Représentation et construction des systèmes d'information par l'image

Guillaume Moreau

▶ **To cite this version:**

Guillaume Moreau. Représentation et construction des systèmes d'information par l'image. Interface homme-machine [cs.HC]. Université de Nantes, 2009. tel-00488770

HAL Id: tel-00488770

https://theses.hal.science/tel-00488770

Submitted on 2 Jun 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université de Nantes
CERMA - IRSTV

Mémoire

présenté en première version en vue d'obtenir
l'habilitation à diriger des recherches

par

Guillaume Moreau

Représentation et construction des systèmes d'information par l'image

Habilitation soutenue le 3 juillet 2009 devant le jury composé de :

M.	Pascal Guitton	Université Bordeaux I	(Rapporteur)
M.	Jean-Pierre Jessel	Université Paul Sabatier	(Rapporteur)
M.	Xavier Pueyo	Université de Gérone	(Rapporteur)
M.	Jacques Tisseau	Ecole Nationale d'Ingénieurs de Brest	(Examinateur)
M.	Patrick Chedmail	Ecole Centrale de Nantes	(Président)
M.	Gérard Hégron	LCPC	(Directeur de recherche)

à Camille, Carla et Héloïse

Remerciements

Je souhaite remercier Pascal Guitton, Jean-Pierre Jessel, Xavier Pueyo, et Jacques Tisseau de m'avoir fait l'honneur et l'amitié de participer au jury. Je n'oublie pas évidemment Gérard Hégron qui a accepté de diriger ce travail et qui m'a accueilli au CERMA en septembre 2002. Ma gratitude va également à Patrick Chedmail pour la confiance qu'il m'a accordée en me recrutant à l'Ecole Centrale de Nantes la même année puis en me confiant certaines responsabilités au sein de l'Ecole. Il a également accepté de présider le jury de cette habilitation.

Encore plus qu'une thèse, une habilitation à diriger des recherches suppose un travail collectif et une insertion dans une ou plusieurs équipes de recherche. Les travaux présentés dans ce mémoire sont les miens mais aussi ceux de nombreux collègues et étudiants. Parmi eux, je souhaite citer particulièrement :

- à L'IRISA : Bruno Arnaldi, Stéphane Donikian, Gwenola Thomas, Dominique Adde et Richard Kulpa ;
- à l'Ecole des mines de Paris : Philippe Fuchs, Laure Leroy, Alexis Paljic, Janos Boudet, Nicolas Lemmer, Mohamed Benali Khoudja, Panagiotis Stergiopoulos ;
- au CERMA et à l'IRSTV : Pascal Joanne, Souha Tahrani, Jihen Jallouli, Myriam Servières, Nicolas Bioret, Chamseddine Zaki, Davy Salerne, Erwan Bocher, Thomas Leduc, Gwendall Petit ;
- à l'Université Keio : Hideo Saito, Hideaki Uchiyama, Naoyuki Yazawa, Ryo Mitsumori ;
- à l'Université des Antilles et de la Guyane : Andrei Doncescu (parti au LAAS depuis), Sébastien Régis.

Ces remerciements seraient incomplets si je n'avais pas une pensée pour mes collègues du département Informatique et Mathématiques – en plus de ceux déjà cités – et en particulier pour ceux qui m'ont enlevé pas mal d'épines du pied et qui ont donc permis l'aboutissement de ce projet d'HDR : Jean-Sébastien Le-Brizaut, Jean-Yves Martin, Morgan Magnin, Vincent Tourre, Camille Constant et Bénédicte Girard.

Enfin, ma reconnaissance va très particulièrement à celles qui m'ont soutenu, encouragé et qui ont supporté mes (nombreuses) absences, soirées et weekends de travail : Patricia, Camille, Carla et Héloïse.

Table des matières

Table des matières **vi**

Liste des figures **ix**

Introduction générale **1**
- Contexte et enjeux . . . 1
  - Positionnement de notre travail . . . 2
- Du système d'information à la réalité virtuelle . . . 3
- Les 5A . . . 4
- Problématiques . . . 5
  - Perception des environnements virtuels . . . 5
  - Création des environnements virtuels . . . 6
  - Vers une boucle fermée entre images et Système d'information . 7
- Classification des travaux . . . 9
- Organisation du document . . . 9

1 Perception des mondes virtuels **11**
- 1.1 Perception physique : la vision stéréoscopique . . . 11
  - 1.1.1 Contexte : le projet IHS10 . . . 12
  - 1.1.2 Fatigue visuelle et problèmes de perception en vision stéréoscopique . . . 12
  - 1.1.3 Tests de mesure de la fatigue visuelle . . . 16
  - 1.1.4 Stratégies d'amélioration du confort . . . 16
  - 1.1.5 Approche par les ondelettes et perspectives . . . 17
  - 1.1.6 Perception des formes . . . 18
  - 1.1.7 Conclusion . . . 20
- 1.2 Perception des objets virtuels : revue de projet et ergonomie . . . 20
  - 1.2.1 Introduction . . . 20
  - 1.2.2 Génération des images en vision stéréoscopique . . . 22
  - 1.2.3 Retour d'effort pour les études d'ergonomie . . . 23
  - 1.2.4 Détection de collisions point-modèle . . . 25
  - 1.2.5 Détection de collision main-modèle . . . 25
  - 1.2.6 Conclusion . . . 26
- 1.3 Perception des mondes virtuels de grande taille . . . 28
  - 1.3.1 Ambiances lumineuses urbaines . . . 29
  - 1.3.2 Analyse des impacts paysagers des éoliennes . . . 36
  - 1.3.3 Conclusion . . . 43
- 1.4 Conclusion sur la perception des mondes virtuels . . . 43

2 Construction des environnements virtuels **45**

2.1 Perception humaine et vision par ordinateur 45
2.1.1 La transformée en quart d'ondelettes 47
2.2 Construction d'environnements virtuels à l'aide d'images 51
2.2.1 Etat de l'art 51
2.2.2 Shape from Depth 54
2.2.3 Gantom : reconstruction 3D d'objets muséographiques . . 55
2.3 Au-delà de la géométrie : fidélité comportementale des environnements virtuels 58
2.3.1 État de l'art 58
2.3.2 Conclusion sur les modèles de contrôle du mouvement . 62
2.3.3 Pré-requis à la modélisation comportementale 64
2.3.4 Le modèle HTPS 66
2.3.5 Application à la simulation de trafic 71

3 Les images comme lien entre le monde réel et le SIG3D+t **79**
3.1 Introduction : contexte applicatif 79
3.2 Abstraction 81
3.3 Archivage 81
3.4 Analyse 82
3.5 Acquisition 84
3.5.1 Géo-référencement à partir d'images 84
3.5.2 Approche alternative pour le géo-référencement : utilisation de bases d'images existantes 87
3.5.3 Vers la construction de SIG3D en réalité augmentée . . . 88
3.6 Affichage 90
3.6.1 Affichage d'informations 3D+t 90
3.6.2 Environnements virtuels urbains : le problème du fond de carte 92
3.6.3 Rendu expressif 93
3.6.4 Réalité augmentée 93
3.7 Vers l'interaction avec le SIG3D+t à l'aide d'images . . 95

Bibliographie **97**

Liste des figures

1 Interfaçage entre un être humain et un système d'informations via la réalité virtuelle 4
2 Les 5A des SIG, d'après [39] 4
3 Les différents niveaux de perception des mondes virtuels étudiés, par degré de complexité 6
4 Exemple de lien entre images et SI : le cas du SIG urbain . . 7
5 Principe de la réalité augmentée en environnement urbain : surimpression d'informations du SIG sur des images réelles 8
6 Chronologie des travaux 9

1.1 Les différents niveaux de perception des mondes virtuels . 11
1.2 Différence entre vision naturelle et vision stéréoscopique . . 13
1.3 Mouvement pseudoscopique parallèle à l'écran 14
1.4 Fonction de confort de Wöpking [139] en fonction de la disparité et du contenu fréquentiel de l'image 15
1.5 Fonction de confort de Perrin [188] 15
1.6 Forme inconnue réelle et virtuelle 19
1.7 Influence de la position relative par rapport à l'écran (à gauche pour la sphère, à droite pour la forme inconnue) . . 20
1.8 Le dispositif IMAVE 21
1.9 Tout premier dispositif de mesure de la discrimination en raideur et en course d'un bouton de commande de planche de bord (à gauche) et pourcentages de bonnes réponses en discrimination en fonction de la variation de raideur pour deux courses différentes (à droite) 24
1.10 Modèle géométrique de la main 26
1.11 Schéma algorithmique de détection de collisions 27
1.12 Résultats expérimentaux : (a) force théorique, (b) force calculée, (c) variation de la force, (d) nombre de triangles en contact 27
1.13 Phénomènes d'ambiances 29
1.14 La notation graphique des 5 éléments caractéristiques de la ville selon [104] 30
1.15 Les processus de la perception visuelle située des ambiances lumineuses [165] 32
1.16 Exemples d'effets solaires dans la rue de la Barillerie à Nantes 32
1.17 Conditions d'expérience pour les expériences 2 et 3 34
1.18 Caractéristiques historiques du parcours choisi 34
1.19 Caractéristiques morphologiques du parcours choisi 35
1.20 Rôle de l'Homme dans le complexe «paysage» 37
1.21 Éolienne : des dimensions particulières 39

1.22 Différentes lectures en zone de perception intermédiaire . . 39
1.23 Répartitions des zones d'influence sonore sur les différents parcours 42
1.24 Utilisation d'un vélo instrumenté 42

2.1 Géométrie épipolaire : C, C', x, x' et X sont coplanaires . . . 46
2.2 Les points p_l et p_r correspondent bien qu'ils soient près d'une occlusion, mais ils ne peuvent être mis en correspondance avec une transformée en ondelettes traditionnelle . . 48
2.3 Comparaison à l'aide de deux demi-transformées : bon appariement avec Ψ^+, pas d'appariement avec Ψ^- 49
2.4 Les résultats de la transformée en ondelettes continue et d'une transformée en quart d'ondelettes sur une image ligne (ondelette utilisée : DOG) 49
2.5 Image aléatoire (image droite uniquement) et carte de disparité associée 50
2.6 Résultats sur une paire d'images de synthèse : l'image gauche d'origine, la carte de disparité avec la méthode de Perrin, la carte de disparité avec la transformée en quart d'ondelettes 50
2.7 En bleu, les zones où l'erreur est supérieure à un pixel, en rouge les zones où aucun appariement n'est trouvé 51
2.8 Principe des algorithmes de Shape From Silhouettes, d'après [28] 52
2.9 Principe de la reconstruction 54
2.10 Dispositif d'acquisition conçu et réalisé pour l'acquisition d'objets contenus dans un cube de 200mm de côté 56
2.11 La statuette sur le plateau tournant prêt pour la modélisation 57
2.12 Modèles lissés et non lissés de la statuette 57
2.13 Règles de comportement : une nuée d'oiseaux évitant un obstacle 61
2.14 Iowa Driving Simulator : extrait de l'automate de contrôle du conducteur virtuel 62
2.15 Contrôle en robotique et en animation 63
2.16 Architecture d'une entité comportementale 65
2.17 HPTS : architecture 66
2.18 HPTS : automate de fonctionnement 67
2.19 HPTS : exemple de fonction d'intégration 68
2.20 Outils de conception de simulation de l'équipe Siames . . . 72
2.21 VUEMS : modélisation de l'environnement urbain 72
2.22 Architecture du conducteur virtuel 73
2.23 Intégration du tramway dans le carrefour de la Croix-Bonneau à Nantes 74
2.24 automate de supervision d'un feu tramway 74
2.25 Projet DIATS : déploiement de moyens télématiques pour l'optimisation des infrastructures autoroutières à fort trafic (à droite le *ramp metering*) 75
2.26 Architecture d'une simulation type pour le projet DIATS . . 77
2.27 Scénario de calibrage pour le projet DIATS 77
2.28 Comparaison de différents dispositifs AICC sur le scénario 2 78

3.1 Les 5A des SIG, d'après [39] 80
3.2 Vers une boucle fermée entre SIG3D+t et images : réalité augmentée en milieu urbain – les images sont utilisées pour représenter et enrichir le SIG (1) mais le SIG est également utilisé pour géo-localiser les images (2). 80
3.3 Capture d'écran de l'outil OrbisGIS 82
3.4 Utilisation de différents types de maillages dans urbSAT . . 84
3.5 Premier résultat de calcul de poses possibles sur une paire de façades 86
3.6 Second cas plus standard 86
3.7 Détermination de la position et de l'orientation d'une image dans une base de données de panoramas 88
3.8 Exemple factice d'annotation en réalité augmentée 89
3.9 *Fond de carte* généralisé 90
3.10 Rendu expressif d'environnements virtuels urbains (©iParla [144]) 93
3.11 Cartes augmentées : localisation robuste à partir de la distribution des intersections 95

Introduction générale

Contexte et enjeux

Contexte professionnel

Les travaux relatés dans ce document se sont déroulés dans 3 laboratoires distincts dans lesquels j'ai eu la chance de travailler :

- tout d'abord l'équipe SIAMES (devenue Bunraku) de l'IRISA (Unité de Recherche Bretagne Atlantique de l'INRIA) de 1995 à 1999 ;
- ensuite, l'équipe «Réalité virtuelle et Réalité augmentée» du Centre de Robotique des Mines - Paristech de 1999 à 2002 ;
- enfin le CERMA, membre de l'IRSTV depuis septembre 2002.

Naturellement, chacun de ces laboratoires a eu une incidence sur mes travaux grâce à la spécificité de son projet de recherche. Néanmoins, ceux-ci, commencés dans le domaine de l'animation par ordinateur, s'inscrivent tous dans le cadre du lien existant entre les images et un système d'information. Le lien se fait en utilisant les techniques de la réalité virtuelle et en utilisant essentiellement la modalité visuelle. Le système d'information considéré représente dans la plupart des cas un objet (et le plus souvent un produit de l'homme) appartenant au passé, au présent ou au futur. Ce produit peut prendre différentes formes allant du produit manufacturé en grande série à un concept abstrait comme un aménagement de carrefour. Dans ce dernier cas, le système d'information est relatif à un territoire, on parle alors de système d'information géographique.

Enjeux de société

Une «société de l'image» ? Ce titre provocateur tiré d'un numéro hors-série du magazine «Sciences Humaines»[1] illustre un dossier qui se demande si l'image a acquis un statut privilégié dans notre société. En effet, les moyens de sa création et de sa consommation n'ont jamais été si nombreux. Certes, l'image considérée ici va bien au-delà de la notion d'image numérique qui nous intéresse directement, mais ce dossier met en valeur la prééminence de l'image dans nos sociétés développées et s'intéresse également à leur création, à leur compréhension et au lien qu'elles peuvent avoir avec le monde qui nous entoure. Bien sûr, la pertinence de ce lien et son rôle sociétal sont largement évoqués.

Une société de l'instantané Parallèlement à cette consommation d'images, les rythmes de nos sociétés se sont considérablement accélérés dans tous les domaines, y compris bien sûr celui de la production/consommation des images : la photo argentique a progressivement

[1] Sciences humaines : le monde de l'image. Hors-série 43, déc. 2003, jan-fév. 2004

remplacé la peinture figurative, elle-même a été supplantée par l'image numérique. Il n'est plus concevable aujourd'hui qu'un événement se produise quelque part sans la présence de caméras diffusant instantanément des images à l'ensemble de notre planète. Cette accélération se ressent partout, la concurrence entre entreprises s'est accrue et la réduction des coûts et du fameux *time-to-market* ou temps de mise sur le marché d'un nouveau produit sont aujourd'hui le cheval de bataille de toute entreprise de production. Ce changement s'est fait grâce à l'introduction du numérique – de l'informatique – à tous les étages de la production. Cela passe bien évidemment par les logiciels de CAO mais aussi l'utilisation des techniques de la réalité virtuelle.

Une société numérique La plupart des informations, qu'elles soient relatives à un produit, un homme ou un territoire sont aujourd'hui numériques ou en voie d'être numérisées. Le débat fait rage sur l'opportunité de cette numérisation et de l'interconnexion des bases de données ainsi constituées (syndrome *Big Brother*) mais aussi sur la représentativité de ces données numériques. Les images numériques sont aujourd'hui si puissantes qu'elles permettent de mélanger des vues du monde réel avec des éléments purement synthétiques de façon quasi-instantanée sans qu'il soit possible de faire la différence. Certains effets spéciaux du cinéma ne sont pas détectables pour un œil non-averti. Cela conduit évidemment les gens à s'intéresser à la crédibilité de ce qui est représenté sur l'image.

Une société plus informée La numérisation des informations ainsi que la mise à disposition rapide des informations via les images dont l'accès a été facilité par le développement de la télévision et d'Internet fait qu'aujourd'hui la population de notre société a accès à une quantité d'information beaucoup plus importante qu'auparavant. Notre société est plus informée et plus rapidement, mais pas toujours mieux. Le législateur a également veillé à augmenter la mise à disposition des informations. Ainsi un projet d'aménagement d'un espace de taille significative doit aujourd'hui recueillir l'avis du public. Malheureusement, de par l'immense quantité d'information disponible, la difficulté à rendre intelligible certaines données techniques pour le public, l'intérêt de certaines parties à n'en avoir qu'une vision partielle et partiale, le débat public n'est pas toujours de qualité.

Positionnement de notre travail

Notre travail se positionne à l'intersection des points évoqués plus haut. Nous nous intéressons au lien entre l'image numérique et un ensemble de données numérisées dans un cadre de conception de produit centré sur l'usager. La notion de produit est à prendre ici au sens large, nous considérerons aussi bien les produits manufacturés en grande série que les produits uniques comme un aménagement de l'espace. Les images constituent une représentation d'un produit présent, passé ou futur. Les informations concernant ce produit sont disponibles sous forme numérique. Notre travail s'intéresse donc aux questions suivantes :

- Comment créer des images à partir des informations numériques disponibles sur un produit ?
- Comment s'assurer que l'image du produit est bien fidèle à celui-ci ?
- Inversement, comment créer des informations numériques à partir d'une image ?

Du système d'information à la réalité virtuelle

Un système d'information se définit comme un « ensemble organisé de ressources permettant d'acquérir, de stocker, de structurer et de communiquer des informations». Ces informations sont liées au monde réel qu'elles représentent. Lorsqu'elles sont relatives à un endroit de l'espace, on parle alors de *système d'information géographique* (SIG). Un des objectifs des SIG est la production de cartes. Ces cartes sont des représentations calculées d'un SIG visualisées sous forme d'images. Ces images sont le plus souvent bi-dimensionnelles et font appel à une sémiologie graphique, c'est-à-dire à la représentation symbolique graphique la mieux adaptée à la communication d'une information.

Les systèmes d'information modernes comportent aujourd'hui des informations beaucoup plus riches que celles qui peuvent être représentées par des cartes : Les ingénieurs conçoivent les plans d'un objet manufacturés, les urbanistes ceux d'une ville... Grâce à la Conception Assistée par Ordinateur (CAO), ces plans sont maintenant représentés sous forme numérique 3D. L'image de synthèse, fixe ou animée, en est bien sûr le représentant le plus fréquent. Cette image constitue alors l'interface entre le système d'information entre les êtres humains qui vont concevoir, renseigner et consulter ce système d'information. Les données contenues dans les logiciels de CAO constituent aujourd'hui des systèmes d'information extrêmement riches puisqu'ils sont également capables de simuler le fonctionnement d'un objet, voire de permettre une interaction avec l'objet à travers les techniques de la réalité virtuelle.

En paraphrasant Fuchs [56], on peut alors considérer la réalité virtuelle comme «une expérience sensorimotrice et cognitive d'un ou plusieurs êtres humains dans un système d'information à travers des interfaces» et en déduire que ces interfaces forment le lien entre l'utilisateur et le système d'informations. En réalité, le lien ne se fait qu'avec une partie du système d'information (en base de données on parlerait de *vue*) qu'on appelle alors monde virtuel. Nous restons cohérents avec la définition technique donnée dans ce même ouvrage qui considère que la réalité virtuelle est un moyen technique qui permet de changer de lieu, de temps et de type d'interaction.

En nous référant au schéma de la figure 1, nous pouvons constater que l'être humain s'interface via un monde virtuel avec le Système d'information. Ses actions sur le monde virtuel peuvent bien évidemment modifier celui-ci mais peuvent aussi engendrer des modifications sur le système d'information. Bien évidemment, les modifications du SI ne sont pas forcément liées à l'environnement virtuel, elles peuvent être effectués directement.

En début de paragraphe, nous avons fait une analogie avec les Systèmes d'information géographique. Denègre et Salgé [39] définissent les

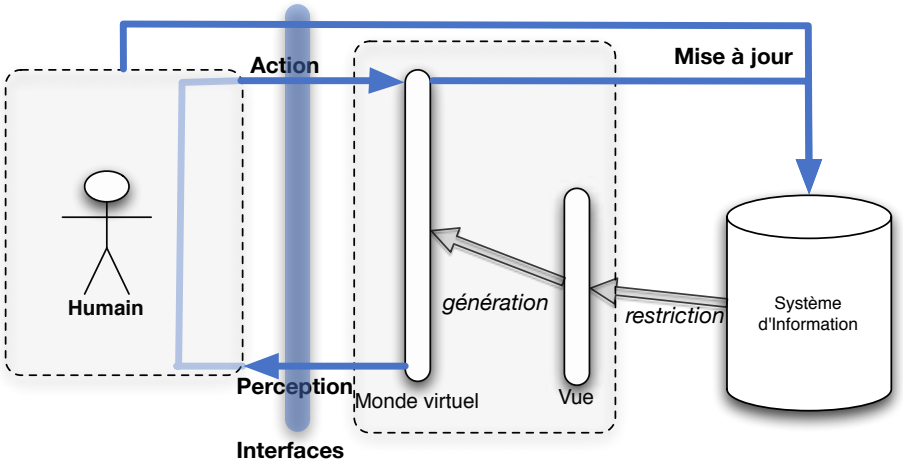


FIG. 1 – *Interfaçage entre un être humain et un système d'informations via la réalité virtuelle*

capacités des SIG par les 5A : abstraction, archivage, affichage, analyse et acquisition. Ces capacités sont reprises figure 2. Nous proposons de reprendre cette classification afin de voir comment elle s'adapte au problème de l'interfaçage avec un système d'information en réalité virtuelle.

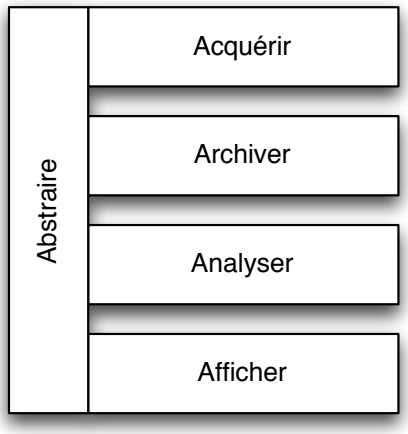


FIG. 2 – *Les 5A des SIG, d'après [39]*

LES 5A

Abstraction : dans un SI(G), il s'agit du modèle associé à l'information, sa structuration comme des modèles liés à l'analyse de données ou à leur représentation ; grâce à sa généralité, nous proposons de conserver cette notion de façon similaire ;

Archivage : cela concerne le stockage et l'organisation des informations de façon persistante. Dans la mesure où cette partie est spécifique

au SI, elle n'est pas directement concernée par les problématiques de l'interfaçage ;

Analyse : L'analyse dans un SI peut être vue comme un ensemble de traitements appliquées à des données afin de construire de nouvelles informations qui seront ensuite archivées et/ou affichées. L'intérêt de la réalité virtuelle est de replacer l'homme au cœur du SI et ainsi s'appuyer sur ses capacités d'analyse. Cette partie est donc moins pertinente dans notre cadre, même si on peut l'envisager sous la forme suivante : des traitements peuvent être appliqués au SI afin de générer le monde virtuel.

Acquisition : dans les SI(G), l'acquisition recouvre le plus souvent deux sous-problèmes : l'ajout d'information provenant du monde réel à travers leur conversion vers le modèle de représentation défini ainsi que la mise à jour de ces informations lorsqu'elle ne coïncident plus avec une perception du monde réel. Dans notre cadre, nous prenons en compte cette conversion d'informations réelles vers le système de représentation interne du SI (le modèle) ;

Affichage : dans un environnement virtuel, c'est bien sûr la partie la plus souvent mise en exergue. L'affichage consiste non seulement produire l'environnement virtuel à partir des données disponibles mais également à s'intéresser à la manière dont il sera perçu par l'observateur. Comme nous l'avons vu précédemment, il y a des restrictions importantes entre le SI et la représentation qui en est faite à l'observateur. Il est fondamental de s'intéresser à la manière dont l'homme perçoit l'environnement virtuel (et donc le SI) à travers les interfaces.

Problématiques

Au vu de la classification précédente, nous allons maintenant définir les grandes problématiques abordées par nos travaux.

Perception des environnements virtuels

La plupart des données contenues dans les systèmes d'information ne sont pas représentables intégralement dans un seul environnement virtuel. Ce dernier n'est pas non plus intégralement représenté à l'observateur, ne serait-ce que par les limitations technologiques des artefacts que constituent les interfaces. On peut les voir comme des filtres dont on ne connaîtrait pas les paramètres. Il est donc nécessaire d'aborder les questions suivantes :

- Comment perçoit-on un environnement virtuel à travers des interfaces ? Cette question s'aborde à plusieurs niveaux :
  - d'abord au niveau psychophysique et technologique : par exemple, si l'on connaît les limites physiques de la vision humaine, on peut en déduire les capacités minimales que les interfaces devraient avoir pour ne pas être un facteur limitant ;
  - cependant, il est facile de constater que d'une part, les limites de la technologie sont souvent très inférieures aux limites perceptives de l'être humain, mais d'autre part, on constate également les très

importantes capacités d'adaptation de l'esprit humain et sa capacité à s'immerger dans un monde virtuel dans des conditions d'interfaçage très loin d'être optimales. La question de la perception peut alors être abordée sous un angle cognitif.
  - comme l'évaluation cognitive de la perception est difficile, il est intéressant de se placer au niveau fonctionnel ou comportemental en s'interrogeant sur le fait que l'être humain dans un environnement virtuel y effectue bien la tâche qu'il est supposé y faire. On peut alors dire que l'interprétation qu'il fait de sa perception est *suffisant* du point de vue fonctionnel, même si cette dernière n'est pas optimale du point de vue cognitif ou sensori-moteur.
- Les systèmes d'information tels que les environnements de CAO sont des outils destinés à représenter une réalité passée, présente ou future. La seconde série de questions posées autour de la perception peut s'énoncer ainsi : comment peut-on comparer la perception d'un SI représentant un objet particulier (passé, présent ou futur) et sa perception réelle ? autrement dit, dans le domaine de la CAO par exemple, qu'est-ce qui garantit que le design d'un objet validé en virtuel sera perçu de façon identique une fois construit ? De façon similaire au problème précédent, cette question peut s'aborder à différents niveaux : psychophysique bien sûr mais aussi sensible au sens où l'on aborde cette fois la dimension subjective de la perception.

Ces différents niveaux de perception sont repris sur la figure 3.

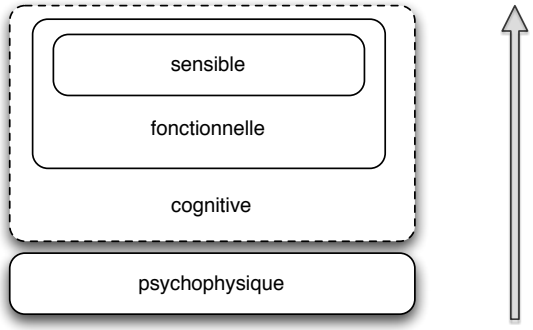


FIG. 3 – *Les différents niveaux de perception des mondes virtuels étudiés, par degré de complexité*

Les réponses à ces questions ont bien entendu une incidence sur la création des environnements virtuels dans lesquels seront immergés les utilisateurs.

Création des environnements virtuels

Le premier schéma général de la figure 1 pose les questions suivantes :
- Comment créer des environnements virtuels à partir des SI ? il s'agit ici de recréer des environnements virtuels qui soient compatibles avec l'immersion de l'observateur de façon à rester dans le cadre de perception énoncé ci-dessus. Il va de soi qu'une automatisation

maximale de la création des environnements virtuels est un objectif important.
- La création automatique de l'environnement virtuel est une première étape, mais il s'avère encore plus intéressant d'alimenter directement le SI. Si l'on s'appuie sur l'hypothèse de la résolution de la question précédente, l'alimentation du SI est un processus plus général que la création d'environnements virtuels. La question posée porte donc plus sur la création des éléments constitutifs du SI à partir du monde réel.

De la même manière que l'interfaçage avec un environnement virtuel se fait le plus souvent à l'aide d'images, nous nous intéressons également à l'utilisation automatisée des images du monde réel pour la création des environnements virtuels et pour l'alimentation des systèmes d'information. En quelque sorte, cette partie constitue la partie duale de la perception des mondes virtuels évoquée plus haut. Il aurait été possible de l'appeler *perception du monde réel pour la construction des mondes virtuels*.

Vers une boucle fermée entre images et Système d'information

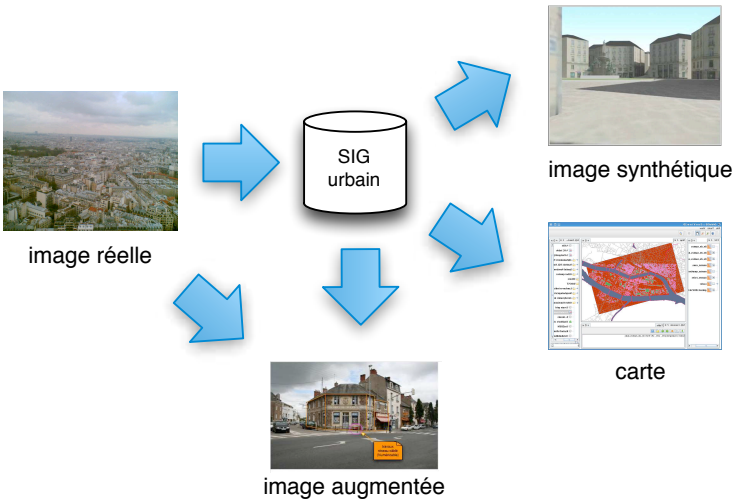


FIG. 4 – *Exemple de lien entre images et SI : le cas du SIG urbain*

Considérant que les techniques de la réalité virtuelle sont utilisées pour accéder au Système d'information essentiellement à travers des images de synthèse et que l'environnement virtuel peut être créé partiellement à partir de ces images (cf. Figure 4), il est pertinent de chercher à fermer cette boucle en considérant que le Système d'Information peut lui-même être créé ou modifié à partir d'images réelles.

