exemple, pour la figure 24 (b), les navires b, c et d doivent être considérés comme un groupe. L'emprise de ce groupe intersectant la zone "at" et la zone "Ext". La zone la plus défavorable : "at", est étendue afin que l'ensemble du groupe soit inclus dans une même zone. Qui plus est, les manœuvres d'évitement découlant du raisonnement sont modifiées afin de contourner l'ensemble du groupe et non pas uniquement l'élément le plus dangereux.

Ces travaux en termes de perception de l'espace et de raisonnement spatial sont détaillés dans (Fournier et al. 2007).

3.2.2 Modèle Agent-Groupe-Rôle (AGR) étendu à la dimension spatiale

Le deuxième apport de la recherche est orienté vers la définition d'un modèle agent flexible qui permet la représentation de comportements multiples et changeants à travers la notion de **rôle**. Cette notion a été définie dans la méthode Aalaadin (Ferber J. et Gutknecht O. 1998) et la méthode GAIA (Wooldrige et al. 2000) comme une description de la fonction abstraite qu'une entité peut avoir au sein d'un système. Ce modèle a été retenu afin que les navires simulés puissent jouer des comportements multiples simultanément (suivre une route, éviter les collisions, éviter échouement, etc.). De plus, il a été étendu pour prendre en compte la dimension spatiale. Un **agent spatial** est donc défini comme un agent situé dans l'espace géographique. Ce type d'agent est capable de raisonner dans l'espace. Il infère des actions et des interactions selon le positionnement des autres agents. En ce qui concerne l'application

¹⁰ Le niveau mésoscopique est un niveau intermédiaire entre le niveau microscopique (les agents) et le niveau macroscopique global.

maritime, l'ensemble des navires naviguant dans une même zone va former un groupe. De même, une flottille de navires militaires est aussi un groupe. Un agent peut donc faire partie de plusieurs groupes et jouer simultanément plusieurs rôles dans chacun de ces groupes.

La figure 25 décrit le métamodèle Agent-Groupe-Rôle (AGR) étendu (Fournier et *al.* 2003b). Les propriétés de ces rôles sont :

- le caractère permanent ou temporaire des rôles,
- les incompatibilités entre les rôles,
- le caractère exclusif de certains rôles,
- le caractère inconciliable entre des rôles et des agents d'un certain type,
- les priorités des rôles.

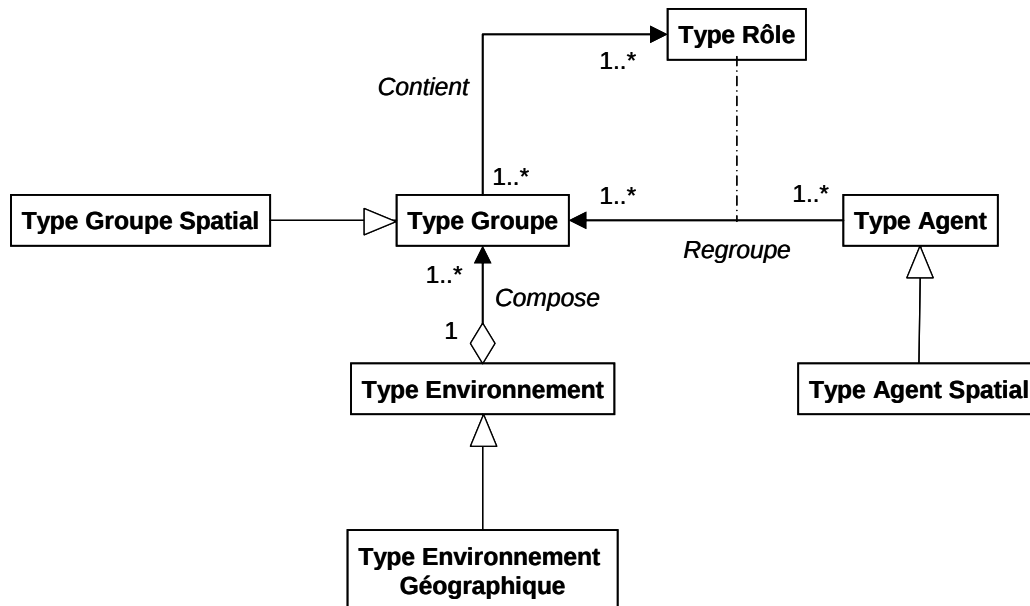


Figure 25 : Métamodèle multi-agents Agent-Groupe-Rôle étendu

A un instant t , chaque agent joue donc un ensemble de rôles. Le comportement de l'agent est fonction de la priorité de ces rôles. Il peut aussi acquérir ou perdre des rôles en fonction d'évènement, de sa position et de la position des agents environnant.

3.2.3 Prototype TRANS

Ce modèle est validé par la réalisation du prototype **TRANS** (Tractable Role Agent prototype for concurrent Navigation Systems) dont l'objectif est de modéliser et de simuler les trajectoires simultanées de navires (Fournier et *al.* 2003a). Dans ce prototype, un large espace maritime est défini à partir de données cartographiques numériques maritimes. Du point de vue de la performance, jusqu'à 200 navires peuvent être simulés, avec un facteur d'accélération de la simulation de 10.

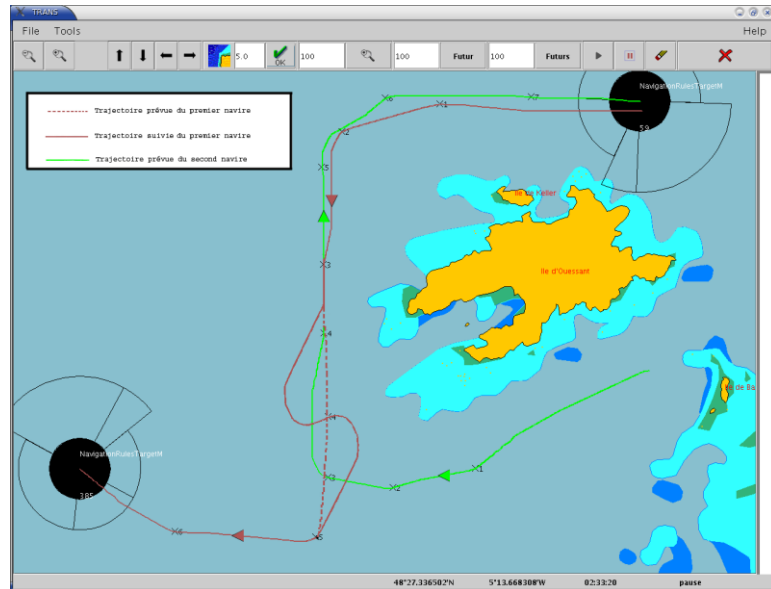


Figure 26 : Capture d'écran de la fenêtre principale de TRANS. Deux navires réalisant une manœuvre d'évitement sont représentés.

La figure 26 fournit une copie d'écran de TRANS. Dans cet exemple deux navires suivent des routes (défini par un ensemble de points) et contournent en sens inverse l'île d'Ouessant. Le trajet déjà effectué par chacun des navires est affiché. La route du navire descendant a été déviée afin de respecter les règles de navigation. La route initialement prévue est en pointillé. Cette déviation est intervenue lorsque les deux navires se sont croisés pour éviter la collision. Le navire descendant étant non privilégié, il a dévié de sa route. Il a ensuite repris la route initialement prévue. Ce changement de trajectoire est obtenu par la prise du rôle manœuvre d'évitement. Ce rôle a une priorité supérieure aux rôles que l'agent jouait déjà, et il a inhibé le rôle "suivi de la route". Cette inhibition oblige le navire à manœuvrer au lieu de suivre la route prévue. Dès que les deux navires ne sont plus en route de collision, ce rôle est perdu et le navire joue à nouveau le rôle "suivi de la route".

3.3 Les systèmes multi-agents cognitifs intégrant des concepts de patrons

Ce premier prototype : TRANS est très satisfaisant du point de vue temps de calcul et prise en compte ou non du règlement. Par contre, il souffre d'un manque de **réalisme** en ce qui concerne des manœuvres complexes tels les évitements de collision. En ce qui concerne une simulation informatique de phénomènes du monde réel, une simulation est dite **crédible** si le comportement observé des objets simulés est proche de celui des objets du monde réel. Dans le cadre du trafic maritime, cette propriété peut être vérifiée à l'aide de tests proche de ceux proposés par Turing (Turing 1950). Un opérateur est devant deux écrans de contrôle un avec le rejou d'un trafic réel et l'autre avec un trafic simulé, si l'opérateur n'est pas capable de distinguer le trafic simulé sur une longue période, alors la simulation est dite crédible. Qui plus est, la simulation est **réaliste** si en plus d'être crédible, elle est capable de reproduire les lois internes basées sur les règles formelles et informelles qui régissent le modèle.

Pour la navigation maritime, Chauvin (Chauvin 2000) a montré qu'environ 20% des décisions d'anticollision ne suivaient pas le règlement stricto sensu. Cette étude a été menée dans le Pas-de-Calais. Un simulateur réaliste doit donc disposer de modules de prise de décision plus complexe que la seule prise en compte ou non du règlement. Une règle informelle est liée à l'expérience et à l'a priori vis-à-vis d'autres types d'objets. Elle entraîne des adaptations du

règlement. Par exemple : "Les petits cargos ne manœuvrent jamais donc je vais manœuvrer très tôt même si je suis privilégié " est une règle informelle courante employée par les ferries. Elle doit être prise en compte pour rendre la simulation réaliste. Cette règle informelle servira d'illustration dans ce chapitre.

Cette section va décrire maintenant le mécanisme permettant de reproduire des raisonnements complexes d'expert basé sur un modèle issu de la psychologie cognitive. Ces travaux sont réalisés dans le cadre de la thèse de M. Le Pors (Le Pors et *al.* 2009). Ils s'appuient sur l'étude de la prise de décision dans des situations d'anticollision réalisées par la MCF Christine Chauvin (Chauvin 2000) (Chauvin et *al.* 2008) (Chauvin C, Lardjane S. 2008) à l'Université de Bretagne Sud, co-encadrante de cette thèse. Le modèle informatique associé est ensuite détaillé. Finalement le nouveau prototype CogTRANS qui intègre dans TRANS ce modèle est présenté et les résultats analysés.

3.3.1 Rôle à base de patrons

Des études dans le domaine de la psychologie (Kobus et *al.* 2001) ont montré qu'en situation naturelle dynamique les experts n'ont pas le temps d'envisager un grand nombre d'options. Ils s'appuient sur des modèles de prise de décision prototype, qui ne reflète pas la complexité de la situation naturelle. Dans le modèle de prise de décision de Klein (Klein 1997) ; **Recognition Primed Decision Model** ou RPD, un expert, confronté à une situation, pour raisonner dans 90% des cas reconnaît qu'elle est l'instance d'un prototype. Ce dernier décrit :

- Le type de la situation,
- Les indices pertinents,
- Les attentes,
- Les buts pertinents,
- Les cours d'action typiques,

Ce modèle conduit l'expert à choisir une décision « satisfaisante » liée au prototype. L'expert fait donc un raisonnement en appariant une situation courante à une situation prototypique. Ces situations prototypiques sont schématiques. Ils sont décrits à l'aide de symboles qualitatifs.

Dans le cadre de la thèse de M. Le Pors, il a donc été décidé de transposer ces résultats dans le cadre du modèle AGR des SMA. Dans ce but, deux types de rôles sont définis dans notre modèle :

- **Rôles à base de règles.** Ces rôles s'appuient sur des cycles simples classiques (perception – décision – action) des agents. Les agents réagissent à une perception simple vis-à-vis d'un autre agent ou de l'environnement. En fonction d'une règle, de type règle d'inférence, ils décident de manière réactive de réaliser telle action. Pour une situation, une seule règle est possible. L'incertitude n'est pas gérée.
- **Rôles à base de patrons.** Ces rôles s'appuient sur des cycles légèrement plus complexes. Les agents perçoivent leur environnement, ils en dégagent les indices pertinents à partir de ces indices, ils identifient des patrons (situation prototypique) proches. Parmi les patrons possibles, un patron est choisi. Plus un patron répond aux critères d'optimisation de l'agent, plus il a de chance d'être retenu.

L'objectif de ce deuxième type de rôle est d'obtenir des agents qui simulent la prise de décision d'expert dans des situations complexes. Ainsi, pour notre application maritime, des comportements liés aux facteurs humains (règles informelles) sont pris en compte. Par exemple, les navires, lors d'un risque de collision est détecté, peuvent décider d'anticiper la manœuvre ou de forcer le passage (Le Pors et *al.* 2007).

3.3.2 Modélisation des décisions d'experts basée sur les patrons en situation naturelle (DPSN)

Ces rôles à base de patrons sont modélisés à l'aide de plusieurs phases. Ce modèle appelé DPSN (Décisions d'experts basées sur les Patrons en Situation Naturelle) est générique et peut être adapté pour simuler le comportement d'expert humain en situation naturelle dynamique. Les informations de type incomplètes, incertaines ou imprécises doivent être prises en compte, l'expert ne disposant pas forcément de l'ensemble des informations. Qui plus est ces informations sont souvent imprécises.

Des méthodes informatiques diverses sont employées dans DPSN:

- sous-ensembles flous,
- distances sémantiques,
- roue de la fortune.

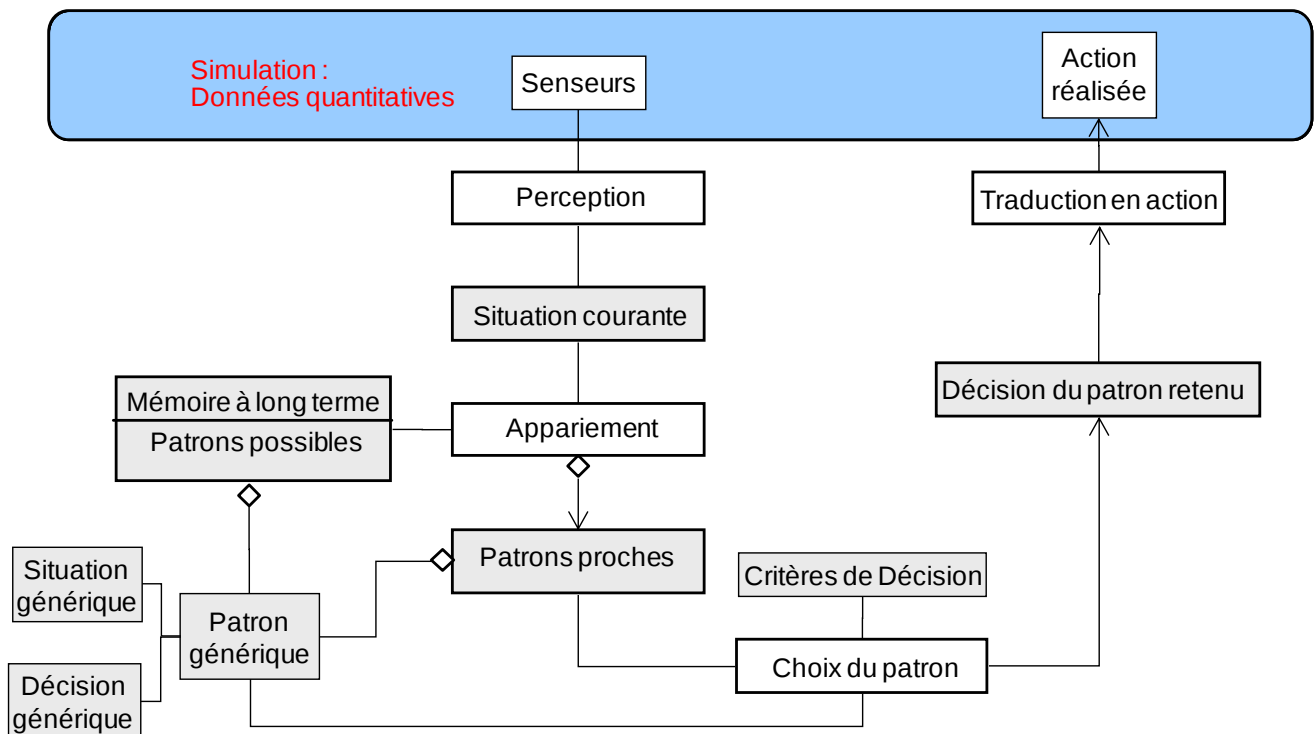


Figure 27: Modèle cognitif de décision d'expert basé sur des patrons

La figure 27 décrit l'enchaînement de ces phases. Premièrement, des données quantitatives (vitesses, directions, distances entre objets) sont envoyées à l'agent. Cette phase simule les données que les senseurs envoient à l'expert. Les experts emploient dans le modèle DPSN des données qualitative (loin, à droite, etc.). Un module de **perception** est donc mis en place. Il a pour objectif de traduire les données quantitatives en données qualitatives perçues. Des sous-ensembles flous (Zadeh, 1965) sont employés dans cet objectif. Ces derniers ont été introduits afin de modéliser la représentation humaine des connaissances. Il est naturel d'employer ce mécanisme classique dans cette phase. L'ensemble des indices pertinents perçus va former la situation courante. Ces indices qualitatifs sont par nature flous,

Pour modéliser des ferries naviguant dans le Pas-de-Calais, il est nécessaire de simuler leur perception des cargos qu'ils croisent en Manche. Le commandant a besoin d'identifier le type de cargo (petit ou gros) afin d'appliquer ou non la règle informelle. Le type de cargo est donc un indice pertinent. Pour obtenir cet indice, la vitesse mesurée par le RADAR des navires environnant est employée. La figure 28 fournit les fonctions d'appartenance des sous-

ensembles flous de cet indice pertinent (petit cargo, grand cargo ou inconnu) en fonction de la vitesse exprimée en nœuds¹¹. Par exemple, un navire naviguant à 13 nœuds est plutôt un petit cargo. Un navire ayant une vitesse de 17 nœuds est perçu de manière plus vraisemblable comme un grand cargo.

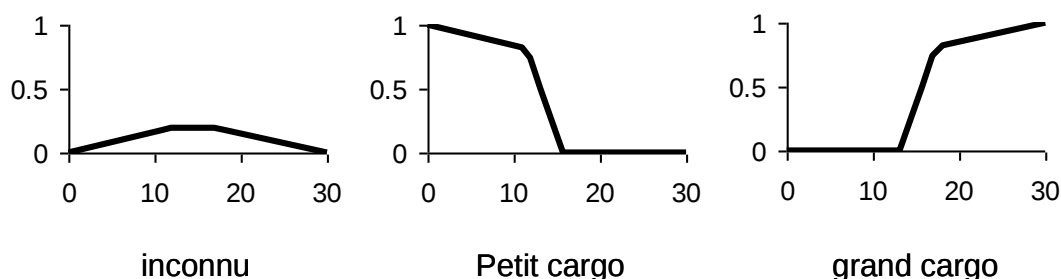


Figure 28 : fonction d'appartenance en fonction de la vitesse en nœuds pour les sous-ensemble flous "inconnu", "petit cargo" et "grand cargo"

De façon plus formelle, une fonction, appelée fonction d'appartenance, est définie. Elle associe à une variable v le degré d'appartenance à un sous-ensemble flou A . Cette fonction est notée $\mu_A(v)$ ($\mu_A(v) \in [0, 1]$ et $\sum \mu_{Ai}(v) = 1$). D'un point de vue statistique, Il y a donc $\mu_A(v)$ chance d'associer le sous-ensemble flou A à la valeur v . Le tableau 6 fournit quelques exemples de valeur de la fonction d'appartenance des 3 sous-ensembles flous ("inconnu", "petit cargo" et "grand cargo") de l'indice "type de cargo" en fonction de la vitesse.

	8 nœuds	13 nœuds	15 nœuds	17 noeuds	22 noeuds
inconnu	0,15	0,2	0,2	0,2	0,15
petit cargo	0,85	0,8	0,4	0	0
grand cargo	0	0	0,4	0,8	0,85

Tableau 6 : Fonction d'appartenance des 3 sous-ensembles flous de type de cargos en fonction de la vitesse

Pour l'indice pertinent "position de croisement", les 5 valeurs possibles sont :

- croisement par l'arrière : la distance lors du futur croisement : DCPA (voir page 48) est relativement faible, mais le ferry passe derrière, environ entre 0,4 et 0,2 mille marin¹²,
- collision : le DCPA est proche de 0 mille marin,
- croisement par l'avant : le DCPA est relativement faible, mais le ferry passe devant, environ entre 0,2 et 0,5 mille marin,
- inconnu,
- pas de risque de collision, quand la distance minimum lors du croisement est relativement importante.

La figure 29 détaille les fonctions d'appartenance de ces 5 sous-ensembles flous. Une dissymétrie par rapport à l'axe des ordonnées est due au fait qu'il est plus dangereux de passer devant que derrière. Les risques de collision ne sont pas appréhendés par tous les navires de la même manière. Les distances seuils doivent donc être adaptées pour chaque type de navires.

¹¹ Un nœud = un mille marin par heure soit 1,852 km/h.

¹² Un mille marin = 1,852 km, Elle a initialement désigné la longueur d'une minute d'arc sur un grand cercle de la sphère terrestre. Autre appellation mille ou nautique.

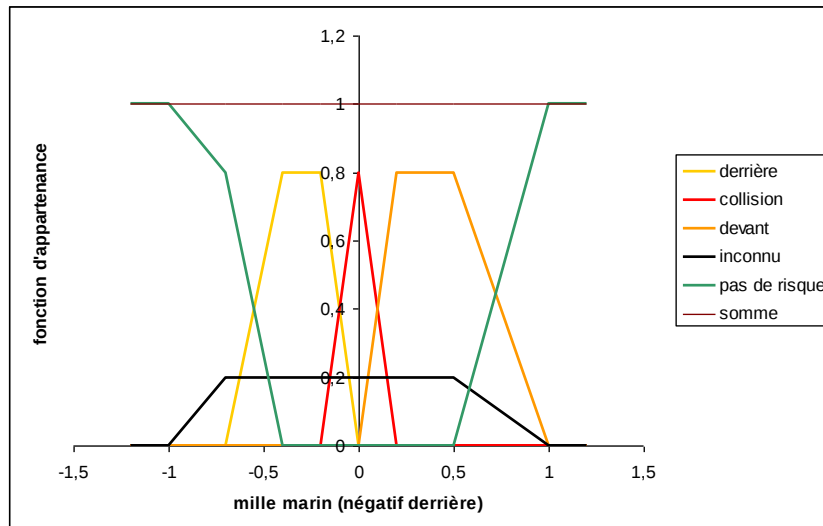


Figure 29: fonction d'appartenance en fonction de la distance minimum en mille marin pour les 5 sous-ensembles flous.

De même, pour l'indice pertinent "zone", 4 zones floues sont définies : (1) Zone de Distance Minimale de Passage, (2) Zone de Distance Minimale d'Urgence, (3) Zone de Distance Minimale de Manœuvre, (4) Zone Extérieur (voir page 48 pour plus de détail sur ces zones). Pour chaque indice pertinent de la situation courante, chaque valeur qualitative a une probabilité d'être retenue égale à la fonction d'appartenance.

Pour l'exemple, nous supposons que l'agent ferry reconnaît en fonction de sa position et des informations fournies par les capteurs simulés et des fonctions d'appartenance qui lui sont associées, la situation courante suivante :

"Situation de croisement, d'un petit cargo, je suis non privilégié, croisement par l'avant, actuellement je suis en zone extérieur (zone 4)"

La situation courante doit alors être comparée avec l'ensemble des patrons formant la mémoire à long terme de l'agent. Ces patrons génériques sont composés d'une part d'indices pertinents avec les valeurs possibles et d'autre part d'une décision associée. Ils sont issus principalement du recueil d'information auprès des experts. Néanmoins, les observations doivent corroborer ces recueils. Les informations similaires recueillies sont alors regroupées en un patron générique. Pour comparer la situation courante avec les patrons génériques, des **distances sémantiques** sont employées. Pour chaque type d'indices pertinents, une métrique est définie. Les valeurs pour chaque indice n'étant pas toujours connues, la valeur inconnue doit être supportée. Par convention, la distance sémantique varie entre 0 (valeurs identiques) et 1 (valeurs opposées). La valeur "inconnue" est à une distance de 0,5 de toutes les autres valeurs. Pour définir les distances entre une situation courante et les situations prototypiques de chaque patron, une somme pondérée de l'ensemble des distances sémantiques est employée. La situation courante est alors appariée avec les patrons les plus proches.

Pour chaque type d'agents, et chaque type de situations complexes (risque de collision dans notre cas), la mémoire à long terme doit être composée d'un nombre raisonnable de patrons (environ de 5 à 20). Cette approche a aussi l'avantage de rendre la base de patrons évolutive sans changer le mécanisme de prise de décision. De même, la base de patrons de chaque type d'agents peut être différente. Elle dépend de la perception de ce type d'agent des autres agents et des règles informelles.

Patt ern	Type of ship		Situat ion	Zone of ship presence				priority		Position in CPA			Decision					
	N°	Small		Big	Cross ing	1	2	3	4	G W	S A	Ahead	Astern	Col	No	Port	Starbd	Slow
1																		
2																		
3																		
4																		
5																		
6																		

Tableau 7 : ensemble des patrons des ferries croisant un cargo prioritaire en zone extérieure (4) sans être privilégiés (GW pour Give Way). Les cases grisées correspondent aux valeurs possibles pour les indices et aux décisions associées

Pour un agent de type ferry, les patrons du tableau 7 sont possibles pour les situations de croisement en zone 4 d'un cargo sans être privilégié (GW pour Give Way). Pour les petits cargos non privilégiés qui croisent par l'avant, deux patrons ont une distance sémantique proche (pour cet exemple, les distances sont même nulles) et sont possibles : les numéros 1 et 2. La situation courante est donc appariée avec ces deux patrons (voir figure 30).

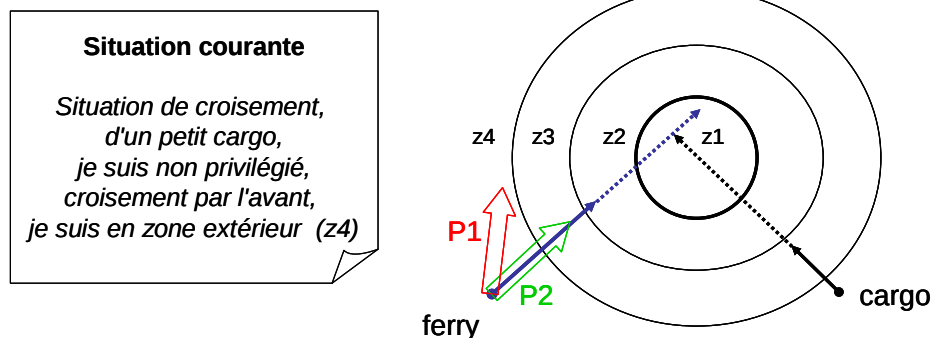


Figure 30 : Situation courante détectée et décisions associées aux deux patrons proches P1 et P2

L'indice "type de cargo" n'est pas pertinent pour cet exemple, ces deux patrons s'appliquant aussi bien au petit et au grand cargo. Par contre, pour les croisements par l'arrière (Astern) cet indice est pertinent. Pour le patron 1, la décision associée est de ne rien faire. Le navire pouvant attendre d'être en zone 3 pour manœuvrer réglementairement et passer par l'arrière en tournant à droite. Pour le patron 2, la décision est de tourner à gauche (bâbord ou port en anglais). Dans ce cas, le navire anticipe la manœuvre, au lieu d'attendre d'être en zone 3 et d'être obligé de tourner à droite pour passer derrière le cargo. La manœuvre réglementaire a le défaut de rallonger le trajet.

Une fois les patrons proches connus, un patron doit être choisi. Pour cela des critères de décisions sont employés. Ces critères représentent les objectifs de l'agent. Plus un patron respecte les objectifs plus ce patron a de chance d'être retenu. Pour chaque critère une pondération est définie en fonction de l'importance accordée à ce critère par l'agent. Chaque agent peut disposer d'un profil différent avec des pondérations qui lui sont propres afin de définir des comportements hétérogènes. Pour choisir le patron, une roue de la fortune est employée. Pour chaque patron apparié, un nombre de parts de la roue, proportionnel à la somme pondérée, est affecté. Une part est tirée aléatoirement. Le patron associé à cette part est alors retenu.

Dans le domaine maritime, pour les manœuvres d'anticollision, trois critères principaux sont retenus :

- le respect du règlement,
- le gain de temps qui a pour objectif de respecter l'heure d'arrivée prévue du navire,
- la sécurité : moins un navire prend de risque plus la manœuvre est sûre. De même, un navire va chercher à maîtriser la situation afin de ne pas dépendre d'une manœuvre de l'autre navire.

Le tableau 8 fournit les valeurs des deux patrons 1 et 2 en fonction de ces trois critères. Les ferries accordant une grande importance au troisième critère, le patron numéro 2 est donc plus souvent choisi.

	Respect du règlement	Gain de temps	sécurité
Patron 1	bon	faible	moyen
Patron 2	faible	moyen	bon

Tableau 8 : valeur des deux patrons pour chaque critère

La dernière phase de la DPSN est la traduction en action de la décision. La décision qualitative associée au patron retenu est traduite en quantitatif. Pour des objets mobiles, cela entraîne par exemple, des changements de cap ou de vitesse.

Pour l'exemple, la traduction du patron 1, consiste à définir un nouveau cap plus à gauche. Le navire étant en zone 4. Il n'a pas besoin de manœuvrer de manière significative, il suffit de manœuvrer de telle sorte que le ferry passe suffisamment loin du cargo. Par contre, si le navire est en zone 3, la manœuvre doit être suffisamment franche pour que l'autre navire perçoive que le ferry manœuvre.

3.3.3 Prototype CogTRANS

Afin de valider la DPSN, le prototype CogTRANS (pour cognitive TRANS), est développé. Il est basé sur TRANS. Il bénéficie donc des avancées de TRANS (perception de l'espace géographique, emploi du modèle AGR, interface cartographique, etc.). CogTRANS, ajoute un module cognitif DPSN. Ce module pour l'instant traite uniquement des manœuvres d'anticollision entre navires. Les patrons génériques des cargos et des ferries ont été définis à partir des travaux de Chauvin et Lardjane (Chauvin et Lardjane 2008). Depuis peu, les patrons des navires de pêche ont été ajoutés (Sarazin et Malini 2009).

Les résultats simulés ont été validés en les confrontant aux résultats de 62 cas observés à bord de ferries opérant dans le détroit du Pas-de-Calais (Le Pors et *al.* 2009). Les manœuvres respectant le règlement sont conformes (le navire qui manœuvre est le même et il choisit la même direction pour le changement de cap). Plus de 90% des manœuvres qui respectent le règlement ont été simulées de manière similaire. Cependant, pour les autres cas, la simulation donne des résultats réalistes, hélas, ils ne sont pas toujours conformes à l'observation. Actuellement, uniquement un peu plus de 50% des simulations sont identiques aux observations. Plusieurs raisons expliquent ces différences :

- pour les manœuvres tardives, les deux navires peuvent manœuvrer indifféremment,
- ce module ne tient pas encore compte du contexte spatial (côte, dispositif de séparation de trafic, barrières, groupes spatiaux)

Ce module cognitif DPSN défini dans CogTRANS est donc adapté à la simulation de prise de décision d'expert en situation naturelle. Il est basé sur la confrontation rapide d'une situation

courante avec une base de situations prototypiques. Il est similaire au raisonnement humain employant le paradigme de la NDM (Naturalistic Decision Making) à base de cas. Ces travaux montrent qu'il est possible d'intégrer dans des systèmes multi-agents, des processus cognitifs complexes d'expert s'appuyant sur ce paradigme et ainsi de bénéficier des travaux issus de la psychologie cognitive. Néanmoins, ces travaux méritent d'être élargis à d'autres types de situations complexes afin de démontrer la généralité de cette méthode. De plus, l'ensemble des apports proposés par TRANS telle la prise en compte du contexte spatial, n'a pas pu encore être pris en compte. Pour améliorer les simulateurs de navigation, il est fondamental de bénéficier de base de données conséquente portant sur les trajectoires des objets mobiles et d'outils d'analyse puissant de ces déplacements, c'est l'un des objets du chapitre suivant.

4 Les bases de données spatio-temporelles

Les SIG ont été conçus initialement pour gérer une représentation du monde réel à un instant donné. Depuis 1988, de nouvelles modélisations de l'espace ont permis de représenter les évolutions du monde réel. Ces premières représentations étaient soit des suites de représentations complètes associées à une date : snapshot (Armstrong 1998) ou des suites de valeurs associées à des estampilles temporelles pour des objets ayant une géométrie fixe (Langran et Chrisman 1988). Les modèles des années 90 ont évolué vers la gestion d'événements modifiant les valeurs thématiques et géométriques des objets (Peuquet et Duan 1995, Worboys 1992, Raper et Livingstone 1995). Néanmoins, ces modèles ne sont pas optimaux pour la gestion d'objets évoluant rapidement tels les objets mobiles.

De plus en plus d'objets mobiles sont équipés de capteurs qui facilitent la surveillance de leur évolution. Parallèlement, les Systèmes de Gestion de Base de Données (SGBD) actuels permettent de gérer des informations issues de capteurs fournissant en temps réel les trajectoires d'objets mobiles. Dans le cadre du GDR SIGMA (MAGIS depuis 2009), en collaboration avec l'université de La Rochelle, une plate forme de gestion du suivi des déplacements des navires (Claramunt et *al.* 2007, Ray et *al.* 2007, Servigne et *al.* 2009) a été modélisée et mise en place. Elle repose sur une architecture n-tiers. Dans un premier temps, elle recueille les positions GPS, envoyées par des transpondeurs AIS (Automatic Identification System) via VHF. Ces données sont filtrées, réorganisées et gérées par un SGBD géographique : POSTGIS. Des clients peuvent alors se connecter en local ou via Internet pour interroger la base sur les données actuelles ou historiques.

Les résultats des requêtes peuvent être visualisés sous la forme de carte dynamique (symboles représentant les objets mobiles superposés sur un fond de carte) à l'aide de Google Maps¹³ ou Google Earth¹⁴. La figure 31 illustre la visualisation produite par cette plate forme. Un navire-citerne : le LNG RIVER NIGER arrive au port de Brest. En cliquant sur le symbole associé à ce navire, l'ensemble des informations disponibles sur cet objet est affiché.

¹³ <http://maps.google.fr/>

¹⁴ <http://earth.google.com>

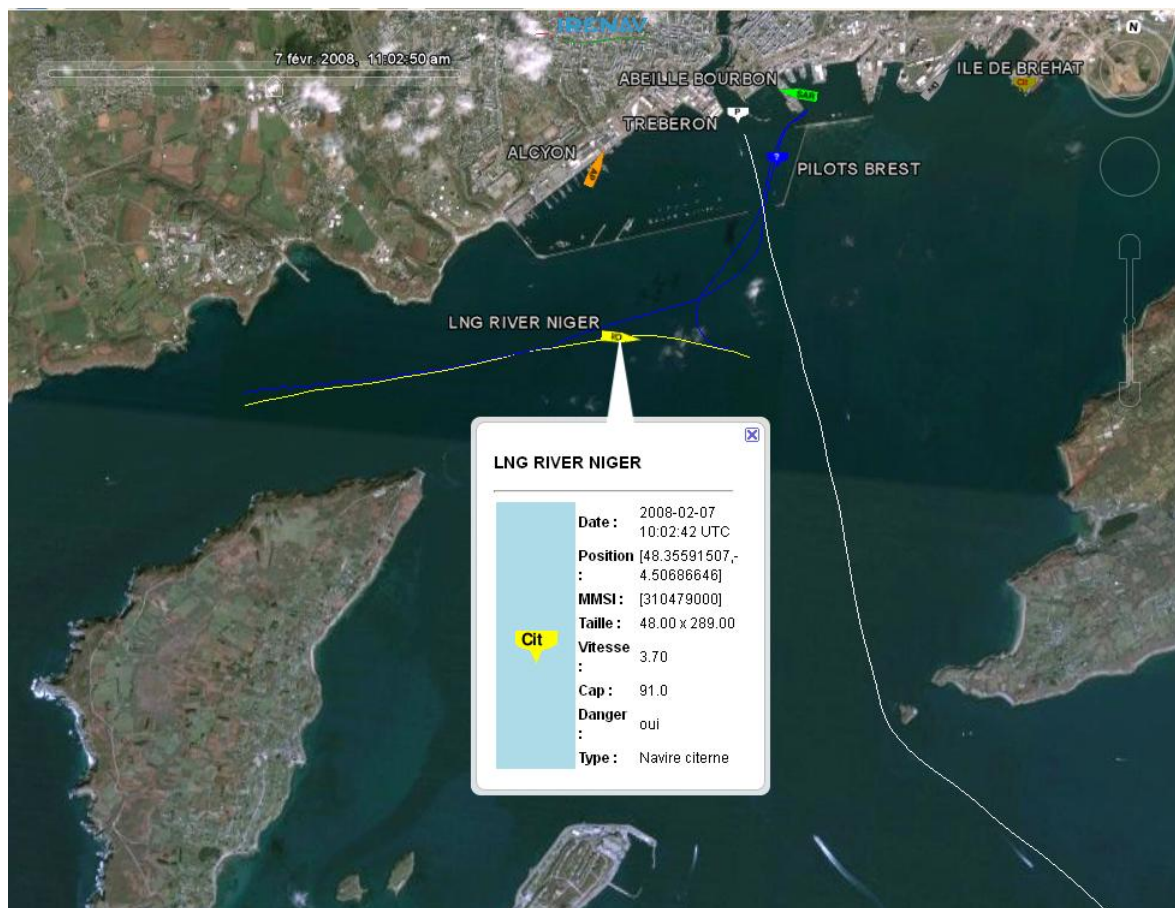


Figure 31 : Une situation issue de la plate-forme de gestion du suivi des déplacements des navires

À partir de ce socle, plusieurs activités de recherche et travaux d'élèves ont été initiés. Dans ce cadre, j'ai co-encadré deux thèses. La première porte sur la définition de vues relatives spatio-temporelles des déplacements d'objets mobiles. La seconde concerne, la détection en temps réel de comportements inhabituels à l'aide de technique de fouille de données.

4.1 Les vues relatives spatio-temporelles

Les trajectoires des navires sont généralement modélisées et visualisées de deux façons différentes. Les positions des navires peuvent être affichées au-dessus d'une carte numérique dans un repère absolu. Elles peuvent aussi être représentées de manière relative autour du navire-référent. Le RADAR utilise cette dernière représentation. Ces 2 types représentations complémentaires favorisent l'observation du positionnement des navires en temps réel. Par contre, elles ne permettent pas une compréhension fine de la dynamique du système et des processus liés (changement de cap, de vitesse, manœuvre, etc.). Cette compréhension est d'autant plus nécessaire qu'il est indispensable d'anticiper les possibles situations critiques.

L'objectif de ces travaux est de développer un nouveau modèle générique conceptuel de représentation des trajectoires géographiques pour l'aide à l'analyse de comportements d'objets mobiles. La représentation proposée est égocentrique et relative. Elle est complémentaire de la représentation absolue traditionnelle. Elle doit permettre de mieux comprendre et percevoir les processus dynamiques de déplacement. Elle prend en référence un objet mobile (objet référent) et représente les évolutions temporelles vis-à-vis des autres objets mobiles (objets cibles). Un raisonnement spatio-temporel est associé. Il doit favoriser la

description des processus dans un langage proche du langage naturel et être traduit graphiquement. Finalement, ces travaux génériques (modèle et raisonnement) sont appliqués pour validation aux déplacements des navires.

4.1.1 Modélisation de la représentation relative spatio-temporelle égocentrique

La modélisation relative des déplacements d'objets mobiles vis-à-vis d'un objet référent qui est aussi mobile que nous proposons, est basée sur deux mesures :

- la position relative,
- la vitesse relativisée.

La **position relative** mesure la distance minimale à parcourir pour que les objets se touchent. La géométrie des objets peut être ponctuelle, linéaire ou surfacique. Deux objets se touchent si uniquement leurs frontières extérieures sont en contacts (voir figure 32 (b)). Si leur intérieur sont en contact, ils sont en relation topologique d'intersection (voir figure 32 (a)). Si les objets sont distants alors ils sont en relation topologique de disjonction (voir figure 32 (c)). La position relative est par convention positive quand les objets sont distants, sa mesure est égale à la distance minimale $d_{\min}(A,B)$ entre les deux objets.

$$d_{\min}(A,B) = \min (d_E(a,b) / a \in A \wedge b \in B)$$

Pour deux objets en intersection, cette mesure est négative et elle représente la distance minimale à parcourir pour que les objets se touchent uniquement. Si A est l'objet référent et B l'objet cible la position relative entre A et B notée $p(A,B)$ est définie de la manière suivante :

$p(A,B) = d_{\min}(A,B)$ si A et B disjoint,

$p(A,B) = 0$ si A et B se touche,

$p(A,B) = -\max(d_{\min}(\overline{A \cap B}, \overline{A^-}))$ si A et B s'intersectent.

Pour le troisième cas, la position relative est calculée à l'aide de la distance minimale entre l'intersection de A et B et l'extérieur de A. Des fermetures sont employées afin de s'assurer que cette valeur est atteinte.