Nous avons également abordé la problématique de la fidélité fonctionnelle et sensible d'un environnement virtuel représentant un SI. Une hypothèse intéressante pour améliorer cette fidélité sensible est de mélanger les informations du monde réel et celles du système d'information. C'est le domaine de la réalité mixte et plus spécifiquement de la réalité augmentée (cf. figure 5. Par rapport à la réalité virtuelle, la réalité augmentée introduit

quelques problèmes scientifiques supplémentaires comme la localisation de l'observateur et le rendu graphique des informations du SIG.

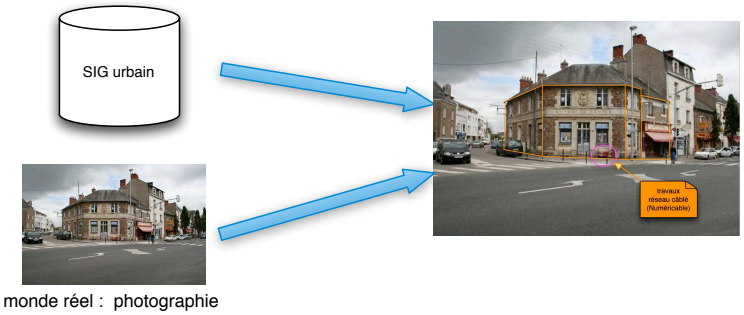


FIG. 5 – *Principe de la réalité augmentée en environnement urbain : surimpression d'informations du SIG sur des images réelles*

Dans le cadre essentiellement architectural et urbain qui est le nôtre maintenant, nous considérons maintenant uniquement des Systèmes d'Information Géographique (SIG). Comme nous l'avons mentionné en début de paragraphe, les données contenues dans les SIG sont relatives à un endroit de l'espace ; c'est cette localisation qui permettra de faire le lien avec la réalité augmentée. Les informations 2D d'un SIG classiques ne sont pas forcément suffisantes : la réalité augmentée permet de faire le lien avec un SIG 3D. De plus, la mise à jour d'un SIG devant se faire continûment, nous souhaitons la faire directement sur le site grâce à la réalité augmentée et être capable de prendre en compte la notion du temps en allant vers un SIG3D+t sur site.

Dès lors, nos travaux abordent ici les problématiques classiques des SIG abordées plus tôt (les 5A) généralisées aux SIG 3D+t :

- Abstraction et Archivage : comment modéliser, stocker et accéder à ces informations 3D+t ?
- Acquisition et Analyse sont intimement liées par notre hypothèse de boucle fermée : l'analyse des images permettra l'enrichissement du SIG à condition de connaître la position de ce qui aura été détecté dans l'image ;
- Pour superposer les images des mondes réel et virtuel et pour enrichir le SIG 3D+t, il est indispensable de connaître la position de l'observateur dans le monde réel avec une bonne précision. Nous faisons l'hypothèse que cette position peut être retrouvée à l'aide des images et du SIG lui-même ;
- Enfin, l'affichage d'informations provenant d'un SIG 3D+t suscite d'autres questions :
  - Comment sélectionner, agréger des informations hétérogènes en nature et en échelle variant dans le temps ?
  - Comment afficher ces informations dans un environnement virtuel qui pose déjà des soucis d'appropriation ?

Classification des travaux

Nos travaux cherchent à répondre aux problématiques abordées dans le paragraphe précédent. Celles-ci ont été abordées de façon non chronologique et un certain nombre de travaux sont évidemment en cours afin de compléter la cohérence de notre activité de recherche. Le diagramme 6 présente l'organisation temporelle des travaux abordés dans la suite de ce document.

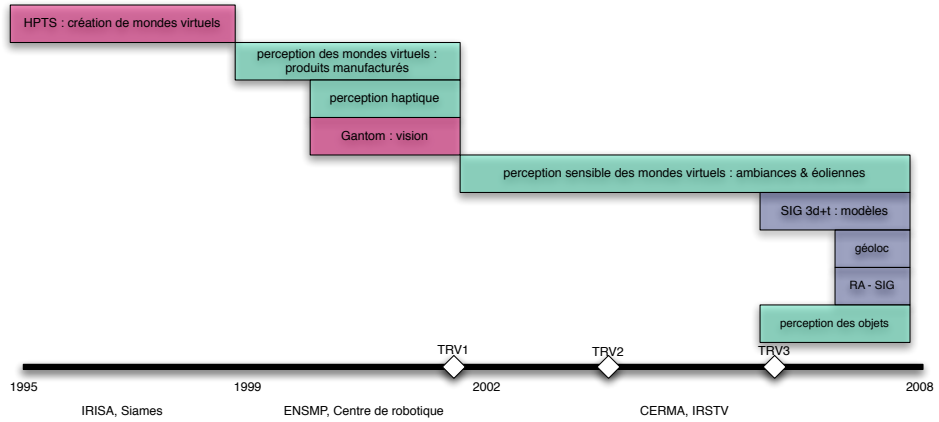


Fig. 6 – *Chronologie des travaux*

Organisation du document

Dans la suite de ce document, nous n'avons pas cherché à présenter les choses de façon chronologique mais plutôt à les organiser autour des grandes problématiques énoncées plus haut. De fait, certaines actions pourraient se retrouver dans plusieurs classes de problèmes, certains choix ont été effectués pour éviter une forme de redondance.

La suite de ce texte comporte alors 3 grandes parties : la première est consacrée à la perception des environnements virtuels selon des niveaux de *complexité* croissante. D'abord effectuée selon des critères psychophysique, l'analyse met progressivement en jeu les différents niveaux de perception de l'être humain.

Les résultats obtenus dans cette partie sont partiellement utilisés pour la construction d'environnements virtuels. Nous nous intéresserons dans cette seconde partie d'une part à des algorithmes liés à la vision par ordinateur utilisés pour construire des objets virtuels à partir d'objets réels, mais aussi à la reproduction comportementale de certains objets (en l'occurrence du couple humain+véhicule) pour les insérer dans des environnements virtuels crédibles.

Nous verrons que ces deux premières parties mettent à profit implicitement ou explicitement la notion de système d'information géographique. Dans la troisième partie, nous aborderons des travaux spécifiquement liés aux SIG et aux images. Cette dernière partie vise à fermer la boucle entre le système d'information géographique et la monde réel via la modalité image. Elle est à la fois un résumé d'activités actuelles (thèses en cours

notamment) et un projet de recherche s'appuyant sur les travaux évoqués dans les deux premières parties.

1 Perception des mondes virtuels

Les travaux évoqués dans ce chapitre s'intéressent à la fidélité de la représentation d'un système d'information à travers un monde virtuel. La question sera abordée par complexité croissante, du niveau psychophysique au niveau sensible : d'abord, nous nous intéresserons à l'interface de restitution et aux contraintes qu'elles imposent. Ensuite, nous serons à même déterminer comment sont perçues les formes géométriques puis les objets qu'elles représentent. Cette étude sera menée sous un angle fonctionnel : dans quelle mesure l'objet conçu sera-t-il à même de remplir sa fonction une fois fabriqué ? Enfin, nous finirons par l'étude de la perception non pas d'un objet mais d'un monde virtuel complet représentant de très nombreux objets et en y incluant la nécessaire dimension subjective de cette perception. Les paragraphes suivants sont donc ordonnées par niveau d'implication perceptive croissante de l'observateur de cet environnement virtuel (cf. figure 1.1). D'un point de vue méthodologique, ces différents travaux sont menés sur le principe d'une comparaison entre un environnement réel et une forme approchante en environnement virtuel. Les détails de construction de cet environnement virtuel approchant sont donnés pour chaque partie.

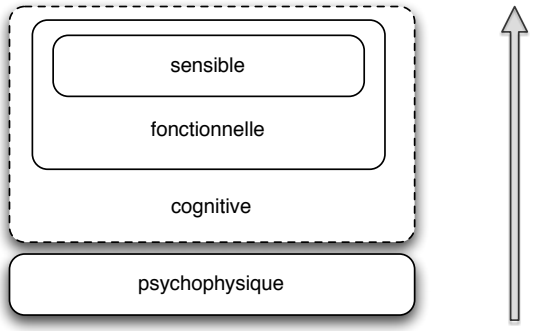


Fig. 1.1 – *Les différents niveaux de perception des mondes virtuels*

1.1 Perception physique : la vision stéréoscopique

Cette première section s'intéresse non pas tant aux problèmes de la compréhension des environnements virtuels mais plus directement à la reproduction d'un signal visuel sur une interface de restitution comme un

écran. Après une brève introduction au système visuel humain et aux problèmes de la vision stéréoscopique, nous aborderons successivement le problème de la fatigue liée à cette vision stéréoscopique puis celui de la perception des formes. Pour commencer, nous évoquerons rapidement le contexte de ces travaux.

1.1.1 Contexte : le projet IHS10

Le projet IHS10 a pour but de créer un intérieur de voiture virtuel avant de le construire. La vision en relief est indispensable pour les constructeurs, les concepteurs et les ergonomes : elle leur permet de bien apprécier l'intérieur, d'une façon similaire au réel. Ils pourront ainsi mieux apprécier les distances, les dimensions et les formes. De cette façon, ils peuvent s'apercevoir très vite si une pièce n'est pas là où elle devrait être. Les personnes qui utiliseront cet outil travailleront pendant plusieurs heures. L'outil ne devra provoquer que très peu de contraintes oculaires sous peine d'être abandonné. De plus, il faut s'assurer que la forme validée dans le monde virtuel correspond bien à une forme identique dans le monde réel.

Certaines causes de fatigue visuelle peuvent être évitées grâce à une disposition et une projection adéquate de l'image (cf. 1.2.2). D'autres peuvent être évitées avec un traitement particulier de l'image. C'est l'objet de cette partie qui présente nos premiers travaux et résultats dans le domaine. Nous commencerons par exposer les difficultés liées à la vision stéréoscopique, les moyens de mesurer la fatigue oculaire liée à ces difficultés. Enfin, nous verrons quelles sont les pistes d'amélioration possibles en utilisant différentes techniques et en nous appuyant sur une étude complémentaire sur la perception comparée des formes en environnements réel et virtuel.

1.1.2 Fatigue visuelle et problèmes de perception en vision stéréoscopique

Pour mieux percevoir en relief les dimensions et les volumes corrects des objets, il faut que la vision restituée soit isomorphe à la vision du monde réel lorsque c'est techniquement possible. On parle alors de vision orthostéréoscopique : elle doit respecter la perception de la profondeur par les lois de la perspective et par les lois de la stéréoscopie. Ceci implique les conditions suivantes :

- l'écartement entre les deux caméras virtuelles doit être égal à la distance interoculaire (DIO) ;
- la distance de tout objet observé par rapport aux caméras doit être égale à la distance de son image par rapport à l'observateur.

L'apport de la stéréoscopie est important en réalité virtuelle pour la perception de petits détails ou pour la perception de la profondeur dans des zones où il y a peu d'indices monoculaires. Malgré tout, de nombreuses études la considèrent comme une cause importante de fatigue visuelle [189, 184, 169, 159, 160]. Les symptômes rapportés incluent maux de tête, fatigue des yeux, vision confuse voire trouble, nausées la vision en double. Les causes de ces fatigues sont les suivantes :

Mauvaise qualité de l'image Les problèmes apparaissant déjà en vision monoscopique seront tout aussi apparents en vision stéréoscopique. Il est important qu'aucun matériel/logiciel du système qui produit les images n'induise de déformation ou de modification de couleur. De plus, dans le cas d'utilisation de lunettes actives, les projecteurs doivent être suffisamment rapides pour éviter les images fantômes.

Propriétés différentes entre image gauche et image droite Aucune différence de propriété entre les deux images ne pourra être tolérée. Elles peuvent être de multiples ordres : luminosité, forme, temps (problème de synchronisation des images), hauteur, orientation...

Contradictions entre indices de profondeur Les contradictions entre les indices stéréoscopiques et les autres indices de relief sont néfastes. Sur les petits écrans, représenter un objet *sortant* de l'écran trop proche du bord est une erreur courante. Le cerveau n'arrive plus à savoir si l'objet est devant le bord de l'écran. De même, tous les indices de lumière et de texture doivent être pris en compte correctement lors de la génération des images.

Rupture de la relation entre accommodation et convergence La relation convergence-accommodation peut être mise en défaut artificiellement lorsque nous regardons des images stéréoscopiques sur un écran : un objet virtuel positionné en arrière de l'écran fait converger les axes optiques sur lui tandis que les yeux accommodent sur l'écran (figure 1.2). Cette modification de la relation convergence-accommodation est susceptible de créer des difficultés visuelles et de la fatigue pour l'observateur, sauf évidemment quand l'objet se situe au niveau de l'écran.

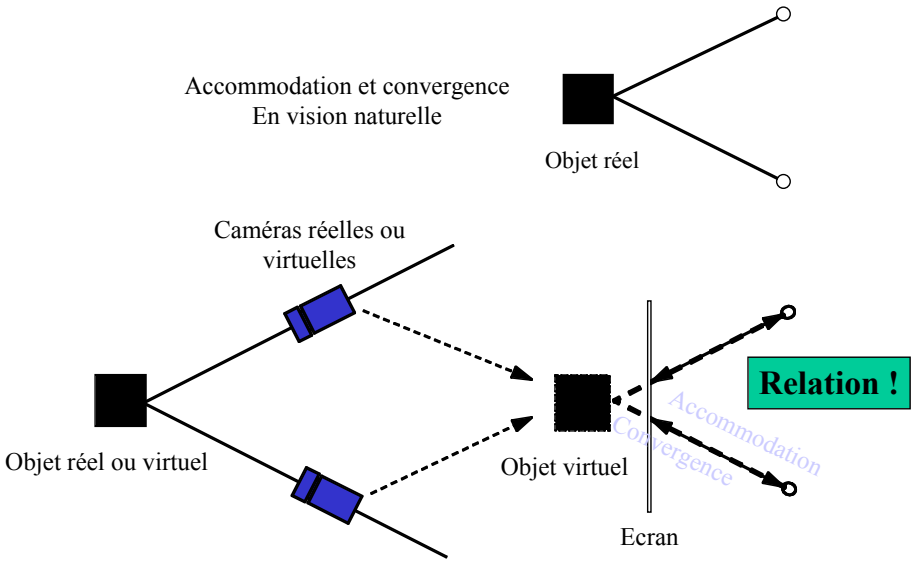


FIG. 1.2 – *Différence entre vision naturelle et vision stéréoscopique*

Disparités horizontales trop importantes Les disparités horizontales ne doivent pas être trop importantes : sinon, soit le cerveau est incapable de fusionner les images (diplopie), soit il le fait au prix d'une fatigue importante. Certaines études expérimentales [182] ont montré que la parallaxe horizontale devait être limitée à 1,2 voire 1,5 degrés pour un stimulus de deux secondes en parallaxe positive. Les conditions sont plus restrictives pour des objets en mouvement rapide. Elle peut également dépendre du contenu fréquentiel de l'image (cf. supra).

Disparités verticales Comme nous le verrons en 1.2.2, les parallaxes verticales sont créées dès que les axes des caméras ne sont plus parallèles (ce qui correspond à des yeux regardant à l'infini). Les parallaxes verticales sont beaucoup plus difficile à supporter pour l'œil : en pratique, une parallaxe verticale supérieure à 20′ d'angle est très difficile à supporter.

Mouvements pseudoscopiques Lorsque l'observateur se déplace parallèlement à un écran et qu'il n'a pas la tête traquée (et donc pas d'adaptation des images à sa nouvelle position), il voit les objets se déplacer en sens opposé à son déplacement, comme le montre la figure 1.3 : les points affichés à l'écran restent fixes. A noter qu'en cas de déplacement dans la direction de l'écran, c'est la taille perçue de l'objet qui sera modifiée. Si les images sont mises à jour parce la tête de l'observateur est traquée, un décalage entre le mouvement de la tête et l'affichage effectif de l'image supérieur à 75ms induit un risque de nausée ou de vertige [106].

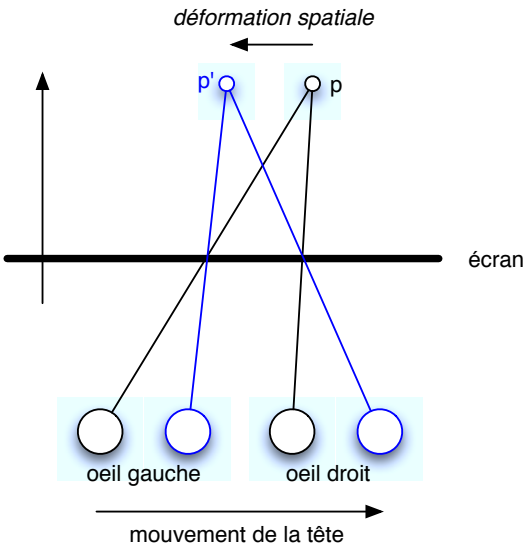


FIG. 1.3 – *Mouvement pseudoscopique parallèle à l'écran*

Le contenu de l'image Roumes [151] a montré que lorsque les fréquences spatiales dans les images sont trop importantes, celles-ci deviennent plus difficile à fusionner. Se basant sur les travaux de Wöpking [188] dont les résultats sont reproduits figure 1.5, Perrin [139] a établi une fonction de confort liant le confort $C(f,d)$, les disparités horizontales

d et les fréquences spatiales f (figure 1.4) :

$$C(d, f) = a(d - d_0 - kf^{k'}) \tag{1.1}$$

avec $a = -0.010, d_0 = 18.9, k = 221.1, k' = -0.74$. D'autre, il a également été montré que l'aire de Panum (zone où la stéréopsie est possible) dépend du contenu fréquentiel de l'image [72]. Cela corrobore les résultats de Wöpking et Roumes.

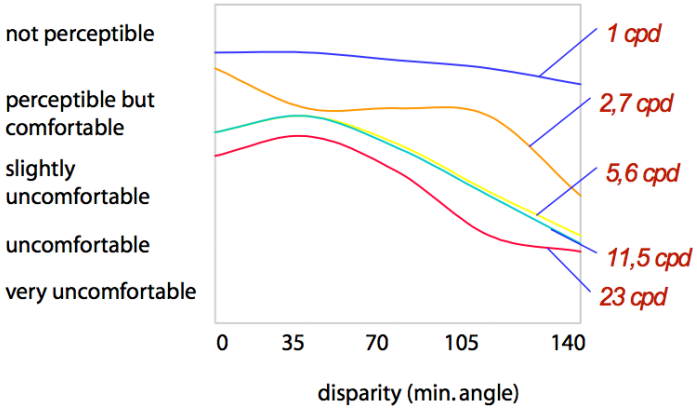


FIG. 1.4 – *Fonction de confort de Wöpking [139] en fonction de la disparité et du contenu fréquentiel de l'image*

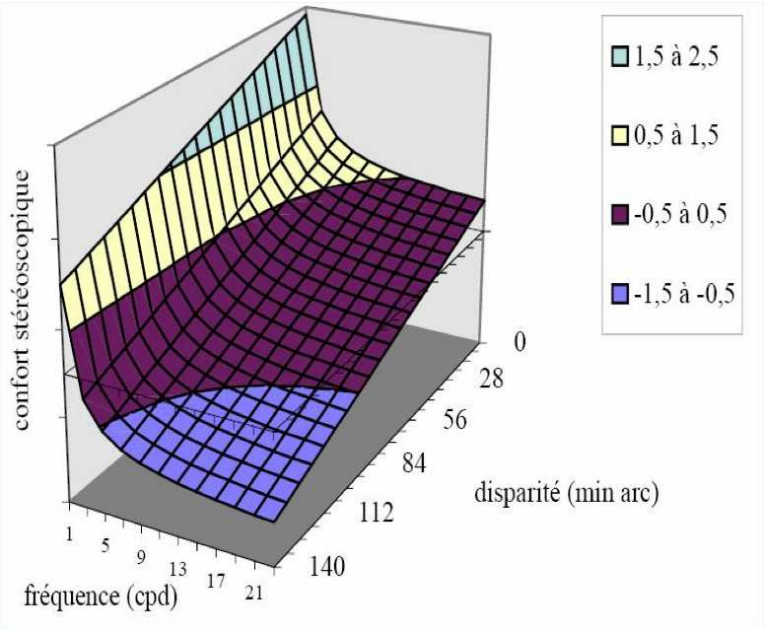


FIG. 1.5 – *Fonction de confort de Perrin [188]*

Comme souvent lorsque qu'on rencontre des phénomènes psychophysiologiques, l'enjeu des travaux de recherche porte autant sur la mesure du phénomène que sur l'amélioration des valeurs de cette mesure. La mesure de la fatigue visuelle est alors un phénomène sur lequel il est difficile de créer une mesure.

1.1.3 Tests de mesure de la fatigue visuelle

Il existe deux grands types de tests de mesure de la fatigue visuelle : les tests objectifs basés sur des mesures optométriques et des tests à base de questionnaire. Ces tests doivent être effectués avant et après expérimentation stéréoscopique afin d'avoir un élément de comparaison. Les grands types de tests objectifs sont les suivants :

- L'acuité monoculaire est définie comme la séparation angulaire minimum entre deux lignes qui est nécessaire pour qu'elles soient perçues comme étant séparées. Elle diminue avec la fatigue ;
- L'hétérophorie est une déviation pathologique des globes oculaires qui n'apparaît que lorsque la vision des yeux est dissociée. La relation entre l'hétérophorie et les modifications de la scène sont mesurées. Allié à d'autres tests, cela constitue un indicateur de fatigue visuelle [34].
- La fixation des disparités utilise la formation de l'image sur la fovéa de l'œil pour mesurer le stress visuel. Elle ne peut être utilisée seule [34].

1.1.4 Stratégies d'amélioration du confort

Il existe différents types d'algorithmes dont le but est l'amélioration du confort. Ils peuvent être classés selon les paramètres qu'ils contrôlent. Une première classe cherche à limiter l'apparition de valeurs de disparités inconfortables tandis qu'une seconde classe s'intéresse à la modification du contenu fréquentiel des images. Le travail sur les disparités peut s'effectuer de plusieurs manières [56] :

- en éliminant les disparités verticales : Julesz [81] a montré que les disparités verticales sont plus difficile à fusionner que les disparités horizontales. L'élimination des disparités verticales peut s'effectuer en utilisant des caméras parallèles et en effectuant un décalage des images [120] comme nous l'avons fait dans le projet IMAVE ;
- en déterminant le point de convergence des yeux et en faisant en sorte de n'avoir aucune disparité en ce point : si l'observateur s'intéresse à un objet unique dans un environnement vide, le point de convergence des yeux s'induit rapidement, mais c'est nettement plus ardu dans un environnement complexe et le processus nécessite de coûteux appareils supplémentaires. Notons que cette technique peut être combinée avec la précédente ;
- en éloignant l'observateur de la scène, ce qui réduit automatiquement la disparité (en rendant de fait les caméras quasiment parallèles). Si c'est une bonne idée théorique, elle est souvent inapplicable en raison du besoin d'échelle 1 :1 exprimé en réalité virtuelle.

La seconde classe de méthodes cherche à obtenir un effet de profondeur de champ (généralement absent des images de synthèse temps réel) : en pratique, un objet trop éloigné ou trop proche du point de convergence sera flouté de façon à limiter les hautes fréquences lorsque la disparité est importante. Blohm [18] a utilisé le tampon d'accumulation de la carte graphique pour simuler un effet de profondeur de champ. Perrin [139, 140] a proposé 3 techniques respecter la fonction de confort de Wöpking : le rideau virtuel (les zones hors critères de confort sont simplement grisées),

le calque virtuel (les mêmes zones sont floutées), le flou adaptatif (adaptation du contenu fréquentiel local).

Ces méthodes ne sont pas exclusives les unes des autres : l'utilisation de caméras parallèles est facile (voir 1.2.2), le suivi de la direction du regard est évidemment plus complexe en raison de la complexité de la chaîne mise en œuvre (matériel, tracking et prise en compte sur les images).

1.1.5 Approche par les ondelettes et perspectives

Suivant le principe énoncé par Perrin dans sa thèse [139], dans leurs stages respectifs que j'ai encadrés, Boudet et Lemmer ont utilisé la technique de la transformée en ondelettes pour respectivement extraire des cartes de profondeur à partir d'une paire d'images (cf. chapitre 2) et améliorer le confort spatial : Perrin a montré que la transformée en ondelettes pouvait être utilisée comme une méthode de localisation des hautes fréquences spatiales dans les images. En liant cette information à la disparité locale entre deux images, on peut en déduire la valeur de la fonction de confort en un point et proposer plusieurs techniques d'amélioration de la fonction de confort.

Malheureusement, les algorithmes utilisés par Perrin, conjugués à la puissance de calcul des PC de l'époque rendait ses approches inutilisables dans le cadre de la réalité virtuelle. Il n'a pu les appliquer qu'à des images fixes. Avec Lemmer [95, 96], nous avons proposé des améliorations à l'algorithme, tout en rendant son exploitation parallèle et en limitant les calculs aux fréquences les plus pertinentes [22]. Conjuguées à l'utilisation d'un matériel évidemment plus performant en 2003 qu'en 1998, nous sommes arrivés à des temps de calcul très inférieurs : 6 secondes pour une image 512x512 et 12 secondes pour une image 1024x512. Néanmoins, ces résultats restent encore 300 fois supérieurs au temps réel.

Grâce à la thèse de Laure Leroy (débutée en octobre 2006 et financée par le projet IHS10), nous avons pu ouvrir de nouvelles pistes de recherche en abordant deux questions : la perception des formes est-elle la même dans les environnements réels et virtuels et peut-on limiter la fatigue dans les environnements immersifs stéréoscopiques ? La première question est abordée au paragraphe 1.1.6, pour la seconde (en cours d'étude), les grandes tâches définies sont les suivantes :

- Qualifier et mesurer la fatigue visuelle en vision stéréoscopique : comme nous l'avons vu précédemment, la mesure de fatigue visuelle est peu adaptée à nos besoins. Pour cette phase, nous allons travailler tant avec des spécialistes de la psycho-physique de la vision humaine qu'avec des ergonomes connaisseurs du domaine de la réalité virtuelle afin d'évaluer la charge spécifique due au travail en environnement stéréoscopique ;
- Étudier la perception de la profondeur en vision stéréoscopique : pour cette étape, nous allons effectuer des tests faisant le plus possible appel à des indices de profondeur binoculaires en utilisant des images dénuées (autant que possible) d'indices monoculaires. Ainsi, nous allons pouvoir mesurer les plus petite différences perceptibles entre deux représentations stéréoscopiques d'un même volume, ce

qui nous donnera un cadre pour modifier les images afin d'améliorer le confort visuel. Cette partie est abordée au paragraphe 1.1.6 sur la perception des formes ;
- Enfin, améliorer le confort visuel dans un cadre réalité virtuelle : pour une amélioration d'un facteur 300 des temps de calcul actuellement rencontrés, la loi de Moore ne suffira pas. Les pistes d'amélioration portent sur quatre thèmes : la parallélisation des algorithmes existants, l'utilisation de nouveaux algorithmes, l'utilisation des capacités de calcul des cartes graphiques (GPU) et enfin la réduction de l'espace de travail des algorithmes utilisés.

Les travaux en cours abordent essentiellement la question des nouveaux algorithmes, en parallèle du problème de leur adaptation aux capacités de calcul des GPU. En particulier, nous étudions l'algorithme SURF [10]. Celui-ci parait prometteur en raison de sa rapidité (largement temps-réel sur de grandes images), de sa capacité à travailler sur des zones d'intérêt plutôt que sur l'image entière, mais aussi sur des gammes de fréquence limitées. En première approximation (par quantification des différences, formalisation en cours), SURF donne les mêmes résultats que l'utilisation des ondelettes de Haar. En revanche, travaillant sur qu'on appelle des images intégrales, sa capacité de parallélisation est largement inférieure à celle des algorithmes utilisant des ondelettes. De même son implantation sur carte graphique par le biais des *shaders* est compromise par le fait que les opérations *shaders* appliquent le même type d'opération sur chaque pixel, ce qui ne correspond guère au principe des images intégrales.

Ces travaux ne peuvent être décorrélés de la partie consacrée à la perception des formes elles-mêmes. En effet, à quoi cela servirait-il d'améliorer le confort en vision stéréoscopique s'il s'avérait que cela induise une modification des formes perçues d'une part, ou pour aller plus loin, que l'objet conçu dans l'environnement virtuel ne *ressemble* pas à l'objet une fois celui produit.

1.1.6 Perception des formes

Beaucoup d'études se sont intéressées à la perception des tailles relatives des environnements virtuels, mais peu se préoccupent directement de la perception des formes et de l'influence des divers paramètres de la vision stéréoscopiques sur cette perception. Nous nous sommes attachés à étudier l'influence de ces paramètres sur deux études de cas où nous comparons une forme réelle et son équivalent virtuel [97]. Ces deux études portent sur une sphère et sur une forme dite inconnue (figure 1.6) au sens où elle ne ressemble à aucun objet connu et ne présente pas d'indice monoscopique évident. Les paramètres que nous avons fait varier au cours de ces expériences sont les suivants :
- position relative de l'écran et de l'objet virtuel ;
- réglages des paramètres de vision stéréoscopique ;
- utilisation d'un traqueur de position de la tête ;

Les expériences se sont déroulées avec 18 personnes pour chaque forme en utilisant la méthode des stimuli constants : le sujet doit juste dire si les stimuli présentés dans un ordre aléatoire sont identiques en taille.

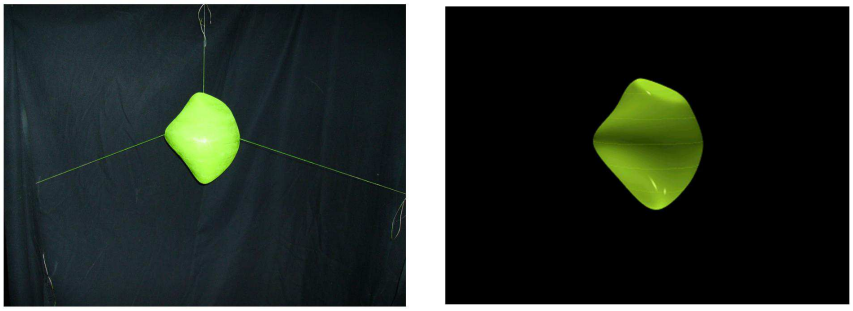


FIG. 1.6 – *Forme inconnue réelle et virtuelle*

test #	IPD stéréo	tracking	position
0	normale	oui	écran
1	normale	non	écran
2	normale	oui	derrière
3	normale	non	derrière
4	normale	oui	devant
5	normale	non	devant
6	mono	oui	écran
7	mono	non	écran
8	2X	oui	écran
9	2X	oui	derrière
10	2X	oui	devant

TAB. 1.1 – *Séquences de tests*

Nous nous appuyons pour cela sur les contraintes de mise en œuvre définis dans [20]. 7 variations de 0.6cm sont utilisées à chaque fois. L'organisation des séquences de test est présenté dans le tableau 1.1. Les protocoles sont adaptés en fonction de la séquence : par exemple, le sujet a le droit de bouger lorsque sa tête est traquée mais doit rester sur un repère fixe au sol en l'absence de tracking.

A partir des données collectées, on calcule le PSE (Point d'Egalité Subjective), c'est-à-dire le ratio entre la taille de la forme virtuelle et celui de la forme réelle ainsi que les JND [57] (Différences juste remarquables) soit la plus petite différence perceptible entre deux formes. La figure 1.7 montre notamment que la perception des différences de taille est meilleure lorsque la sphère est devant l'écran, ce qui est assez intuitif. Le PSE est quant à lui très proche de 1.

De façon générale, les observations faites sur la sphère se retrouvent de façon amplifiée sur la forme inconnue. C'est notamment le cas pour l'évaluation de l'influence de la position de la forme comme le montre les graphes de la figure.

La principale conclusion de cette étude est que l'utilisation du tracking de la tête de l'utilisateur est le paramètre le plus important de tous ceux que nous avons fait varier. A noter que la position par rapport à l'écran a également une influence et notre évaluation peut être biaisée par le fait que nous cherchons pas à évaluer des tailles absolues d'objets (relativement) connus mais bien des tailles relatives en l'absence d'autres indices. Nous

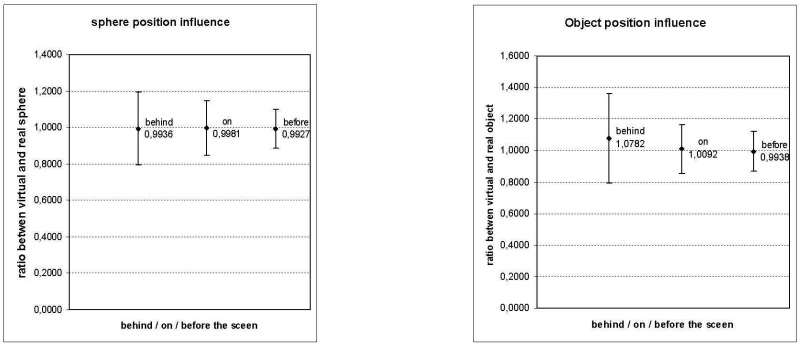


FIG. 1.7 – *Influence de la position relative par rapport à l'écran (à gauche pour la sphère, à droite pour la forme inconnue)*

allons continuer ce travail en faisant varier cette fois non plus la taille de la forme inconnue mais ses formes elles-mêmes. L'objectif est de faire modifier la forme virtuelle modifiée afin de retrouver la forme réelle. Les performances seront évaluées en fonction des différences modes de rendu stéréoscopiques. Comme la tâche utilisée pour la mesure sera elle-même de durée significative, nous couplerons cette étude à des mesures sur la fatigue visuelle liée à l'utilisation de la vision stéréoscopique, en fonction des résultats obtenus dans nos études sur l'amélioration du confort visuel.

1.1.7 Conclusion

Il serait vain d'aborder la question de la fidélité des environnements virtuels sans avoir envisagé les conditions dans lesquels se font la restitution et la vision des mondes virtuels eux-mêmes. Dans cette première partie, nous avons montré les limites de la perception sur un système comme un écran (quelle que soit la nature de l'écran). Nous avons également étudié le problème de la perception des formes dans cette environnement de restitution. Connaissant ces limites, nous allons maintenant aborder la perception de ce que les formes servent à constituer, c'est-à-dire des objets.

1.2 Perception des objets virtuels : revue de projet et ergonomie

La seconde partie de ce chapitre est consacrée à la perception des objets manufacturés n'existant que sous forme de maquette numérique, c'est-à-dire d'éléments du système d'information. Cette partie est un peu particulière puisqu'elle ne s'intéresse pas qu'au problème de la perception visuelle mais elle aborde également la restitution haptique.

1.2.1 Introduction

L'application principale de la réalité virtuelle en conception des produits manufacturés est la revue de projet. Dans [119] et [120], nous nous intéressons à la revue de projet chez un concepteur de planches de bord automobile. Les fonctions identifiées (liées aux décisions prises lors de la revue) sont les suivantes :

- visualisation des différentes solutions de conception à l'échelle 1 en situation de conduite ;
- visualisation des jeux entre surfaces plastiques qui conditionnent la qualité perçue de la planche de bord ;
- vérification de la non-visibilité des zones techniques (par exemple les vis de montage) ;
- vérification d'accessibilité à la fois des organes de commande mais aussi des zones techniques évoquées ci-dessus.

L'objet de cette étude était de montrer que la revue de projet porte sur une évaluation non-numérique du prototype, par opposition aux logiciels de CAO qui effectuent des simulations numériques. Cette évaluation *qualitative* est néanmoins effectuée à l'aide de la maquette numérique. A cette fin, nous avons conçu le système IMAVE (IMersion Adaptée à un VEhicule) avec les exigences suivantes :

- vision identique à celle d'un conducteur ou d'un passager, stéréoscopique en raison de la proximité de la planche de bord et du niveau de détail requis ;
- échelle 1 ;
- tête traquée et vision adaptée aux mouvements de l'utilisateur ;
- résolution adaptée à la visualisation des jeux entre pièces (pixel < 1mm) ;
- position de conduite adaptable ;
- manipulation simplifiée des commandes grâce à un capteur de localisation ;
- éclairage variable et contrôlable.

Le système est alors composé d'un grand écran rétro-projeté, situé devant un siège de voiture comme le montre la figure 1.8. Nous avons préféré l'utilisation de l'écran rétro-projeté au casque après avoir montré que le simple positionnement d'un casque sur la tête d'un observateur n'était pas une opération répétable [117].

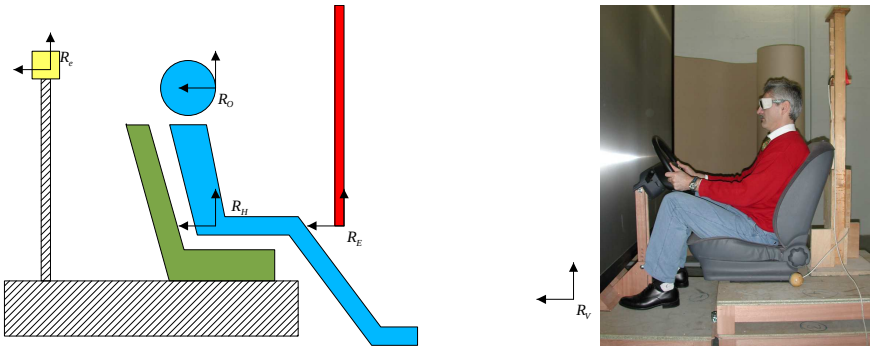


FIG. 1.8 – *Le dispositif IMAVE*

1.2.2 Génération des images en vision stéréoscopique

La visualisation stéréoscopique a fait l'objet d'un traitement particulier que nous allons résumer dans les paragraphes qui suivent[1]. Pour atteindre les objectifs d'IMAVE, nous nous sommes fixés les contraintes suivantes :

C_1 L'objet d'intérêt principal doit avoir une parallaxe très faible (idéalement nulle) afin de limiter la rupture entre accommodation et convergence ;

C_2 La parallaxe horizontale est limitée à 1.5 degrés [182] ;

C_3 Il n'y a pas de parallaxe verticale, ce qui implique des caméras parallèles ;

C_4 Les dimensions de la planche de bord dans les espaces réels et virtuels sont identiques, l'écran doit alors être un peu plus large que la planche de bord.

Pour respecter C_1, nous avons choisi de positionner la planche de bord au niveau de l'écran et d'utiliser un volant réel. Nous pouvons alors calculer les images stéréoscopiques : Pour un point (x, y, z) de l'espace 3D dans le repère de la caméra (le repère déterminé par le barycentre du repère des deux caméras), on détermine aisément les points images I_l et I_r correspondants :

$$I_l = \left(\frac{f\left(x + \frac{e_c}{2}\right)}{z}, \frac{fy}{z} \right), I_r = \left(\frac{f\left(x - \frac{e_c}{2}\right)}{z}, \frac{fy}{z} \right) \tag{1.2}$$

où f est la focale de la caméra et e_c la distance entre les caméras. Si on ne translate pas les images obtenues, tous les objets seront vus avec une parallaxe négative, c'est-à-dire devant l'écran, en contradiction avec C_1. La parallaxe actuelle est donnée par (1.3) :

$$P_h = -\frac{fMe_c}{zd_0} \tag{1.3}$$

où d_0 représente la distance entre l'écran et l'observateur et M le facteur d'échelle (égal à 1 d'après C_4). En translatant les images d'une distance D_h, la nouvelle parallaxe est donnée par (1.4) :

$$P_h = -\frac{fMe_c}{zd_0} + \frac{2D_h}{d_0} \tag{1.4}$$

Nous connaissons maintenant z_{p_0} la profondeur du plan de parallaxe nulle et sommes à même de calculer les maxima de parallaxe en fonction de la scène (qui ne contient que la planche de bord). En utilisant C_4, nous en déduisons que :

$$P_h^{max} = e_c.(\frac{1}{z_{p_0}} - \frac{1}{z_{max}}) = e_c.(\frac{1}{d_0} - \frac{1}{z_{max}}) \tag{1.5}$$

$$P_h^{min} = e_c.(\frac{1}{z_{p_0}} - \frac{1}{z_{min}}) = e_c.(\frac{1}{d_0} - \frac{1}{z_{min}}) \tag{1.6}$$

[1]Ces travaux ont été menés avant les travaux sur le confort visuel que nous avons évoqués en début de ce chapitre. Ces derniers ont été placés avant dans le document en cohérence avec le schéma général de la figure 1.1, page 11

où $z_{min} < 0$ est la distance entre les yeux du conducteur et l'avant de la planche de bord. $z_{max} > 0$ est égal à z_{min} augmenté de la profondeur de la planche de bord. Il est alors possible de déterminer d_0 d'avoir des valeurs identiques pour P_h^{max} et P_h^{min}. Ce n'est pas le choix que nous avons fait, partant du principe que l'observateur regarde plus l'avant de la planche de bord que l'arrière, qui est plus difficile à voir. Compte tenu des autres paramètres de l'application, les valeurs obtenues sont $P_h^{min} = -1.3°$ et $P_h^{max} = -2°$, ce qui n'a pas posé de problème aux testeurs. D'autres travaux ont été menés afin d'adapter la vision stéréoscopique lorsque l'observateur de rapproche de l'écran. L'ensemble est présenté plus en détail dans [118].

Les résultats d'évaluation montrent que, sur une centaine de testeurs, seuls 8% demandaient des améliorations avant mise en service (auxquels il faut probablement ajouter les 8% de sans opinion). De façon générale, les demandes d'améliorations (y compris celles des personnes prêtes à utiliser le système en l'état) portaient sur la représentation des panneaux de porte (conçus et fabriqués en même temps que les planches de bord), la visualisation des grains du plastique (problème extrêmement difficile puisque le comportement du grain sur les surfaces gauches n'est que rarement connu et que celui ne présente généralement pas de symétrie, ce qui complexifie énormément la difficulté de la représentation informatique) et le calcul des reflets.

1.2.3 Retour d'effort pour les études d'ergonomie

Le système IMAVE présenté dans les paragraphes précédents a permis de mener des études d'accessibilité simples (peut-on toucher telle commande ?). Les logiciels d'ergonomie vendus dans le commerce sont aussi capables de calculer de tels résultats à partir des bases de données anthropométriques dont ils sont dotés. Pour autant, ils présentent des limites (au-delà du fait qu'ils transforment des sensations de confort en valeurs numériques appréhendables uniquement par des spécialistes) :

- les collisions avec des objets mobiles et les collisions en plusieurs points ne sont pas prises en compte : il n'est pas possible de toucher un objet pour en atteindre un autre (par exemple : appuyer un doit sur la planche de bord pour atteindre un petit bouton de l'autoradio) ;
- la dynamique n'est généralement pas prise en compte, l'analyse posturale est purement statique, dans le meilleur des cas, les mouvements sont simplement calculés par cinématique inverse ;
- l'apprentissage et l'appropriation des logiciels d'ergonomie est un processus long et complexe ;
- enfin, les ergonomes pensent que la seule validation ergonomique possible se fait en conduisant réellement.

Un autre inconvénient majeur de ces logiciels est l'absence de sensations et notamment de retour haptique qui rend impossible la validation de la sensation de confort associé aux commandes d'une planche de bord. C'est dans ce contexte que nous avons proposé un projet autour du retour d'effort dans les études ergonomiques avec le CEA et PSA Peugeot-Citroën dans le cadre de la plate-forme PERF-RV du RNTL.

Pour cela, nous avons cherché à vérifier que la sensation haptique (le retour d'effort et pas le tactile) ressentie au niveau d'un doigt appuyant sur un bouton dont on peut modifier à la fois la course et la raideur peut être reproduite. Ce paramétrage permet de comparer une configuration nominale avec d'autres en nous appuyant sur la notion de JND (Just Noticeable Differences) définie dans [57]. Si l'on trouve des travaux sur des expériences réelles [170, 80], peu s'intéressent à la comparaison entre les environnements réels et virtuels : Tzafestas [177] propose une expérience mais uniquement dans le mode virtuel. Notre expérience [85], dont le tout premier dispositif est présenté figure 1.9 présente l'avantage de n'adresser que des variables d'expérimentation très limitées. La dispositif a ensuite évolué pour utiliser un système à retour d'effort un axe dédié, le système 6 axes dont 3 motorisés créant trop de vibrations pour être exploitable.

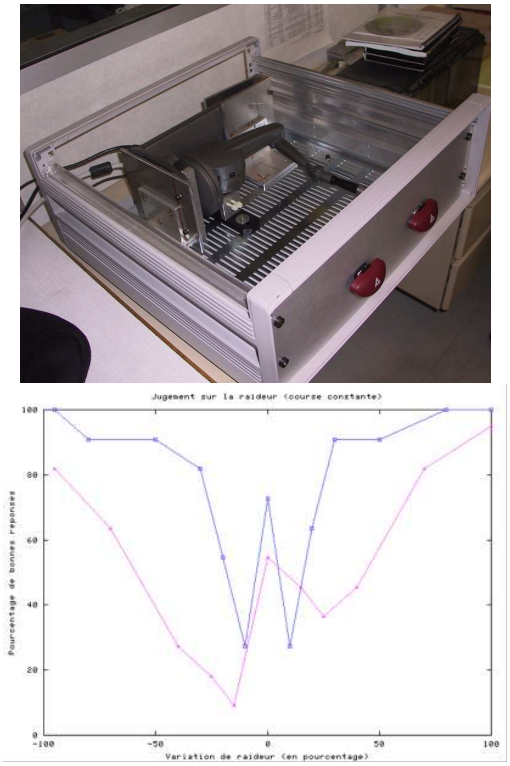


FIG. 1.9 – *Tout premier dispositif de mesure de la discrimination en raideur et en course d'un bouton de commande de planche de bord (à gauche) et pourcentages de bonnes réponses en discrimination en fonction de la variation de raideur pour deux courses différentes (à droite)*

Cette étude, menée pendant le stage stage de DEA de M. Benali Khoudja [84] a abouti aux conclusions suivantes :

- pour une variation de course, plus la course nominale est faible, moins les résultats sont reproductibles ;
- La discrimination en raideur est meilleure qu'en course ;
- La variation conjuguée de la raideur et de la course est très difficile à détecter ;
- Le manque de retour visuel est une des plus grandes difficultés (ceci corrobore les résultats d'A. Lécuyer sur le *pseudo-haptique* [105]) ;

– la fatigue des testeurs doit être prise en compte, la discrimination et le nombre de réponses correctes diminuant au cours du temps.

Ces travaux ont été suivis par une intégration de cette application dans le poste de conduite virtuel. Cette intégration s'est faite grâce à l'utilisation d'un exosquelette à deux doigts afin d'intégrer les retours d'efforts en couple dédiés aux commandes à mouvement rotatif.

1.2.4 Détection de collisions point-modèle

L'intégration des systèmes à retour d'effort a nécessité un travail sur la détection de collision entre un être humain réel et un environnement virtuel, première composante informatique du retour d'effort. En se basant sur l'hypothèse émise par Brooks [27] que la boucle de retour haptique doit fonctionner à une fréquence d'environ 1kHz, on peut considérer que cette partie est capitale dans le bon fonctionnement du système.

Nous nous sommes d'abord intéressés à la détection de collisions entre l'environnement virtuel et un point. Ce point peut représenter une extrémité d'une sonde comme le stylet d'un système à retour d'effort 3D comme le Phantom Desktop. Nous avons proposé un algorithme [162] séparant classiquement la détection du premier contact et le suivi de surface (qui est possible en temps constant, voir [99, 98]). Cet algorithme est composé de 3 étapes :

– une limitation rapide de la surface examinée : elle mixe un nombre limité de volumes englobants orientés (OBB [60]) et une grille régulière dont l'organisation est plus efficace (table de hachage) en conservant la notion d'objet. Cela permet de résoudre les problèmes d'objets en mouvement en convertissant les coordonnées de la sonde vers le repère du volume englobant de l'objet.
– Détection des polygones candidats pour l'intersection en utilisant la notion de point CP du modèle le plus proche de la sonde (il n'y a pas forcément contact puisqu'en temps discret, la sonde peut être passée à l'intérieur de l'objet).
– Détermination du polygone intersecté parmi les candidats en effectuant une recherche rapide par projection de CP sur les polygones candidats.

Une fois le premier contact déterminé, une méthode similaire à [70] est appliquée pour exploiter la continuité du mouvement. C'est également dans [70] qu'a été introduite la notion de *force shading* utilisée pour le rendu haptique.

L'algorithme se montre moins gourmand en mémoire que la méthode classique à base d'OBB [60] et limite surtout le temps de pré-calcul nécessaire. Néanmoins, il présente l'inconvénient de nécessiter une densité de triangles homogène pour fonctionner de manière optimale. Ceci pourrait être contourné par l'utilisation de grilles de résolutions différentes par OBB mais n'a pas été testé.

1.2.5 Détection de collision main-modèle

La détection de collision point modèle est limitée puisqu'elle ne permet le rendu de sensation qu'un point (en force uniquement). La détection

de collisions entre surfaces est un problème plus difficile du point de vue algorithmique. Dans le contexte des études d'ergonomie en environnement virtuel, nous nous sommes restreints à la détection de collisions entre la main et le modèle virtuel [161]. Popescu [142] se contente d'appliquer un modèle de détection point-modèle pour chaque point du maillage, d'autres approximent la main par des sphères [175] ou par des NURBS [127].

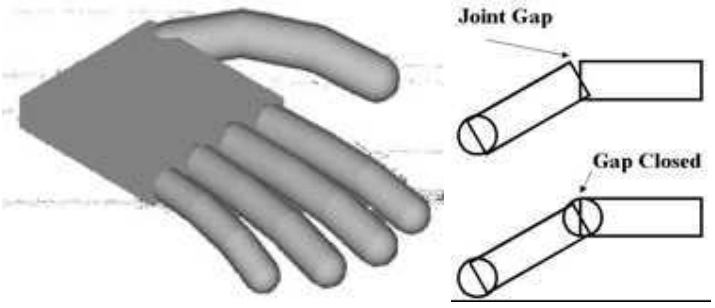


FIG. 1.10 – *Modèle géométrique de la main*

Dans notre approche, la main est représentée par un plan pour la paume (ne change rien) tandis que les phalanges sont représentées par des cylindres joints par des sphères pour éviter les trous comme le montre la figure 1.10. L'avantage de cette représentation est qu'on peut utiliser des algorithmes de précision différente pour les différentes parties de la main : on peut ainsi restreindre la précision maximale au bout des doigts qui sont l'endroit le plus sensible, les phalanges n'étant utilisées que pour les tests de collision et la préhension d'un objet. Pour la détection de collisions, nous n'utilisons que la partie visible des sphères jouant le rôle d'articulations.

Le schéma algorithmique que nous avons défini est présenté figure 1.11. Les résultats expérimentaux permettent de montrer que les forces et couples rendus sont continus et sans à-coups (figure 1.12). Une des limites de l'algorithme est que le temps de calcul est proportionnel au nombre de contacts. A l'époque de ce travail (2002), les temps de calcul permettaient d'obtenir un rendu haptique dans un environnement virtuel de 10000 triangles à 2kHz pour une quinzaine de contacts et à 700Hz pour une cinquantaine de contacts.

A mon départ du Centre de Robotique fin 2002, cet algorithme a été intégré dans le système de rendu haptique développé autour de l'exosquelette manuel conçu par Panagiotis Stergiopoulos pour sa thèse de doctorat.

1.2.6 Conclusion

Dans ces deux premières parties, nous nous sommes intéressés au niveau le plus bas de la perception des environnements virtuels : le niveau psychophysique. Nous l'avons abordé selon la modalité visuelle et selon la modalité haptique. Ces approches ont été validées au niveau fonctionnel en vérifiant la capacité de l'utilisateur à effectuer la même tâche dans l'environnement virtuel que celle qu'il aurait effectués dans un environnement réel. Dans la partie qui suit, nous allons généraliser cette notion à un niveau de perception incluant une part de subjectivité et prenant en

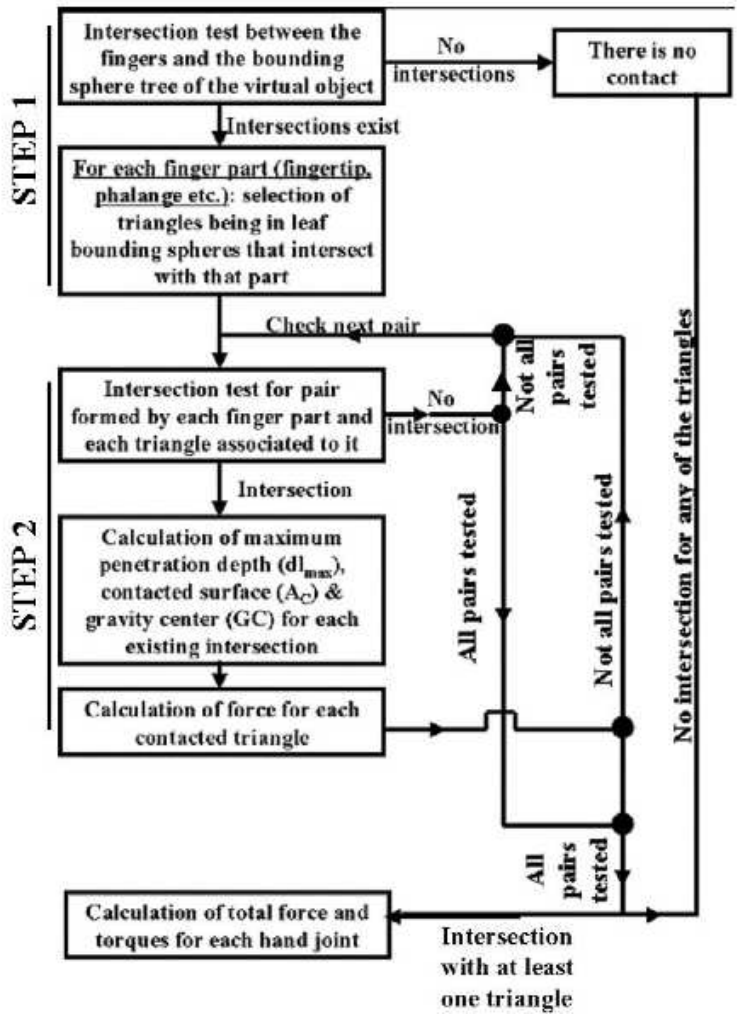


FIG. 1.11 – *Schéma algorithmique de détection de collisions*

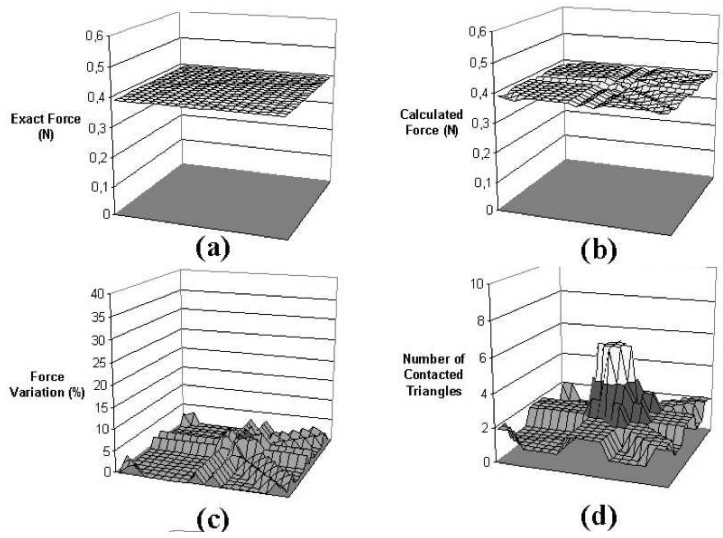


FIG. 1.12 – *Résultats expérimentaux : (a) force théorique, (b) force calculée, (c) variation de la force, (d) nombre de triangles en contact*

compte la dimension sensible de l'individu. Parallèlement, on notera que le domaine d'application passera des produits manufacturés en grande série à des produits humains de très grande taille produits en un exemplaire unique, mais qui dépendent largement de l'environnement dans lequel ils sont mis en œuvre.

1.3 Perception des mondes virtuels de grande taille

Tous les travaux cités dans les paragraphes précédents concernent le processus de conception des produits manufacturés en grande série. On peut remarquer que la taille de ceux-ci est du même ordre de grandeur que l'être humain qui les conçoit (1,45m en largeur pour une planche de bord, quelques centimètres pour un bouton de commande). Dans cette section, nous abordons, toujours dans le domaine de la conception, les problèmes liés aux environnements de grande taille (plusieurs centaines de mètres) et pour lesquels l'évaluation sensible est une question fondamentale : le contexte devient architectural, urbain voire paysager. Les autres différences portent sur le nombre d'exemplaires construits qui passent de la série à l'unité, le vécu des usagers (notamment à cause de la différence de durée de vie) et sur la problématique de l'évaluation de l'application au niveau fonctionnel.

Éléments méthodologiques

Du point de vue méthodologique, nous avons vu dans les sections précédentes que l'objectif de la revue de projet en environnement immersif était d'éliminer le plus tôt possible les erreurs de conception et de limiter (plutôt que de remplacer) le nombre de maquettes physiques. Cela permet, in fine, de valider aussi bien l'utilisation de la réalité virtuelle pour la revue de projet que le projet lui-même. A l'échelle urbaine, le problème ne peut pas être abordé de la sorte, faute de maquette possible à l'échelle 1. De plus, dans le cas particuliers des représentations virtuelles de l'espace urbain, il est clair qu'il existe une différence d'échelle majeure entre cette représentation et l'environnement réel dans lequel l'expérience de réalité virtuelle a lieu.

Nous proposons quelques éléments méthodologiques supplémentaires :

- l'étude des potentialités de la réalité virtuelle sera effectuée en comparant un environnement réel existant et un environnement virtuel construit pour l'occasion. Les similitudes et les biais de cette comparaison seront déterminés et validés sur une seconde étude dans un contexte différent ;
- la comparaison entre les deux environnements se fera à l'aide des outils classiques d'évaluation des projets dans le contexte concerné, en particulier dans le domaine sensible ;
- compte tenu des différences de taille entre le projet à évaluer et le lieu d'expérience, les méthodes utilisant l'hypothèse de l'enrichissement de la perception par le mouvement [58, 62] seront privilégiées.

1.3.1 Ambiances lumineuses urbaines

Contexte

L'espace urbain est un complexe déployant, à différentes échelles, les dimensions sociales, historiques, culturelles et politiques, etc. Il constitue un élément essentiel de la vie quotidienne puisqu'il constitue un espace d'échanges, de rencontre, de circulations, de relations et de communication [165]. Les ambiances urbaines (figure 1.13) naissent à la convergence des caractéristiques spatiales et physiques, au travers de l'interprétation humaine. La compréhension de l'espace urbain nécessite donc une approche immersive et dynamique qui s'appuie sur le vécu de l'usager. Nous avons proposé, à travers le travail de thèse de Souha Tahrani, d'aborder l'analyse des ambiances lumineuses par les techniques de la réalité virtuelle, en nous appuyant sur la notion d'effet solaire.