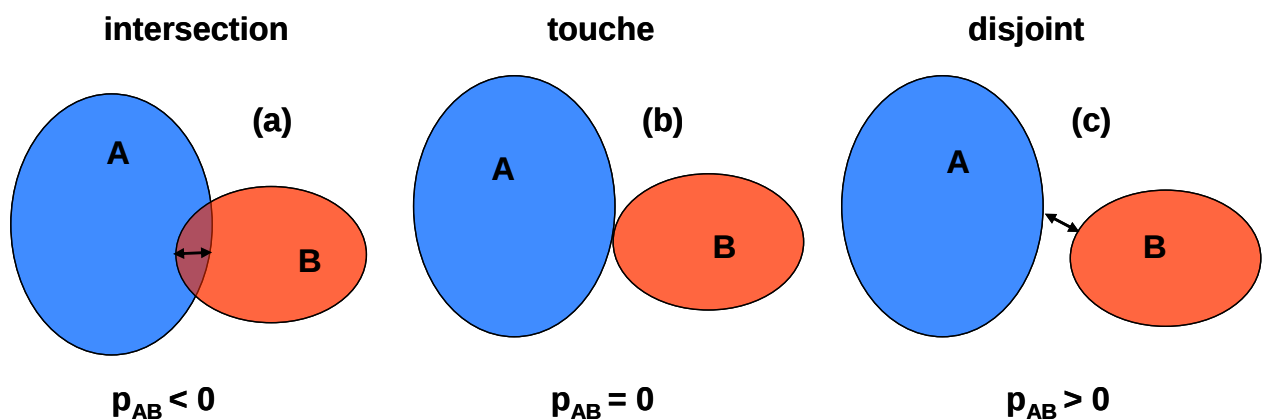


Figure 32: positions relatives en fonction de la relation topologique

Pour les objets mobiles référents ponctuels ou linéaires, la position relative n'est jamais négative. La figure 33 illustre les évolutions temporelles possibles entre deux objets mobiles surfaciques (a). En t_0 , t_1 , t_7 et t_8 les objets sont disjoints et la position relative positive. En t_5 et t_6 les objets se touchent et la position relative est nulle. Cette figure permet aussi de constater que la position relative n'est pas symétrique pour les valeurs négatives en t_2 , t_3 et t_4 (voir $p(A,B)$ de (c) et $p(B,A)$ de (d)). Cette mesure est continue.

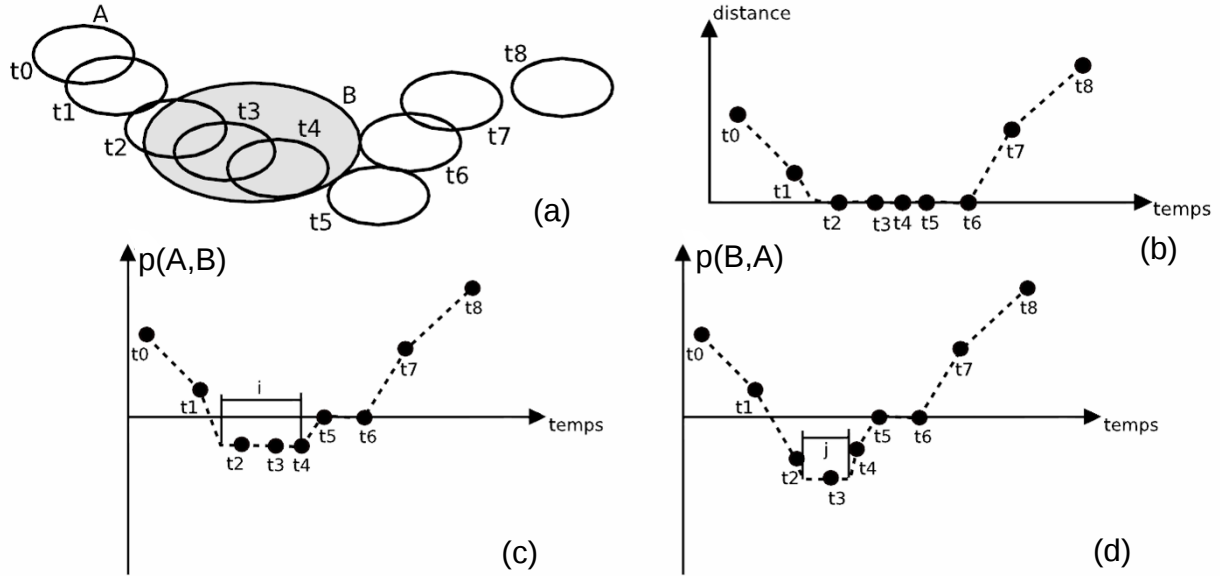


Figure 33: Evaluation temporelle des positions relatives entre deux objets mobiles surfaciques

Le deuxième critère est la **vitesse relativisée**. Cette mesure reflète le différentiel de vitesse entre l'objet référent et l'objet cible et distinguant l'objet le plus rapide. La vitesse relativisée entre les objets mobiles A (référent) et B (cible) est défini par :

$$v(A,B) = (\|\vec{v}_A\| - \|\vec{v}_B\|) \times \|\vec{v}_A - \vec{v}_B\|$$

La vitesse relative classique ($\|\vec{v}_A - \vec{v}_B\|$) n'a pas été retenue, car elle ne distingue pas l'objet le plus rapide. La vitesse relativisée est positive quand l'objet référent est plus rapide et négative sinon ($v(A,B) = -v(B,A)$). Une vitesse relative importante signifie que la vitesse relative est grande et que les normes des vecteurs vitesses sont très différentes. La figure 34 illustre l'évolution temporelle de la vitesse relativisée entre deux objets mobiles (a). L'évolution du vecteur vitesse est représentée en (b). La vitesse relative est fournie en (c). Les parties (d) et (e) illustrent la symétrie de la vitesse relativisée par rapport à l'axe des abscisses.

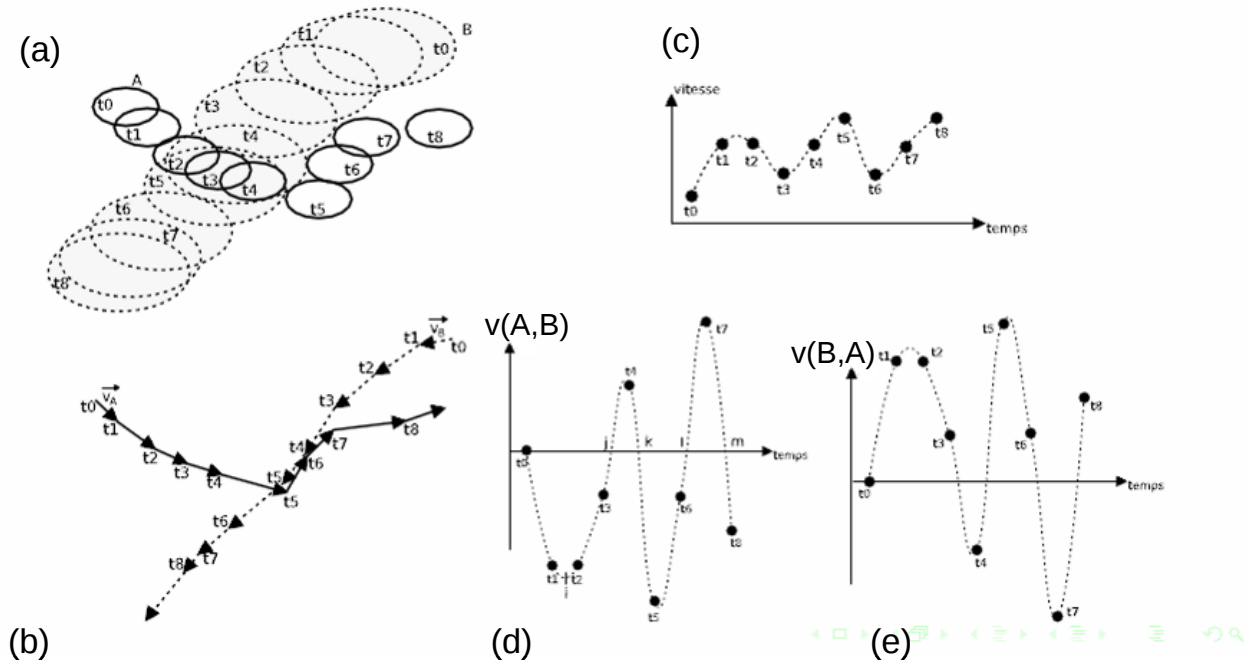


Figure 34: Evolution temporelle vitesse relativisée entre deux objets mobiles surfaciques

Une fois ces deux critères définis, il est possible de les combiner dans un espace à deux dimensions. L'abscisse représente la vitesse relativisée et l'ordonnée la position relative. La figure 35 illustre cette combinaison, les déplacements des deux objets mobiles (a) sont ceux de la figure 34. La composition de la position relative (b) et de la vitesse relativisée (c) donne la représentation (d) appelée **vue relative**. Cette représentation est très riche sémantiquement et elle est très complémentaire de la représentation absolue. Cependant, elle souffre de ne pas être intuitive.

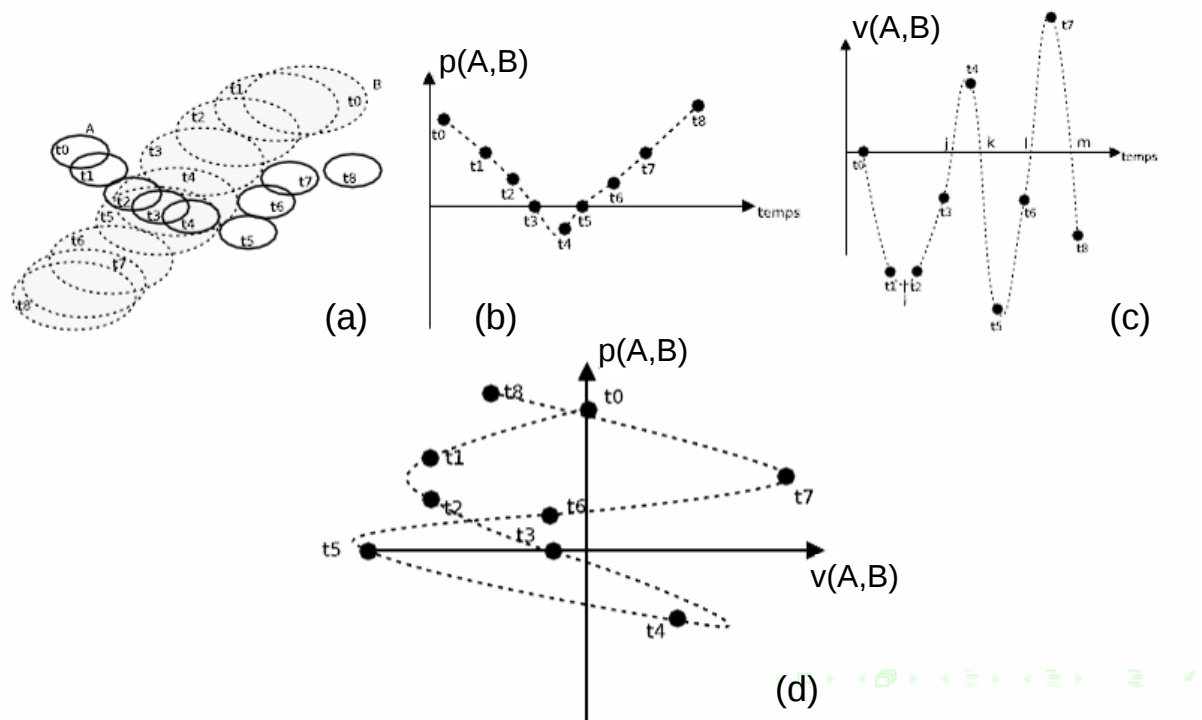


Figure 35 : Combinaison des mesures de position relative et de vitesse relativisée

4.1.2 Raisonnement spatio-temporel égocentrique relatif

Afin de faciliter la compréhension par l'utilisateur de ce modèle spatio-temporel, il est nécessaire de joindre à ce dernier des mécanismes de raisonnement proche du langage naturel qui étendent les relations définies par le raisonnement spatial (voir chapitre 3.2.1).

Ces mécanismes sont basés sur la description qualitative :

- des états relatifs,
- des processus relatifs.

Les **états relatifs** décrivent de manière qualitative d'une part la vitesse relative :

- l'objet B est à l'extérieur de l'objet référent A,
- l'objet B touche A,
- l'objet B intersecte A.

D'autre part, pour la vitesse relative, trois états sont aussi définis :

- A est plus rapide que B,
- A a la même vitesse que B,
- B est plus rapide que A.

Ces états peuvent être combinés. Ainsi, neuf états sont possibles.

Parallèlement, des processus qui permettent de passer d'un état à un autre, sont définis pour la position relative :

- B se déplace de l'extérieur à la frontière de A,
- B se déplace de la frontière à l'extérieur de A,
- B reste à l'extérieur de A,
- B reste à la frontière de A,
- B reste à l'intérieur de A,
- B se déplace de l'intérieur à la frontière de A,
- B se déplace de la frontière à l'intérieur de A,

Il est possible d'affiner ces processus. Par exemple pour "B reste à l'extérieur", ce processus peut être divisé en "B se rapproche de A", "B reste à la même distance de A" et "B s'éloigne de A". De même pour la vitesse relativisée, sept processus sont possibles. Pour chacun de ces processus, ils sont liés à une action de A et/ou de B :

- décélération de A et/ou accélération de B pour atteindre une vitesse relativisée nulle,
- accélération de A et/ou décélération de B à partir d'une vitesse relativisée nulle,
- décélération de A et/ou accélération de B à partir d'une vitesse relativisée nulle,
- accélération de A et/ou décélération de B pour atteindre une vitesse relativisée nulle,
- B reste moins rapide que A,
- A et B reste à la même vitesse,
- B reste plus rapide que A.

A l'aide de ces états et ces processus, les déplacements relatifs peuvent être décrits. Par exemple pour la figure 35 (d), en t_1 , B est à l'extérieur de A et B est plus rapide que A. Entre t_1 et t_2 , B se rapproche de A et leur vitesse reste stable. En t_2 l'état est le même quand t_1 . Entre t_2 et t_3 , B se déplace de l'extérieur à la frontière de A, B est plus rapide, mais la différence se réduit.

4.1.3 Prototype dans le contexte maritime

Ce modèle est employé dans le contexte d'un prototype portant sur la surveillance maritime. Cette activité peut s'exercer à terre dans des centres tels les CROSS ou en mer à bord des navires. Dans le deuxième cas, le navire-référent est implicitement le navire sur lequel le système est embarqué. A terre, les opérateurs de surveillance à partir de la vue absolue choisissent des navires qu'ils veulent suivre plus particulièrement. Ce choix se fait en fonction de la dangerosité de certains navires et de la densité du trafic dans certaines zones. Des tests ont été réalisés à partir de rejeux de données réelles obtenues à l'aide de transpondeurs AIS situés à la Rochelle ou à Brest.

L'objectif du prototype dans ce contexte, est d'aider les opérateurs à identifier les risques de collision. Effectivement, les navires manœuvrant difficilement, il est important de faciliter la détection très en amont des déplacements relatifs dangereux. Il faut par exemple, être capable de distinguer deux navires qui se mettent à couple¹⁵ de deux navires qui risquent l'abordage.

La vue relative étant complémentaire de la vue absolue, le prototype rassemble côte à côte ces deux vues. Les navires suivis étant actuellement vus comme des objets ponctuels, seule la partie haute de la vue relative est présente. De même, la représentation relative n'étant pas intuitive, il est primordial de soigner la sémiologie graphique (Bertin 1973) associée. Ainsi, les zones dangereuses sont identifiées en rouge afin de faire ressortir les risques. Ces zones sont les extrémités droites et gauches proches de l'axe des abscisses. Dans cette partie, les objets sont proches avec une vitesse relativisée très importante en valeur absolue. Inversement, les zones sûres sont en vertes. De même, il est fondamental de mettre en valeur les processus relatifs. Dans cet objectif, l'intérieur et l'extérieur de la courbe de la vue relative, décrivant de manière concomitante la position relative et la vitesse relativisée, sont colorées de deux couleurs. La couleur intérieure de la courbe est fonction de la position relative :

- rouge si elle diminue, la distance se réduisant ce processus peut s'avérer dangereux,
- orange si elle est stable,
- verte si elle augmente.

¹⁵ Se mettre à couple, c'est coller son bateau contre un autre, bord à bord.

De manière analogue, la couleur extérieure est fonction de la vitesse relativisée :

- rouge si elle augmente,
- orange si elle est constante,
- verte si elle tend vers 0.

Ainsi pour deux navires qui se rapprochent en accélérant, situation la plus critique, la courbe est entièrement rouge.

La figure 36 illustre l'interface du prototype développé pour visualiser les déplacements maritimes de navires. Dans cet exemple, des données réelles issues d'un transpondeur AIS, sont visualisées simultanément à l'aide de la représentation absolue (fenêtre de gauche) et la représentation relative (fenêtre de droite). Le petit carré noir en bout de courbe localise la dernière position relative affichée. Les deux navires ont un différentiel de vitesse quasi constante, et se rapprochent avant de s'éloigner. Pour la représentation relative, l'abscisse fournit la vitesse relativisée qui est pratiquement stable. L'ordonnée montre que la distance diminue sans être dangereuse (la courbe ne rentre pas dans les parties ayant un fond rouge). Pour la figure 37, les déplacements relatifs sont plus complexes. Les deux navires partent simultanément de Brest. Une phase dangereuse apparaît quand ils se rapprochent en accélérant. Puis la distance entre les deux navires est stable (partie horizontale). Finalement, ils s'éloignent avec une vitesse relativisée stable.

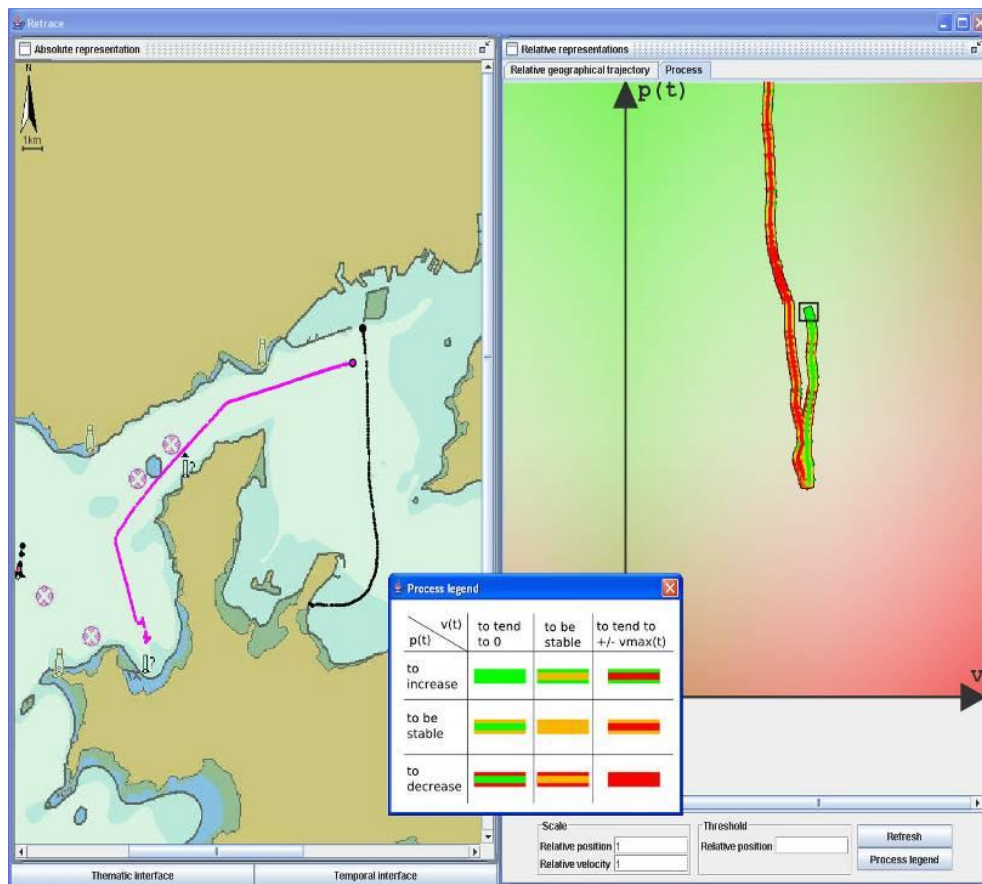


Figure 36 : Copie d'écran du prototype de suivi absolu et relatif des déplacements de navires non dangereuse

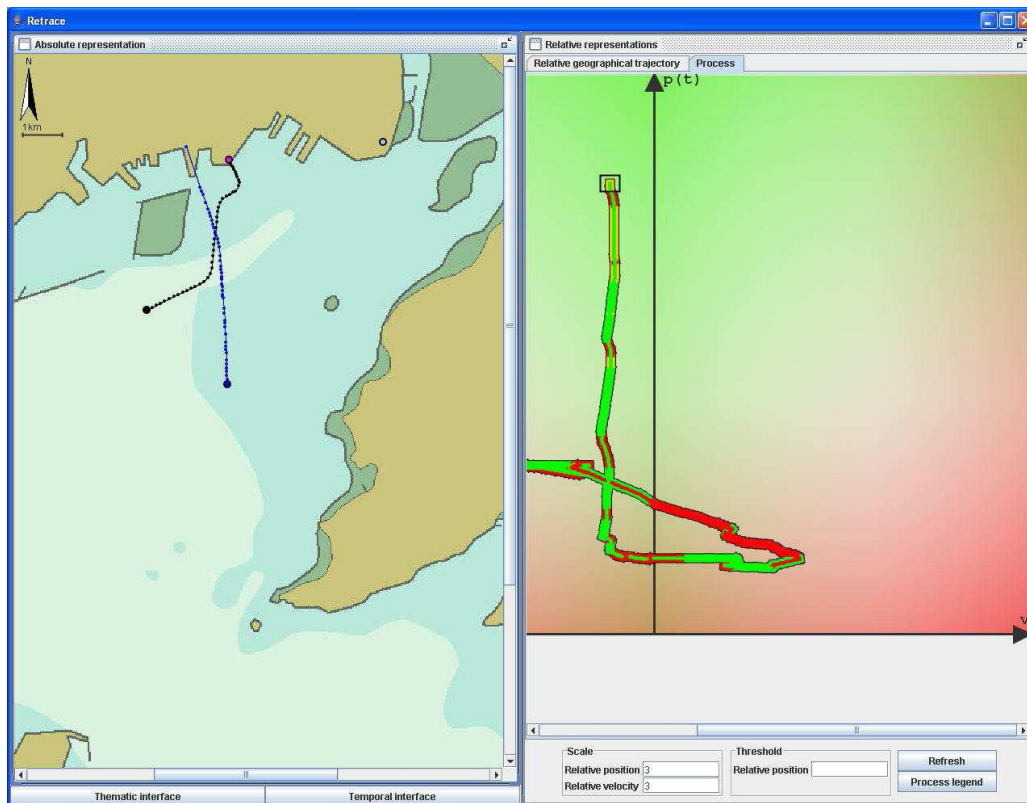


Figure 37 Copie d'écran du prototype de suivi absolu et relatif des déplacements de navires partiellement plus dangereux

4.1.4 Expérimentation du modèle et du prototype

Ce premier prototype a été expérimenté à l'aide de plusieurs cas d'école auprès de marins, d'ingénieurs et de doctorants. Ces premiers tests montrent le grand intérêt de cette représentation relative pour compléter la représentation absolue. Des situations potentiellement dangereuses difficilement détectables à l'aide de la représentation absolue sont mieux appréhendées à l'aide de cette représentation relative. La figure 38 illustre ce propos. Elle regroupe deux cas d'école. Les situations 1 et 2 où les navires ont la même route. L'unique différence se situe au niveau du navire noir qui ralentit pour la situation 1 et qui accélère pour la situation 2. L'intervalle de temps entre 2 points de chaque trajectoire de navire est constant sur la représentation absolue. Ces deux situations sont dangereuses, elles ne sont pas dans la zone verte de la vue relative. Par contre, les utilisateurs en disposant uniquement de vue absolue distinguent très mal que la situation 2 est plus dangereuse. En bénéficiant de la vue relative, la différence de niveau de dangerosité est au contraire bien évaluée.