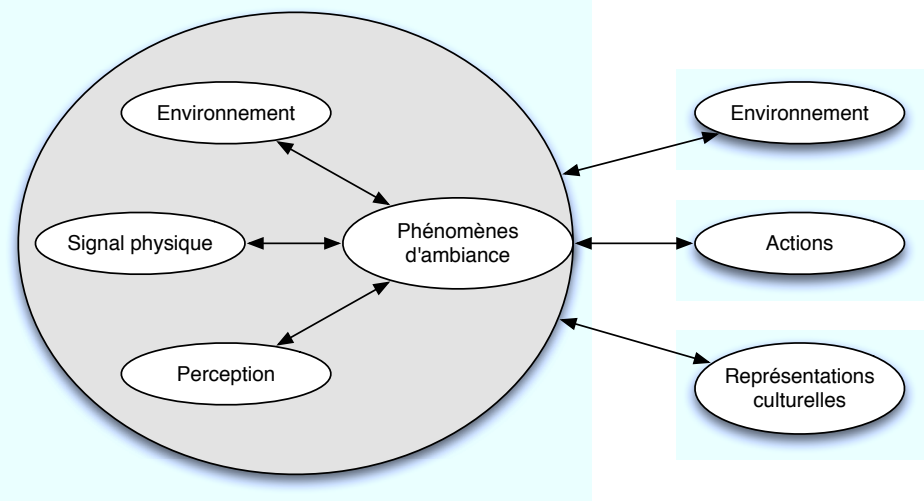


FIG. 1.13 – *Phénomènes d'ambiances*

Plusieurs études s'intéressent à la perception dynamique dans les environnements architecturaux et urbains et montrent que la marche peut être un outil de lecture et d'évaluation de l'espace public. Ces méthodes sont toutefois limitées à l'environnement bâti existant et donc impossibles à appliquer lors de la phase de conception.

Analyse de l'espace urbain

Nous nous intéressons ici au problème de l'évaluation de l'espace du point architectural et urbain en focalisant notre étude sur les aspects visuels et sur les apports de la perception en mouvement. Moles et Rohmer [113] définissent la psychologie de l'espace en disant que l'espace pur n'existe pas, qu'il n'existe que par la référence à un sujet, un groupe, un contenu ou un point de vue. Deux modes de perception sont alors employés :

– une vision cartésienne de l'espace qui est une vue d'un observateur extérieur ayant une vision illimitée où aucun point de vue ne peut être privilégié et tous sont alors équivalents. C'est le point de vue de la carte. Nous reviendrons sur cette notion dans le chapitre 3.

– une perception centralisée qui est en réalité le point de vue d'un individu immergé dans l'espace à un instant donné. Ce mode est plus adapté à la dimension sensible de l'être humain et se fonde sur les 3 éléments principaux que sont l'individu, l'espace et le temps. Il permet une prise en compte de la subjectivité.

La notion d'immersion est centrale dans notre étude, parce qu'elle constitue un des caractéristiques principales de la réalité virtuelle. Les *parcours urbains* ont été particulièrement étudiés dans la littérature et constituent une approche intéressante qui prend en compte l'aspect dynamique de la perception (et donc l'aspect temporel). Le temps est transformé, classiquement, en mouvement dans l'espace 3D, d'où cette notion de parcours. L'hypothèse faite par ces méthodes est que le meilleur moyen de comprendre une ville est de marcher le long d'une rue et de tout regarder (les arbres, la largeur des trottoirs et des rues, la signalisation, les façades, le mobilier urbain...) afin de s'en construire une représentation subjective qui évolue dans le temps (le parcours vécu et le parcours mémorisé). Les méthodes peuvent se résumer en deux grandes types :

– l'analyse picturale qui nécessite de représenter la ville sous forme de séquences de vues [180, 155] comme des peintures, des croquis ou même des photos. Cullen [37] développe une *conscience de l'espace* qui peut s'interpréter par la ville définie par un ensemble de points de vue.
– les notations graphiques : certains symboles visuels sont dessinés en fonction des émotions ressenties. Ils permettent de construire une image mentale de la ville [104]. Lynch affirme que les résidents perçoivent de façon identique (au sens du symbole) ce qu'ils vivent de façon similaire.

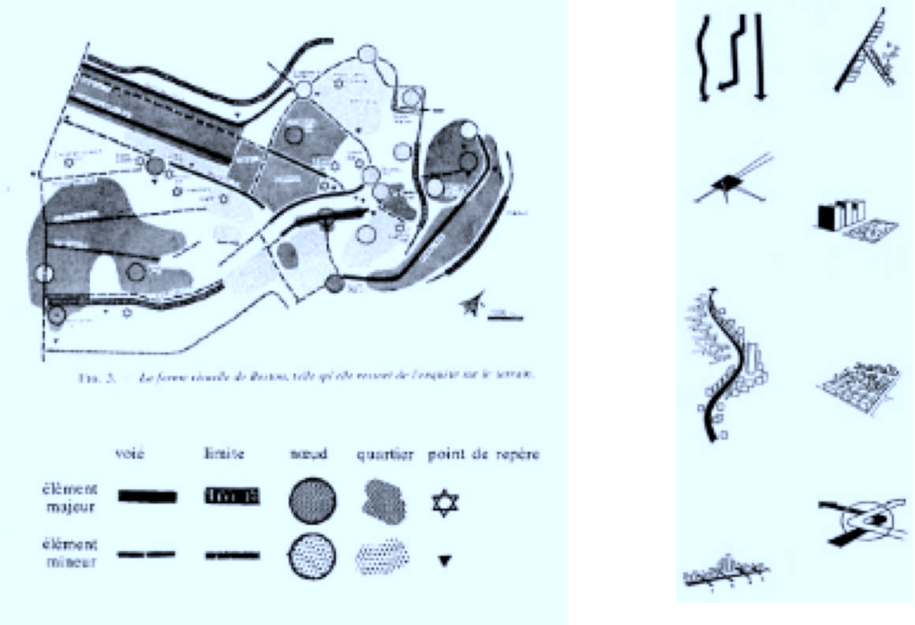


FIG. 1.14 – *La notation graphique des 5 éléments caractéristiques de la ville selon [104]*

La limite de ces deux types d'approche réside dans leur caractère discret : les séquences ne sont continues ni dans le temps ni dans l'espace.

Certains chercheurs ont alors créé des villes virtuelles qui sont de plus en plus plausibles. Cependant, la difficulté majeure liées à ces applications de plus en plus répandues et de plus en plus photo-réalistes réside dans le manque de mouvement naturel et/ou le manque de liberté pour l'observateur. L'approche sensible n'est généralement pas prise en compte.

Les approches sensibles prennent en compte de leur côté à la fois les paramètres physiques d'ambiance et l'analyse subjective des individus. Elles sont conçues pour étudier la cité à l'intersection de sa morphologie, des paramètres physiques et des facteurs personnels et culturels. Les travaux de recherche associés étudient les facteurs d'ambiance dans un contexte immersif et définissant les liens entre l'endroit analysé et son environnement. Thibaud et coll.[172] ont développé la méthode du parcours commenté pour exhiber la relation entre l'être humain, l'espace sensible et l'espace physiques dans leurs contextes respectifs. L'objectif est d'accéder à l'expérience sensible par la perception de l'espace public tout en bénéficiant de l'enrichissement de la perception par le mouvement. Ses principales étapes sont :

- les sujets suivent un parcours dans l'environnement urbain et décrivent ce qu'ils voient. Ils ont ici 3 types d'activité : marcher, percevoir et décrire ;
- L'ensemble des descriptions des sujets est synthétisée en un parcours unique qui est un produit intermédiaire de l'analyse ;
- Les phénomènes décrits par les sujets sont observés pour examiner les conditions de leur émergence. Les descriptions sont analysées et interprétées pour exhiber les propriétés sociales, spatiales et physiques des lieux.

Cette méthode est considérée comme un moyen inter-disciplinaire et dynamique d'étudier les ambiances lumineuses. Elles cherchent à mettre en place une lecture sensible de l'espace par ses propres usagers en étudiant leurs perception par rapport à leur environnement (et la manière dont ils le perçoivent, le décrivent et leur comportement dans celui-ci). Ces approches qualitatives se focalisent sur l'usager de l'espace urbain qui est à la fois l'acteur et le public de l'espace public. Les inconvénients majeur de ce type d'approche portent d'une part sur le fait qu'elles ne sont applicables qu'à des environnements existants et que le chercheur accompagnant le sujet lors de son parcours peut influencer celui-ci.

Les effets solaires

Le langage utilisé dans les méthodes sensibles, basées sur la description de ce qui est perçu par les usagers est très important. C'est pourquoi, le travail de thèse de Souha Tahrani a visé à formaliser cette description autour de la notion *effet solaire*. La notion d'effet en architecture n'est bien entendu pas neuve, elle a été considérablement discutée dans les études pour améliorer et comprendre l'esthétique des villes [155, 180, 37]. La notion d'effet d'ambiance a été proposée par Augoyard et Torgue [8] dans leurs travaux sur les*effets sonores*. De façon similaire, considérant que la vision humaine est conditionnée par la présence de lumière produisant des effets tant physiques que psychologiques, les effets solaires sont générés par la combinaison des 3 facteurs que sont les sources lumineuses

naturelles, l'espace (architectural et urbain) et l'observateur (perception visuelle, images mentales). En d'autres termes, la perception située des effets solaires (et des ambiances lumineuses) se fait en deux étapes 1.15 :

- l'une est physique : l'interaction entre la lumière et la morphologie urbaine est interprétée par la théorie de la forme *gestalt* à travers ses deux composants que sont le fond et la forme ;
- l'autre est sensible : le passage du phénomène physique vers le phénomène sensible interprété se fait par la construction d'une image mentale (une représentation se faisant via l'émergence des effets solaires).

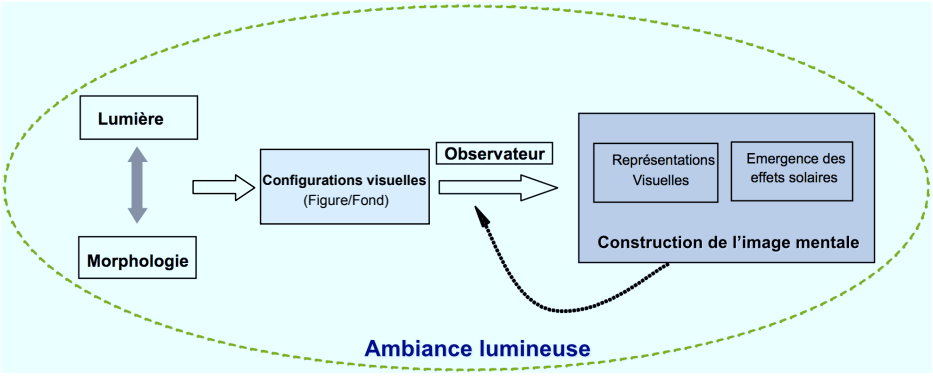


FIG. 1.15 – *Les processus de la perception visuelle située des ambiances lumineuses [165]*

Dans notre travail, nous utilisons les effets solaires comme un outil perceptuel pour l'interprétation de l'espace public. Par exemple, la figure 1.16 montre une observation d'une rue où nous avons relevé plusieurs effets comme une *ouverture* qui met en évidence l'ouverture de l'espace en fonction de la position de l'observateur, une *empreinte* qui met en évidence la forme urbaine et qui peut avoir une incidence sur la perception de l'échelle, et enfin un effet de *silhouette* qui donne une vision de la morphologie urbaine. Tahrani [165, 168] a par ailleurs proposé un répertoire et une classification détaillées de ces effets.

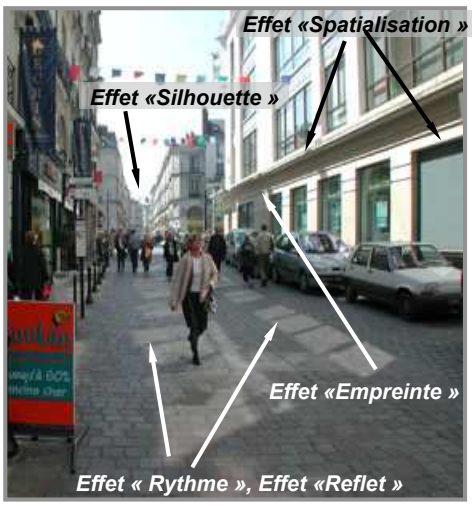


FIG. 1.16 – *Exemples d'effets solaires dans la rue de la Barillerie à Nantes*

Vers une application de réalité virtuelle pour l'évaluation des ambiances lumineuses

Pour pouvoir combiner les atouts des approches dynamiques (perception en mouvement) et de l'informatique qui permet d'appliquer ces méthodes dès la phase de conception (là où les méthodes classiques ne s'appliquent qu'à un environnement bâti), nous avons cherché à évaluer l'intérêt des techniques de la réalité virtuelle en y adaptant la méthode des parcours commentés. La conception de l'application (au sens évoqué dans [56]) montre rapidement que la principale difficulté provient de la marche du sujet compte tenu de la taille relative du parcours virtuel retenu (500m) et du dispositif expérimental. Les solutions permettant la marche à échelle 1/1 sont limitées ergonomiquement, mais sont aussi très complexes et difficilement abordables financièrement. On trouve notamment les tapis roulants bi-dimensionnels [181] où la cybersphere de l'Université de Warwick. .

Pour contourner cette difficulté et pour évaluer l'apport de la marche dans la méthode proposée, nous avons mis en place 3 expériences différentes :

- *l'une à partir d'un jeu d'images fixes* : cette pré-expérience a servi à déterminer des éléments de comparaison en faisant usage de photographies et d'images de synthèse fixes [164] ;
- *la seconde à partir d'un modèle numérique animé* : il s'agit d'un parcours totalement imposé d'un environnement virtuel projeté à l'échelle 1 sur un dispositif immersif. Aucune interaction n'est possible. La caméra virtuelle a été placée à 1,65m de hauteur avec un champ de vision horizontal de 61 degrés correspondant à un observateur placé à 1,5m d'un écran de 2,40m sur 1,92m. Cette situation a été considérée comme nettement plus immersive que la précédente en raison de la taille des images projetées et du champ de vision important (cf. figure 1.17). Bien que les observateurs aient regretté le manque de liberté de mouvement, cette expérience leur a néanmoins permis de se concentrer leur perception de l'espace et de sa lumière [166].
- *la troisième expérience* a cherché à améliorer la sensation d'immersion et la capacité de perception en introduisant une certaine capacité de mouvement. Comme nous l'avons montré dans les paragraphes précédents, la difficulté vient de la différence de taille entre l'environnement virtuel et la salle dans laquelle se déroule effectivement l'expérience. Pour cela, nous nous sommes appuyés sur le principe de la métaphore du guide virtuel défini en fondant notre choix sur le fait que la méthode des parcours commentés est basée sur la marche à pied et qu'il était préférable de ne pas modifier ce paramètre. L'idée est la suivante [167] : l'observateur suit le guide quand le guide marche et se déplace autour du guide lorsque celui-ci s'arrête. Cela lui permet d'observer les alentours. Le biais par rapport à un guide réel réside dans le fait que l'observateur (et le chercheur accompagnateur) peut arrêter et relancer le guide à sa guise.

Le site expérimental choisi est situé au centre ville de Nantes. C'est un axe piéton d'une longueur approximative de 500m choisi pour sa diversité : on y trouve différents tissus urbains relatifs différentes caractéris-

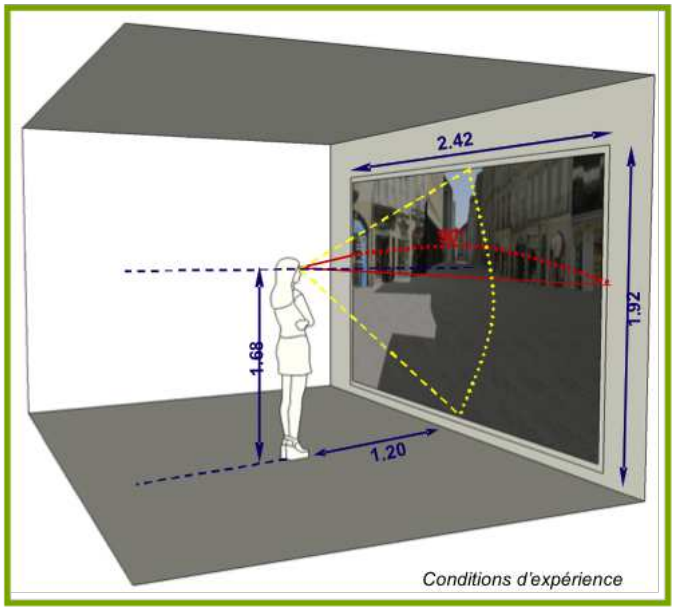


FIG. 1.17 – *Conditions d'expérience pour les expériences 2 et 3*

tiques historiques, morphologiques, spatiales mais aussi pratique (facilité d'accès). Ces différentes caractéristiques sont résumées sur les figures 1.18 et 1.19.

FIG. 1.18 – *Caractéristiques historiques du parcours choisi*

Les conditions expérimentales étaient les suivantes : l'environnement virtuel a été bâti à partir de données réelles (SIG), les modèle 3D ainsi bâti a été calculé en radiosité afin d'obtenir une simulation correcte des ombres pour le 1er avril 2002 à 13h30. L'hypothèse de radiosité a permis de pré-calculer des textures d'ombres rendu en temps réel en faisant l'hypothèse que ni l'observateur ni le guide ne projettent d'ombre sur l'environnement. Le rendu s'est fait en vision monoscopique, les formes

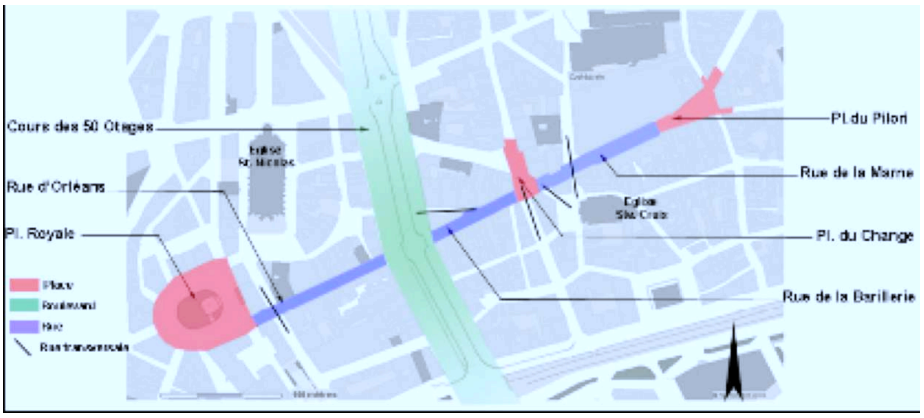


FIG. 1.19 – *Caractéristiques morphologiques du parcours choisi*

urbaines n'étant pas compatibles avec les conditions de confort stéréoscopiques (voir 1.1.2).

Quelques résultats

Dans ce paragraphe, nous présentons les principaux résultats obtenus en nous focalisant essentiellement sur les différences entre les expériences et l'apport de l'interaction. Des résultats plus complets sont disponibles dans [165] et [168]. La dernière expérience, dite «immersive dynamique» a confirmé les résultats obtenus dans les études précédentes. Néanmoins, elle a montré une meilleure corrélation entre la perception visuelle des effets solaires dans les mondes réel et virtuel. L'analyse des parcours commentés (in situ et in vitro) a montré que les deux parcours ont engendré des effets analogues dans l'ensemble. La perception de ces effets s'est basée sur l'interaction entre la morphologie urbaine et la lumière ; dans le virtuel, cette perception était pseudo-naturelle dans la mesure où l'usager pouvait s'arrêter et avoir un déplacement proche de celui du réel, notamment dans les espaces de croisement des rues.

L'analyse des impressions globales sur le parcours a montré que, dans un premier temps, les sujets décrivent la morphologie de l'espace pour établir ensuite ils établissent le lien avec la présence de la lumière dans le parcours. Les analyses des jugements sur les ambiances lumineuses du parcours ont dévoilé que les sujets décrivent les ambiances lumineuses principalement au travers d'expressions génériques qui :

- lient les dimensions spatiales à l'espace lumineux (rue étroite sombre, espace large lumineux) ;
- déterminent le sentiment de confort ou d'inconfort (je me sens refermé, on respire) ;
- traduisent un mouvement (j'ai envie d'accélérer jusqu'à la zone lumineuse, on retombe dans une rue sombre).

L'analyse des réponses aux entretiens a confirmé l'importance de l'interaction lumière/morphologie dans une lecture des formes spatiales, tandis que le mouvement est un élément complémentaire que met en valeur cette relation (lumière/morphologie).

Même si des améliorations techniques peuvent toujours être appor-

tées, cette étude a prouvé qu'il était possible d'étudier les ambiances lumineuses urbaines par les techniques de réalité virtuelle : pour cela, elle s'appuie sur les descripteurs que sont les effets solaires dont elle a déterminé les conditions d'émergence dans le monde virtuel comme dans le monde réel. Une des extensions possibles de ce travail porte sur l'analyse paysagère qui présente des caractéristiques similaires mais aussi des difficultés supplémentaires : l'environnement est dynamique et la dimension multimodale doit être prise en compte. En particulier, nous allons voir dans les paragraphes qui suivent une étude portant sur les impacts paysagers des éoliennes, aussi bien sur le mode visuel que sonore.

1.3.2 Analyse des impacts paysagers des éoliennes

Contexte

Ce travail, qui fait l'objet de la thèse de Jihen Jallouli, s'inscrit dans le problématique de l'analyse des impacts du projet paysager, c'est-à-dire les effets du projet sur l'objet complexe paysage. Ces impacts doivent être mesurés à l'aune de deux étapes du projet : avant celui-ci afin de qualifier le paysage qui sera impacté par le projet et après la phase de conception afin d'étudier différentes alternatives et permettre ainsi la concertation publique, telle que prévue par la loi. Nous nous sommes focalisés sur le problème du projet éolien parce qu'il nous paraissait représentatif du problème paysager tout en soulevant des questions très actuelles.

En effet, le contexte énergétique international encourage les énergies renouvelables et fixe des objectifs ambitieux à travers le protocole de Kyoto. La France, deuxième potentiel éolien d'Europe, s'est notamment donné comme objectif d'atteindre une puissance installée de 10 000 MW en 2010. En mars 2006, la puissance installée n'était que de 918.6MW, elle atteignait les 1300MW à la fin de l'année. Devant cette pression, les projets se sont multipliés dans les régions littorales qui présentent les gisements les plus importants. Cette accélération constructive a devancé la mise en place d'une réglementation rigoureuse et diffuse vis-à-vis des opérateurs et d'une bonne communication en concertation avec le grand public.

Un projet éolien implique l'implantation des aérogénérateurs et de leurs locaux annexes. Son incidence sur l'environnement croise plusieurs impacts :

- Impact paysager : impacts visuel, sonore, faunistique et floristique.
- Impact économique : taxe professionnelle, tourisme, création d'emplois.
- Impact social : proximité des habitations, acceptabilité par les populations.
- Impact structurel : aménagements et équipements secondaires, routes d'accès, parkings...

Dans ce contexte multicritères complexe, l'intégration paysagère des éoliennes suscite l'intérêt des spécialistes et des populations. Entre dégradation et nouveau paysage, les avis sont partagés. Pourtant, la prise de conscience environnementale dans les projets éoliens est de plus en plus accrue et réglementée. «L'étude d'impact» et «l'enquête publique» sont des étapes clés du dossier permis de construire où les opérateurs sont appelés à communiquer le plus en amont possible avec les spécialistes de

l'environnement et les populations concernées pour favoriser un meilleur impact sur le paysage. Ceci est parfois loin d'être appliqué. Les autorités locales et régionales sont souvent indécises devant des exploitants insistants et quelques riverains endurcis, l'argumentation étant convaincante des deux côtés : de la part des opérateurs, on cite l'urgence de penser aux énergies propres et les parcs éoliens réussis alors que de la part des populations, on parle pollution visuelle, bruit, perturbations radioélectriques et insécurité.

Dans cette partie, nous nous intéressons à l'évaluation de l'impact paysager des éoliennes en émettant l'hypothèse que les techniques de la réalité virtuelle vont permettre de mieux évaluer ses impacts lors de l'instruction des dossiers par l'Etat comme participer à une meilleure acceptabilité par les populations en constituant un élément d'objectivation supplémentaire. La méthode des parcours commentés comparés en environnements réels et virtuels sera adaptée aux problématiques paysagères, une fois que nous aurons trouvé une mesure de ces impacts afin d'étudier les différences entre les deux environnements.

Le paysage

Le paysage est un concept hybride qui croise, selon Y. Michelin (cité dans [79]), trois dimensions :

- *«Le paysage a une dimension **objective** : le paysage comme le résultat d'activités humaines, d'un système territorial et de processus naturels [...] Le paysage est un objet, un produit. C'est une réalité ;*
- *Le paysage a une dimension **sensible**. Parler de paysage renvoie à un acte de perception : un paysage est toujours vu (entendu, senti, touché). Un paysage est toujours lié à un point de vue. Un même paysage objectif, aura une réalité différente en fonction du point de vue choisi. Le paysage est une vue, fondée sur une distance ;*
- *Le paysage est aussi **une interprétation**. La même vue sera lue, interprétée différemment par les acteurs en fonction de leur culture, de leur âge, de leur itinéraire... Au-delà du paysage perçu, se trouve le paysage pensé. Le paysage est un regard.»* (Figure 1.20)

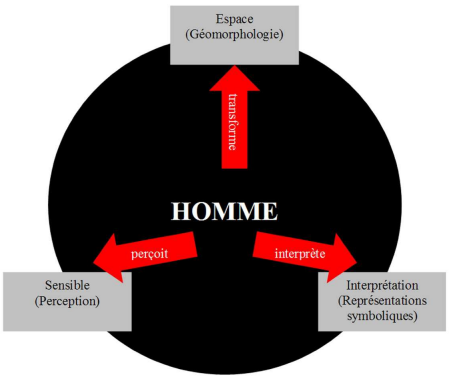


FIG. 1.20 – *Rôle de l'Homme dans le complexe «paysage»*

L'interaction des sens humains avec l'environnement immédiat nécessite aussi une communication entre ces sens pour la perception du monde [58]. L'information échangée entre les sens (et donc la cohérence sensorielle) garantit que lors de la perception d'une éolienne, un seul et même objet est perçu et non pas deux.

La perception du paysage est le résultat d'une interaction entre l'individu et son environnement. Cette étude s'intéresse aux impacts paysagers, c'est-à-dire aux changements qui sont introduits sur un site par une activité développement, produite et évaluée par des individus.

Les méthodes d'évaluation du paysage

Les années soixante-dix ont vu se développer de nombreuses méthodes d'évaluation paysagère se développer afin de produire des éléments quantitatifs et objectifs au problème des réponses nécessairement subjectives liées à l'esthétique paysagère. Certains travaux ont identifié certains attributs de la qualité esthétique paysagère comme [33] qui montre que la verticalité engendre de la tension tandis que l'horizontalité engendre la tranquillité. Ces modèles *objectifs* sont limités à des cas précis et ne permettent pas de prendre en compte la complexité croissante de l'espace. Un point de vue immergé (défini comme centralisé dans [112]) est plus adapté à l'expérience humaine comme nous l'avons vu pour l'espace urbain en 1.3.1. Dans le cas particulier des éoliennes, certaines informations sensorielles ne peuvent pas être perçues en dehors de leur contexte : l'échelle impressionnantes, les pales en mouvement, le son produit par celles-ci.

Les méthodes dédiées à la perception du paysage s'intéressent surtout aux aspects visuels [87, 33]. Cela reste vrai dans le cas des éoliennes [16]. Bien sûr, l'aspect visuel est prédominant dans la perception mais les autres sens participent également (en particulier dans le cas des éoliennes), voire dominent cette perception : la perception est multi-sensorielle et, toujours dans le cas des éoliennes, la perception visuelle et sonore peuvent exister séparément ou simultanément en fonction de la distance à l'éolienne.

Les méthodes actuelles d'évaluation paysagère posent à peu près les mêmes difficultés que celles d'évaluation de l'espace urbain : soit elles sont immersives et dynamiques mais s'adressent à un environnement existant, soit elles sont limitées par leur côté discret dans l'espace comme dans le temps. Par exemple, les représentations proposées (notamment par les opérateurs éoliens) sont basées sur des photo-montages, des cartes, des croquis qui figent à la fois un point de vue et un instant à une échelle qui ne permet pas la prise en compte de la dimension particulière d'une éolienne (cf. Figure 1.21).

Impacts paysagers des éoliennes

En France, les études d'impact et la participation publique sont obligatoires pour tous les nouveaux parcs éoliens mais leur implantation reste largement controversée. L'impact visuel est bien sûr la principale caractéristique d'une éolienne mais cet impact dépend fortement de la distance. Un rapport de l'Université de Newcastle [135] définit des niveaux de perception pour des zones d'influence visuelle (ZVI) :

FIG. 1.21 – *Éolienne : des dimensions particulières*

- la zone de perception lointaine (dans un rayon supérieur à 10km) : les éoliennes ne sont pas toujours visibles et les objets les plus proches dominent la perception. Cependant, dans un paysage étendu et relativement vide, la vision peut se focaliser sur les éoliennes comme éléments structurants de l'espace ;
- la zone de perception intermédiaire (entre 1 et 10 km) : c'est la zone la plus étudiée dans les photo-montages parce qu'elle constitue une vision d'ensemble du parc. Deux lectures différentes peuvent être effectuées en fonction du point de vue : une lecture frontale et horizontale ou une lecture latérale et verticale (cf. Figure 1.22). A ce niveau, les éoliennes dominent visuellement l'espace en raison de leur hauteur. Les pales en mouvement créent une attraction visuelle ;
- la zone de perception immédiate (moins d'un kilomètre) : les éoliennes dominent encore plus la perception visuelle en raison de leur taille. Les impacts acoustiques sont également importants dans cette zone.

FIG. 1.22 – *Différentes lectures en zone de perception intermédiaire*

D'après cette étude, les impacts visuels doivent être étudiés de potentiellement n'importe quel point de toutes ces zones en fonction du site choisi, alors que les photo-montages utilisés actuellement ne présentent qu'un nombre réduit de vues, toutes situées dans la zone de perception intermédiaire.

Les éoliennes produisent aussi des bruits mécaniques et aérodynamiques mais c'est le bruit aérodynamique (la friction entre l'air et la pale) qui se propage à plusieurs centaines de mètres et qui peut gêner le voisinage. Depuis le décret du 19/04/1999, ce sont des ratios de bruit par rapport au bruit ambiant qui sont utilisés pour déterminer les normes acceptables. Selon l'équipe de Gamba [154], il faut également prendre en compte la vitesse du vent :

- en dessous de 15km/h, les éoliennes ne tournent pas ;
- entre 15 et 20 km/h, le bruit de l'éolienne est le plus significatif puisque la bruit du vent dans la végétation n'est pas suffisant ;
- au-delà de 20km/h, le bruit résiduel, lié à la végétation exposée au vent devient plus significatif que le bruit de l'éolienne.
- encore au-delà, à partir d'une certaine vitesse de vent importante, les pales sont bloquées et l'éolienne ne produit plus qu'un bruit d'obstacle statique.

La question acoustique est actuellement prise en compte sur la base d'études techniques qui cherchent à réduire le bruit aérodynamique. Peu d'études s'intéressent à la perception acoustique des éoliennes. Pedersen [137] a utilisé des questionnaires pour montrer que la gêne était plus importante chez les habitants qui voyaient les éoliennes depuis chez eux ainsi que chez les ruraux par rapport aux personnes qui vivent en banlieue. De plus, Pedersen a montré que la perception se focalisait sur le bruit des éoliennes en raison de son incongruité avec le fond sonore.

Notre approche : une méthode immersive basée sur les parcours

Notre approche s'est divisée en deux phases : dans un premier temps, nous avons proposé une méthode immersive sur site en nous appuyant sur l'expérience acquise dans les environnements urbains. Ensuite, nous nous sommes attachés à vérifier les potentialités de cette méthode dans un environnement virtuel. Pour cette seconde étape, nous nous sommes appuyés sur deux séries d'expérimentation.

La méthode (détaillée dans [74]), s'appuyant largement sur les parcours commentés [172], comprend 3 étapes principales :

1. l'étude *objective* : elle utilise des éléments objectifs comme la topographie, les études d'accessibilité, de visibilité etc. pour déterminer les parcours qui seront mis en œuvre. La sélection se fait en les parcourant et en les filmant pour en effectuer une analyse. Pour ne pas ennuyer les futurs participants, les parcours seront limités à 10 minutes [29]. L'enregistrement vidéo sera réutilisé dans l'étude subjective pour corréler les discours des sujets avec leur progression sur le parcours ;
2. l'étude *subjective* : elle se compose d'une partie libre et d'une partie dirigée. La partie libre utilise le discours du sujet avec une intervention minimum de l'enquêteur et sert à caractériser la perception

instantanée des participants (voir 1.3.1). Au contraire, la partie dirigée se compose de questionnaires cherchant à classer les éléments qui ont marqué les participants et à enrichir le résultat du parcours commenté précédent.

3. le *croisement des résultats*.

Pour valider cette méthode, nous nous sommes appuyés sur une étude de cas, le parc de Plouguin dans le Finistère nord. Il présente le double avantage de posséder plusieurs parc éoliens construits à des époques différentes et de bénéficier d'éoliennes spécifiques dessinées par un architecte. L'étude de cas a été menée sur deux parcours différents. Les résultats ont été comparés avec la méthode non immersive classique et ont permis de conclure à une richesse d'information bien supérieure [74]. En outre, elle a remis en perspective un résultat classique des études de l'espace urbain, à savoir l'importance du mouvement pour la perception.

Fort de ces conclusions, nous sommes engagés dans la même démarche en mettant en œuvre cette fois un environnement virtuel reprenant les deux parcours identifiés à l'étape précédente.

Comparaison des parcours réels et virtuels

De la même façon que pour le travail relatif aux ambiances lumineuses, le problème de l'interaction (ici du déplacement) est le problème principal. Il faut néanmoins y ajouter le problème de l'ambiance sonore en incluant les nuances sur les seuils évoqués dans le paragraphe précédent. Nous nous sommes appuyés sur des enregistrements sur site effectués par une équipe d'acousticiens de l'Université de Bretagne Occidentale. A parti de ces enregistrements, nous avons défini des zones d'influence de ces enregistrements des éoliennes mais aussi des sons environnants (oiseaux, bruissements de feuilles, passage de véhicules). Ces sons et leurs zones d'influence (reprises Figure 1.23) ne constituent pas une simulation mais une tentative de restitution de l'ambiance sonore dont l'objectif est bien de déterminer les biais par rapport à l'environnement réel. Comme pour les ambiances lumineuses la détermination des biais et des conditions d'émergence des effets constitue l'objectif expérimental, le réalisme *physique* n'en fait pas partie.

Nous avons effectué deux série d'expérimentations :

- la première série s'est déroulée au Centre Européen de Réalité Virtuelle : s'appuyant sur le logiciel ARéVi [68] et la salle immersive du CERV, le déplacement s'effectuait au moyen d'une Wiimote. Celle-ci était utilisée à la fois pour déplacer le sujet sur un parcours prédéfini et pour orienter le regard de celui-ci lorsqu'il était arrêté. Si l'utilisation de la Wiimote n'a pas posé de problèmes particuliers contrairement à ce qu'on aurait pu craindre, elle a montré les limites du parcours imposé, même en présence d'un guide virtuel [75]. Cependant, les résultats sont intéressants par rapport à perception de l'environnement réel.
- pour remédier à ce problème de déplacement, nous avons utilisé un vélo que nous avons instrumenté à très bas coût (Figure 1.24) [76, 153] : nous avons posé le vélo sur un socle afin de faire tourner la roue arrière dans le vide et utilisé 3 souris dont les

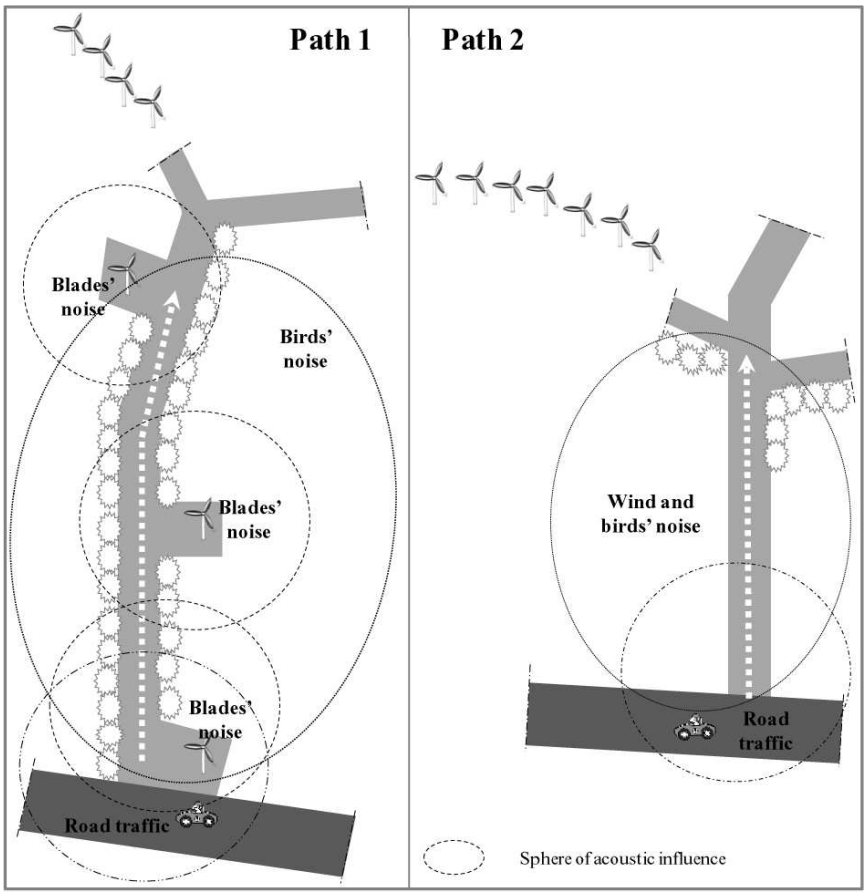


FIG. 1.23 – *Répartitions des zones d'influence sonore sur les différents parcours*

capteurs opto-mécaniques ont été utilisés respectivement pour mesurer la vitesse de rotation de la roue arrière, l'angle du guidon et l'inclinaison donnée à la vue. Ce dispositif a été réalisé par Davy Salerne pendant son stage de Master [153] et intégré directement dans Virtools. Il en résulte un parcours parfaitement libre (dans les limites des zones accessibles physiquement) et une interaction tout à fait naturelle. La roulette de souris utilisée pour l'inclinaison de la vue a été également considérée comme naturelle. De plus, la position penchée sur le vélo empêchant le sujet de lever la tête, la limitation à un seul écran vertical a été levée.

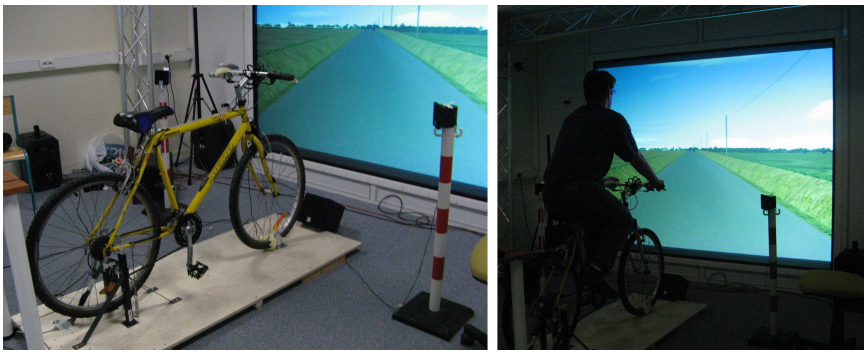


FIG. 1.24 – *Utilisation d'un vélo instrumenté*

Les résultats de cette étude sont en cours d'analyse et seront publiés prochainement.

1.3.3 Conclusion

Dans les paragraphes précédents, nous avons abordé l'étude des potentialités de la réalité virtuelle dans un contexte particulier où :

- les maquettes physiques à l'échelle 1 :1 sont impossibles à réaliser ;
- le produit sera réalisé à un unique exemplaire ;
- le produit étudié est d'échelle très supérieure à celle de l'être humain ;
- l'appréciation sensible du produit est la dimension la plus importante ne serait-ce qu'en raison de la durée de vie de ces produits.

Nous avons, dans chaque cas, comparé un produit existant à un environnement virtuel reproduisant partiellement le réel en s'appuyant sur l'hypothèse que le mouvement enrichit la perception. Cette comparaison a permis de conclure qu'une majeure partie partie de la dimension sensible des projets est reproductible par les techniques de la réalité virtuelle. Les travaux futurs liés à ces études incluent :

- l'intégration de l'outil réalité virtuelle dans le projet (urbain ou paysager) ;
- la confrontation des résultats sensibles aux données existantes : dans les deux cas évoqués ci-dessus, les environnements virtuels utilisés s'appuient sur des données SIG même si aucun lien explicite n'est fait. A l'avenir, il faudra effectuer le lien avec non seulement les SIG, mais aussi les outils de simulation disponibles qui apportent une information supplémentaire nécessairement intéressante (par exemple pour le décideur). Ce lien sera discuté au chapitre3.
- Enfin, pour l'architecture comme pour l'urbanisme et le paysage, la problématique de la perception *réelle* de l'être humain n'a pas été abordée : pour intégrer un élément de conception virtuel dans un projet réel, il est nécessaire de s'assurer que la forme validée dans le projet virtuel correspond bien à celle attendue dans l'esprit de l'observateur. Cette question a déjà été abordée dans les paragraphes précédents, le lien avec la perception sensible reste à effectuer.

1.4 Conclusion sur la perception des mondes virtuels

Ce premier chapitre a exposé les travaux que nous avons abordés dans le domaine de la perception des mondes virtuels. Contrairement à d'autres chercheurs du domaine, nous avons cherché à avoir une approche verticale de cette question en s'intéressant à différents de perception, allant du psychophysique au sensible qui cherche à faire ressortir la part de subjectivité (qui peut être pour partie due à des critères sociaux ou culturels) liée à la perception des environnements très grands et complexes comme le paysage. Au cours de ces travaux, notre démarche méthodologique a été majoritairement fondée sur la comparaison entre un environnement réel et un environnement virtuel tentant de reproduire cet environnement réel. La mesure de comparaison a été majoritairement effectuée par l'exécution

comparée de tâches à un certain nombre d'observateurs. Ces éléments méthodologiques ont également été appliqués dans l'incursion faite dans le domaine de l'haptique pour la partie consacrée aux études d'ergonomie des boutons de commande des planches de bord automobiles.

Dans le chapitre suivant, nous allons nous intéresser à la perception du monde réel afin de construire des environnements virtuels convaincants. Là encore, nous ne nous limiterons pas à la construction d'environnements virtuels géométriquement corrects mais nous chercherons à aller plus loin en nous posant la question de la fidélité comportementale des mondes virtuels.

Construction des environnements virtuels

2

Ce chapitre est en quelque sorte le dual du chapitre précédent : il s'intéresse à la perception du monde réel pour la construction des environnements virtuels. Cette fois, la perception est synthétique au sens où ce sont essentiellement des algorithmes de vision par ordinateur qui vont être mis à profit pour la construction de la géométrie des objets virtuels. Là encore, nous commençons par le niveau physique d'analyse de l'image avant d'aborder la reconstruction d'objets à l'aide de ces images. Dans ce chapitre, nous abordons également la question de la production d'environnements virtuels fidèles d'un point de vue comportemental et pas seulement géométrique. Cette dernière partie nous servira à effectuer le lien avec le chapitre suivant consacré aux Systèmes d'Information Géographique représentés et construits à partir d'images.

2.1 Perception humaine et vision par ordinateur

Afin de créer une scène virtuelle à partir d'une scène réelle, une des approches les plus courantes consiste à extraire des informations sur la troisième dimension à partir de photographies. La stéréovision consiste à utiliser une paire d'images à cette fin. C'est d'ailleurs une des activités principales du système visuel humain, ainsi que nous l'avons vu au chapitre précédent.

La géométrie épipolaire permet de reconstruire cette information de profondeur à partir de deux images dont on connaît l'emplacement de prise de vue : le lieu géométrique d'un point d'une image est une demi-droite passant par le centre de projection. Lorsqu'on dispose deux représentations d'un même point 3D sur une paire d'images, l'intersection de ces droites fournit les coordonnées dans l'espace 3D de ce point (cf. figure 2.1). Bien sûr, ce n'est pas la seule information disponible comme le montrent les nombreux travaux effectués dans ce domaine.

Nous avons choisi de nous intéresser au problème de la mise en correspondance des points 2D sur les images. Deux grands types de techniques sont utilisées : les algorithmes basés sur les caractéristiques des points utilisés (coins, lignes...) [109, 133] qui fournissent, pour un coût de calcul généralement modéré, des listes de points appariés et les méthodes denses qui cherchent à apparier tous les points des images et à fournir des cartes de profondeur denses. Bien que de coût supérieur, les méthodes denses présentent l'avantage de pouvoir tenir compte de certaines discontinuités et occlusions (lorsqu'un point visible dans une image est masqué dans

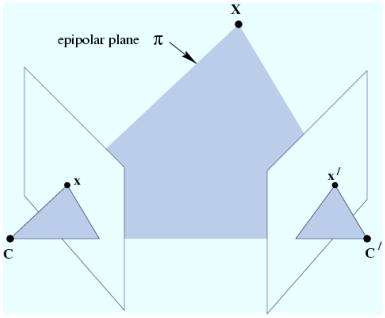


FIG. 2.1 – *Géométrie épipolaire : C, C′, x, x′ et X sont coplanaires*

l'autre). On trouve là encore deux types de méthodes : les méthodes globales cherchent à minimiser une fonction de coût combinant les disparités entre images et des termes de lissage tandis que les méthodes locales utilisent implicitement ce lissage en utilisant des fenêtres de recherche autour d'un point à apparier en agrégeant les informations disponibles.

Bien qu'il existe de nombreuses méthodes denses locales, nous nous intéresserons ici principalement aux méthodes dites de corrélation qui résument bien l'approche mise en œuvre dans la plupart des algorithmes : Pour un pixel $p_l = (x, y)$ de l'image gauche, le problème consiste à trouver le point p_r dans l'image droite qui minimise une distance appelée entre un voisinage de p_l et un voisinage de p_r. Les questions soulevées par ce type de méthode portent donc sur :

- le choix de la taille et de la forme du voisinage ;
- le choix de la taille et de la forme de la fenêtre de recherche ;
- la fonction de distance à utiliser ;
- l'optimisation du code, le calcul de l'appariement d'un seul point nécessitant p^2w^2 calculs si p est le côté du voisinage et w le côté de la fenêtre de recherche.

De notre côté, à la suite des travaux menés sur les ondelettes pour l'amélioration du confort en vision stéréoscopique évoqués en 1.1.5, nous sommes intéressés à l'utilisation de celles-ci en stéréovision. L'utilisation des ondelettes se justifie sur au moins 3 plans :

- leur propriétés de localisation en temps et fréquence d'un signal ;
- leur capacité à capturer des niveaux de détail très différents dans une scène (échelles) ;
- parce qu'elles constituent un bon modèle du système visuel humain [174, 150].

Il existe déjà de nombreuses méthodes utilisant les ondelettes en stéréovision ; elles varient essentiellement sur la nature et les propriétés des ondelettes utilisées. Celles-ci peuvent être utilisées dans les fonctions de lissage des méthodes d'extraction de points caractéristiques [108]. Pour les méthodes denses, les ondelettes permettent de travailler sur des pyramides d'images de résolution croissante.

Notre méthode se base sur les travaux de Perrin [139, 138]essentiellement consacrés à la vision binoculaire humaine. Il a défini une fonction de cor-

rélation localisée de la façon : si f et g sont deux fonctions de $\mathcal{L}^2(R)$, la corrélation localisée de f et g notée $s(b,b')$ s'écrit :

$$s(b,b') = \int_0^{+\infty} WT(f)(a,b)\overline{WT(g)(a,b')}\frac{da}{a} \tag{2.1}$$

où $WT(f)$ est la transformée en ondelettes de f. L'avantage de cette approche est que la transformée en ondelettes en un point contient des informations sur tout le voisinage, remplaçant ainsi le calcul de la fonction de corrélation. Perrin a montré que la fonction de corrélation localisée est effectivement locale (elle ne prend en compte qu'un voisinage restreint du point) et localisée (discriminante). Elle va au delà de la corrélation classique (donc le noyau est un Dirac) puisque son noyau est une fonction de b et b' et qu'elle permet ainsi de mesurer la similarité entre f et la translatée de g. Perrin a testé son approche avec plusieurs types d'ondelettes, il s'est également intéressé au nombre d'échelles à utiliser. Cependant, la reconstruction 3D n'étant pas l'objet principal de son travail, cette méthode n'a pas été totalement exploitée. Nous avons conservé cette idée en y apportant quelques améliorations pour mieux prendre en compte les discontinuités.

Les discontinuités et les occlusions sont un problème majeur des méthodes de stéréovision parce qu'elles peuvent résulter en des erreurs importantes. Nous proposons une méthode [121] qui n'est pas basée sur la modification de la fenêtre de recherche mais sur la décomposition en ondelettes.

2.1.1 La transformée en quart d'ondelettes

Motivations

On considère Ψ une fonction ayant toutes les propriétés d'une ondelette. L'ondelette multi-échelle $\Psi_{a,u,v}$ est définie par :

$$\Psi_{a,u,v} : (x,y) \in R^2 \mapsto \frac{1}{a} \cdot \Psi\left(\frac{x-u}{a}, \frac{y-v}{a}\right) \tag{2.2}$$

où $(a,u,v) \in (I \times \times J_x \times J_y)$ avec I borné. Si f est le signal bidimensionnel (l'image), la transformée s'écrit alors :

$$WT_f = < \Psi_{a,u,v}, f > = \int_{R^2} f(x,y)\Psi_{a,u,v}(x,y)dxdy \tag{2.3}$$

Cette transformée peut aussi être vue comme la convolution à différentes échelles de l'ondelette et du signal. Toutefois, la transformée en ondelettes en un point dépend du voisinage de ce point. L'inconvénient est donc que la transformée prend en compte l'ensemble du voisinage du point : s'il existe une discontinuité de disparité (due à une occlusion par exemple), l'information qui peut ne pas correspondre à la même profondeur pourra nous induire en erreur.

Nous proposons donc de restreindre la transformée en ondelettes à un quadrant : il s'agit d'essayer de trouver des correspondants pour les points qui sont proches des zones d'occlusion mais dont la transformée contient suffisamment d'information. La figure 2.2 illustre ce principe sur

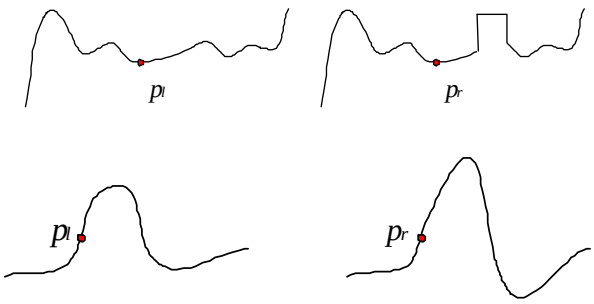


FIG. 2.2 – *Les points p_l et p_r correspondent bien qu'ils soient près d'une occlusion, mais ils ne peuvent être mis en correspondance avec une transformée en ondelettes traditionnelle*

un exemple unidimensionnel simple : les points p_l et p_r peuvent correspondre, le problème étant que p_r se trouve proche d'une discontinuité. Si on calcule la transformée en ondelettes des deux signaux au voisinage de ces points, les coefficients obtenus ne permettront pas de conclure à un appariement. Décomposons l'ondelette en deux contributions $\Psi^+_{a,b}$ et $\Psi^-_{a,b}$:

$$\Psi^+_{a,b}(x) = \begin{cases} \Psi_{a,b}(x) & \text{for } x - b \geq 0 \\ 0 & \text{sinon} \end{cases} \tag{2.4}$$

$$\Psi^-_{a,b}(x) = \begin{cases} \Psi_{a,b}(x) & \text{for } x - b \leq 0 \\ 0 & \text{sinon} \end{cases} \tag{2.5}$$

Ces contributions nous permettent d'analyser le signal des deux côtés du point b et d'obtenir des résultats sur les possibilités d'appariement entre p_l et p_r comme le montre la figure 2.3. Pour les images, les mêmes principes sont appliqués sur chacun des axes, d'où le nom de transformée en quart d'ondelettes. En ne prenant en compte qu'un quart du domaine de l'ondelette, la transformée devient très sélective dans les hautes fréquences comme le montre la figure 2.4. Dans [146], nous montrons que la transformée en quart d'ondelettes tend à accentuer les discontinuités mais également la sensibilité au bruit. L'utilisation des 4 quarts d'ondelettes permet de résoudre partiellement ce problème en fournissant 4 résultats simultanément.

Algorithme

Le schéma algorithmique est simple : il suffit de calculer les 4 transformées en quart d'ondelettes puis de rechercher les appariements et en déduire la carte de disparités.

La transformée en ondelettes est calculées à l'aide de convolutions 2D utilisant elles-mêmes la transformée de Fourier rapide (FFT). Pour deux fonctions bi-dimensionnelles $\hat{f}$ et $\hat{g}$:

$$\hat{f} * \hat{g} = FT^{-1}(FT(\hat{f} \cdot FT(\hat{g})) \tag{2.6}$$

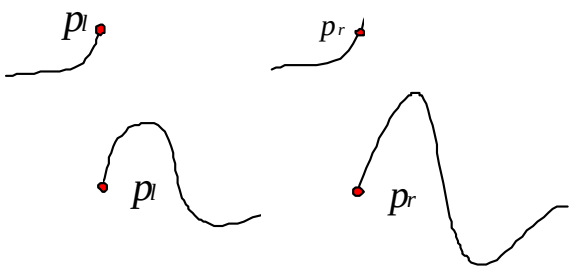


FIG. 2.3 – *Comparaison à l'aide de deux demi-transformées : bon appariement avec* Ψ^+*, pas d'appariement avec* Ψ^-

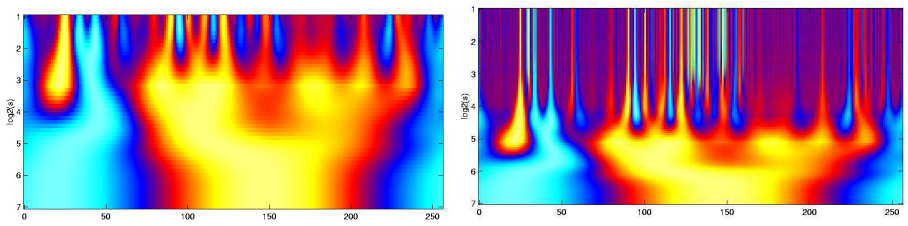


FIG. 2.4 – *Les résultats de la transformée en ondelettes continue et d'une transformée en quart d'ondelettes sur une image ligne (ondelette utilisée : DOG)*

où $\cdot$ est la multiplication terme à terme. Cela permet de gagner du temps dans la mesure où une partie importante de la FFT peut être pré-calculée et utilisée sous forme de cache. Le processus d'appariement consiste à appliquer les 4 transformations aux 2 images. Nous faisons les hypothèses que les caméras sont parallèles, restreignant ainsi l'espace de correspondance possible à une seule dimension[1].

Après de nombreux tests, il s'avère que l'application d'un lissage Gaussien (pour suivre l'hypothèse de continuité globale émise par Marr et Poggio[109]) suivie d'une sélection du candidat dont la disparité d minimise la quantité suivante :

$$\min_d \left(\min_{i \in \{1.2.3.4\}} \| QWT_i^l(x+d,y) - QWT_i^r(x,y) \|_2 \right) \tag{2.7}$$

où les QWT_i sont les transformées en quarts d'ondelettes dans les 4 cadrans, est la solution la plus efficace. Le problème suivant consiste à distinguer les bons appariements. Pour cela, nous nous basons sur l'algorithme *Match Goodness Jumps* décrit dans [50] qui est une forme de classifieur non supervisé appliqué à l'ensemble de l'image.

Résultats

Le choix de la base d'ondelettes est important : les fonctions gaussiennes ne sont pas adaptées car elles ont tendance à flouter les arêtes. Les dérivées les plus populaires sont les DOG (différences de gaussiennes) ou les LOG

[1] on peut aussi effectuer une rectification préalable des images comme dans [53].

(dérivées secondes de gaussiennes). Les ondelettes à plus fort coefficient d'évanescence sont généralement préférés parce qu'elles nécessitent moins de coefficients et donc moins de calcul. Toutefois, la tendance des LOG et des DOG à accentuer les discontinuités nous convient davantage.

Sur des images de synthèse aléatoires avec disparité uniformes (cf. figure 2.5), tous les pixels sont parfaitement appariés en dehors des zones d'occlusion (le bord de l'image et le bord du disque).

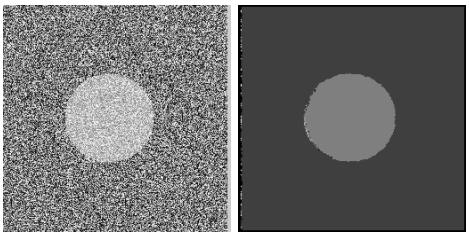


FIG. 2.5 – *Image aléatoire (image droite uniquement) et carte de disparité associée*

La seconde batterie de tests a consisté à utiliser des images de synthèse pour lesquelles nous pouvions obtenir facilement une carte de disparité exacte. Cela a été rendu possible grâce à une version modifiée du logiciel de lancer de rayon *povray* développée par Perrin. La figure 2.6 présente une carte de disparité extraite par cette méthode. Comparé aux résultats obtenus par Perrin [139, 138], sont visuellement meilleurs aux voisinages des zones de disparité, on distingue par exemple beaucoup plus aisément les contours du pion.

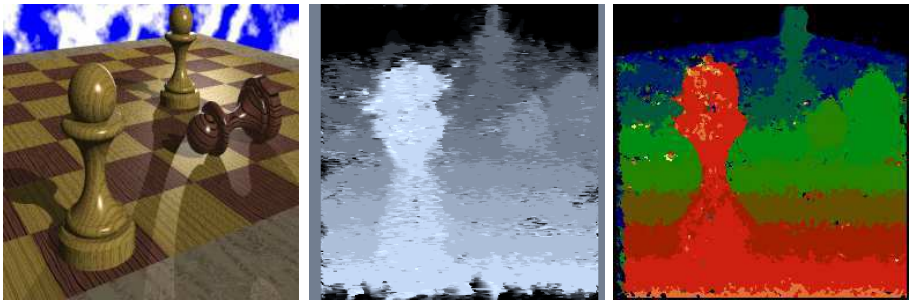


FIG. 2.6 – *Résultats sur une paire d'images de synthèse : l'image gauche d'origine, la carte de disparité avec la méthode de Perrin, la carte de disparité avec la transformée en quart d'ondelettes*

Pour les tests sur images réelles, nous avons utilisé la fameuse base d'images de l'Université de Tsukuba à l'instar de la comparaison menée par Szelisky [163]. Les résultats (comme par exemple sur l'image de la figure 2.7) sont meilleurs qu'en utilisant les ondelettes seules. Ils atteignent 88% d'appariements corrects[2] et 91% si supprime les zones de discontinuités et d'occlusions. Comparée à d'autres techniques prenant en compte occlusions et discontinuités, notre méthode donne de bons résultats pour une méthode locale. En ajoutant un bruit gaussien, la qualité des résultats obtenus diminue toutefois rapidement [146].

[2] en utilisant le même critère que Szelisky [163].

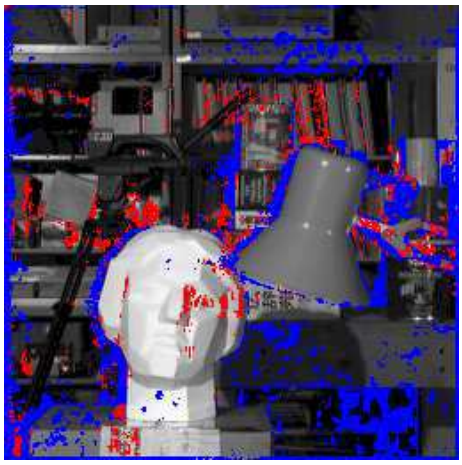


FIG. 2.7 – *En bleu, les zones où l'erreur est supérieure à un pixel, en rouge les zones où aucun appariement n'est trouvé*

Dans le paragraphe qui suit, nous allons utiliser cette méthode dans un contexte plus global de construction de modèles géométriques d'objets réels.