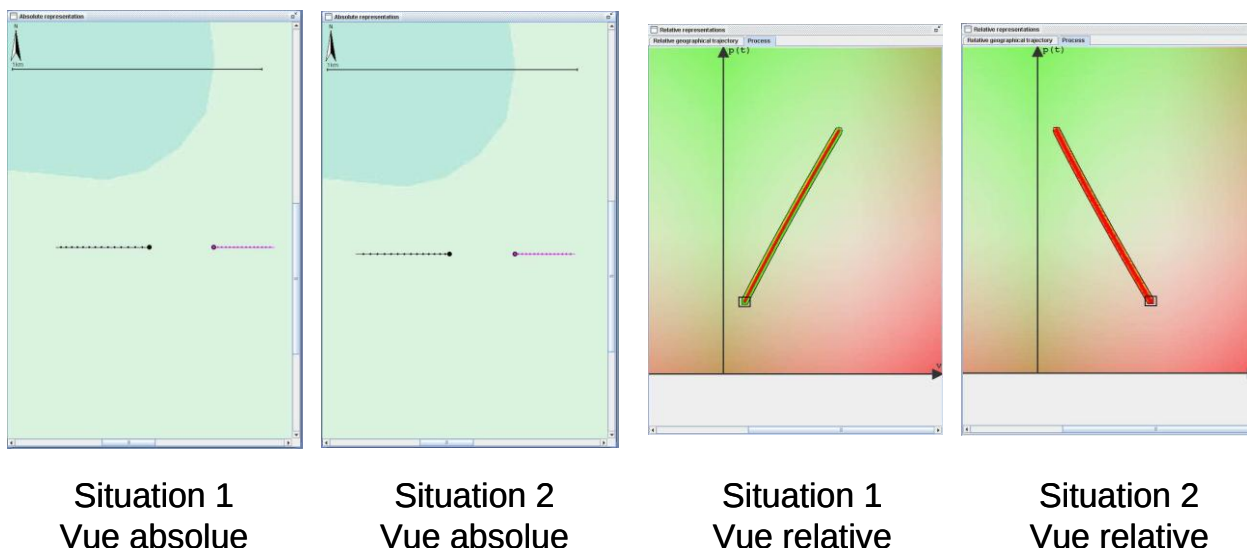


Figure 38 : Deux des cas d'école visualisés à l'aide de la vue absolue et de la vue relative (situation 1 le navire noir ralentit, situation 2 : il accélère)

Ces travaux ont été réalisés dans le cadre de la thèse de Valérie Noyon. Ils ont donné lieu à plusieurs publications (Noyon et al. 2007) (Noyon et al. 2005a) (Noyon et al. 2005b) (Noyon et Devogele 2004). Par ailleurs, ils ont été récompensés par le prix 2008 de la meilleure thèse du GDR SIGMA. Les perspectives majeures pour cette recherche se situent au niveau de l'interface. Une étude plus poussée au niveau de la sémiologie graphique est indispensable. De même pour rendre cette représentation plus intuitive, il est nécessaire de modifier la visualisation. L'objectif est de transposer les zones dangereuses au centre de la figure.

4.2 Analyse de similarité de trajectoires d'objets mobiles

La plate-forme de gestion des trajectoires peut être employée aussi pour déduire les comportements habituels des objets mobiles. La fouille de données (data mining) est employée classiquement pour extraire des connaissances à partir de grandes quantités de données par des méthodes automatiques ou semi-automatiques. Cette fouille dans les historiques volumineux de déplacements, offre des perspectives intéressantes dans le domaine de l'étude du comportement de ces objets. Les analyses proposées dans la littérature concernent la densité ou la distance entre des trajectoires afin de définir des clusters de trajectoires et en déduire des trajectoires aberrantes. Lee et al. (Lee et al. 2008) ont réalisés une bibliographie importante sur la détection de trajectoires inhabituelles en data mining. Il est à noter que les travaux mentionnés prennent en compte essentiellement l'aspect spatial. La partie temporelle de l'analyse est peu ou pas traitée. Les travaux que nous menons s'appuient sur l'axiome suivant : "les objets d'un même type suivant le même itinéraire ont des comportements similaires et suivent une route optimisée". La section suivante présente les concepts que nous allons employer. Puis les méthodes d'extractions des données sont élaborées afin d'obtenir des clusters homogènes de trajectoires. L'analyse de ses clusters permet de définir des patrons spatio-temporels : les routes types. Ces patrons sont des aides précieuses pour détecter en temps réel des situations inhabituelles qui s'éloignent de ces patrons.

4.2.1 Modélisation des itinéraires et des trajectoires

Les données issues des capteurs AIS sont destinées à la gestion de l'anticollision principalement. Elles fournissent essentiellement des positions (latitude, longitude, estampille

temporelle) de navires. Afin d'analyser finement l'historique de ces données, la première étape consiste à modéliser les informations implicites que sont les **itinéraires** et les **trajectoires**. Contrairement aux objets mobiles se déplaçant sur un réseau, les navires, comme les avions, ou les piétons, se déplacent dans un espace ouvert avec des obstacles. Cependant, un réseau implicite existe. Il est formé des destinations des objets mobiles et des lieux de passage obligatoire ou réglementé. Dans notre cas, les ports, les détroits, les dispositifs de séparation de trafic (DST) tel le rail d'Ouessant, les zones d'attente forment les sommets du graphe maritime. Ces zones d'intérêt doivent être délimitées géographiquement. Ils sont reliés par des arcs orientés pour former le **graphe de zones** (G_Z). La figure 39 fournit le graphe de zones de la région de Brest.

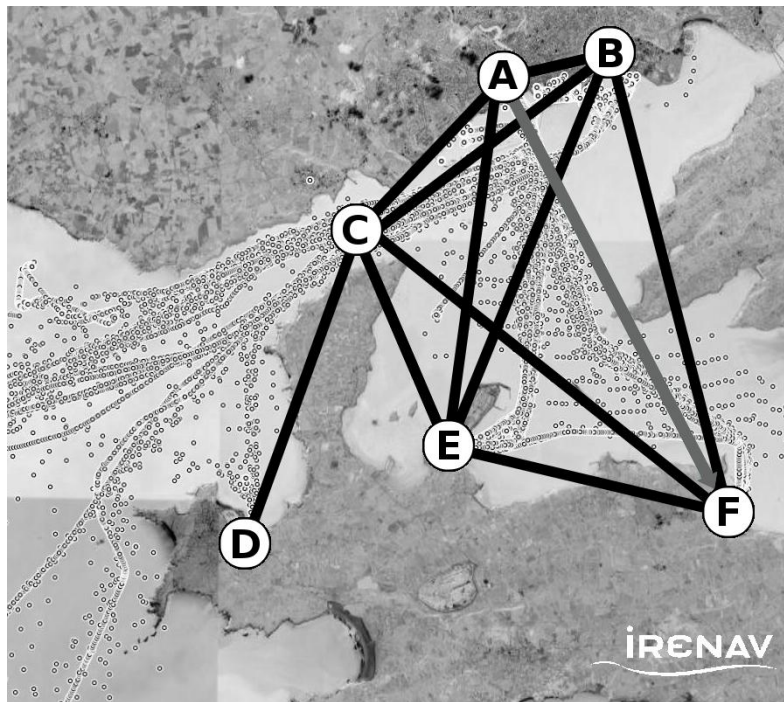


Figure 39 : Graphe de zones dans la région de Brest

Un **itinéraire** peut ainsi être alors défini comme une suite ordonnée de zones reliées deux à deux par des arcs, correspondant à une chaîne du graphe de zones. Un itinéraire étant composé d'au moins un arc, il dispose par conséquent d'une zone de départ Z_D et d'arrivée Z_A .

De même, une **trajectoire** est définie comme une suite ordonnée temporairement de positions d'un même objet mobile. Les trajectoires peuvent suivre des itinéraires. Une trajectoire To est considérée comme suivant un itinéraire I passant par des sommets du graphe G_Z si :

- pour chaque zone de l'itinéraire I , il existe au moins une position Po de la trajectoire To dans cette zone respectant la relation d'ordre temporelle de I ,
- aucune position Po de la trajectoire To n'appartient à une zone du graphe autre que celles de l'itinéraire,
- seule la première position Po_d de la trajectoire appartient à la zone de départ de l'itinéraire Z_D ,
- de même, seule la dernière position Po_a de l'itinéraire appartient à la dernière zone de l'itinéraire Z_A .

Une fois ces définitions posées, il est possible construire la **route type** des navires d'un certain type suivant un même itinéraire. Cette route type doit être déduite statistiquement à partir de l'ensemble des trajectoires répondant aux critères (itinéraire et type de navires

commun). Nous avons choisi de modéliser ces patrons spatio-temporels à l'aide de suites ordonnées de positions relatives. Une position relative est une position qui a pour estampille temporelle une durée depuis un évènement. Dans notre cas, cet évènement est la sortie de la zone de départ. Cette suite est complétée par un couloir spatio-temporel qui délimite dans l'espace géographique et temporel les trajectoires qui suivent de manière normale cet itinéraire.

Le schéma UML des positions, du graphe de zones, des trajectoires, des itinéraires et des routes types est fourni par la figure 40.

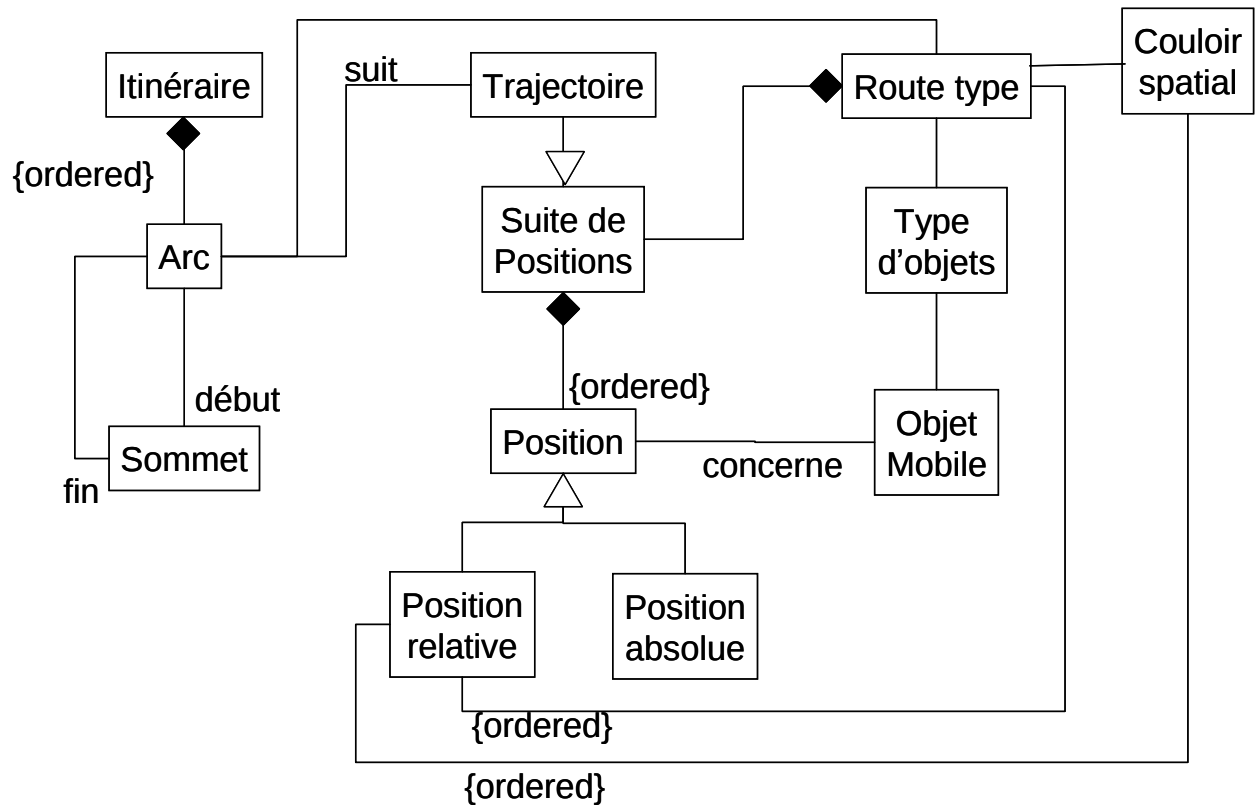


Figure 40 : Schéma UML des trajectoires et itinéraires

4.2.2 Extraction et filtrage des données

L'objectif de cette partie est d'extraire de la base de données spatio-temporelle des positions AIS, les trajectoires d'un même type T d'objets mobiles suivant le même arc A d'un itinéraire I . Cet ensemble est appelé **groupe homogène de trajectoires** d'objets mobiles de même type suivant le même arc d'un itinéraire (GHT_{AIT}). Le premier critère est le type d'objets mobiles. Les objets d'un même groupe font tous partie d'un même type. Le second critère est géographique, seules sont concernées les trajectoires suivant l'arc A . Finalement, le dernier critère employé est temporel, certains des objets mobiles pouvant effectuer ces itinéraires périodiquement. Il est nécessaire de définir pour ces objets, une trajectoire pour chaque aller. Finalement, pour chaque trajectoire il faut s'assurer qu'elle ne passe pas par une autre zone du graphe. L'ensemble des trajectoires ainsi reconstitué compose le GHT_{AIT} à analyser.

Cette sélection a permis d'extraire les 554 trajectoires de navires à passagers suivant l'itinéraire de Brest Arsenal à Lanvéoc École navale sur une période de 18 mois (l'arc AF de la figure 39).

Une fois le GHT_{AIT} extrait de la base de données, il est nécessaire d'effectuer plusieurs filtrages :

- Les trajectoires incluant des ruptures de communication longues où des positions erronées (l'objet mobile ne peut pas être à cette position au vu de ces caractéristiques) sont supprimées. Dans le cadre de notre exemple, parmi les 554 trajectoires, 506 trajectoires ont été conservées après filtrage des trajectoires erronées
- Les trajectoires sont tronquées des parties à l'intérieur des zones du graphe. Afin de supprimer tous les biais liés à l'émission non continue des données, la figure 41 illustre la troncature des 506 trajectoires dans la zone de départ pour notre exemple. Sans cette opération, des trajectoires pouvaient mesurer jusqu'à 200m supplémentaires et induire des biais importants dans les futurs calculs.

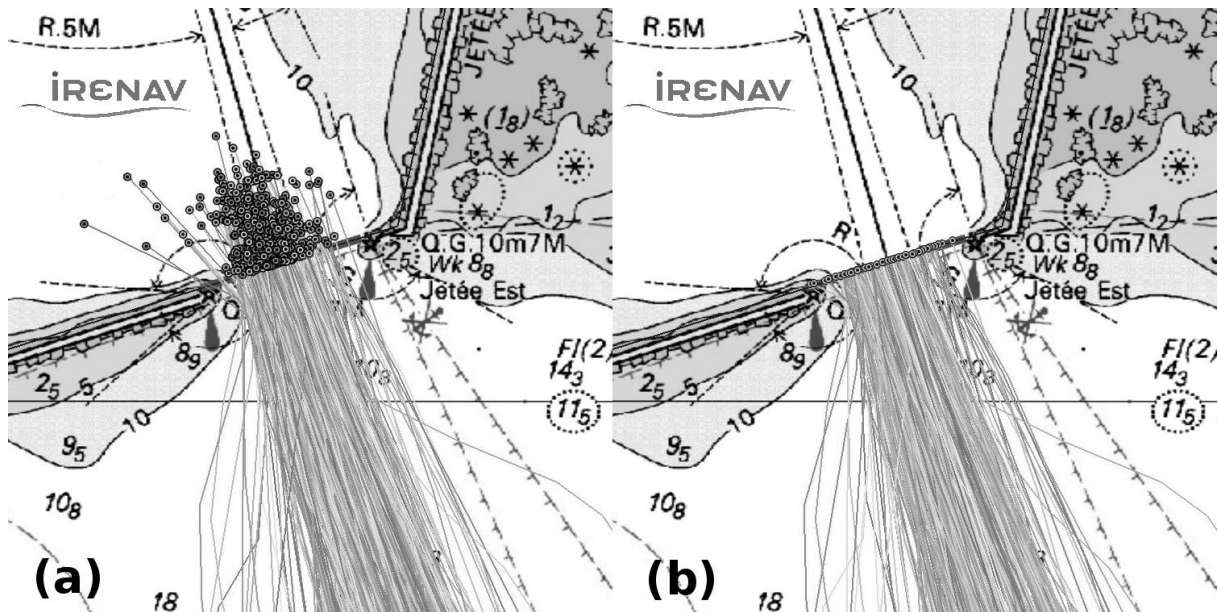


Figure 41 : filtrage des parties des trajectoires dans la zone de départ (a) sans troncature, (b) avec troncature

- Finalement, les trajectoires sont filtrées à l'aide de l'algorithme de Douglas et Peucker spatio-temporel (Bertrand et al. 2007). L'objectif de ce filtrage est de ne conserver que les positions significatives et non les petites variations qui peuvent être des erreurs liées aux mesures GPS. Qui plus est, les futurs temps de calcul sont ainsi fortement réduits. Pour notre exemple, l'application du filtrage de Douglas et Peucker spatio-temporel a permis de réduire le nombre nécessaire de positions de 104 201 à 16 110 positions soit un taux de compression de 84,54 % pour un seuil de 10m (correspondant à la précision d'un GPS).

Une fois les données extraites de la base et filtrées, il est possible de les analyser afin d'en déduire des patrons spatio-temporels.

4.2.3 Analyse des trajectoires suivant un même itinéraire

L'objectif de cette étape est d'obtenir par analyse statistique spatiale et temporelle une route type d'un itinéraire pour les navires d'un même catégorie. Cette route se compose d'une trajectoire médiane et d'un couloir l'encadrant. Une approche médiane a été préférée à une analyse en moyenne. En effet, cette approche est moins sensible aux écarts importants qui ne sont pas symétriques. D'un point de vue temporel, il est en effet plus facile de prendre du retard de manière significative que de l'avance. De même, d'un point de vue spatial, lors des virages, il est plus aisé de s'écarter vers l'extérieur de manière importante que de serrer un obstacle.

Une fois le GHT_{AIT} temporellement normalisé, la route type est construite statistiquement. Pour chaque position de chaque trajectoire du GHT_{AIT} , les positions interpolées des autres trajectoires du groupe sont calculées. Ce calcul de la **route type** génère des nuages de positions interpolées, ordonnés selon le temps normalisé. En sélectionnant les valeurs médianes des latitudes et longitudes des différents nuages de positions, des positions médianes sont obtenues pour chaque temps normalisé. En reliant ces positions médianes, la **trajectoire médiane** (tr_{med}) de l'itinéraire est définie. Cette méthode génère un grand nombre de positions. Afin de ne conserver que les positions significatives, le filtrage de Douglas et Peucker spatio-temporel est à nouveau employé. La figure 42 illustre cette construction statistique pour notre exemple. Sur la partie (a) la trajectoire médiane est représentée en blanc et les 506 trajectoires du GHT_{AIT} sont tracées en noir. La partie (b) détaille un nuage de points (noirs) des trajectoires (blanc) près d'un virage ayant la même estampille temporelle. Le point blanc cerclé de noir correspond à la position médiane du nuage.

Pour commencer, les trajectoires du GHT_{AIT} sont transposées en **temps relatif** et normalisées temporellement. Ainsi, leur première position a pour estampille temporelle 0 et leur dernière $t_{médian}$. La durée $t_{médian}$ est la durée médiane qu'un objet mobile de ce type met pour parcourir cet arc. L'ensemble des estampilles est donc normalisé.

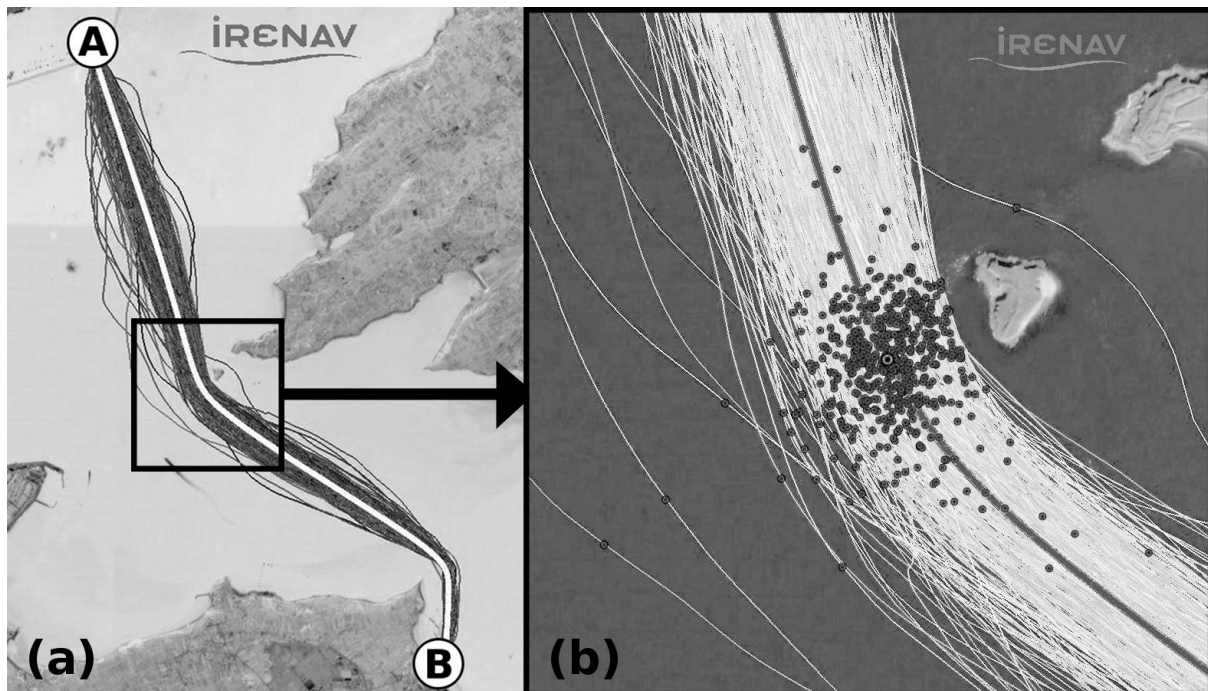


Figure 42: Calcul de la trajectoire médiane et nuages de positions au même temps normalisé

L'observation de la figure 42 permet de constater que les objets mobiles d'un même type effectuant un même itinéraire se déplacent globalement sur une route type. L'espace dans lequel ces objets évoluent étant ouvert, il est normal de constater des écarts spatiaux. Ceux-ci doivent être distingués d'écarts plus importants considérés comme inhabituels. Ces positions inhabituelles sont isolées et à la périphérie du nuage. Afin de séparer les comportements normaux des situations inhabituelles, un couloir spatial est calculé à l'aide d'une analyse statistique des positions des trajectoires du GHT_{AIT} vis-à-vis de la tr_{med} .

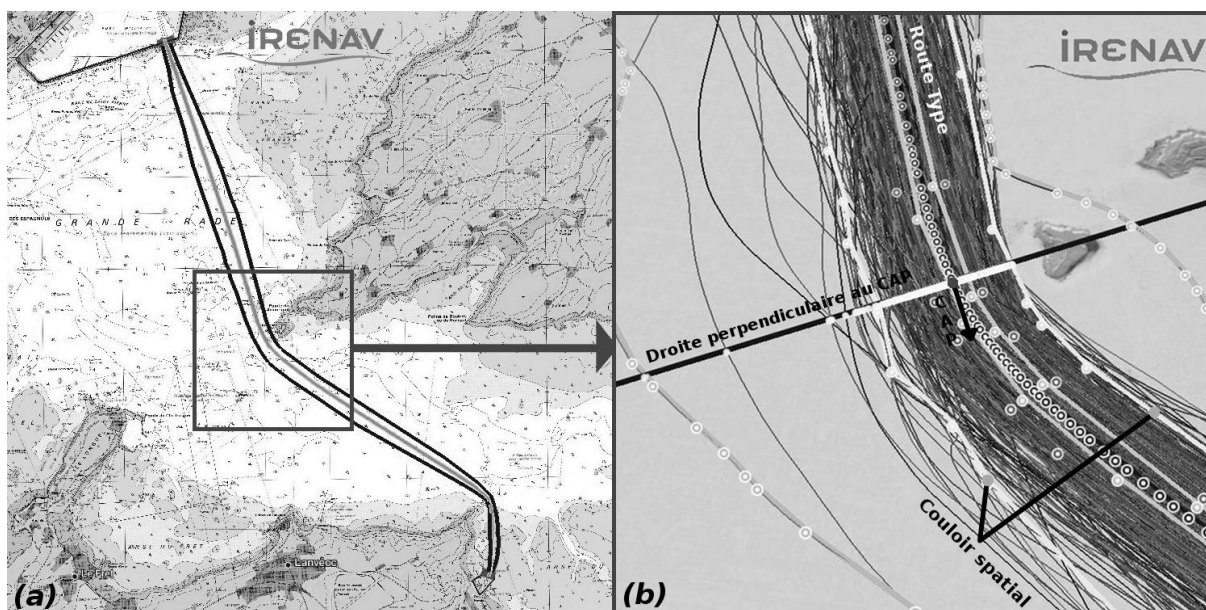


Figure 43: Couloir spatial autour de la trajectoire médiane

Deux analyses statistiques vont être menées pour définir un **couloir spatio-temporel** englobant les positions habituelles autour de tr_{med} . La première étude est géographique. Elle s'appuie sur la droite perpendiculaire au cap (DPC) de chaque position de tr_{med} . Pour chaque trajectoire du GHT_{AIT} , le point d'intersection avec la DPC est calculé, classé en deux parties (à droite ou à gauche) et ordonné par ordre de distance à la position de tr_{med} . Pour chaque partie, le neuvième décile¹⁶ de chaque côté de la route type est sélectionné. En reliant ces positions ordonnées, le couloir spatial de navigation est obtenu. La figure 43 illustre le calcul de ce couloir pour notre exemple.

Il faut noter que la distance entre la route type et le couloir de droite et de gauche n'est pas identique. Effectivement, dans un espace ouvert, il est plus facile pour un objet mobile de dévier vers l'extérieur lorsqu'il tourne. De même, la largeur de ce couloir fournit une indication concernant l'écartement des trajectoires vis-à-vis de la tr_{med} . Dans notre exemple, comme le montre la figure 43, cette largeur est plus étroite au départ, à l'arrivée et dans les virages. Par contre, dans les lignes droites, celle-ci a tendance à augmenter.

La deuxième étude temporelle va délimiter les avances ou les retards habituels des trajectoires dans le couloir spatial. De manière similaire au couloir spatial, il est possible de calculer un **couloir temporel**. Pour cela, les positions d'intersection avec les DPC à l'intérieur du couloir spatial sont prises en compte. Les positions à l'extérieur ne sont pas considérées. En effet, il n'est pas possible de dire si elle est en avance ou en retard. En sortant de la route, un objet peut couper un virage ou éviter un autre objet et prendre du retard. Les positions à l'intérieur sont classées en deux parties (en avance, en retard) et ordonnées en fonction du temps relatif (durée depuis le départ). De même, les positions en avance et en retard au neuvième décile sont sélectionnées. Un retard et une avance acceptable sont ainsi obtenus. Ils forment les bornes temporelles de notre couloir, pour chaque position de la trajectoire médiane. La tr_{med} et le couloir spatio-temporel obtenu forment la route type du GHT_{AIT} .

¹⁶ Un décile est en statistique descriptive, chacune des 9 valeurs qui divisent un jeu de données, triées selon une relation d'ordre, en 10 parts égales, de sorte que chaque partie représente 1/10 de l'échantillon de population.

Il faut noter que les écarts temporels entre une position de la route type et les bornes associées du couloir temporel ne sont pas identiques. Effectivement, il est plus facile pour un objet mobile de prendre du retard que de l'avance. Graphiquement, la combinaison du couloir spatial et temporel génère une suite de quadrilatères irréguliers dans un espace 3D dont l'axe z représente le temps. La liaison de ces quadrilatères génère un couloir spatio-temporel représenté pour notre cas d'étude sur la figure 44 par un polyèdre allongé de couleur foncée.

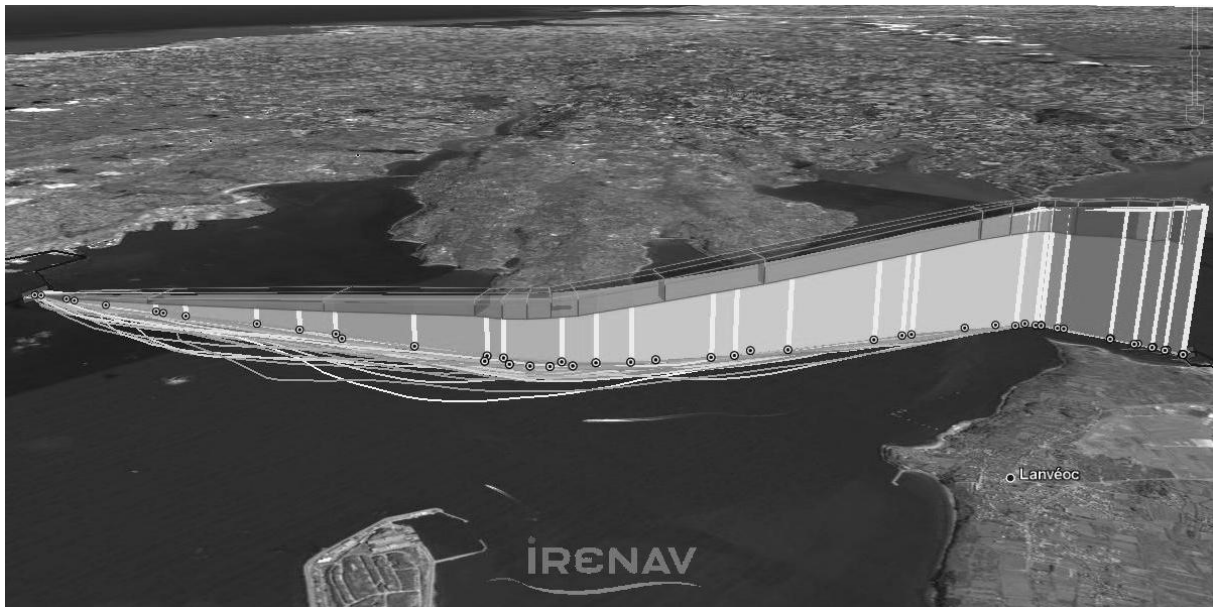


Figure 44: Couloir spatio-temporel

Le temps de calcul de ces patrons spatio-temporels est d'environ 20 secondes à 1 minute pour un PC performant pour chaque GHT_{AIT} de notre base de données. Il est donc préférable de les conserver. Dans notre cas, ils sont stockés dans une base de connaissances associée à la base de données des déplacements. Ces patrons peuvent être recalculés régulièrement (tous les mois par exemple) afin d'affiner leur précision spatio-temporelle.

4.2.4 Détections et qualification des situations inhabituelles

En s'appuyant sur cette base de connaissances, il est alors possible de détecter en temps différé ou en temps réel (temps réel souple) des situations inhabituelles et de qualifier le comportement d'un objet mobile.

Pour une nouvelle position reçue, ce processus de détection des situations inhabituelles suppose préalablement d'associer cette position à un des patrons spatio-temporels via la trajectoire associée à l'objet mobile lié à cette position. Dans cet objectif, sa trajectoire doit être appariée avec une des trajectoires médianes de la base de connaissances. Cet **appariement** s'effectue en fonction :

- de la position actuelle de l'objet mobile,
- du type de l'objet mobile,
- de l'historique de sa trajectoire,
- de l'éventuelle information concernant sa destination prévue.

Cet appariement est plus ou moins précis. Si les tr_{med} sont proches et la destination inconnue, l'appariement est peu fiable. Par contre, si les destinations des objets mobiles sont connues ou les tr_{med} éloignées alors l'appariement est fiable.

Si la destination est connue, à l'aide du graphe de zone, il est possible de déterminer la tr_{med} à apparier à l'aide de la dernière zone traversée par la trajectoire. Pour notre exemple, si

- la position est à l'extérieur de la rade de Brest,
- la dernière zone traversée est le port de Camaret (zone D du graphe de la figure 39),
- la destination est Lanvéoc Poulmic / Ecole navale (zone F),

il est alors possible d'apparier la trajectoire actuelle de l'objet mobile à l'arc D-C. En effet pour aller en zone F, il faut passer par la zone C (goulet de Brest).

Si la destination est inconnue, il faut alors apparier la trajectoire actuelle avec un des arcs partant de la dernière zone du graphe traversée : z. Une **distance linéaire partielle** est appropriée à cette tâche (voir page 27). L'objectif est de mesurer les distances linéaires entre la trajectoire (tr_a) partant de z jusqu'à la position actuelle et les trajectoires médiane (tr_{med}) associées aux arcs partant de z. Cette distance peut bénéficier de l'estampille temporelle relative. Effectivement, cette dernière permet de fixer les couples de positions homologues. Les positions relatives de tr_a sont apparées avec les positions des tr_{med} aux mêmes instants. De la sorte, il est possible de tenir compte des accélérations et des ralentissements des objets mobiles. La formule de la distance maximale linéaire entre la trajectoire tr_a et une trajectoire médiane tr_{med} est donc :

$$d_{tp}(tr_a, tr_{med}) = \max_{0 \leq t \leq d} \|tr_a(t) - tr_{med}(t)\|$$

Cette distance est discrète, seules les positions relatives stockées de tr_a et des tr_{med} sont prises en compte. Pour chaque position entre l'instant de départ de la dernière zone ($t = 0$) et la durée depuis le départ ($t = d$), la distance avec la position interpolée de l'autre ligne est calculée. Il faut noter que les positions associées ne sont plus homologues selon des critères visuels, mais prennent en compte uniquement l'aspect temporel de la trajectoire.

La tr_{med} , la plus proche et à une distance linéaire partielle inférieure à un seuil défini empiriquement, est apparée avec la trajectoire. Une fois la tr_{med} sélectionnée, la dernière position reçue de l'objet mobile peut être qualifiée à l'aide du couloir spatio-temporel. L'information qualitative associée peut-être :

- sur la route à l'heure (situation normale),
- sur la route en retard,
- sur la route en avance,
- à droite de la route,
- à gauche de la route.

Ainsi, pour la position P de la figure 45, le couloir spatio-temporel de la route type à ce temps relatif est employé. La position P est donc à droite de la route.

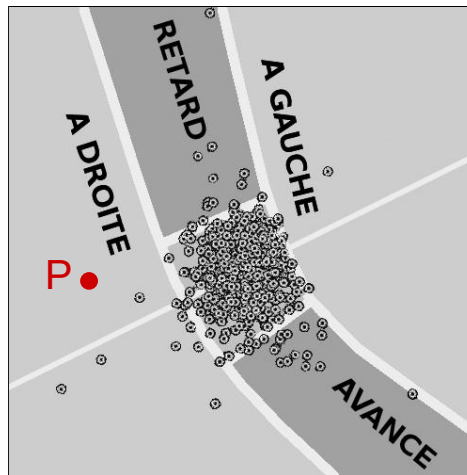


Figure 45 : Partition de l'espace à un instant relatif. Les positions ayant servi à l'analyse spatio-temporelle sont cerclées de blanc.

En ce qui concerne la navigation maritime, les opérateurs de surveillance du trafic peuvent avoir un grand nombre de navires à suivre. La transmission des informations qualitatives concernant la trajectoire du navire leur permet de se focaliser sur un petit nombre de navires. De plus, ce module d'analyse est capable, une fois la trajectoire appariée, d'indiquer la destination du navire ou du prochain nœud du graphe ainsi que le temps nécessaire pour arriver à destination par comparaison avec les données de la route type.

4.2.5 Résumé des processus de fouille de données et d'analyse temps réel des trajectoires

Les travaux portant sur la fouille de données spatio-temporelle et l'analyse temps réel des trajectoires ont été initiés début 2008 par des travaux d'élèves (Jonglez et Songy 2008) et se poursuivent dans le cadre de la thèse de M. Etienne (Etienne et *al.* 2008) et (Etienne et *al.* 2009).

La figure 46 résume les différentes étapes génériques nécessaires à la détection des positions inhabituelles. Dans un premier temps, les données issues de capteurs des positions sont acquises en temps réel (phase 1). Dans notre cas, c'est le système AIS qui est employé. Ces données sont gérées dans la base de données spatio-temporelle (phase 2). Pour chaque arc du graphe et chaque type d'objets mobiles, les positions entre la zone de départ et d'arrivée de cet arc sont extraites. Elles sont alors filtrées, transformées en position relative afin de former le groupe de trajectoires homogène (phase 3). Suite à cette phase pour chaque GHT_{AIT} la trajectoire médiane et le couloir spatial puis temporel sont calculés (phase 4). Ces patrons spatio-temporels sont stockés dans la base de connaissances (phase 5). Cette base est employée par la suite pour qualifier en temps réel les positions reçues et ainsi détecter les situations inhabituelles (phase 6).

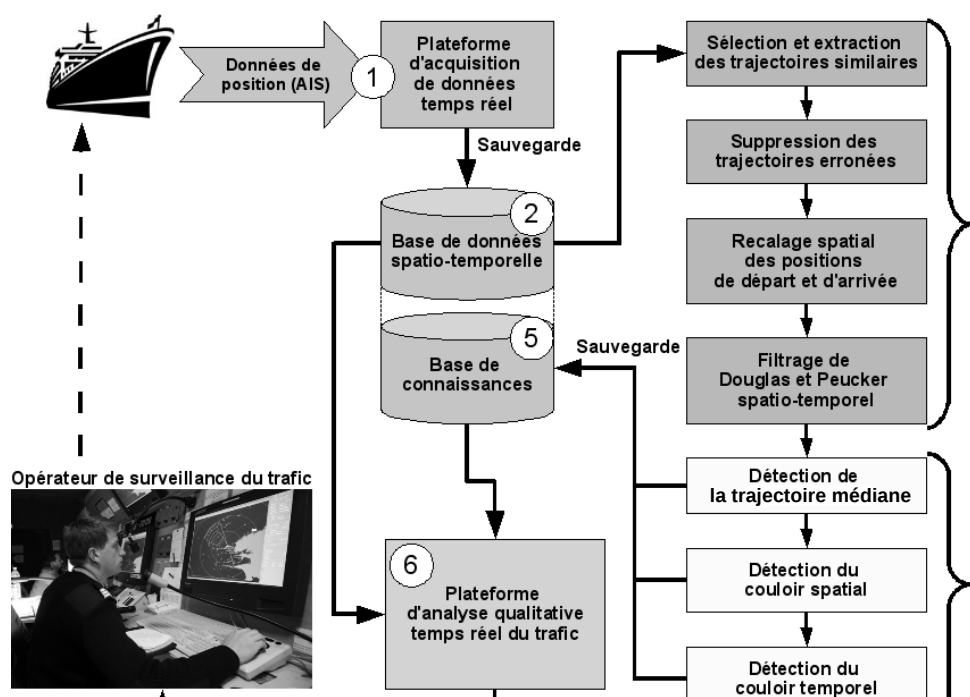


Figure 46 : Les différentes étapes de la détection de positions inhabituelles

Ces travaux s'intéressent à la problématique particulière du déplacement d'objets mobiles dans un espace ouvert. L'originalité de cette démarche d'analyse de similitude réside dans l'introduction de métaconnaissances (GHT_{AIT} , trajectoire médiane, couloirs spatio-temporels) permettant de qualifier les comportements d'objets mobiles suivant un itinéraire dans un espace ouvert. Elle a été appliquée plus particulièrement au domaine maritime comme exposé dans notre exemple basé sur un jeu conséquent de données réelles. Cependant, la méthodologie est générique. Un prototype fonctionnel a été développé pour les cinq premières phases. La partie traitant de l'appariement en temps réel d'une trajectoire à un itinéraire a été modélisée, mais n'a cependant pas encore été implémentée dans ce prototype. Il est désormais important de valider cette dernière phase et de mettre à la disposition d'opérateurs de surveillance du trafic maritime l'ensemble du prototype.