2.2 Construction d'environnements virtuels à l'aide d'images

Les photographies (prises séparément ou dans leur ensemble) comportent de nombreux indices visuels donnant des informations tridimensionnelles sur le contenu. Dans cette partie, on s'intéresse à la construction de maillages d'objets contenus dans des images qu'on trouve généralement sous le nom anglais générique de *shape from* *, le caractère * pouvant prendre de nombreuses formes. Dans un premier temps, nous effectuerons un survol rapide des différentes techniques classiques. Pour une étude plus approfondie, le lecteur pourra se référer à [115]. Nous décrirons ensuite nos contributions et leur application dans le cadre d'un projet PRIAMM du CNC dédié à la reconstruction 3D d'objets muséographiques.

2.2.1 Etat de l'art

Shape from Silhouette

La technique la plus simple et la plus facile à décrire s'appelle le *Shape From Silhouette* [88]. A partir d'un ensemble d'images $I = (I_i), i \in \{1, ..., n\}$ pour lesquelles on connaît la position de la caméra P_i, on extrait de chaque image la silhouette de l'objet à modéliser (son contour extérieur). La silhouette est ici une courbe fermée notée $\mathcal{C}_i$. L'objet est donc contenu dans le cône généralisé formé par son sommet P_i et la courbe $\mathcal{C}_i$. Il est donc contenu dans l'intersection de ces cônes ainsi que le montre la figure 2.8

D'emblée, on peut faire les remarques suivantes :

– Cette opération n'est possible qu'avec un modèle de caméra purement *pinhole*, c'est-à-dire à projection perspective parfaite. Cette difficulté peut-être résolue en effectuant une distorsion inverse préa-

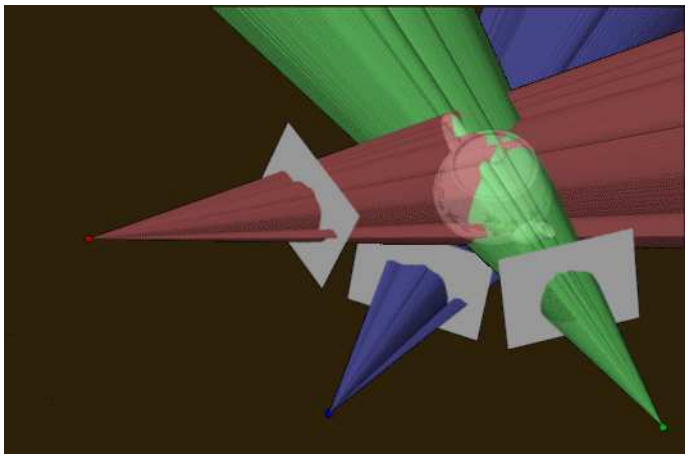


FIG. 2.8 – *Principe des algorithmes de Shape From Silhouettes, d'après [28]*

lable de l'image dans le cas où la caméra présente une distorsion connue ;
- Utilisée telle quelle, cette technique ne permet pas de modéliser de façon exacte des objets comportant des trous. Elle construit que l'*enveloppe convexe* d'un objet ;
- Aucune information n'est disponible sur le nombre d'images nécessaires à la reconstruction de l'objet.

L'enveloppe visuelle $VH(O,P)$ d'un objet O et d'un ensemble de points de vue $P = \{P_1, ..., P_n\}$ est une région de l'espace telle pour pour chaque point $p \in VH(O,P)$ et chaque point de vue $v \in P$, la demi-droite issue de p et passant par v contienne un point de O. Laurentini montre alors :
- que $VH(O,P)$ est la meilleure approximation possible de O en utilisant le Shape from Silhouettes avec les points de vue P ;
- un objet est reconstructible de façon exacte par la technique du Shape from Silhouette si et seulement si $O = VH(O,v)$;
- L'enveloppe visuelle contient l'objet mais est contenue dans l'enveloppe convexe de celui-ci.

Il existe très peu de résultats sur le nombre de vues nécessaires à la reconstruction d'objets : Laurentini [89] montre simplement que dans certains cas, on peut reconstruire avec $\mathcal{O}(n^5)$ vues, un objet polyédrique à n faces ! Du point de vue la précision de la reconstruction, la situation n'est pas bien meilleure : Niem [132] distingue 3 grands types d'erreurs : la quantification de l'image 2D et donc de la silhouette mais aussi de l'espace 3D, la précision du calibrage de la caméra (le modèle d'erreur est décrit dans [176]) et enfin du nombre et du choix des vues. Il arrive à montrer que l'erreur de reconstruction croît linéairement avec le rayon de courbure local à la surface mais de façon quadratique avec l'angle entre deux vues successives d'un même point. Enfin, Slabaugh [156] cherche à minimiser l'erreur de reprojection de l'objet construit sur les images de départ en rendant transparentes certaines zones masquées mais l'erreur est liée de façon très indirecte à cette la transparence, rendant l'optimisation très difficile.

Les implémentations de cette technique sont essentiellement volu-

miques à base de voxels. Pour chaque vue, il suffit d'éliminer les voxels situés à l'extérieur du cône (P_i, C_i). Au delà des problèmes de complexité (cubique en espace et en temps), les agrégats de voxels ne sont guère seyants : ils sont donc généralement transformés en maillages surfaciques grâce à l'algorithme des Marching Cubes [103] qui génère malheureusement beaucoup de triangles mais aussi et surtout quelques artefacts. La complexité spatiale peut être améliorée en utilisant des octrees comme le suggèrent [32] et [143], au détriment de la simplicité d'implémentation. Niem [131] propose une reconstruction par piliers en représentant ceux-ci uniquement par leurs extrémités ce qui permet un gain en espace et de temps sans difficulté de programmation.

Autres techniques

Il existe d'autres techniques, que nous listons brièvement ici :

- Le *Shape from Shading* consiste à inverser les lois d'éclairement. Si on considère une surface uniforme, l'intensité lumineuse réfléchie par celle-ci ne dépend alors que du type de surface, des caractéristiques de l'éclairage et de l'angle sous lequel est vue cette surface : la loi de Lambert s'intéresse au cas des surfaces réfléchissant la lumière de façon uniquement diffuse. Pour pallier cet inconvénient majeur, il est courant de prendre également en compte les réflexions spéculaires en utilisant la loi de Phong qui permet les réflexions spéculaires non-ponctuelles. D'après Zhang [192], les algorithmes fournissent des résultats de mauvaise qualité sur les images de synthèse, mais ceux-ci sont encore pires sur des images réelles. En outre, les résultats sur les images de synthèse ne permettent aucune prédiction sur l'utilisation d'images réelles.
- Les techniques de *Shape from Stereo* ont pour objectif d'exploiter les informations de profondeur obtenues par l'étude des différences entre deux images proches. La stéréo peut venir de deux caméras situées à faible distance l'une de l'autre (stéréo anthropomorphe) ou d'une caméra en mouvement linéaire (à l'instar de la vision des oiseaux). La plupart des approches, à l'instar de celle que nous avons présentée en 2.1 se contentent de construire des cartes de profondeur à partir de 2 ou plus images. Les papiers traitant de la reconstruction de formes à partir d'une paire stéréo sont plus rares [152]. Eisert et Steinbach [51, 158] proposent une approche complète commençant par une estimation de forme à partir de deux vues.
- les techniques de *Shape from Defocus* s'intéressent quant à elle au flou de profondeur de champ présent dans les images. Ces approches visent à la fois à améliorer la netteté de certaines zones des images mais permettent aussi de reconstruire des cartes de profondeur à partir de l'information de flou [40].

Comme l'on pouvait s'en douter, les techniques énoncées précédemment ont été combinées de façon à essayer de se compléter mutuellement. L'approche la plus classique consiste à effectuer une passe de shape from silhouette puis une autre passe d'une des techniques des paragraphes précédents. Matsumoto e.a. [110] proposent de réaliser une phase de shape

from silhouettes pour déterminer la zone de présence de l'objet puis d'appliquer l'algorithme de stéréo-vision sur la zone déterminée.

2.2.2 Shape from Depth

Notre méthode s'appuie également sur la notion de *Shape from Silhouette* associée à un appariement stéréoscopique de paires d'images. Elle sera appliquée dans le paragraphe dédié au projet Gantom. Elle diffère des autres méthodes en ce sens que les images sont habituellement utilisées pour extraire une forme 3D qui servira de contrainte pour le volume final. L'utilisation de la seule information de silhouette ne permet de générer que des cônes. En réalité, on peut générer des formes plus complexes grâce à l'utilisation de cartes de profondeur. Nous proposons de couper les cônes construits à l'aide de la carte de profondeur, comme le montre la figure 2.9.

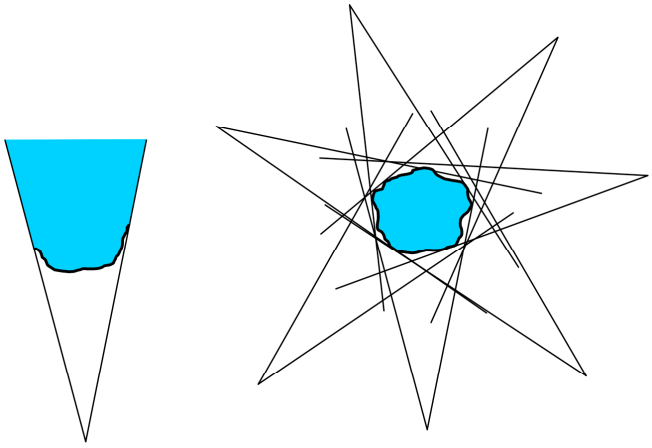


FIG. 2.9 – *Principe de la reconstruction*

Dans la pratique, une approximation de l'intersection des cônes est obtenue sous la forme d'un octree dont les feuilles représentent des parties de l'objet. L'utilisateur a ainsi potentiellement la possibilité de choisir la précision à laquelle l'objet virtuel sera calculé, en imposant une limite sur la taille des feuilles de l'octree. Les informations de densité liée à la structure de l'octree nous permettent de calculer une fonction de lissage qui affine visuellement la forme de l'objet avant et après passage d'un algorithme de Marching Cubes (interpolation locale dans les cartes de profondeur là où c'est possible, lissage anisotropique sur le modèle triangulé [7]).

Pour des questions de robustesse, l'algorithme *shape from silhouette* classique est ensuite relâché en utilisant le principe du vote selon le principe présenté dans [86] : là où l'algorithme d'origine effectue une intersection de tous les cônes, nous utilisons un système où nous sommons le nombre d'occurrences de la présence d'un voxel à l'intérieur d'un cône. Au delà d'un certain seuil, le voxel est conservé.

2.2.3 Gantom : reconstruction 3D d'objets muséographiques

Gantom [6] est un projet PRIAMM du CNC-Ministère de l'Industrie que nous avons mené en partenariat avec la société SimTeam.

Introduction

En 1999 (date de lancement du projet), il existait des solutions industrielles de scanners 3D pour acquérir la géométrie d'un objet de manière plus ou moins précise. Parmi ces dispositifs, certains sont capables d'acquérir en même temps des informations de texture, augmentant le réalisme des objets reconstruits. Néanmoins, dans les cas les plus favorables, les textures obtenues ont une faible résolution et sont imparfaites du point de vue colorimétrique. Ces technologies sont principalement utilisées pour des applications liées à l'Internet. Elles restent cependant inadaptées aux exigences de qualité liées aux objets d'art : respect de l'intégrité des couleurs et des détails des objets numérisés.

L'originalité de notre projet est de proposer une chaîne complète de numérisation 3D, intégrant à la fois :

- L'acquisition de la géométrie d'un objet en recourant à des techniques conventionnelles de vision passive (stéréovision et shape-from-silhouette.) avec un calibrage géométrique très précis ;
- La construction d'un maillage comportant des informations géométriques précises ;
- L'obtention d'informations de texture en haute définition et cohérentes en colorimétrie pour un rendu synthétique plus fidèle ;
- La maîtrise et le calibrage des sources de lumières dans la réalisation du dispositif d'acquisition et de restitution des textures.

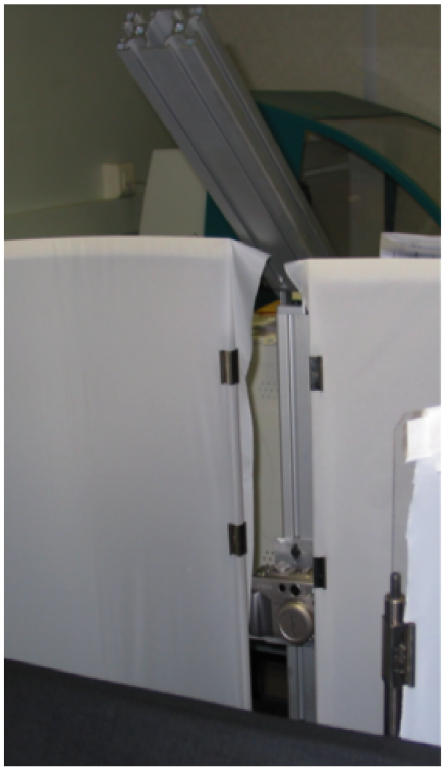
L'apparition des appareils photographiques numériques de type «méga-pixels» sur les marchés grand public et professionnel à la fin des années quatre-vingt-dix laisse entrevoir de nouvelles perspectives quant à la numérisation 3D de qualité pour un coût raisonnable. Notre approche consiste donc à utiliser la géométrie 3D d'un objet pour extraire les informations colorimétriques de clichés haute-résolution, après recalage précis des modèles sur les images réelles. Par rapport à une caméra vidéo standard, l'emploi d'appareils à haute-définition permet d'augmenter la quantité d'informations disponibles pour constituer une texture pouvant représenter les détails très fins d'objets complexes.

Notre projet propose en outre de respecter l'intégrité des couleurs des objets numérisés, en veillant à maîtriser les conditions d'éclairages utilisées lors des prises de vues : les prises de vue seront calibrées colorimétriquement, afin de pouvoir éliminer par traitement informatique, leur influence sur les couleurs des objets numérisés. Le contrôle des sources lumineuses implique des contraintes d'emploi : les prises de clichés sans lumière parasite (dans une *boîte noire* ou dans une pièce occultant la lumière du jour), éliminant les réflexions entre l'environnement et l'objet.

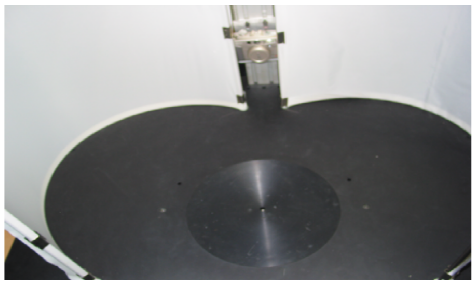
Résultats obtenus

Le dispositif d'acquisition réalisé est présenté figure 2.10 : il s'agit d'un plateau tournant dont le moteur pas-à-pas (de précision de l'ordre du $\frac{1}{10}$

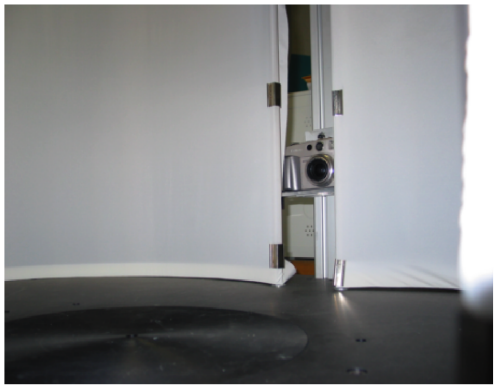
degré) est instrumenté de façon à limiter à un seul degré de liberté les différences entre les différentes prises de vue, une fois le premier calibrage effectué. 2 appareils photos numériques très haute-résolution (en 2001) de 4 millions de pixels sont attachés à ce dispositif. Ils sont pilotés directement par l'ordinateur, de même que le moteur de rotation. Les effets spéculaires de lumière sont supprimés grâce à un éclairage rendu diffus par la toile diffusante photographique que nous avons utilisée [3]. La figure 2.11 montre le dispositif avec la statuette prête pour la phase de modélisation.



Un appareil photo entre les écrans de diffusion de la lumière



Vue du plateau circulaire tournant, support de l'objet



Le plateau tournant, un appareil photo et les écrans de diffusion de la lumière

FIG. 2.10 – *Dispositif d'acquisition conçu et réalisé pour l'acquisition d'objets contenus dans un cube de 200mm de côté*

Le calibrage est classiquement effectué à partir d'une mire rectangulaire plane inclinée à 45 degrés et de l'algorithme de Tsai [176]. Une douzaine de prises de vue à des angles différents du plateau tournant sont utilisées afin de déterminer à la fois les paramètres de la caméra et l'axe de rotation du plateau, en confortant celle-ci par les positions du plateau déterminés par le capteur optique du moteur pas-à-pas. Sur des images de résolution 2272x1704 pixels, la précision de reprojection est inférieure au pixel partout dans la zone du cube d'acquisition pour une distance

[3] les spécularités perçues sur la figure sont le fait de la prise de vue du du dispositif qui a nécessité un démontage partiel de celui-ci et l'emploi d'une flash.

FIG. 2.11 – *La statuette sur le plateau tournant prêt pour la modélisation*

d'acquisition moyenne de 350mm. Les erreurs résiduelles proviennent très probablement des distorsions des images non prises en compte par l'algorithme de Tsai.

Les résultats obtenus seront présentés à partir d'un exemple : une statue de l'époque mochéenne. Cette statue peut s'inscrire dans un cube de 200mm de côté. Le premier maillage obtenu à partir de la technique du shape from depth utilise des voxels dont la dimension est d'environ 0,35mm. Le premier maillage surfacique est lourd d'environ 1,7 millions de triangles. Pour obtenir ce modèle, nous avons utilisé 64 paires stéréo différentes, soit 128 images, ce qui a abouti à un temps de traitement de 6 heures sur une machine du commerce (de 2001). Le résultat est présenté figure 2.12.

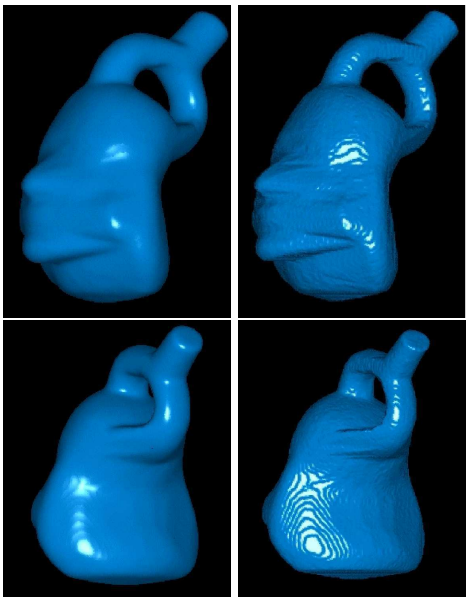


FIG. 2.12 – *Modèles lissés et non lissés de la statuette*

2.3 Au-delà de la géométrie : fidélité comportementale des environnements virtuels

Un des critères de crédibilité des mondes virtuels est celui de la fidélité du mouvement, bien plus que celui de la fidélité géométrique. Les travaux évoqués au chapitre 1 cherchent d'ailleurs à déterminer des conditions minimales de restitution de certains phénomènes comme par exemple les effets solaires. Dans cette partie, nous nous intéressons à la restitution du mouvement. Nous verrons dans un bref état de l'art que la reproduction du mouvement peut se faire à différents niveaux d'abstraction. Nous montrerons comment on peut créer des environnements virtuels comportant des objets animés reproduisant des comportements similaires à ceux d'êtres humains du monde réel. Nous décrirons ensuite une application des ces travaux à la simulation du couple conducteur-véhicule dans le cadre de simulations de trafic autoroutier et piéton.

2.3.1 État de l'art

G. Hégron et B. Arnaldi [69] donnent une classification à trois niveaux des modèles de contrôle du mouvement fondée sur le niveau d'abstraction de la description du mouvement. Le plus bas-niveau est nommé *niveau descriptif* et reproduit les effets décrits sans se préoccuper de ce qui les a produits ; par contre, le niveau immédiatement supérieur, dit niveau *générateur*, s'intéresse aux causes physiques du mouvement. Mais la description du mouvement uniquement à l'aide de la physique s'avère rapidement compliquée à mettre en œuvre pour la simulation des organismes vivants : un niveau d'abstraction supplémentaire est nécessaire, il s'agit du niveau *comportemental*. Le développeur cherche par exemple à donner un ordre concret à son personnage de synthèse : «gare-toi derrière la pendule de la gare». Cet ordre induit un comportement, qui se traduit par des actions physiques (le déplacement d'une voiture). Ce déplacement physique est représenté à l'écran par un ensemble d'images décrites les unes après les autres par le simulateur physique. La suite de ce paragraphe est consacré à l'étude des modèles comportementaux.

Modèles de transformation internes

Ces modèles s'attachent à décrire de manière réaliste l'évolution interne d'entités vivantes (plantes, êtres humains...) en fonction de paramètres extérieurs. Il s'agit cependant de rendre compte des modifications de l'enveloppe externe de l'entité à animer comme par exemple la croissance d'une plante en utilisant des systèmes de réécriture de règles [73, 17]. Les autres aspects concernent la visualisation et l'animation d'humanoïdes de synthèse.

Humanoïdes de synthèse Du point de vue de l'animation, l'être humain se compose essentiellement d'un squelette, de muscles et d'une enveloppe extérieure, la peau. Pour animer un humanoïde, il faut définir un modèle d'un ou plusieurs de ces composants. La structure utilisée généralement est celle d'une structure osseuse (solide rigide articulé) sur laquelle est

plaquée une peau déformable. Deux types de problèmes se posent donc : le mouvement du squelette (par exemple la reproduction de la marche humaine) et la déformation de la peau (par exemple, l'animation faciale).

Le cas de la locomotion humaine est un cas particulier de locomotion. En effet, contrairement à la nage d'un poisson ou au vol d'un oiseau, la locomotion humaine se distingue par les interactions avec un sol situé à altitude variable. Il n'est donc pas possible d'ignorer complètement les contraintes cinématiques. Différentes techniques sont utilisées [126] :

- Dans le cadre du projet Jack [141, 93] de l'université de Pennsylvanie, X. Zhao [193] propose une méthode fondée sur la dynamique et la minimisation d'un critère énergétique. La minimisation de l'effort ou de l'énergie ne correspond pas toujours à une vérité biomécanique, à cause d'une corrélation très forte avec la morphologie du sujet. De plus, l'apprentissage de la marche, ou même la marche au pas des militaires sont des exemples de marche *non-économique* ;
- J. Hodgins [71] utilise des techniques classiques de contrôle : il y a un contrôleur différent pour chaque phase de la marche et ces contrôleurs sont reliés entre eux par des automates finis. Ces modèles sont extrêmement coûteux en temps de calcul ;
- Dans le cadre de l'animation interactive, des approches purement cinématiques sont proposées [23].Plus récemment, R. Boulic [24] a intégré des notions de dynamiques en regroupant les articulations dans un arbre de distribution des masses. Ces approches sont intéressantes du point de vue du coût mais permettent difficilement la personnification des démarches obtenues. Une amélioration consiste à utiliser des mouvements acquis à l'aide de systèmes vidéo et à les rejouer [54]. Le coût mémoire est alors plus élevé mais la démarche nettement plus réaliste ;
- F. Multon [125] s'intéresse à la construction de modèles biomécaniques de la locomotion humaine. Le cycle de marche est décomposé en différentes phases reliées par des automates. Pour chacun des cycles des trajectoires articulaires sont interpolées par de simples polynômes. Cette méthode permet l'identification simple de paramètres naturels comme la longueur ou la fréquence des pas.

L'animation faciale consiste à animer le visage d'un être humain. C'est une tâche complexe qui nécessite la modélisation, l'animation et le rendu de la géométrie du visage mais aussi d'éléments distinctifs du visage comme la peau, les cheveux, les lèvres ou la langue. On peut facilement imaginer le niveau de complexité du visage en se remémorant que la tête ne contient qu'une seule articulation solide, celle de la mâchoire. Tout le reste doit être représenté par des modèles déformables ! P. Fua [55] classe les modèles d'animation faciale en deux catégories : les modèles à peu de degrés de liberté, qui se concentrent sur les mouvements de la tête. et les modèles plus sophistiqués qui doivent en premier lieu extraire la géométrie à partir d'images réelles [52] ou la générer automatiquement [94], puis définir les paramètres de déformation pour contrôler les mouvements du visage [49].

Modèles de transformation externes

Les modèles dits de transformation externes définissent le comportement extérieur d'un être, c'est-à-dire ses actions et réactions vis-à-vis de son environnement, que ce soit de façon individuelle ou collective.

La classification des modèles de transformation externes n'est pas simple. Dès lors il existe une classification fondée sur le niveau d'abstraction [191], mais aussi des classifications plus techniques [42]. Une classification technique paraît plus pertinente ici dans la mesure où les techniques employées sont plus faciles à classer.

Approche stimulus-réponse L'approche dite *stimulus-réponse* définit le comportement des objets à partir d'un ensemble de capteurs et d'effecteurs ; ils sont reliés entre eux par un réseau dont les nœuds transforment l'information passée [187]. Le comportement d'un objet dépend alors à la fois de la perception qu'il a de son environnement et des transformations que va subir l'information perçue. Les réseaux SAN (Sensor Actuator Networks) [183] relèvent typiquement de ce type d'approche. Les *Space-time constraints* de Thomas Ngo [130] optimisent le comportement produit à l'aide d'un algorithme génétique. Enfin, les SCA Sense Control Action) [61] ajoutent une certaine hiérarchie à ces réseaux de comportement.

L'approche stimulus-réponse à base de réseaux est une première brique de l'animation comportementale et notamment de l'approche «Perception-Décision-Action» telle que nous la décrirons plus loin. La différence entre les approches que nous avons cités porte plus sur l'algorithme d'identification des paramètres que sur la nature réelle du réseau utilisé. Les mouvements générés peuvent être très différents, mais le niveau d'abstraction reste très bas. La partie décisionnelle est très limitée, tant dans sa portée que dans son adaptabilité. De plus, une modification d'une réaction à un stimulus quelconque de l'environnement nécessite souvent un recalcul complet des paramètres du réseau.

Règles de comportement Les approches dites à base de règles tentent de répondre au problème du niveau d'abstraction de la description du comportement : la structure reste la même, les données perçues sont transformées en actions. Simplement, il existe un ensemble de règles explicites qui spécifient l'action en fonction de la perception. De fait, les différents comportements possibles d'un objet peuvent être représentés par un arbre de décision. Cette approche est celle qui se rapproche le plus des systèmes multi-agents [59, 26] de type cognitif.

Historiquement, C. Reynolds [149] a été un des premiers à introduire des règles de comportement pour les objets à animer en s'intéressant au vol des nuées d'oiseaux ; il définit à cette fin un vol géométrique qui est un mouvement le long d'une trajectoire. Trois règles de comportement influent sur cette trajectoire : *éviter les collisions*, *harmoniser sa vitesse à celle de ces voisins* et *rester proche de ses voisins*. La figure 2.13 présente un exemple de nuée d'oiseaux dont le comportement respecte les règles précédentes.

T. Widyanto [186] propose une approche fondée sur la psychologie *éco-*

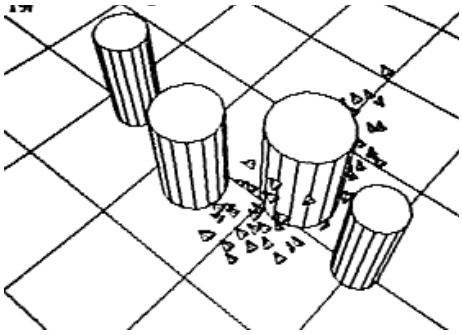


FIG. 2.13 – *Règles de comportement : une nuée d'oiseaux évitant un obstacle*

logique[4] issue des travaux de J. Gibson dont nous reparlerons à la section suivante. L'intérêt de ce travail réside donc dans la sémantique apportée aux données perçues qu'au moteur de décision. In VitrAm [124, 123] reprend partiellement cette approche en spécifiant plusieurs niveaux de comportement : les objets ne font que subir les forces auxquelles ils sont soumis, les personnages font preuve d'un comportement qui peut être réactif (au sens de l'approche stimulus-réponse décrite précédemment) ou plus cognitif faisant appel à des règles et à un moteur d'inférences. La simulation se fait de manière discrète, en tenant compte des différentes fréquences internes.

Les systèmes à base de règles apparaissent plus simples à manipuler que les systèmes stimulus-réponse que nous avons vu précédemment. Malheureusement, la spécification du comportement d'un système à animer en un petit nombre de règles est une tâche difficile à mener à bien. De plus, il existe un double problème de gestion du temps [145] :

- les algorithmes des moteurs d'inférences ne sont pas compatibles avec le temps réel et donc avec l'animation réellement interactive ou la réalité virtuelle ;
- la spécification à base de règles ne rend pas compte des fréquences de fonctionnement différentes [102] des activités menées en parallèle dans le cerveau humain.

Automates à états finis Certains systèmes d'animation comportementale ont une partie décisionnelle représentée par un ou plusieurs automates. L'Iowa Driving Simulator est un simulateur de conduite temps réel où l'on cherche à proposer à l'utilisateur un environnement virtuel réaliste tant statiquement que dynamiquement : il faut donc inclure dans cet environnement des véhicules virtuels dont le comportement se rapproche le plus possible de la réalité. Le contrôle de ces véhicules [21] s'effectue grâce à un automate d'états finis. À chaque état de l'automate (figure 2.14) sont associés des attributs qui permettent d'obtenir une commande qui s'adapte à la situation. Les conducteurs sont dotés d'un certain nombre de paramètres (tels que temps de réaction, agressivité...). Le scénario distribue statistiquement tous ces paramètres et fournit la répartition des véhicules sur le réseau routier.

[4] au sens biologique du terme : étude des êtres vivants au sein des milieux où ils vivent, ainsi que de leur interaction avec ces milieux.