Les trajectoires sont actuellement découpées en s'appuyant sur le graphe de zones. Parallèlement, il est intéressant de scinder les trajectoires non plus selon leur itinéraire, mais selon des tronçons homogènes (ligne droite avec vitesse constante, virage pour changer de cap, etc.) afin de pouvoir décrire qualitativement les trajectoires d'objet mobile ou médiane et dans un deuxième temps de comparer qualitativement des trajectoires. L'objectif est donc de décrire les trajectoires à l'aide de variable linguistique. Une description possible est pour aller de la zone A à la zone B pour un objet de tel type, il faut aller tout droit avec une vitesse d'environ 20 km/h et un cap au sud-est, puis au bout de 10 minutes changer de cap pour aller à l'est et aller tout droit pendant 2 minutes.

Finalement, une deuxième perspective à cette analyse est de prendre en compte les spécificités dues à l'environnement dans lequel il évolue (les saisons et les marées dans le cadre de notre exemple). Il est intéressant d'appliquer d'autres critères de sélection du GHT_{AIT} ainsi qu'enrichir cette analyse en la couplant avec des travaux portant sur la prise en compte de l'environnement (Richard 2007), des statistiques d'accidents (Degré 2007) et des comportements illicites (Morel et al. 2008).

5 Conclusion

5.1 *Résumé des recherches effectuées*

Les travaux de recherche en informatique qui sont présentés portent sur les Systèmes d'Information Géographique. L'intégration dans ces systèmes de grands volumes de données liés aux modèles 3D très précis ou aux objets mobiles engendre de nouveaux verrous scientifiques et technologiques. Parallèlement, cette intégration ouvre des perspectives prometteuses en termes de compréhension du milieu naturel et des déplacements. Les contributions apportées par ces travaux ont trait les domaines informatiques de la géométrie algorithmique, de l'intelligence artificielle et des bases de données spatiales. En ce qui concerne la géométrie algorithmique, de nouvelles distances linéaires sont dérivées. Ces dernières adaptent la notion de distance linéaire maximale aux lignes fermées et aux parties homologues de lignes d'emprises différentes. Elles sont à la base de techniques innovantes d'intégration de modèles numériques de terrain maritime et terrestre. Des processus d'appariement de lignes caractéristiques, de fusion de ces lignes et de déformations élastiques des MNT sont modélisés et mis en place.

Mes contributions dans le domaine de l'intelligence artificielle appliquée aux SIG touchent les systèmes multi-agents et les algorithmes génétiques. Afin de simuler les déplacements maritimes de navires, un modèle à base d'un système multi-agents reprenant le paradigme Agent-Groupe-Rôle est défini. La réglementation maritime est modélisée, dans un premier temps, sous la forme d'un ensemble de règles d'inférence. Puis, nous avons intégré un modèle cognitif plus évolué qui reproduit la prise de décision naturelle d'experts. Il est fondé sur une base de patrons reprenant les réglementations formelles et les connaissances informelles. Dans ces travaux, les notions propres aux Systèmes d'information géographique ou temporelle comme les proximités, les distances et les vitesses sont incluses. Des mécanismes de logique floue sont employés, afin de déduire les données qualitatives indispensables à la modélisation du raisonnement cognitif. Les notions spécifiques aux SIG (proximités, voisinages) sont aussi intégrées à travers de nouveaux types de codage, de croisement et d'évaluation des solutions dans le domaine des algorithmes génétiques.

Finalement, en ce qui concerne les bases de données spatiales, les apports consistent d'une part à définir de nouvelles vues spatio-temporelles qui offrent des perspectives prometteuses en termes de mise en évidence de comportements spatio-temporels. Ces vues relatives sont complémentaires de la vue absolue. L'évolution des distances et des vitesses difficilement appréhendables à l'aide de la représentation absolue sont mises en évidence. Parallèlement, les déplacements relatifs des objets mobiles peuvent être facilement décrits de manière qualitative. D'autre part, des techniques de fouille de données spatio-temporelles sont mises en place sur les bases de données des trajectoires. Pour des trajectoires d'objets mobiles de même catégorie et ayant un itinéraire commun, des routes types sont définies. Chacun de ces patrons spatio-temporels regroupe une trajectoire médiane et un couloir spatio-temporel. Ils sont issus d'un processus d'extraction, de filtrage et d'analyse statistique des groupes de trajectoires homogènes. Ces métaconnaissances peuvent être par la suite employées pour faciliter la détection en temps réel des comportements inhabituels des objets mobiles. Finalement, Plusieurs de ces travaux portant sur les objets mobiles reprennent des avancées issues du raisonnement spatial.

La thématique du domaine maritime est une composante forte de mes recherches. La manipulation des MNT (maritimes ou côtiers), la modélisation et l'analyse des trajectoires de navires, sont les deux principaux sujets d'étude. Néanmoins, les travaux qui sont présentés

sont génériques dans le sens où ils peuvent être appliqués à d'autres types d'objets mobiles, à d'autres reliefs numériques, ou à l'appariement d'autres polygones. Le tableau 9 synthétise les domaines informatiques traités en fonction des thématiques abordées.

thématiques	domaines informatiques	Géométrie Algorithmique			Intelligence Artificielle appliquée aux SIG					Bases de Données Spatio-temporelles	
		distances linéaires	déformation élastique	distances spatio-temporelles	algorithme génétique	Système multi-agents	logique floue	raisonnement spatial	raisonnement à base de patrons	vue relative	fouille de données
Données 3D : MNT	Généralisation de MNT										
	Segmentation de MNT										
	Appariement de lignes caractéristiques										
	Fusion de lignes caractéristiques										
	Fusion de MNT										
Données spatio-temporelles : objets mobiles	Modélisation des déplacements relatifs										
	Visualisation relative										
	Filtrage des trajectoires										
	Définition de routes types										
	détection temps réel de trajectoires inhabituelles										
	Simulation des déplacements d'objets mobiles										
	Simulation cognitive des déplacement d'objets mobiles										

Tableau 9: Synthèse des activités de recherche par domaines informatiques et thématiques de recherche

La plupart de ces travaux ont été réalisés dans le cadre de thèses. Deux thèses ont été achevées (M. Fournier¹⁷ et M^{elle} Noyon¹⁷) et deux thèses doivent être soutenues début 2010 (M^{elle} Mascaret¹⁸ et M. Le Pors¹⁹) et une est en cours (M. Etienne¹⁷). La figure 47 synthétise ces travaux de recherche. Des projets d'élèves de l'Ecole navale sont régulièrement définis en lien avec ces recherches afin de débroussailler un sujet ou d'approfondir un point connexe à un sujet de recherche. Des travaux d'étudiants d'autres écoles sont aussi associés à ces recherches (M. Riffault, M. Dommy, M. Brosset, M. Brocarel, M. Le Lamer).

¹⁷ Financement Marine Nationale

¹⁸ Financement Brest Métropole Océane

¹⁹ Financement Région Bretagne

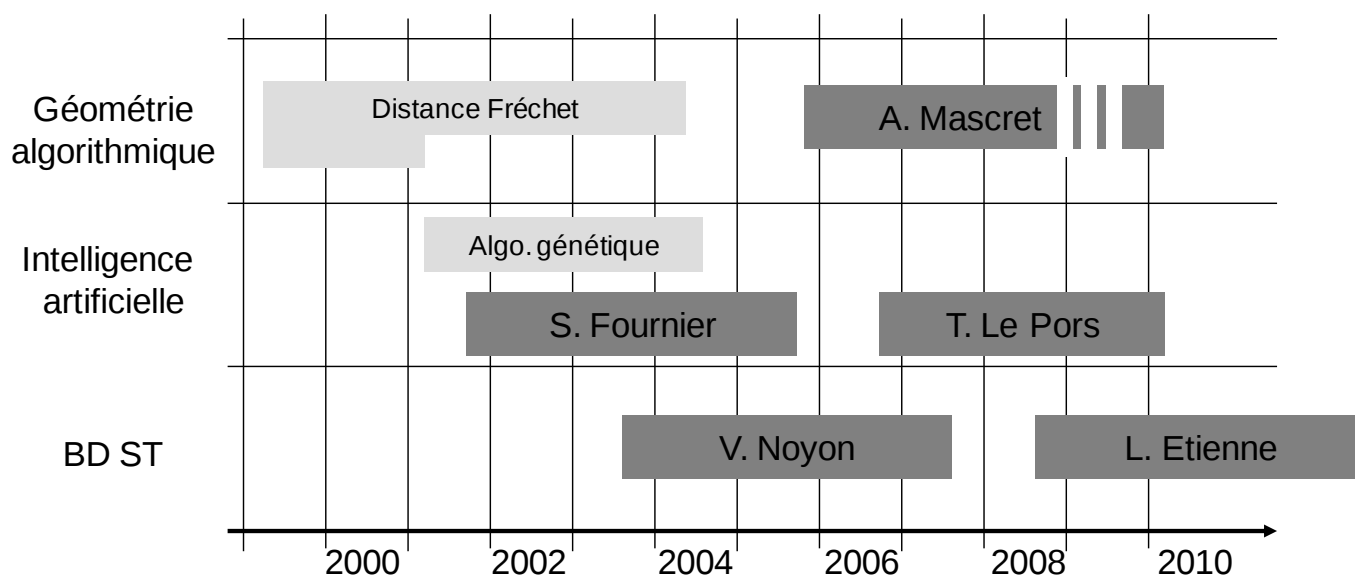


Figure 47 : Résumé chronologique des travaux de recherches

Parallèlement de collaborations sont recherchées dans le cadre du Groupe de recherche en géomatique (GDR CASSINI – SIGMA - MAGIS) ainsi, des études communes ont été menées avec :

- le L3i de l'université de La Rochelle (M. Bouju et Bertrand),
- le laboratoire COGIT de l'IGN (M. Mustière),
- l'équipe Géomer de l'IUEM (M. Cuq, Le Berre, Hénaff).

De même, des coopérations ont été mise en place avec des sociétés ou des établissements publics (SHOM, Ifremer, Geomod, Magic Instinct Software). Finalement, des collaborations de recherches ont été nouées avec des universités étrangères :

- Université de Glamorgan, Pays de Galles²⁰,
- Université Laval, Québec,
- Université technique de Nottingham.

5.2 Projets de recherche à 4 ans

Les travaux de recherche en système d'information géographique présentés portent donc sur un spectre large de compétences informatiques (IA appliquée aux SIG, BD spatio-temporelles, géométrie algorithmique). Plusieurs améliorations ont été évoquées à la fin de chaque chapitre. Néanmoins, deux thématiques maritimes sont fortement associées : les modèles numériques de terrain et les objets mobiles. Les projets de recherche pour ces 4 prochaines années se focalisent sur les objets mobiles.

²⁰ Financement EGIDE

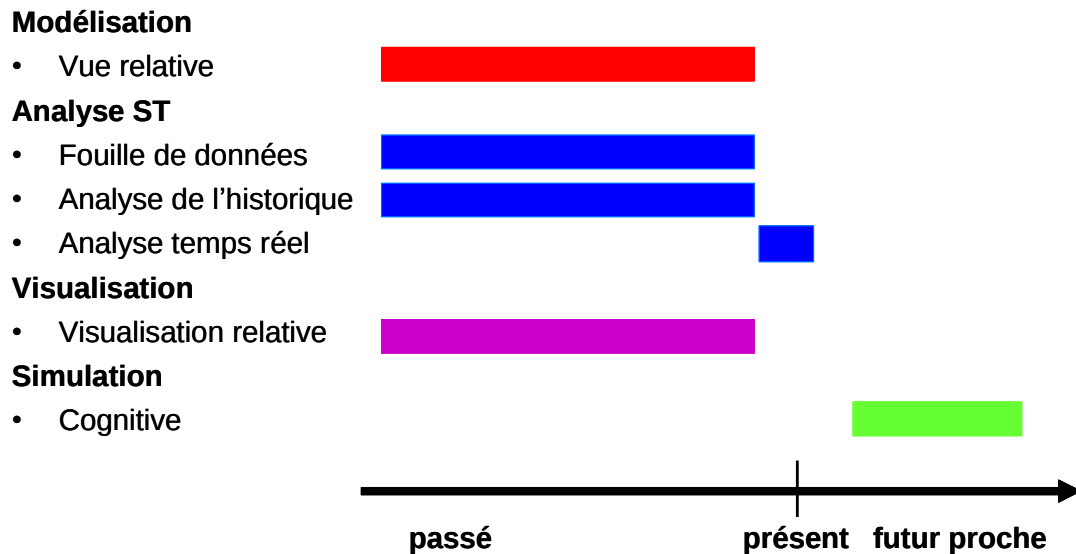


Figure 48 : Classification des différentes phases de mes travaux de recherche (modélisation, analyse, visualisation, simulation) en fonction du type de données traitées (passées, présentes ou futures).

L'intérêt principal est de mettre à profit les premiers résultats complémentaires en faisant converger plusieurs de ces travaux. Ces derniers ont été étudiés jusqu'à maintenant relativement séparément. Ils se situent à différentes phases (modélisation, analyse, visualisation, simulation) et traitent des données passées, présentes ou futures. La figure 48 résume le positionnement des actions entreprises. Ce rapprochement doit permettre de renforcer chacune des phases et de combiner des données concernant des périodes différentes. L'objectif global est de mieux comprendre le comportement des objets mobiles. La figure 49 présente les apports liés à l'intégration des recherches déjà réalisées.

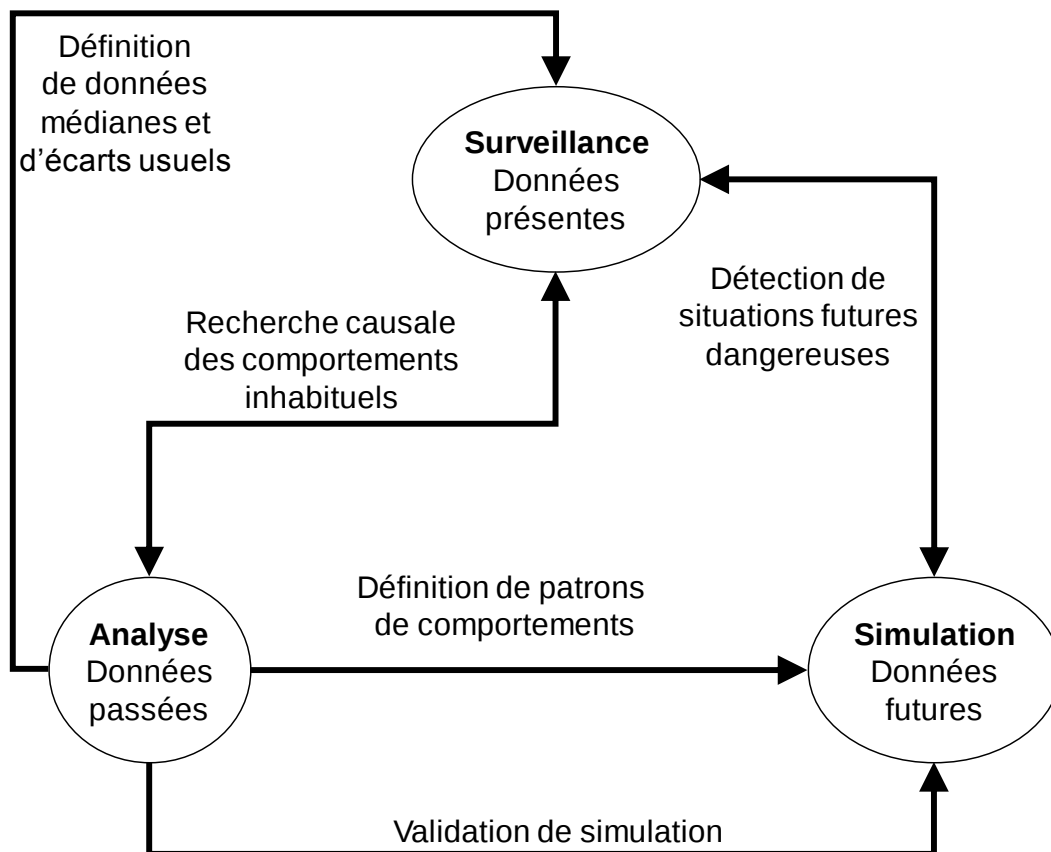


Figure 49 : Interactions futures entre les travaux de recherche réalisés

Le premier projet de recherche se situe au niveau de l'analyse spatio-temporelle et combine l'analyse des données présentes en temps réel et l'analyse de l'historique dans un passé proche. Les positions sont actuellement qualifiées en temps réel de manière qualitative à l'aide de cinq états (sur la route à l'heure, en retard, en avance, à droite et à gauche de la route). Pour les quatre états inhabituels, il doit être possible de justifier la qualification à l'aide des informations passées. Ce processus permet de séparer les cas relevant d'un comportement normal de l'objet mobile, des cas relevant d'un comportement dangereux ou inexpliqué. Ces objets devant être suivis attentivement par les opérateurs de surveillance. Ce projet consiste donc à rechercher la cause (recherche causale) passée d'une situation présente inhabituelle en fonction des événements et du comportement passé de l'objet mobile.

Par exemple, pour le qualificatif en retard, cet état peut être dû par exemple à :

- une vitesse plus faible que les objets du même type,
- un trafic important,
- un évitement de collision passé.

Ces trois cas sont sans danger. Par contre, cela peut être aussi le fait de comportement erratique (changement de vitesse ou de direction sans raison apparente). Une procédure complexe doit pouvoir être formalisée et mise en œuvre pour :

- extraire le comportement des objets mobiles des bases de données historiques (métaconnaissances),
- établir le comportement d'objet mobile en temps réel,
- expliquer la plupart des états inhabituels.

Pour illustrer cette approche dans le domaine maritime, un exemple de position qualifiée "en retard" est détaillé. Pour identifier la cause du retard (le comportement passé qui justifie le retard présent), plusieurs comportements sont recherchés. Par exemple, une sortie de la route

due à l'évitement antérieur d'un autre objet mobile est l'une des possibilités qui doivent être cherchées à l'aide des positions passées. Ainsi, si dans la base de données des trajectoires, une sortie du couloir de la route type et la présence d'un navire en route de collision à cette période sont trouvées simultanément, le comportement évitement de collision est fort probable. La figure 50 illustre cette recherche de manœuvre d'anticollision. La position P de l'objet O est qualifiée, en temps réel, comme " en retard ". Or, l'extraction de la trajectoire de O depuis la zone de départ, montre que O est sortie de la route à la position P'. O à cet instant, était en route de collision avec deux navires. Le retard de O est donc très probablement dû à une manœuvre d'anticollision.