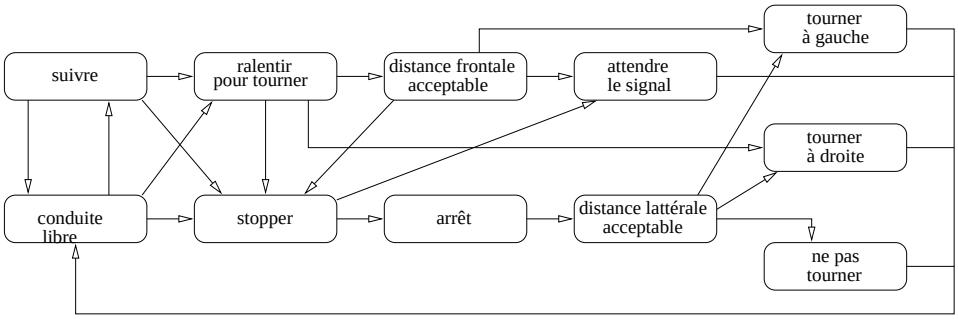


FIG. 2.14 – *Iowa Driving Simulator : extrait de l'automate de contrôle du conducteur virtuel*

Les auteurs du conducteur virtuel d'IDS [21] reconnaissent eux-mêmes un certain nombre de limitations dues à l'utilisation d'un automate unique : difficulté de modification de l'automate (taille et non-modularité), gestion quasi-impossible des interruptions, gestion des activités concurrentes. Par la suite, plusieurs solutions ont été proposées pour résoudre ce problème : Thalmann [107, 171] décompose les actions en un ensemble d'automates séquentiels qui sont gérés sous forme de piles, améliorant ainsi la modularité. Ils peuvent également interagir avec un système à base de règles [134]. En combinant SCA et Pat-Nets [185, 147] N. Badler[9, 185] propose un modèle composé à la fois d'une partie stimulus-réponse et d'une partie automates parallèles hiérarchisés qui ne prend pas en compte les fréquences de fonctionnement. La gestion des conflits entre comportements concurrents n'est pas non plus abordée. De façon similaire, le formalisme HCSM (Hierarchical Concurrent State Machine) [3, 65] a été défini pour améliorer la définition des comportements dans le simulateur IDS : il s'agit d'un ensemble d'automates hiérarchisés fonctionnant à la même fréquence. Il présente l'avantage d'être également utilisé pour la description de scénarios de simulation [36].

2.3.2 Conclusion sur les modèles de contrôle du mouvement

Cette rapide présentation des modèles de contrôle de mouvement trouve un parallèle dans le contrôle des systèmes robotiques [90]. Comme le montre la figure 2.15, la problématique est similaire sur bon nombre de points. La cinématique et la dynamique sont bien intégrés en matière d'animation / simulation mais il reste beaucoup à faire en ce qui concerne le contrôle de haut-niveau : en robotique, les travaux sur la coordination existent mais restent le plus souvent abordés au cas par cas [31] avec une forte dépendance à l'application. Malgré tout, des différences subsistent entre les deux domaines : en animation, la visualisation est une fin en soi, et non plus simplement un moyen parmi d'autres de contrôle des résultats obtenus en matière de modélisation. Le contrôle en robotique incorpore également des notions de criticité (par exemple pour éviter la destruction du matériel) qui n'intéressent pas le monde de l'animation. Dès lors, le degré de précision dans les calculs peut aussi être inférieur [4].

Il est important de noter pour la suite de ce chapitre que le contrôle de haut-niveau en robotique se fait la plupart du temps à partir de systèmes

réactifs fondés sur des automates : les langages réactifs comme Signal [25] et Esterel [14], même s'ils adoptent une syntaxe déclarative ou impérative génèrent en fait des automates pour leur exécution et leur vérification. Nous avons vu dans cette étude que les approches à base de règles étaient peu compatibles avec les contraintes du temps réel. L'approche à base d'automates peuvent s'avérer intéressante en tant que sur-couche de l'approche stimulus-réponse : c'est d'ailleurs ce qui est proposé dans les travaux de Badler. Cela ne répond pas complètement au problème à cause du manque de concurrence et de modularité. Ces notions sont alors introduites par le biais des automates parallèles hiérarchisés.

Avant d'aller plus dans la description de systèmes de contrôle du mouvement et de description du comportement, il est important de se demander comment fonctionner le mécanisme décisionnel humain (au moins dans ses aspects les plus basiques). C'est l'objet des paragraphes qui suivent.

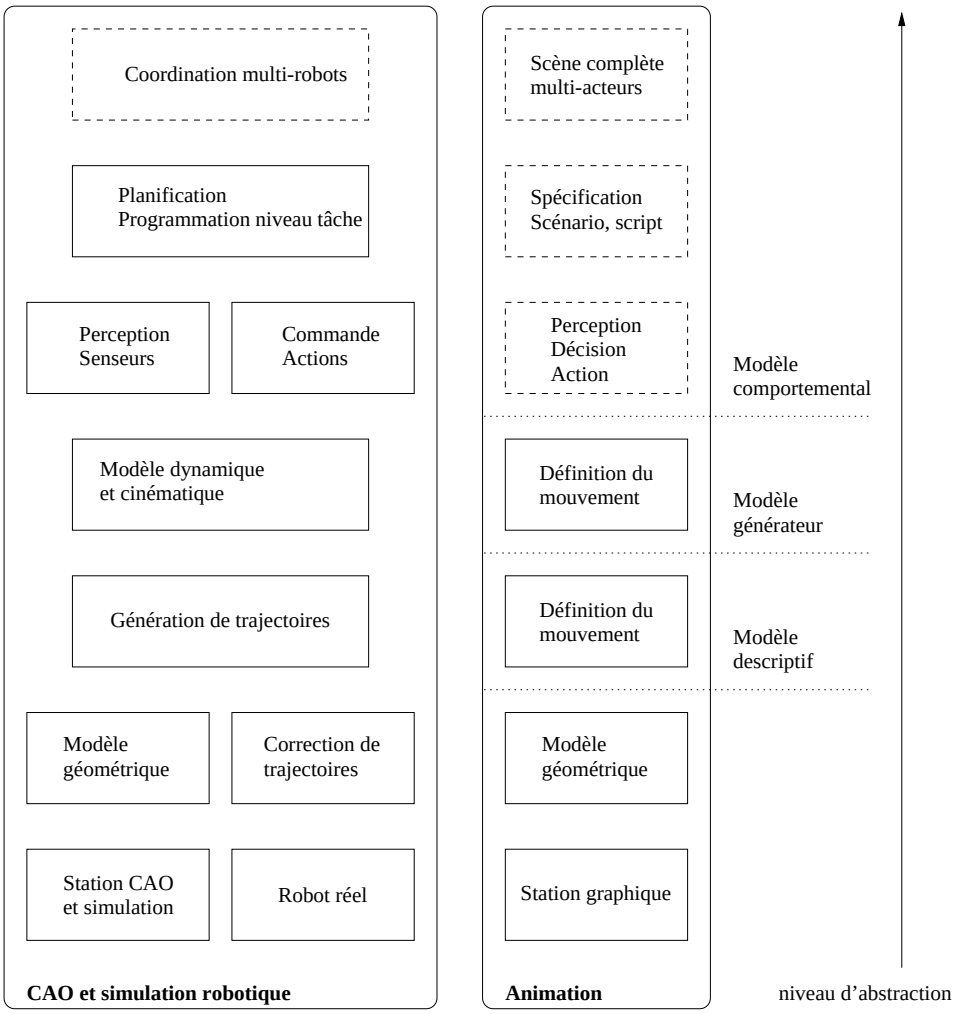


FIG. 2.15 – *Contrôle en robotique et en animation*

2.3.3 Pré-requis à la modélisation comportementale

La boucle «Perception-Décision-Action» apparaît comme une notion relativement intuitive en ce qui concerne l'architecture décisionnelle de l'être humain : en effet, l'humain voit ce qui se passe autour de lui, décide de son action en fonction de sa perception et de ses connaissances, et enfin agit sur lui-même (par exemple en apprenant une nouvelle notion) et sur son environnement (par exemple en poussant la barre pour virer). En nous basant sur l'état de l'art que nous avons effectué [114], nous proposons dans ce paragraphe les conditions de construction d'un *bon* modèle de comportement.

De nombreux chercheurs (biologistes, psychologues et informaticiens notamment) se sont penchés sur la modélisation informatique du comportement humain. Bien sûr, les approches ont été extrêmement diverses, c'est pourquoi Newell [129] s'est attaché à synthétiser les contraintes à respecter pour former un *bon* modèle de comportement. Ce sont les suivantes :

- un comportement flexible et adaptatif[5], c'est-à-dire que l'organisme simulé doit s'adapter à son environnement et à l'évolution de celui-ci afin d'effectuer ce qu'il est supposé faire, sans mettre en danger son existence ;
- des interactions temps réel avec l'environnement. Le temps réel est obligatoire lorsqu'on se place dans le cadre de la simulation interactive et donc de la réalité virtuelle ;
- un environnement riche et complexe : une entité simulée doit pouvoir percevoir son environnement, doit avoir une base de connaissance sur cet environnement (c'est-à-dire qu'elle possède un moyen de se représenter son environnement), et enfin elle doit avoir plusieurs degrés de liberté d'actions sur cet environnement ;
- représentation symbolique et abstraction de la connaissance : une entité doit posséder des informations symboliques sur son environnement ;
- possibilité d'apprentissage : une entité doit être capable de tenir compte de son expérience passée ;
- autonomie dans un environnement social : la «vie» d'une entité se déroule dans un environnement où il y a d'autres entités. Il doit alors y avoir une vie sociale, c'est-à-dire des possibilités de coopération, de conflit...
- conscience de soi, perception de soi : en plus des informations qu'elle possède sur son environnement, une entité doit pouvoir raisonner sur elle-même, et modifier ses buts. C'est ce que Th. Bouron [26] appelle la réflexivité ;
- être réalisable par un réseau neuronal : toutes les activités précédemment citées doivent pouvoir être réalisées par un réseau de neurones ;
- être constructible par un processus embryologique ;
- émergence par évolution.

Bien sûr, les trois dernières contraintes évoquées par Newell ne nous concernent pas directement et nous n'en tiendrons donc pas compte.

[5] le mot adaptatif utilisé par Newell doit plutôt être pris au sens *adapté* qu'au sens *adaptatif* couramment utilisé en robotique

En vue de sa mise en œuvre dans une application de réalité virtuelle, notre modèle devra respecter les contraintes suivantes :

- *réactivité* : le modèle se doit réagir continûment aux changements de son environnement, afin de permettre l'interaction. La réaction comprend la notion d'événements (stimuli envoyés par l'environnement) ;
- *modularité* : la modularité de la description du comportement est une nécessité informatique : des éléments de comportements doivent pouvoir être modifiés, ajoutés, supprimés sans que l'ensemble du travail de spécification soit à reprendre.

Des études précédentes [114] se dégagent également deux grandes notions :

- *La boucle* ***perception-décision-action*** : la structure comportementale d'un être humain plongé dans son environnement (la partie du monde extérieur avec laquelle il est en interaction) se présente sous la forme d'une boucle perception-décision-action reprise sur la figure 2.16. La théorie du contrôle explique les flots d'information circulant en présentant l'avantage de cumuler les avantages des théories biologiques (signaux nerveux) et informatiques (calculs).
- *Structures de décision* : la mémoire de travail s'organise sous forme de structures *condition-action* où la condition est en fait constituée par la reconnaissance d'une situation. Les systèmes de transition utilisent cette même notion (gardes) pour déclencher une action. L'organisation hiérarchique du système de décision humain met en évidence également des différences de fréquence de fonctionnement entre les processus qui régissent l'activité. Les processus les plus lents doivent pouvoir contrôler l'exécution les processus les plus rapides (démarrage, arrêt, suspension, reprise) mais par les processus eux-mêmes. Cela permet d'effectuer plusieurs activités en parallèle tout en gardant la maîtrise de l'exécution en cas d'évènement à caractère exceptionnel.

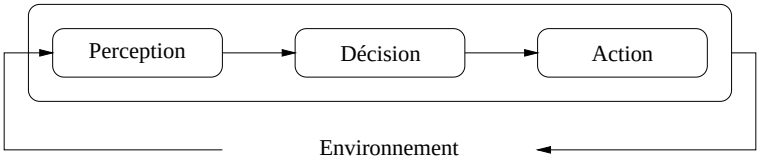


FIG. 2.16 – *Architecture d'une entité comportementale*

La synthèse de ces contraintes est présentée dans le tableau 2.1.

Le tableau 2.2 effectue le bilan des modèles de comportement au regard de la synthèse effectuée ci-dessus. A partir des manques relevés, nous allons présenter dans la section suivante le modèle HPTS que nous avons défini.

Le paragraphe suivant définit le modèle HPTS vérifiant ces contraintes, que nous proposons.

Contrainte	Motivation
Modularité	mise à jour
Réactivité	simulation interactive
Hiérarchie	architecture de contrôle
Concurrence	activités menées simultanément
Préemption	résolution des conflits
Exceptions	traitement des situations exceptionnelles
Gestion du temps	activités cognitives de fréquence différentes
Flots de données	communications entre les différents niveaux

TAB. 2.1 – *Modèles de comportement*

	mod.	réact.	hiérarchie	//	exceptions	temps	flots
stimulus	non	oui	non	non	non	non	oui
règles	non	non	non	non	non	non	non
IDS	non	oui	non	non	non	non	oui
PaT-Nets	oui	oui	oui	+/-	oui	non	oui
Virtual Life	oui	oui	oui	non	non	non	non
HCSM	oui	oui	oui	oui	indirectes	non	oui

TAB. 2.2 – *Adéquation entre les contraintes de modélisation du comportement et les modèles existants*

2.3.4 Le modèle HTPS

Définition

Le modèle décisionnel HPTS (Hierarchical Parallel Transition Systems) est un système réactif, constitué d'une hiérarchie d'automates (qui sont les comportements possibles). Chaque machine à états peut être vue comme une boite noire (cf. figure 2.17) qui comporte des flots d'entrée et de sortie, et des paramètres de contrôle. On définit donc une machine à états comme étant :

$$s =< S, \Gamma, IS, OS, CP, LV, IF, M, \nu > \tag{2.8}$$

où
- S est un ensemble de sous-états ;
- Γ est la fonction d'activité des sous-états ;

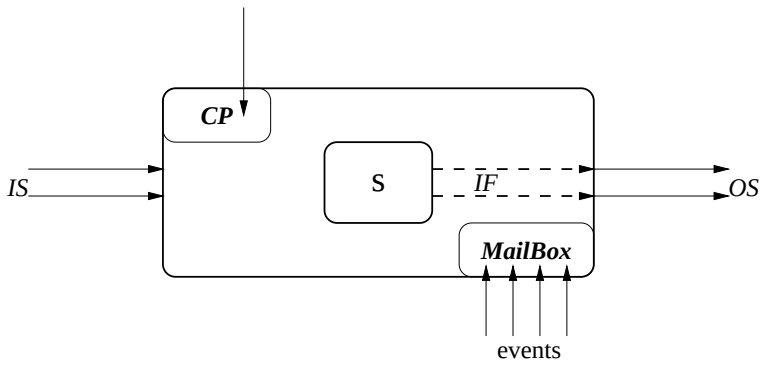


FIG. 2.17 – *HPTS : architecture*

– IS est un ensemble de signaux d'entrée ;
– OS est un ensemble de signaux de sortie ;
– CP est un ensemble de paramètres de contrôle ;
– V est un ensemble de variables locales ;
– IF est la fonction d'intégration ;
– M est la boite aux lettres du système ;
– ν est la fréquence de fonctionnement.

s est soit un état atomique du système, dans ce cas $S = \emptyset$, soit un état composite ($S \neq \emptyset$). À chaque état s est attribué un paramètre d'activité a_s tel que :

$$a_s \in \mathcal{A} = \{idle, active, suspended\} \tag{2.9}$$

La fonction d'activité Γ est donc la fonction qui fait évoluer le paramètre d'activité $a_{s'}$ de tous les sous-états $s' \in S$. Par exemple, après avoir débuté, un état se retrouve actif. La fonction d'activité Γ à valeurs dans $\mathcal{A}$ s'écrit :

$$\forall s' \in S, a_{s'}(t) = \Gamma(a_{s'}(t-1), IS, CP, V, M) \tag{2.10}$$

Il existe une suite logique entre les statuts. La fonction d'activité est contrainte par cette suite, qui peut être elle-même représentée sous la forme d'un automate repris figure 2.18. Le carré situé dans l'état *active* signifie que la fonction d'intégration (décrite plus loin) est exécutée. Les transitions entre états se font à chaque pas de temps de simulation. Il n'est donc pas possible de passer d'*idle* à *suspended* en un seul pas de temps.

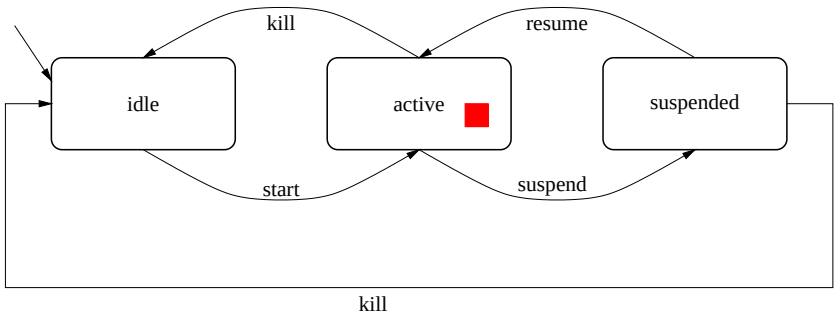


FIG. 2.18 – *HPTS : automate de fonctionnement*

La boite aux lettres M est un système de messagerie qui permet de recevoir des messages (provenant notamment des sous-états ou de l'état père). Un message est ici un événement accompagné de données. On distingue un certain nombre d'événements spécifiques qui commandent la fonction d'activité Γ : ce sont les événements qui apparaissent sur la figure 2.18 et qui permettent de modifier l'état d'une machine HPTS. À chaque primitive est associé un événement renvoyé par la machine HPTS à qui le message était destiné. Cet événement permet à la machine de signaler son nouveau statut.

On distingue deux types de messages :
– les messages globaux : leur diffusion est assurée dans un environnement de broadcasting, c'est-à-dire que toutes les machines HPTS y ont accès, et que toutes peuvent les utiliser ;

– les messages locaux : ceux-ci ont un destinataire particulier spécifié lors de l'envoi. Ce mécanisme permet de mettre en œuvre la communication entre machines HPTS, voire entre entités comportementales. On pourrait imaginer une diffusion générale de tous les messages, où le protocole de communication incorporerait le destinataire (par exemple comme dans le protocole IP), mais ce serait moins efficace d'un point de vue transmission de données.

La fonction d'intégration gère la cohérence entre les différents comportements proposés par les sous-états. On distingue trois cas :

– $S = \emptyset$: s est donc un état atomique, IF est donc une fonction qui ne dépend que des entrées IS, des variables locales et des paramètres de contrôle ;
– *card* $S = 1$: s a un seul sous-état s_1 actif, dont la sortie est le résultat de la fonction d'intégration de s_1 appliquée sur les sous-états actifs de s_1. Typiquement, la fonction IF de s est alors une fonction identité. On peut imaginer que IF soit aussi un filtre ;
– *card* $S > 1$: c'est le cas le plus complexe : IF est alors l'intégration des différents comportements proposés par les sous-états actifs s' de S. Nous étudierons ce problème plus tard. De manière formelle, IF s'écrit simplement :

$$OS(s) = IF(OS(s', \forall s' \in S), V, IS, PC) \tag{2.11}$$

Un exemple relativement simple de fonction d'intégration, dans le cadre de la simulation de conduite, est présenté figure 2.19. Les trois sous-états de l'état *supervisor* proposent chacun une action paramétrée. Aux actions sont associées des priorités, comme le montre la relation d'ordre définie dans la boite *action*. Si deux états proposent la même action, une loi conservative est appliquée sur ses paramètres : en cas de demandes de freinage simultanées, la distance d'arrêt la plus courte est choisie.

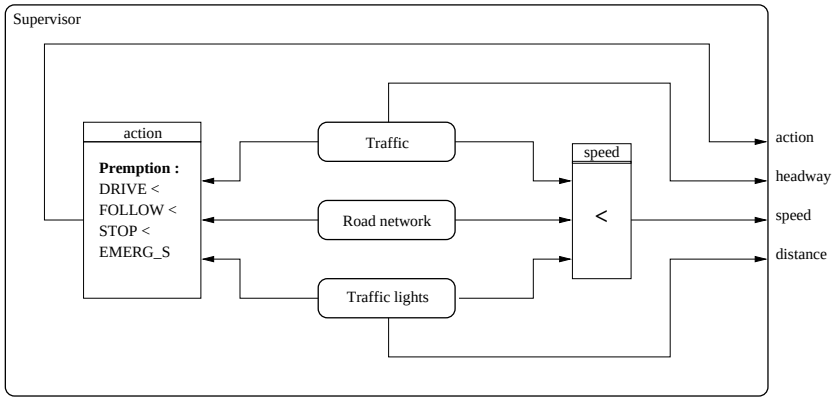


FIG. 2.19 – *HPTS : exemple de fonction d'intégration*

Aspects temporels

Le n-uplet qui définit une machine HPTS incorpore un paramètre ν, défini comme la fréquence d'exécution. En pratique, seul le cas où la fréquence d'exécution d'un automate s est un multiple de la fréquence de son automate père a été envisagé. Tout en évitant les problèmes de déphasage,

cela pose néanmoins un certain nombre de problèmes, que nous allons aborder dans les paragraphes suivants.

Si on se place dans le cas d'un module de simulation comportementale qui doit produire des données pour un contrôleur de mouvement à une fréquence de référence ν_0, les sorties OS de la machine HPTS de plus haut-niveau s_0 doivent être produites elles aussi à une fréquence ν_0. Cela implique que les fonctions d'intégration des automates placés au niveau le plus bas de la hiérarchie doivent être exécutées aussi à une fréquence nominale ν_0, puisque ce sont elles qui ont le temps de réponse le plus faible (cf. [114] première partie, sur la théorie du contrôle). Maintenant, si les fonctions d'intégration des sous-automates produisent leurs données à une fréquence ν_0, il devient nécessaire d'intégrer ces différents comportements proposés à cette même fréquence. On en déduit donc que la fréquence de calcul de IF pour s et tous ses sous-automates est égale à la fréquence de calcul des sorties OS.

Les variables locales pourraient être calculées de deux manières : soit en même temps que la fonction d'activité, soit en même temps que la fonction d'intégration. Pour permettre des calculs temporaires sur les comportements et des calculs liés au temps (donc à la fréquence d'activité), les deux possibilités sont offertes.

Par opposition aux entrées, les événements sont sporadiques. Néanmoins, pour gérer les différences fréquentielles, ils doivent avoir une certaine durée de vie, c'est-à-dire jusqu'à leur consommation effective par l'automate auquel il sont destinés. C'est la fonction de la boite aux lettres M. Celle-ci est donc remplie de façon asynchrone, mais consultée et vidée à chaque exécution de la fonction d'activité, seule consommatrice de ces événements.

Nous avons montré que la gestion des différentes fréquences de fonctionnement de la hiérarchie n'était pas possible au niveau des sorties, elle se fait donc au niveau des fonctions d'activité. La fonction d'activité de l'automate de plus haut niveau s est exécutée à sa fréquence ν, et les sous-automates sont exécutés à des fréquences ν_i multiples de ν. Grâce au mécanisme de boite aux lettres, cela permet à un automate de haut-niveau de contrôler à l'aide d'événements des fonctions de plus bas-niveau.

Les entrées doivent être disponibles à chaque fois qu'un calcul le nécessite, c'est-à-dire :

- lors du calcul de la fonction d'intégration, soit à la fréquence ν_0 ;
- lors du calcul de la fonction d'activité de s et de chacun de ses sous-automates, soit à $\bar{\nu} = ppcm(\nu, ppcm_{i \in S}(\nu_i))$;
- lors de l'évaluation des variables locales, c'est-à-dire à $ppcm(\nu, \bar{\nu})$.

On remarque qu'un choix judicieux de ν et ν_0 simplifie considérablement les calculs.

Exécution

Dans ce paragraphe, nous présentons l'algorithme d'exécution du modèle HPTS. On suppose les algorithmes exécutés à la fréquence $ppcm(\nu, \bar{\nu})$ et que t est un paramètre définissant le numéro du pas d'exécution.

```
run(s : HPTS)
    {
    // lecture des événements (à la racine)
    if (s.pere == ∅)
    {
        lireEvt();
    }
    // Exécution de la fonction d'activité ?
    if (t mod s.ν == 0)
    {
        ∀s' ∈ s.S
        {
            // calcul du nouvel état
            a_s'(t) = Γ(a_s'(t − 1), IS, CP, V, M)
            // Exécution de la machine HPTS fille
            run(s')
        }
    }
    // mise à jour éventuelle des variables locales
    if (t mod s.ν ou t mod s.ν_0)
    {
        ∀v ∈ s.V, v(t) = Ω(v(t − 1), IS, CP, V)
    }
    // si actif, exécution de la fonction d'intégration
    if (s.a == active et t mod s.ν_0)
    {
        OS(s) = IF(OS(s', ∀s' ∈ S), V, IS, PC)
    }
    // écriture des événements
    ecrireEvt();
}
```

Les appels récursifs à la fonction `run` mettent en évidence la hiérarchie tandis que la concurrence est représentées par les itérations sur les machines HPTS filles. Les variables étant locales à chaque machine HPTS, les problèmes de dépendance temporelles des données sont ainsi évités.

Différentes implantations des machines HPTS ont été proposées :

- sous forme de couplage [48] avec le langage SignalGti permettant ainsi de spécifier les propriétés temporelles de HPTS mais au prix d'une définition extrêmement ardue ;
- sous forme de Statecharts [66] par Hahusseau [64] aux dépends de la gestion du temps et de l'expressivité des fonctions d'intégration ;
- directement en C++ dans le cadre de la plate-forme de simulation GASP [45, 46], limitée par la difficulté du langage C++ pour les non-informaticiens ;
- enfin, nous avons proposé un langage de description du modèle HPTS couplé à un générateur de code spécifique pour la plate-forme GASP, conservant ainsi l'expressivité des fonctions d'intégration, la puissance associé à C++ et une description relativement simple des machines HPTS. Le langage associé et le générateur de code sont

décrits plus avant dans [116] et bien sûr dans [114]. Une ébauche d'interface graphique a également été proposée [2].

2.3.5 Application à la simulation de trafic

Dans les paragraphes précédents, nous nous sommes donnés les moyens de spécifier le comportement humain pour un certain nombre d'activités. Nous nous intéressons maintenant aux applications de notre modèle de comportement dans le cadre de la simulation du trafic routier. Deux grandes classes d'application sont présentées :

- *trafic multimodal en environnement urbain* : il s'agit d'étudier les comportements des différents moyens de transport urbains (piétions, véhicules, tramways). Nous verrons qu'il est nécessaire de mettre en œuvre un processus de perception de l'environnement ;
- *trafic autoroutier dense* : il s'agit d'étudier l'influence de nouveaux moyens de signalisation sur un trafic autoroutier dense comportant plusieurs milliers de véhicules. Ce travail a été réalisé dans le cadre du projet européen DIATS.

Classiquement, les simulateurs de conduite effectuent soit une simulation de conduite (un poste de conduite assorti d'un modèle dynamique de véhicule), soit une simulation de trafic (modèles fluides ou modèles à particules très simples), mais très peu sont capables d'effectuer les deux en même temps. On peut noter que le comportement des véhicules simulés (ceux qui ne sont pas dirigés par l'utilisateur) est souvent basique, les interactions entre véhicules manquent généralement de réalisme et diversité. De plus, dans la plupart des cas, seuls les véhicules légers sont traités, piétons et cyclistes sont oubliés alors qu'ils représentent une part importante de l'accidentologie (supérieure à leur représentation sur la route).

Environnement de simulation

L'environnement de simulation de l'équipe SIAMES [6] de l'IRISA dans laquelle se sont déroulés ces travaux, est composé de plusieurs éléments rappelés figure 2.20 qui génèrent des données ou du code pour la plateforme de simulation GASP :

- DREAM [35] est un langage de modélisation des corps solides et articulés permettant la modélisation de systèmes dynamiques complexes comme une voiture ou un bras humain.
- VUEMS [7] [43, 44] est un modeleur/générateur d'environnements urbains réalistes : en effet, une fois la phase de modélisation effectuée, VUEMS génère un environnement 3D complet pour la simulation. Par environnement complet on entend un environnement qui comprend non seulement la géométrie (pour la visualisation), mais aussi la topologie du réseau routier (pour la conduite) et la sémantique (pour le comportement). L'architecture générale de l'outil VUEMS est reprise figure 2.21.
- GASP [45, 46] est un environnement de simulation distribué mis au point dans l'équipe SIAMES qui permet d'agréger par le biais de mo-

[6] devenue Bunraku récemment
[7] Virtual Urban Environment Modelling System

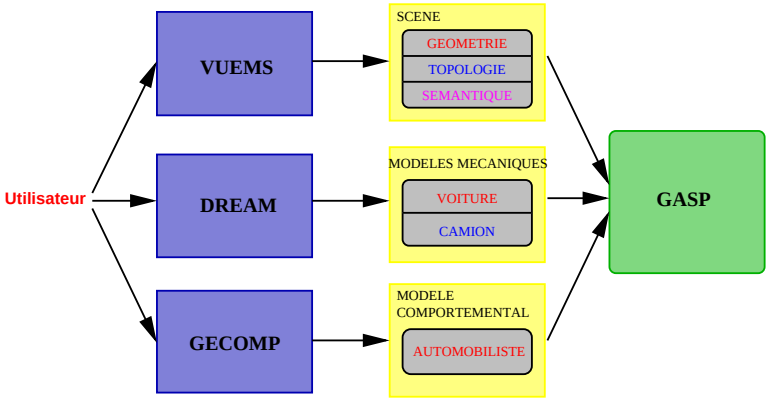


FIG. 2.20 – *Outils de conception de simulation de l'équipe Siames*

dules l'ensemble des travaux de recherche effectués dans l'équipe, chacun profitant ainsi des résultats de recherche des autres. Pour effectuer nos simulations, nous nous sommes naturellement appuyés sur les modules existants.