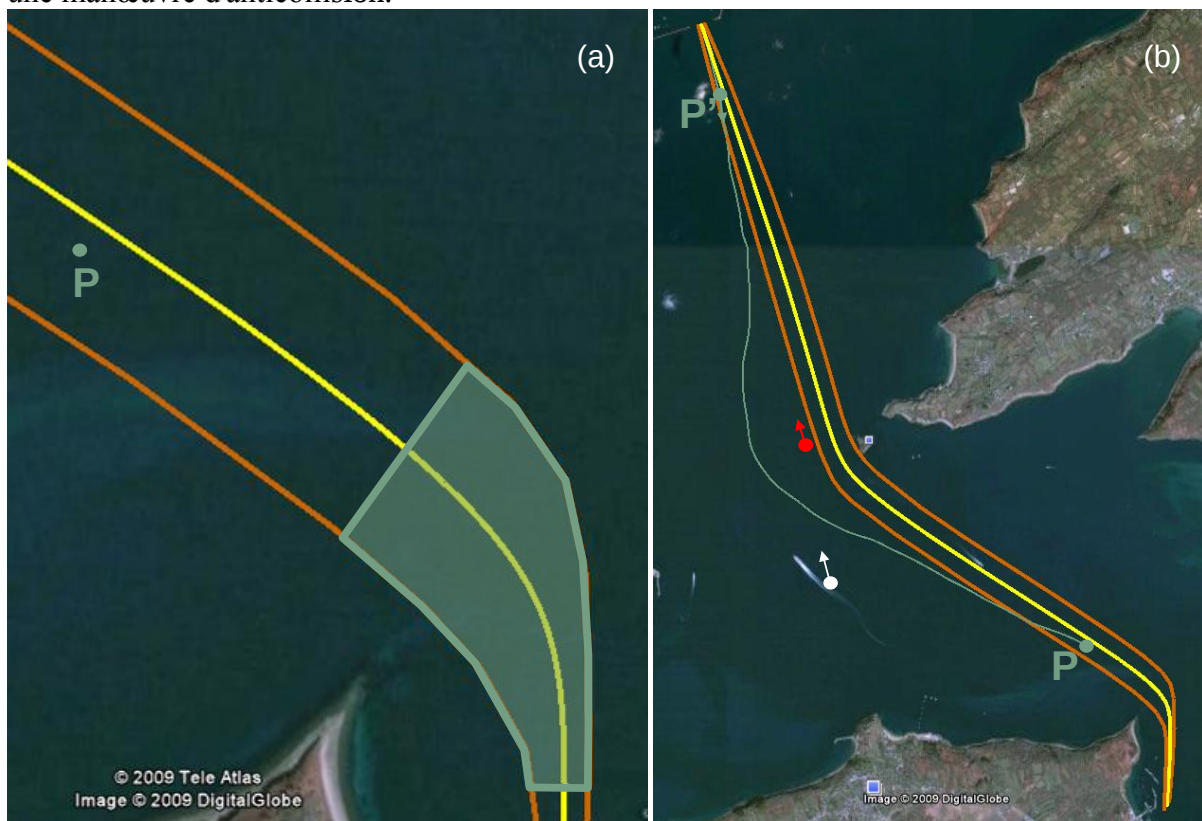


Figure 50 : Position P du navire et couloir spatio-temporel (a). Trajectoire du navire (b), P' position lors de la sortie de la route, positions et directions d'autres navires (en rouge et en blanc) lors de la sortie.

La deuxième perspective concerne la détection des comportements des objets mobiles à partir de la BD historique afin d'affiner la simulation. L'étude des évitements passés et la récupération des paramètres de ces manœuvres (distances des manœuvres, angle de barre, privilégié ou non privilégié, types des navires en route de collision, environnement) doit en effet permettre :

- d'améliorer puis de valider globalement les bases de patrons,
- d'estimer la pondération entre les critères de choix (respect du règlement, gain de temps, maîtrise de la situation) des patrons de chaque type de navire.

De la sorte, la simulation des déplacements futurs dans le cadre d'outils d'aide à la surveillance du trafic serait grandement améliorée.

Le troisième futur sujet de recherche traite de l'aide à la décision pour l'évitement de collision. Effectivement, la simulation des positions futures doit de même permettre de mieux prendre une décision d'évitement présente. Actuellement, seuls les obstacles fixes de l'environnement géographique (côtes, jetées, etc.) sont pris en considération et uniquement dans la phase de perception. Pour choisir parmi les différents patrons, il faut aussi se projeter dans l'avenir

(phase de projection du paradigme Naturalistic Decision Making), pour évaluer les différents patrons. Pour cette phase, il est nécessaire de détecter les conflits possibles entre les trajectoires des navires tiers et la trajectoire d'anticollision proposée par le patron. Si un futur conflit apparaît, il faut alors récuser ce patron ou modifier les paramètres de l'action (modification du cap ou de la vitesse).

Des études sur l'emploi des vues relatives pour suivre l'évolution d'une trajectoire vis-à-vis de la trajectoire médiane appariée sont aussi prometteuses. La vue relative permet en effet de visualiser les écarts entre ces deux trajectoires en termes de différentiel de vitesse et de distance. Néanmoins, il faut revoir la sémantique des résultats. En effet, par exemple, une distance faible est dans cette optique signe d'une trajectoire qui suit normalement la trajectoire médiane. Or, dans le contexte original, c'est un danger.

Finalement, la comparaison des trajectoires à la trajectoire médiane appariée est actuellement limitée. Elle se focalise sur la comparaison d'une position avec la position médiane et le couloir spatio-temporel ayant la même estampille temporelle. Ce processus peut être étendu à des comparaisons de partie linéaire de trajectoires. Pour cela, des nouvelles méthodes doivent être définies. Elles sont basées sur la comparaison d'une section avec la partie de la trajectoire médiane homologue. Les résultats en géométrie algorithmique devront donc être combinés à ceux obtenus en fouille de données afin d'obtenir des résultats pertinents. Simultanément, de nouveaux comparateurs permettant de qualifier la section doivent être définis. Ces derniers doivent qualifier par exemple vis-à-vis la partie homologue de la trajectoire médiane que :

- la section est plus sinueuse,
- la vitesse est plus irrégulière.

Pour définir ces opérateurs, une partie des résultats obtenus lors de la thèse de M^{elle} Noyon, et de M^{elle} Mascret seront réutilisés.

Le développement de ces recherches se situe à quatre niveaux :

- en interne, dans le cadre de la thèse de M. Laurent Etienne qui a commencé fin 2008,
- dans le cadre du GDR MAGIS,
- via un projet industriel,
- à travers des projets universitaires européens.

Les coopérations de recherches dans le cadre du **GDR MAGIS** déjà présentes, ont été renforcées. La thèse de M. Etienne est ainsi co-encadrée par M. Alain Bouju (MCF HDR) de l'université de La Rochelle. Ces travaux font et feront l'objet de présentations lors de journées de travail de l'axe "SIG mobiles, réseau de capteurs localisés, SIG ubiquitaires". Une industrialisation de ce projet fait aussi partie des perspectives proches. Le projet "**Web AIS** : suivi intelligent via l'Internet des déplacements de navires" est en cours de labellisation auprès du pôle Mer Bretagne. Les partenaires sont MAGIC INSTINCT Software société nantaise, Geomod société brestoise, l'Ecole navale, l'université de La Rochelle et le port de Nantes. L'objectif est de fournir à des opérateurs tels les ports des outils Internet de gestion de leur trafic maritime et fluvial.

Parallèlement, une coopération entre des **universités européennes** intéressées par la thématique du suivi des navires se met en place. Dans ce sens, une première demande de finance a été déposée au printemps 2009 auprès de l'association EGIDE par l'Ecole navale et le département de Géoinformatique de l'université technique de Gdansk (programme Polonium 2010). Dans un premier temps, l'objectif de ce projet bilatéral est d'intégrer nos travaux autour de la détection en temps réel de comportements anormaux de navires en zone portuaire. Cette collaboration devrait permettre à l'école navale de travailler sur des données issues de différents capteurs et d'étudier les trajectoires en zone portuaire. Dans un deuxième temps, nous souhaitons élargir rapidement cette collaboration aux équipes universitaires suivantes :

- Geographical Information Systems Research Unit de l'Université de Glamorgan, Pays de Galles, qui développe des travaux sur la visualisation en 3D de ce type de système.
- Department of Product and Systems Design Engineering, de l'Université d'Aegean, Syros, Grèce, qui a de grandes compétences en termes de services web de suivi de navires avec une couverture mondiale (voir <http://www.marinetraffic.com/ais/>).
- Instituto Superior de estatística e Gestão de informação, Université nouvelle de Lisbonne, Portugal qui travaille sur la définition de cluster de trajectoires maritimes.

Ces différents projets universitaires et industriels favorisent la dynamique scientifique nécessaire à ces futures recherches autour du suivi d'objets mobiles.

Les recherches présentées dans ce manuscrit portent sur les Systèmes d'Information Géographique maritime. Elles sont appliquées à des champs d'études larges : Les Modèles Numériques de Terrain et les objets mobiles. Ces champs fournissent des données nombreuses et précises. De même, les domaines informatiques abordés sont vastes : géométrie algorithmique, intelligence artificielle appliquée aux SIG et base de données spatiales. Les résultats produits depuis 10 ans sont probants et complémentaires. L'intégration future de ces recherches s'avère donc prometteuse, néanmoins elle nécessite des efforts en termes de formalisation, d'intégration et de validation.

6 Annexes

6.1 Curriculum vitæ

DEVOGELE Thomas

Maître de conférences Ecole Navale / Ecole des Arts et métiers ParisTech

11 Mars 1970

Marié, 3 enfants



BCRM Brest, Institut de Recherche de Ecole navale
CC 600
29240 BREST CEDEX 9



02 98 23 43 89



thomas.devogele@ecole-navale.fr



Formations

- Doctorat de Méthodes Informatiques " Processus d'intégration et d'appariement de Bases de Données Géographiques, Application à une base de données routières multi-échelles ", Université de Versailles, Laboratoire d'accueil : COGIT / Institut Géographique National, 1997.
- D.E.A de Modélisation et Traitement des Données et des Connaissances (SGBD, IA, Analyse des données symboliques) Université de Dauphine (Paris IX), 1993.

Positions

Année	Fonction	Lieu
Depuis 01/2001	Maître de conférences	Ecole navale / Ecole des Arts et métiers ParisTech
10/1998 – 12/2000	Enseignant / Chercheur (CDD ministère de la défense)	Ecole navale
10/1997 – 08/1998	Attaché Temporaire d'Enseignement et de Recherche (demi-poste)	Université de Créteil

Encadrement de Thèses

Noms	Titres	%	finance ment	Dates de soutenance	Positions actuelles
Fournier Sébastien	Intégration de la dimension spatiale au sein d'un modèle multi-agents à base de rôles pour la simulation : application à la navigation maritime	50%	Marine nationale	11/2005	Maître de conférences à l'Université de Provence, Laboratoire LSIS
Noyon Valérie	Modèle de vue spatio-temporelle et relative de trajectoires d'objets mobiles : application à la navigation maritime.	35%	Marine nationale	10/2007	ATER à l'IUT de La Rochelle, (2 ^{ème} année), qualifiée en 27 ^{ème} section (2008-2009)
Mascret Ariane	Intégration de modèles numériques de terrain application aux données terre mer	70%	Brest Métropole Océane	Prévue en février 2010	ATER à l'Université de Polynésie Française (2008-2009) CDD UBO (2009-2010)
Le Pors Thierry	Modélisation spatio-cognitive des interactions entre objets mobiles ; Application à la simulation du trafic maritime	40%	Région Bretagne	Prévue printemps 2010	ATER à l'INSA Rouen depuis septembre 2009
Etienne Laurent	Analyse de similarité et visualisation de trajectoires d'objets mobiles suivant un même itinéraire ; application aux trajectoires de navires	60%	Marine nationale	Prévue automne 2011	CDD ministère de la défense de 4 ans, avec une charge annuelle d'enseignement

Comité de rédaction de revue

- Revue internationale de géomatique depuis 2003, publiée chez Hermès Lavoisier.

Relecteur régulier de revue

- Revue internationale de géomatique publiée chez Hermès Lavoisier. 2 articles par an depuis 2000
- Revue IJGIS (International Journal of Geographical Information Science) publiée chez Taylor & Francis, d'un article par an, depuis 2005.

Organisation de congrès

- 13th International Symposium on Spatial Data Handling (SDH'2008), co-président du comité d'organisation à Montpellier. Chargé des actes Springer série LNGC (650 pages).
- 6^{ème} journées Cassini du GDR 2340 SIGMA à l'Ecole navale en 2002, co-président du comité de programme avec F. Cuq de l'IUEM (L'institut universitaire de la Mer) et membre du comité d'organisation (38 propositions d'articles, sélection de 15 articles et 13 posters, actes de 458 pages).

Membre de comité de sélection de congrès et relecteur

- Membre du comité de programme de " ISPRS International Workshop on Multidimensional and Mobile Data Model 2009 "
- Membre du comité de programme de MajecSTIC (MANifestation des JEunes Chercheurs en Sciences et Technologies de l'Information et de la Communication) en 2008 et 2009.

- Membre du comité de programme de International Conference on Humans Operating Unmanned systems, Brest, 2008.
- Membre du comité de programme des conférences international SAGEO (International Conference on Spatial Analysis and GEomatics), en 2005 à Avignon, en 2006 à Strasbourg et en 2007 à Clermont-Ferrand. Conférences portées par le GDR SIGMA-CASSINI et par le GDRE S4.
- Membre du comité de programme de IF&GIS-2007 et IF&GIS-2005 Information Fusion and Geographical Information Systems, Saint Pétersbourg, Russie.
- Membre du comité de programme de CoMoGIS'2006 (3rd International Workshop on Conceptual Modeling for Geographic Information Systems). Atelier rattaché à ER (25th International Conference on Conceptual Modeling).
- Relecteur pour W2GIS'2006 (6th International Symposium on Web and Wireless Geographical Information Systems).
- Membre du comité de programme des 8^{ième} journées Cassini 2004, à Grenoble. Journées organisées par le GDR SIGMA.
- Relecteur pour la 6^{ième} conférence Agile en 2003 à Lyon. Conférence de l'Association of Geographic Information Laboratories for Europe.

Jurys de Thèse

- Président du jury de la thèse de S. Balley, Aide à la restructuration de données géographiques sur le web - vers la diffusion a la carte d'information géographique, Thèse en informatique de l'Université de Paris-Est Marne-la-Vallée, laboratoire d'accueil COGIT/IGN, soutenue le 19/9/2007
- Membre de jury de la thèse d'O. Bedel, Application des Systèmes d'Information Logiques (SIL) à la gestion de données géographiques, thèse en informatique de l'Université de Rennes 1(ED. Matisse), soutenue le 22 janvier 2009, à l'IRISA.

Groupes de Recherche

- Co-animateur du groupe de travail concernant les problématiques de " la mobilité, la réactivité et le temps réel " du GDR SIGMA (Systèmes d'Information Géographique - Méthodologies et Applications) GDR 2340 CNRS, de 2005 à 2008.
- Co-animateur du groupe projet : SIG mobiles, réseau de capteurs localisés, SIG ubiquitaires, du GDR MAGIS depuis 2009
- Membre du bureau élargi depuis janvier 2001 du GDR SIGMA.

Commissions de spécialistes

- Membre titulaire élu de la commission de spécialistes mixte (27-61-63) de l'Ecole des Arts et métiers ParisTech de 2005 à 2008.
- Membre suppléant extérieur de la commission de spécialistes 27 de l'Université de La Rochelle.

Coopérations internationales

- Université Laval, Québec, Canada, Laboratoire CRAD (Centre de Recherche en Aménagement et Développement) : séjour en 2002 d'un mois sur la modélisation multi-agents des activités de déplacement dans la ville de Québec.
- Université de Glamorgan, pays de Galles, Royaume Unis, Geographical Information Systems (GIS) group: co-responsable du projet de recherche soutenu par le Ministère des Affaires Etrangères, à travers l'association Egide (programme Alliance), thématique suivi de navires en temps réel dans le cadre d'un environnement 3D (2004 - 2008).

Expertises scientifiques

- Expertise de rapport de projet pour le GEOIDE Network du Canada (www.geoide.ulaval.ca) 2009.
- Expertise pour la région Bretagne d'une demande de financement dans le cadre de l'appel d'offre d'attribution des allocations recherche par la Région Bretagne (Appel à projets ARED 2008)
- Expertise pour le REMPEC (Regional Marine Pollution Emergency Response Centre for the Mediterranean Sea; www.rempec.org). Conseiller scientifique pour la mise en place d'un système d'information géographique destiné à communiquer aux administrations maritimes méditerranéennes via Internet, les statistiques de pollution et de trafic maritime, dans le cadre du projet SAFEMED.

Prime d'encadrement doctoral et de recherche

- Attribuée en octobre 2008

6.2 *Liste détaillée des publications*

2009

- Le Pors T., Devogele T., Chauvin C.** (2009) Multi-agent system integrating naturalistic decision roles: application to maritime traffic, In Proceedings of the IADIS International Conference on Intelligent Systems and Agents, IADIS Press, pp100-107.
- Etienne L., Devogele T., Bouju A.** (2009) Analyse de similarité de trajectoires d'objets mobiles suivant le même itinéraire : Application aux trajectoires de navires, Revue Ingénierie des Systèmes d'Information (ISI), vol. 14, n°5, Hermès, pages 85 à 106.
- Servigne S., Ray C., Bouju A., Devogele T., Bertrand F., Gutiérrez C., Noel G.** (2009) Gestion des masses de données au sein de bases de données capteurs, Revue internationale de géomatique, vol. 19, num. 2, pages 133 à 150.
- Bouju A., Bertrand F., Malle-Noyon V., Servigne S., Devogele T., Ray C., Martin H., Gensel J.** (2009) Gestion de données spatio-temporelles au sein de bases de données capteurs, Première Conférence Francophone sur les Technologies de l'Information, de la Communication et de la Géolocalisation dans les Systèmes de Transports (COGIST 2009), pages 123-130.

2008

- Mustière S., Devogele T.** (2008) Matching Networks with Different Levels of Detail, Geographical Spaces, Geoinformatica, Springer (Publisher), vol. 12, Num. 4, pages 435-453.
- Brazo F., Devogele T., Bergonzo A.** (2008) Mesoscopic approach of the traffic for safety of navigation within the Mediterranean region, Conférence SAGEO, Montpellier, 14 pages.
- Etienne L., Devogele T., Bouju A.** (2008) Outils d'aide aux décideurs concernant le suivi de navires : suivi de trajectoires relatives entre navires et détection de trajectoires inhabituelles, 7^{ième} journées scientifiques et techniques du CETMEF, Paris, 4 pages.

2007

- Bertrand F., Bouju A., Claramunt C., Devogele T., Ray C.** (2007) Web architectures for monitoring and visualizing mobile objects in maritime contexts, In Proceedings of the 7th International Symposium on Web and Wireless Geographical Information Systems (W2GIS 2007), pages 94-105, G. Taylor and M. Ware (eds.), Springer-Verlag, LN series in Computer Science (LNCS 4857), Cardiff, UK, November 2007.
- Servigne S., Devogele T., Bouju A., Bertrand F., Gutierrez C., Laucius S., Noel G., Ray C.** (2007) : Gestion de masses de données temps réel au sein de bases de données capteurs, Conférence Québeco-Française pour le Développement de la Géomatique (CQFD-Géo), SAGEO 2007, 15 pages, Clermont-Ferrand, France.
- Noyon V., Claramunt C., Devogele D.** (2007) A Relative Representation of Trajectories in Geographical Spaces, Geoinformatica, Springer (Publisher), volume 11, num. 4, pages 479-496.
- Fournier S., Devogele T., Claramunt, C.** (2007) Un modèle de raisonnement contextuel pour la navigation maritime, Revue Internationale de Géomatique, Hermès Lavoisier, volume 17, num. 1, pages 63-89.

Le Pors T., Chauvin C., Devogele T. (2007) Etude de systèmes cognitifs pour l'évitement de collisions, journées EPIQUE'2007, congrès de la SFP (Société française de Psychologie) Nantes, 6 pages.

Claramunt C., Devogele T., Fournier S., Noyon V., Petit M., Ray C. (2007) Maritime GIS: From Monitoring to Simulation Systems - Information Fusion and Geographical Information Systems (IF&GIS'07), May 27-29, 2007, St. Petersburg, Russia, invited paper, pages 34-44.

Ray C., Devogele T., Noyon V., Petit M., Fournier S., Claramunt C. (2007) GIS Technology for Maritime Traffic Systems - In European Research Consortium for Informatics and Mathematics News: Special Theme on Traffic Planning and Logistics, vol. 68, pages 41-42, Kuntz, P. (eds.), ERCIM EEIG, January 2007, ISSN 0926-4981.

2006

Mascret A., Devogele T., Le Berre I., Hénaff A. (2006) Coastline Matching Process based on the discrete Fréchet distance - In Proceedings of the 12th International Symposium on Spatial Data Handling (SDH), pages 383-400, W. Kainz, A. Reidl, and G. Elmes (eds.), Vienna, Austria, Springer-Verlag.

Mascret, A., Devogele, T. (2006) Intégration de données Terre/Mer basée sur une approche paysagère - SAGEO 2006. Colloque International de Géomatique et d'Analyse Spatiale, Strasbourg, 13 pages

2005

Noyon V., Devogele T., Claramunt C. (2005 a) A formal model for representing point trajectories in two-dimensional spaces, , in Proceedings of the Conference on Conceptual Modelling and GIS (COMOGIS'05), Lecture Notes in Computer Science, Springer, pages 208-217.

Noyon V., Devogele T., Claramunt C., (2005 b) A relative modelling approach for modelling trajectories, in Proceedings of the 4th ISPRS Workshop on Dynamic and Multi-dimensional GIS, University of Glamorgan, Archives of ISPRS, pages.139-143.

Le Berre I., Hénaff A., Devogele T., Mascret A., Wenzel F. (2005) Spot 5 : un outil pertinent pour le suivi du trait de côte ? - Norois, Presses Universitaires de Rennes, vol. 196, n° 2005-3, pages. 23-35.

2004

Brosset D., Devogele T. (2004) Algorithmes génétiques et généralisation de MNT marins. Actes des journées des 7^{ième} journées Cassini, Grenoble, éditeur MSH Alpes, pages 35 - 40.

Noyon V., Devogele T. (2004) Système distribué de gestion d'objets mobiles. Actes des journées des 7^{ième} journées Cassini, Grenoble, éditeur MSH Alpes: pages 7-12.

2003

Fournier S., Devogele T., Claramunt C. (2003 a) A role-based multi-agent model for concurrent navigation systems, Proceedings of the 6th AGILE Conference on Geographic Information Science, Gould, M. et *al.* (eds.), Presses Polytechniques et Universitaires Romandes, pages. 623-632.

Fournier S., Brocarel D., Devogele T., Claramunt C. (2003 b) TRANS : A Tractable Role-based agent prototype for Concurrent Navigation Systems, The first European workshop on Multi-Agent Systems, 6 pages.

Cuq F., Devogele T. (éditeurs du numéro spécial) (2003) Numéro spécial SIG transport, Revue internationale de Géomatique, Vol. 13 N° 2, 138 pages.

Avant 2003 (sélection)

Cuq F., Devogele T., Populus J. (éditeurs du numéro spécial) (2002) Numéro spécial SIG côtiers, Revue internationale de Géomatique, Vol.12 N° 3, 120 pages.

Desvignes G., Lucas de Couville G., Peytchev E., Devogele T., Fournier S., Claramunt C. (2002) The Share-Loc Project: a WAP-based maritime location system, 3rd International Conference on Web Information Systems Engineering (WISE), IEEE press, pages 88-94.

Devogele T. (2002) A new Merging process for data integration based on the discrete Fréchet distance, 10th International Symposium on Spatial Data Handling (SDH), Ottawa (Canada), D. Richardson et P. van Oosterom, pages 167-181.

Barbe F., Gélébart F., Devogele T., Claramunt C. (2001) A knowledge-bases GIS For concurrent navigation monitoring, chapter of Spatial Information and the Environment, P. Halls (ed.), Taylor & Francis, pages 135-146.

Devogele (2000) Mesure d'exactitude et processus de fusion à l'aide de la distance de Fréchet discrète, Revue internationale de Géomatique, Hermès volume 10 - n°3-4/2000, pages 359-381.

Devogele T., Parent C., Spaccapietra S. (1998) On Spatial Database Integration, International Journal on Geographical Information Systems (IJGIS) Vol. 12, Taylor & Francis, pages 335-352.

Devogele T., Trevisan J., Raynal L. (1996) Building a multi-scale database with scale-transaction relationships, Spatial Data Handling (SDH), Delft (Pays-Bas), Taylor & Francis, pages 337-351.

7 Bibliographie

- Armstrong, M. P.** (1988) Temporality in spatial databases. Proceedings: GIS/LIS'88, 2:880-889.
- Alt H., Godau M.** (1992) Measuring the resemblance of polygonal curves". Actes des 8th Annu. ACM Symposium on Computational. Geometry, 1995, pages 102-109.
- Bavoux J-J.** (1998) Introduction à l'analyse spatiale, Paris, Armand Colin, collection Synthèse, série Géographie, 96 pages
- Bertin J.** (1973). "*Sémiologie graphique*", Mouton/ Gauthier-Villars, 2nd edition, 1973.
- Bertrand F., Bouju A., Claramunt C., Devogele T., Ray C.** (2007) Web architectures for monitoring and visualizing mobile objects in maritime contexts, In Proceedings of the 7th International Symposium on Web and Wireless Geographical Information Systems (W2GIS 2007), pages 94-105, G. Taylor and M. Ware (eds.), Springer-Verlag, LN series in Computer Science (LNCS 4857), Cardiff, UK, November 2007.
- Brunod A., Got D.** (2006) Segmentation de relief et algorithme génétique, rapport d'élèves de l'Ecole navale, projet de voie d'approfondissement Systèmes d'Information et de Modélisation, 37 pages.
- Claramunt C., Devogele T., Fournier, S., Noyon, V., Petit, M., Ray, C.** (2007) Maritime GIS: From Monitoring to Simulation Systems - Information Fusion and Geographical Information Systems (IF&GIS'07), May 27-29, 2007, St. Petersburg, Russia, invited paper, pages 34-44.
- Chauvin, C.** (2000) Analyse de l'activité d'anticollision à bord des navires de commerce: Des marques linguistiques aux représentations mentales, Travail humain, Vol. 63, Presses Universitaires de France, pp. 31-58.
- Chauvin, C., Clostermann, J. and Hoc, J.** (2008) Situation Awareness and the Decision-Making Process in a Dynamic Situation: Avoiding Collisions at Sea", Journal of Cognitive Engineering and Decision Making, Vol. 2, Human Factors and Ergonomics Society, pp. 1-23.
- Chauvin C, Lardjane S.** (2008) Decision making and strategies in an interaction situation: Collision avoidance at sea, Transportation Research Part F: Psychology and Behaviour , Vol. 11, Elsevier, pp. 259-269.
- Coello, C.A., Romero C.E.M** (2002) Evolutionary Algorithms and Multiple Objective Optimization, chapitre 6 de Multiple Criteria Optimization, State of the Art Annotated Bibliographic Surveys, éditer par Ehrgott M. et Gandibleux X, Kluwer Academic Publishers.
- Deb, K., Pratap, A., Agrawal, S. and Meyarivan, T.** (2000) A fast and elitist multi-objective genetic algorithm: NSGA-II. Technical Report No. 2000001. Kanpur: Indian Institute of Technology Kanpur, India.
- Denègre J., Salgé F.** (2004) Les systèmes d'information géographique, collection "Que sais-je ?", PUF, 128 pages.
- Degré T.** (2007) The use of risk concept to characterize and select High Risk Vessels for ship inspections», 12th International Congress of the International Maritime Association of the Mediterranean, vol. 6, p. 37-49

- Devogele T., Riffault Y.** (2000) Alarme de précision pour les lignes de côte des ECDIS, CoastGIS'99, Editeur Populus and Loubersac, Ed. IFREMER, Actes de colloques n°25, p 159-169, 2000.
- Devogele** (2000) Mesure d'exactitude et processus de fusion à l'aide de la distance de Fréchet discrète, Revue internationale de Géomatique, Hermès volume 10 - n°3-4/2000, pages 359-381.
- Devogele T.** (2002) A new Merging process for data integration based on the discrete Fréchet distance, 10th International Symposium on Spatial Data Handling (SDH), Ottawa (Canada), D. Richardson et P. van Oosterom, pages 167-181.
- van Dijk S., Thierens D and de Berg M.** (2002) Using genetic algorithms for solving hard problems in GIS, GeoInformatica, vol. 6 n° 4, p 381-413, 2002.
- Douglas D., Peucker T.** (1973) Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, The Canadian Cartographer 10(2), page 112-122.
- Edus F., Salis J.** (2005) Route de navire optimale et algorithme génétique, rapport d'élèves de l'Ecole navale, projet de voie d'approfondissement Systèmes d'Information et de Modélisation, 24 pages.
- Egenhofer M., Franzosa R.** (1991) Point-Set Topological Spatial Relations, International Journal of Geographical Information Systems, vol. 5, pp. 161-174
- Eiter T., Mannila H.** (1994) "Computing Discrete Fréchet Distance", rapport technique num. CD-TR 94/64. Institute of Information Systems, Technical University of Vienna. 1994.
- Ferber J., Gutknecht O.** (1998) Aalaadin: a meta-model for the analysis and design of organizations in multi-agent systems, ICMAS 98 (International Conference on Multi-Agent Systems), IEEE Press, pp. 128-135, 1998.
- Fournier S., Devogele T., Claramunt C.** (2003 a) A role-based multi-agent model for concurrent navigation systems, Proceedings of the 6th AGILE Conference on Geographic Information Science, Gould, M. et al. (eds.), Presses Polytechniques et Universitaires Romandes, pages. 623-632.
- Fournier S., T. Devogele, C. Claramunt** (2003 b) TRANS : A Tractable Role-based agent prototype for Concurrent Navigation Systems, The first European workshop on Multi-Agent Systems, 6 pages.
- Fournier S., Devogele T., Claramunt, C.** (2007) Un modèle de raisonnement contextuel pour la navigation maritime, Revue Internationale de Géomatique, Hermès Lavoisier, volume 17, num. 1, pages 63-89.
- Fréchet M.** (1906) Sur quelques points du calcul fonctionnel, Rendiconti del Circolo Matematico di Palermo, 22 :1-74.
- Hausdorff F.** (1919) Dimension und ausseres, Mass. Mathematische Annalen, num. 79, pages 157-179.
- Holland J.** (1975) , Adaptation in Natural and Artificial Systems, University of Michigan Press Ann Arbor, 1975.
- Jonglez S., Songy M.** (2008) Etude de la similitude des déplacements de navires, Rapport de Projet de Voie d'Approfondissement de l'Ecole navale, 40 pages.
- Klein G.** (1997) The recognition-primed decision (RPD) model: Looking back, looking forward", Naturalistic decision making , Lawrence Erlbaum, pp. 285-292.

- Kobus D., Proctor S., Holste, S.** (2001) Effects of experience and uncertainty during dynamic decision making, *International Journal of Industrial Ergonomics* , Vol. 28, Elsevier, pp. 275-290.
- Langlois P.** (1994) Une transformation élastique du plan basée sur un modèle d'interaction spatiale, Application à la géomatique », rapport technique 1994, MTG, université de Rouen.
- Langran G., Chrisman N, R.** (1988) A framework for temporal geographic information. *Cartographica*, 25(3):1-14.
- Laurini R.** (1998) Spatial multi-database topological continuity and indexing: a step towards seamless GIS data interoperability”, dans *International Journal of Geographical Information Science: Special issue: Interoperability in GIS*, Taylor & Francis, Vol. 12, n° 4, pages 373-402.
- Le Berre I., Hénaff A., Devogele T., Mascret A., Wenzel F.** (2005) Spot 5 : un outil pertinent pour le suivi du trait de côte ? - Norois, Presses Universitaires de Rennes, vol. 196, n° 2005-3, pages. 23-35.
- Le Roux C., Martinoty G.** (2009) le projet Litto3D, http://www.shom.fr/fr_page/fr_act_Litto3D/index_litto3D.htm
- Le Pors T., Devogele T., Chauvin C.** (2009) Multi-agent system integrating naturalistic decision roles: application to maritime traffic, In *Proceedings of the IADIS International Conference on Intelligent Systems and Agents*, IADIS Press, pp100-107.
- Lee J., Han J., Li X.** (2008) Trajectory Outlier Detection: A Partition-and-Detect Framework, *Data Engineering*, 2008. ICDE 2008. IEEE 24th International Conference on Data Engineering, p. 140-149.
- Mascret, A., Devogele, T., Le Berre, I., Hénaff, A.** (2006) Coastline Matching Process based on the discrete Fréchet distance - In *Proceedings of the 12th International Symposium on Spatial Data Handling (SDH)*, pages 383-400, W. Kainz, A. Reidl, and G. Elmes (eds.), Vienna, Austria, Springer-Verlag.
- Mascret, A., Devogele, T.** (2006) Intégration de données Terre/Mer basée sur une approche paysagère - SAGEO 2006. Colloque International de Géomatique et d'Analyse Spatiale, Strasbourg, 13 pages
- Morel M., Napoli A., Littaye A., George J., Jangal F., Alhadeb B., Scapel C., Lebrevelec J., Dejardin D.** (2008) Surveillance et contrôle des activités des navires en mer ScanMaris, Workshop Interdisciplinaire sur la Sécurité Globale.
- Mustière** (2001) Apprentissage supervisé pour la généralisation cartographique, Thèse de l'Université Paris 6, 246 pages.
- Mustière S., Devogele T.** (2008) Matching Networks with Different Levels of Detail, *Geographical Spaces, Geoinformatica*, Springer (Publisher), Volume 12, Number 4, pages 435-453.
- Noyon V., Devogele T.** (2004) Système distribué de gestion d'objets mobiles. Actes des journées des 7^{ième} journées Cassini, Grenoble, éditeur MSH Alpes: pages 7-12.
- Noyon V., Devogele T., Claramunt C.** (2005 a) A formal model for representing point trajectories in two-dimensional spaces, , in *Proceedings of the Conference on Conceptual Modelling and GIS (COMOGIS'05)*, Lecture Notes in Computer Science, Springer, pages .208-217.
- Noyon V., Devogele T., Claramunt C.,** (2005 b) A relative modelling approach for modelling trajectories, in *Proceedings of the 4th ISPRS Workshop on Dynamic and Multi-dimensional GIS*, University of Glamorgan, Archives of ISPRS, pages.139-143.

- Noyon V., Claramunt, C., Devogele, D.** (2007) A Relative Representation of Trajectories in Geographical Spaces, *Geoinformatica*, Springer (Publisher), volume 11, num. 4, pages 479-496.
- OMI** (1972) Règlement International pour prévenir les abordages en mer, convention, Internationale de l'Organisation Maritime Internationale, Entrée en vigueur le 15 juillet 1977 (décret n° 77-778 du 7.7.1977)
- Oraas, S.** (1975) Sélection automatisée des sondes, *Revue Hydrographique Internationale*, Monaco, vol. LII(2)
- Peuquet D. , Duan N.** (1995) An event-based spatiotemporal data model (ESTDM) for temporal analysis of geographical data. *International Journal of Geographical Information Systems*, 9(1): 7-24.
- Pumain D., Saint-Julien T** (2004) L'Analyse spatiale, 1. Localisations dans l'espace, collection *Cursus*, Armand Colin, 168 pages.
- Ray C., Devogele T., Noyon V., Petit M., Fournier S., Claramunt C.** (2007) GIS Technology for Maritime Traffic Systems - In European Research Consortium for Informatics and Mathematics News: Special Theme on Traffic Planning and Logistics, vol. 68, pages 41-42, Kuntz, P. (eds.), ERCIM EEIG, January 2007, ISSN 0926-4981.
- Rapper J., Livingstone D.** (1995) Development of a geomorphological spatial model using object-oriented design. *International Journal of Geographical Information Systems*, 9(4): 359-384.
- Richard A.** (2007) L'optimisation des transits de navires marchands par l'utilisation des courants de marée, *Canadian Journal of Transportation*.
- Röffer T.** (1999) Route navigation using motion analysis, actes de COSIT, Fresca (eds), pp21-36.
- Servigne S., Ray C., Bouju A., Devogele T., Bertrand F., Gutiérrez C., Noel G.** (2009) Gestion des masses de données au sein de bases de données capteurs, *Revue internationale de géomatique*, vol. 19, num. 2, pages 133 à 150.
- Saux E., Thibaud T., Li K.J. & Kim M.H.** (2004) A New Approach for a Topographic Feature-Based Characterization of Digital Elevation Data, 12th ACM International Symposium on Advances in Geographic Information Systems (ACM GIS 2004), pp. 73-81.
- Spaccapietra S., Parent C., Dupont Y.** (1992) Model Independent Assertions for Integration of Heterogeneous Schemas, *Very Large DataBases Journal*, Vol. 1, Num. 1, pages 81-126.
- Sarazin et Malini** (2009) Modélisation et simulation du comportement des navires de pêche en situation d'anticollision, rapport de projet de spécialité de master, Ecole navale, 22 pages.
- Turing A.** (1950) Computing machinery and intelligence, *Mind*, vol. 59, n° 36, p. 433-460.
- Wooldridge M., Jennings N., Kinny D.** (2000) The Gaia Methodology for Agent-Oriented Analysis and Design. *Autonomous Agents and Multi-Agent Systems*, vol. 3, num.3, pages 285-312.
- Worboys M.** (1992) A model for spatio-temporal information. *Proceedings: the 5th International Symposium on Spatial Data Handling*, 2:602-611.
- Zadeh L.A** (1965) Fuzzy Sets, *Information Control*, vol. 8, page 338-353.
- Zimmerman K., Freksa C.** (1996) Qualitative spatial reasoning using orientation, distance and path knowledge, *Applied Intelligence*, vol. 6, num. 1, pp 49-58.

Résumé :

Les Systèmes d'Information Géographique (SIG) gèrent des informations complexes localisées dans un espace géographique. Durant la dernière décennie, ces systèmes ont vu l'émergence d'une part des SIG en trois dimensions et plus particulièrement des modèles numériques de terrain (MNT) et d'autre part de l'intégration de données temporelles. Les travaux de recherche en informatique présentés dans ce manuscrit s'attaquent à ces deux thématiques et sont appliqués au domaine maritime.

En termes de MNT, ces recherches ont essentiellement porté sur l'intégration de MNT terrestre et maritime. Une méthode de fusion générant un MNT continu a été définie. Elle est basée sur l'appariement d'éléments caractéristiques linéaires (crêtes, talwegs, lignes de rupture de pente). Elle emploie des distances linéaires discrètes introduites par Fréchet. Ces distances fournissent des vecteurs d'appariement autorisant la fusion de MNT à l'aide d'une technique de déformation élastique.

En ce qui concerne les SIG spatio-temporels, ces travaux concernent les déplacements d'objets mobiles et plus particulièrement les navires. Dans le cadre de la navigation maritime, les SIG ont une place prépondérante via les aides à la navigation, les systèmes de surveillance du trafic et les simulateurs de formation. Dans ce cadre, un nouveau modèle de représentation des déplacements, la vue relative, a été formalisé. Ce formalisme s'appuie sur la perception humaine de l'espace spatio-temporel et des objets mobiles. Cette représentation est complémentaire de la représentation absolue. Les déplacements relatifs vis-à-vis d'un objet référent sont ainsi plus facilement analysables. Les évolutions des distances et des vitesses relatives sont mieux perçues. Parallèlement, une simulation des activités de navires à l'aide de systèmes multi-agents a été conçue. Les déplacements simultanés d'un grand nombre de navires dans un environnement maritime sont ainsi simulés. Elle prend en compte les différentes activités concurrentes des objets mobiles et les contraintes géographiques. Qui plus est, elle intègre des raisonnements à base de patrons pour reproduire les raisonnements humains. Finalement, à partir des bases de données historiques des positions d'objets mobiles dans un environnement ouvert, des techniques d'extractions des trajectoires et de fouille de données spatio-temporelles ont été réalisées. Elles permettent pour chaque itinéraire et chaque type d'objets de définir des routes types spatio-temporelles et de détecter à la volée des comportements inhabituels.

Ces travaux, bien qu'appliqués au domaine maritime, sont génériques. Les méthodes de fusion de MNT peuvent être appliquées à tous types de MNT ayant des zones de recouvrement. De même, les travaux de modélisation, d'analyses et de simulation sur les objets mobiles sont adaptables à tous les déplacements d'objets mobiles dans des espaces ouverts.

Abstract:

Geographic Information Systems (GIS) manage information located in geographic space. These systems have included more and more complex information over the past decade. Indeed, they can now store temporal data as well as tree-dimensional data: for example, the digital elevation model (DEM) represents relief. My research in computing science deals with those two fields and is applied to maritime domain.

As far as the DEM is concerned, those works focus on DEM integration. A new approach based on landscape elements (ridges, talwegs, etc.) allows to merge land and maritime geographical datasets. This includes a process matching linear elements. It develops from several extensions of linear distance as defined by Fréchet. Those distances also provide data matching vectors that allow to merge DEMs thanks to rubber sheeting process.

In the spatio-temporal GIS area, my work studies moving object trajectories and more particularly vessel displacements. For maritime navigation, GIS are key systems. Navigation, monitoring and simulation learning systems are very close to Geographic Information System. First, a new conceptual model has been introduced: the relative view, oriented to the representation of trajectories in time and space. This model is based on the human perception of time and space for mobile objects. This representation complements the traditional absolute view. Displacements of moving objects relative to a referent object are qualified in terms of relative speed and distance. With this new representation, evolutions of displacements are better understood. In the same way, a new multiagent system (MAS) is designed to simulate maritime activities. Several vessel displacements can be simulated at the same time in a maritime space. This simulator takes into account the maritime space constraints and different activities of vessels with qualitative spatial reasoning and role paradigm. Moreover, to simulate human reasoning, it integrates a decision making system based on patterns. Finally, data mining on a huge quantity of positions stored in spatio-temporal databases is carried out in order to deduce the behavior of those mobile objects. We can define spatio-temporal patterns for each cluster of trajectories of mobile objects following the same itinerary.

Those works are applied to the maritime domain but can also be extended to other types of data. DEM integration processes can be useful for other DEMs with overlap area. In the same way, research on moving objects can be adapted to other kind of objects moving in an open area.