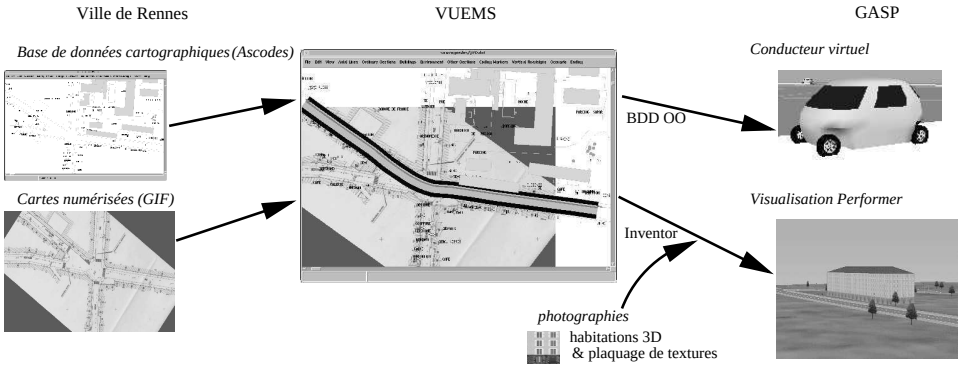


FIG. 2.21 – *VUEMS : modélisation de l'environnement urbain*

Création d'un modèle de conducteur virtuel

Nous avons créé un modèle de conducteur virtuel (en réalité plutôt un couple conducteur-véhicule) reprenant le schéma mis en évidence de perception-décision-action. Les modules qui le composent sont regroupés en une *entité* décrite figure 2.22. Toutes les objets doués d'un comportement dans le monde virtuel seront regroupés ainsi, permettant la communication entre objets à travers une interface uniformisée. Les objets de type capteur, pour la perception, n'auront ainsi pas à se préoccuper des objets recevant l'information. Nous allons rapidement décrire les différents éléments composant la base générique de l'entité conducteur-véhicule, celle-ci sera spécialisée pour les différentes applications :

– un contrôleur bas-niveau permettant l'utilisation du modèle dynamique à trois entrées (couple sur l'accélérateur, couple sur le frein et angle volant) permettant des actions simples comme accélérer, freiner ou suivre un véhicule [5]. Ce contrôleur a également besoin

d'un point cible (la direction du véhicule) qui fait l'objet d'un calcul spécifique (tronçon routier droit ou en courbe, franchissement d'intersection, changement de voie...) décrit dans [114] ;
- un module de perception de l'environnement : l'immersion d'entités autonomes dans un environnement virtuel ne se limite pas à la géométrie [11, 136]. Pour éviter de simuler la vision humaine après production d'images de synthèse, nous avons choisi de l'émuler en utilisant les informations disponibles (et notamment celles générées par Vuems) en construisant différents modèles de capteurs prenant en compte le champ de vision et les occultations entre objets de façon limitée afin de ne pas trop pénaliser les temps de calcul. Les capteurs effectuent un travail d'interprétation (aidés par les bases de données utilisées) dont les résultats sont fournis directement au modèle de conducteur, nous appuyant pour ce faire sur la théorie écologique de la perception. Le conducteur peut ainsi connaître directement la couleur d'un feu de circulation par exemple (dès lors que celui-ci est visible).
- un modèle de comportement de conducteur menant trois grandes activités de front : le suivi du réseau routier selon un itinéraire fourni (avec utilisation des voies appropriées et notamment des présélections le cas échéant), le respect (ou non) de la signalisation et la prise en compte de l'environnement dynamique (les autres). Ce modèle est relativement conservateur au sens où le conducteur virtuel privilégiera le freinage au dépassement en cas de doute (freinage du véhicule précédent à un feu rouge non vu par le véhicule courant).

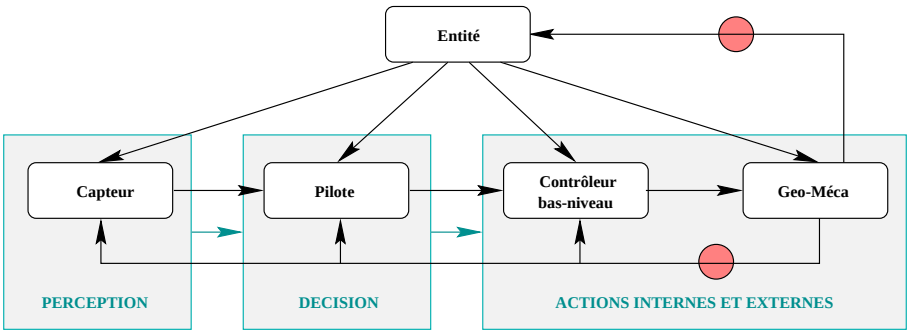


FIG. 2.22 – *Architecture du conducteur virtuel*

Interactions multimodales en environnement urbain

Dans le cadre d'une collaboration avec la société ITI, nous avons étudié l'impact de l'intégration du tramway dans un environnement de circulation urbain. Notre travail a consisté à étudier les possibilités d'interblocage dûs à l'ajout de feux de circulation sur les ronds-points. Ces feux, déclenchés par des capteurs à boucle magnétique, permettent de faire circuler en priorité les tramways. Le cas particulier du carrefour de la Croix-Bonneau à Nantes (reproduit figure 2.23) posait le problème du dédoublement de la ligne de tramway. Cette intégration a nécessité divers travaux :
- création d'un modèle géométrique, cinématique et comportemental du tramway (la dynamique n'a pas été nécessaire) ;

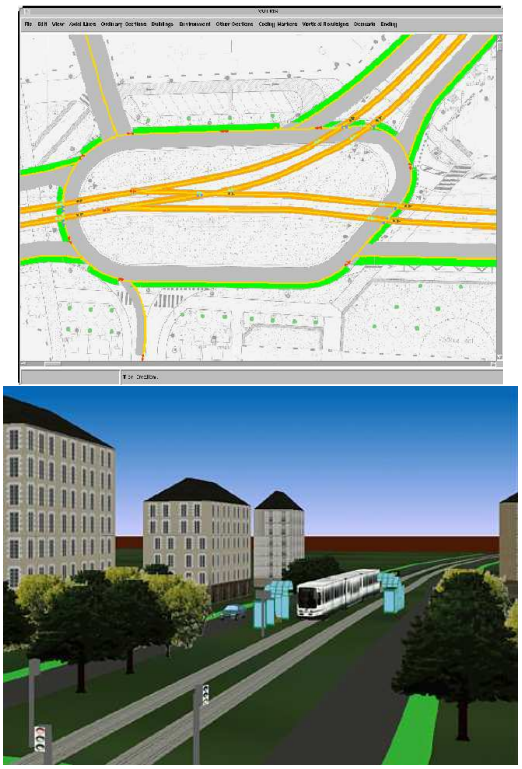


FIG. 2.23 – *Intégration du tramway dans le carrefour de la Croix-Bonneau à Nantes*

- extension du modeleur VUEMS pour la prise en compte des voies ferrées (voies, signalisation, capteurs spécifiques) ;
- feux de signalisation spécifiques reliés à des capteurs et non plus à des horloges comme les feux tricolores classiques.
- feux tramways : la signalisation ferroviaire est assez complexe à décrire, nous avons choisi d'utiliser des machines HPTS, l'automate de supervision d'une feu est repris figure 2.24.

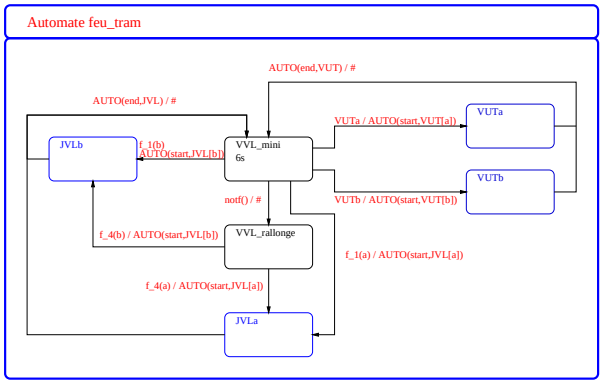


FIG. 2.24 – *automate de supervision d'un feu tramway*

Pour ces projets, nous avons développé les composants d'une voiture virtuelle autonome respectant le principe de la boucle perception-décision-action. Ces composant ont été enrichis par le développement d'un modèle de tramway ainsi que d'un modèle de bus. Ces certains élé-

ments d'interaction avec des piétons virtuels ont également été mis en œuvre par Thomas et Donikian [173]. Un autre projet avec la société ATN a servi à étudier l'emplacement des voies réservées aux bus dans un grand carrefour parisien.

Projet DIATS

FIG. 2.25 – *Projet DIATS : déploiement de moyens télématiques pour l'optimisation des infrastructures autoroutières à fort trafic (à droite le* ramp metering*)*

DIATS (Deployment of Interurban ATT Test Scenarios) est un projet européen (DG VII Transportation Program) dont l'objectif est de tester le déploiement de systèmes automatisés (à l'intérieur des véhicules ou dans la signalisation) sur des autoroutes inter et péri-urbaines. Ces routes se caractérisent par un trafic routier dense de plusieurs milliers de véhicules par heure. Plutôt que de construire de nouvelles autoroutes, les gouvernements européens ont cherché à optimiser les infrastructures existantes : les dispositifs à tester incluent les panneaux à message variables (VMS) et notamment les limitations de vitesse variables, les régulateurs de vitesse intelligents (ACC) ou le contrôle des flots d'entrée sur réseaux saturés (ramp metering). Les partenaires de ce projet étaient le Transport Research Group de l'Université de Southampton, le Transport Research Lab, l'Université Technique de Hambourg, l'INRETS, le SINTEF et Stratec Consultant. L'ensemble des partenaires cherchaient à reproduire différents scénarios et une validation croisée des résultats était effectuée à partir de paramètres de mesure communs.

Par rapport au modèle de conducteur virtuel générique évoqué précédemment, des modifications ont dues être apportées afin de prendre en compte ces nouveaux moyens de signalisation. La quantité de véhicules a impliqué la simplification des modèles géométriques afin de permettre un affichage en temps réel. De plus, les flots de véhicules lancés sur les autoroutes obéissaient à des règles précises de sorte que nous avons dû créer un orchestre pour ordonner cette simulation et récupérer les statistiques macroscopiques des flots de circulation. Cet orchestre a été implanté sous forme de machine HPTS préfigurant ainsi l'extension de HPTS au pilotage de scénario [47, 41] dont l'architecture générale est présentée figure 2.26. Les modules de perception ont été également été largement simplifiés puisqu'ils ne prennent plus en compte les objets fixes comme les arbres et les bâtiments. D'autres faits ont également été intégrés pour simplifier le modèle de perception : tous les véhicules circulent mainte-

nant dans le même sens, il n'y a que de véhicules (voitures ou camions) et les seules intersections sont les voies d'accès et de sortie. Dès lors, seules les informations sur les véhicules les plus proches sont nécessaires. Les modules comportementaux relatifs au dépassements ont été enrichis en fonction de l'état de l'art disponible [111].

Comme mentionné en introduction, nous avons utilisé de nouveaux systèmes embarqués ou de signalisation comme :

- *les VSL* : ce sont des panneaux de limitations de vitesse variables. Les limitations de vitesse évoluent en fonction des données fournies par un capteur situé quelques centaines de mètres plus loin. Si la vitesse moyenne au passage du capteur diminue trop, la vitesse est limitée en amont à un seuil inférieur. Cela permet d'éviter les décélérations trop brutales qui ont pour conséquence une augmentation des risques d'accident, ainsi que la formation de bouchons.
- *le Ramp Metering* : il s'agit ici de contrôler le flot d'accès sur les autoroutes, lorsque le trafic est intense. Il y a un feu de circulation sur la voie d'accès à l'autoroute, qui laisse passer les véhicules un par un. Le comportement du feu est amené à changer en fonction du débit de l'autoroute et de la taille de la file sur la voie d'accès.
- *l'A(I)CC* : c'est un système de régulation de vitesse intelligent embarqué sur les véhicules. Au départ, il s'agit d'un régulateur de vitesse standard dont la fonction est de maintenir la vitesse à une valeur donnée. L'intérêt de l'AICC par rapport aux régulateurs classiques est que, grâce à ses capteurs, il cherche aussi à maintenir une distance de sécurité raisonnable par rapport au véhicule précédent. Les avantages sont triples : en premier lieu un système électronique embarqué a un temps de réaction nettement inférieur à un être humain et ne connaît pas de problèmes de vigilance. Ceci contribue à une amélioration de la sécurité du trafic. Dans un second temps, cette vitesse de réaction augmentée permet de minimiser l'amplitude de la réaction à une modification de l'environnement. De ce fait, le trafic n'est plus soumis aux effets d'accordéon ; la circulation est donc plus fluide, et en sollicitant moins la mécanique de la voiture, la consommation d'essence (et donc la pollution) diminue du même coup. À terme ce système peut permettre de diminuer les distances de sécurité entre les véhicules et donc augmenter le flot maximal admissible sur une autoroute. Nous avons implanté différents types d'AICC dans le cadre de DIATS.

Une première étape de calibrage a été ensuite effectuée selon un scénario simple : une autoroute à 3 voies est perturbée par un rétrécissement à deux voies sur une distance de 800m (figure 2.27). Le trafic est généré de façon régulière et déterminée, simplement les vitesses moyennes et distances de sécurité assignées aux véhicules sont répartis selon une loi normale. Sur le scénario de calibrage, notre modèle SSE fournit des résultats similaires aux autres.

Parmi les études effectuées durant ce projet, nous nous sommes intéressés au problème de la validation des algorithmes de contrôle des AICC selon deux scénarios : une autoroute à trois voies classique d'une part, puis dans le même ordre d'idée que le scénario de calibrage, une neutralisation de la voie de droite. La figure montre le résultat d'une comparaison

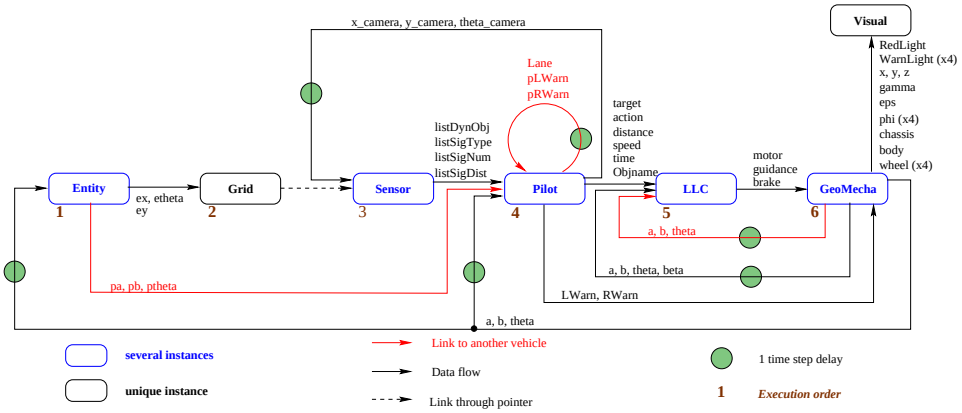


FIG. 2.26 – *Architecture d'une simulation type pour le projet DIATS*

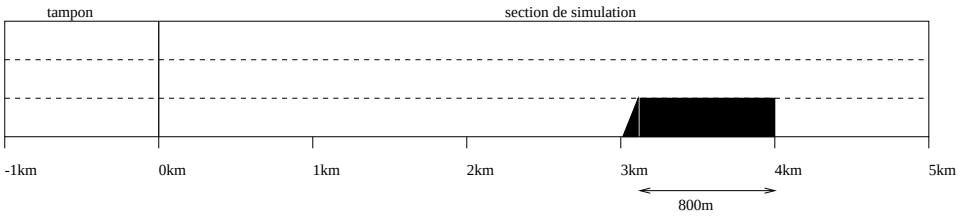


FIG. 2.27 – *Scénario de calibrage pour le projet DIATS*

des temps de traversée en fonction du type d'AICC (TRL, TNO ou INRIA qui est noté NEW sur les graphiques) paramétrés avec des headways [8] différents et un pourcentage d'équipement des véhicules variables : les temps de parcours ont tendance à augmenter et les vitesses moyennes des véhicules diminuent en fonction. Sur le second scénario, l'AICC diminue sensiblement les temps de parcours en raison de son temps de réaction plus rapide que l'être humain. En revanche, ses capacités de décélération sont limitées. En revanche, lorsqu'on diminue les headways des AICC, les véhicules non-équipés induisent une distribution hétérogène menant à des situations dangereuses.

[8] distance de sécurité visée exprimée en secondes.

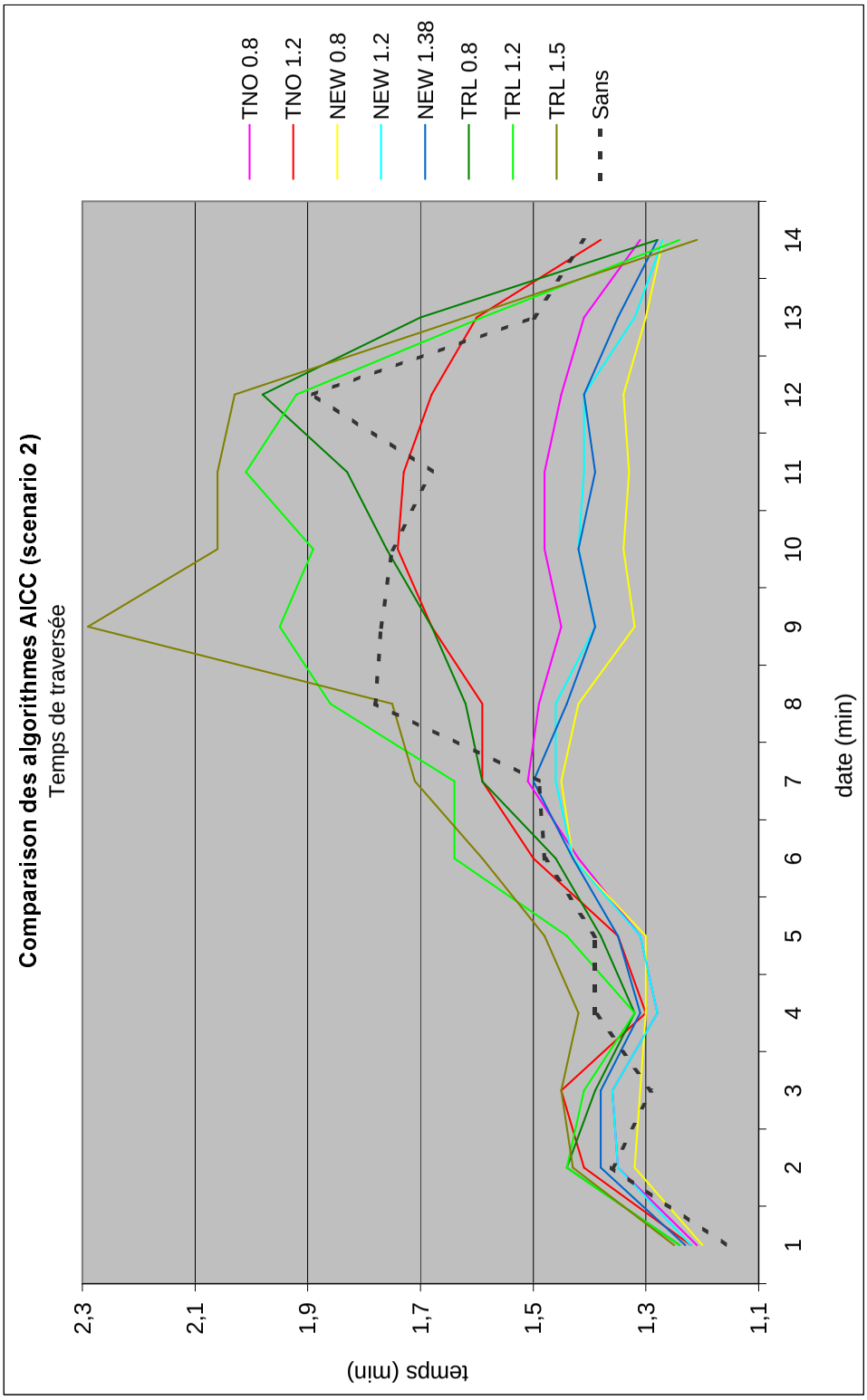


FIG. 2.28 – *Comparaison de différents dispositifs AICC sur le scénario 2*

Les images comme lien entre le monde réel et le SIG3D+t

3

Dans les chapitres précédents, nous avons successivement abordé la perception «réelle» (en quoi est-ce pertinent d'accéder à un système d'informations par l'intermédiaire d'images ?) et la perception «synthétique» (la construction automatisée des environnements virtuels – éléments constitutifs des systèmes d'information – à partir d'algorithmes basés sur une perception réelle ou synthétique). La proposition faite dans ce chapitre, qui couvre majoritairement des travaux en cours et à venir est de refermer cette boucle en considérant que le Système d'Information peut être représenté par des images mais aussi construit et mis à jour directement avec des images, dans un contexte temps-réel[1]. Nous nous appuyons pour cela sur les techniques de la réalité augmentée avec un domaine d'application particulier : les systèmes d'information géographique.

3.1 Introduction : contexte applicatif

Nos travaux actuels se déroulent au laboratoire CERMA, membre de la fédération de recherche IRSTV. Le domaine d'application naturel de nos recherches est donc le domaine architectural et urbain. Assurant la responsabilité du Projet de Recherche Fédératif «Données urbaines» de l'IRSTV ainsi que celle de l'action SIG du projet MEIGEVille (Modélisation Environnementale Intégrée et GEstion durable de la Ville), j'ai orienté mes travaux vers le domaine des Systèmes d'Information Géographique. Ce sont des bases de données intégrant une composante spatiale. Ils sont devenus les outils incontournables de gestion et de développement des territoires. Déployés en milieu urbain, ils sont à la fois des outils de diagnostic, de gestion, de concertation et d'aide à la décision pouvant faire intervenir à la fois des experts et des non-experts. En effet, les évolutions législatives de ces vingt dernières années font intervenir le public dans les toutes les décisions relatives à l'espace.

Comme pour les bases de données, les principaux défis liés à l'exploitation des SIG sont la structuration des données, leur mise à jour et leur croisement afin d'exploiter les connaissances ainsi construites. Les données utilisées sont hétérogènes par nature (relevés qualitatifs in situ, campagnes de mesure type télédétection, simulations numériques...) mais aussi représentées dans les systèmes d'information avec une qualité et une échelle spatiale différente. Elles peuvent également évoluer dans le temps.

[1]On entend ici par «temps-réel» le temps interactif, c'est-à-dire le système qui réagit dans un délai non gênant pour l'utilisateur.

Enfin dans le cas spécifique de la ville, il existe un besoin, aujourd'hui mal adressé par les outils actuels de représentation, de prendre en compte les 3 dimensions de l'espace ainsi que la dimension temporelle.

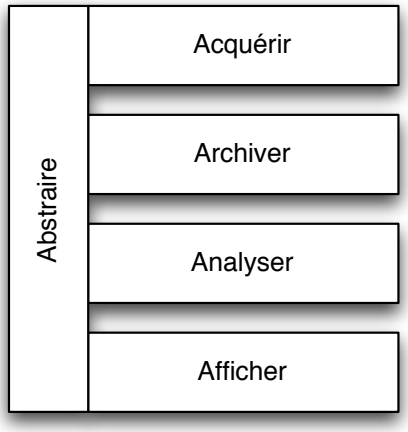


FIG. 3.1 – *Les 5A des SIG, d'après [39]*

Dans la suite de chapitre, nous allons reprendre la classification issue de Denègre et Salgé [39] que nous avons introduite en début de ce document (les 5A, repris figure 3.1) et nous introduirons les travaux en cours et à venir dans le domaine des SIG3D+t. Le lien avec les images réelles ou de synthèse sera le fil conducteur de cette classification dans les parties où cela a un sens : celles sur l'acquisition et l'affichage comme le montre le schéma de principe de la figure 3.2.

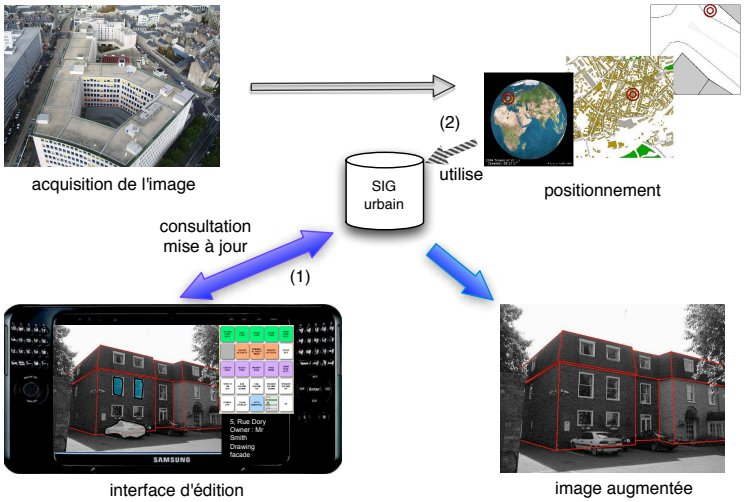


FIG. 3.2 – *Vers une boucle fermée entre SIG3D+t et images : réalité augmentée en milieu urbain – les images sont utilisées pour représenter et enrichir le SIG (1) mais le SIG est également utilisé pour géo-localiser les images (2).*

3.2 Abstraction

Comme nous l'avons mentionné en introduction, les SIG urbains présentent la particularité de mélanger de nombreux types de données : leur nature géométrique peut varier, les données attributaires associées sont également variables, l'échelle d'acquisition est aussi différente en fonction des moyens d'acquisition. Enfin, ce schéma serait incomplet si l'on n'ajoutait pas que la collecte d'information elle-même évolue dans le temps : non seulement les données ont une évolution et une validité temporelle limitée, mais les concepts eux-mêmes évoluent. Par exemple, le POS (plan d'occupation des sols), dont le contenu était remis en question périodiquement, est devenu le PLU (plan local d'urbanisme), notion beaucoup plus riche et non *plongeable* dans un POS.

Nous avons commencé à aborder la problématique de la modélisation spatio-temporelle dans les SIG avec la thèse de C. Zaki (démarrée en octobre 2007) en nous intéressant d'abord au temps et à l'évolution sémantique. Nous avons récemment proposé un couplage entre le modèle MADS et les ontologies afin de permettre l'évolution sémantique des données spatio-temporelles [190].

Les travaux à venir dans ce domaine chercheront à intégrer la troisième dimension dans les modèles, mais aussi la problématique de la multi-représentation des données. Ceci est d'autant plus nécessaire en 3D que nous disposons aujourd'hui de modèles très partiels de l'espace urbain : ces modèles ont été produits à des époques différentes, avec des moyens d'acquisition eux aussi différents qui ont abouti à des modèles de précision différente. La validité temporelle de ces modèles est-elle même bien entendu limitée dans le temps. L'utilisation des ontologies voire des Systèmes d'Information Logique appliquées aux données géographiques [12] comme le propose l'équipe LIS de l'IRISA apparaît comme une piste intéressante nécessitant des travaux supplémentaires puisque les SIL ne prennent à ce jour pas en compte la composante temporelle ni la géométrie en dimension 3.

L'utilisation d'un modèle de données nouveau nécessite, afin de bien l'intégrer dans un SIG (stockage et accès aux données), de pouvoir manipuler ce SIG au niveau du modèle de données interne.

3.3 Archivage

Nous abordons très succinctement le problème de l'archivage dans le cadre du projet MEIGEVille labellisé par la Région Pays de la Loire. Dans ce projet, nous proposons l'utilisation d'un outil unique dans l'ensemble de l'IRSTV pour le partage des données géographiques (éventuellement 3D+t) mais aussi celui des connaissances et des méthodes de gestion de l'espace urbain. Dans ce cadre, l'outil OrbisGIS a été conçu et développé par une partie de l'équipe «Données urbaines» de l'IRSTV (E. Bocher, T. Leduc et Fernando Gonzalés-Cortés) comme un outil support à la recherche, mais aussi parfois comme un outil utilisant des résultats de travaux de recherche. Nous avons par exemple mis au point GMDS [19, 91], une couche d'abstraction (*middleware*) pour l'accès aux données vecteur mais aussi raster indépendamment de la source physique de données :

l'utilisateur, mais aussi le développeur, utilise les données vecteur comme raster à partir d'un langage unique d'accès à celle-ci.

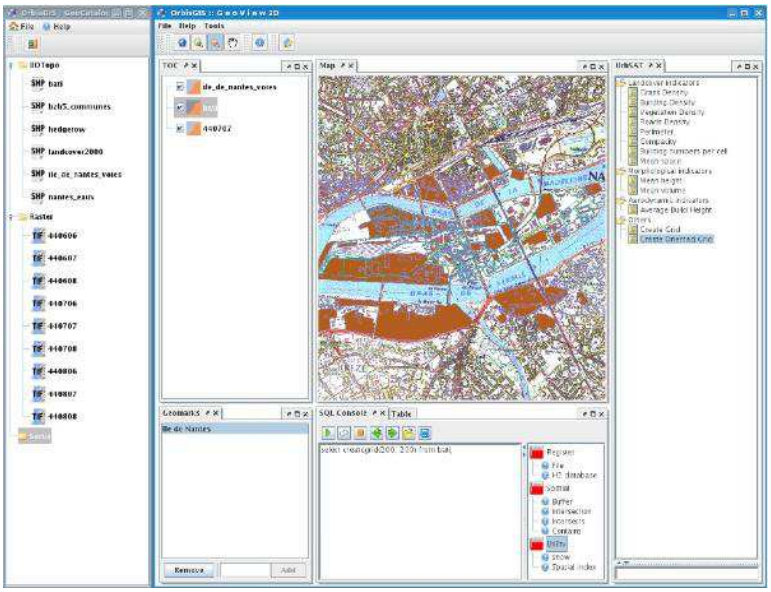


FIG. 3.3 – *Capture d'écran de l'outil OrbisGIS*

OrbisGIS a été développé sur le modèle Open Source et est accessible au téléchargement[2] depuis près d'un an. Tous les travaux évoqués dans ce chapitre utilisent ou utiliseront dès leur aboutissement le cadre de développement fourni par OrbisGIS : le modèle de données issus de la thèse de C. Zaki seront naturellement implantés dans OrbisGIS tandis que la boucle que nous cherchons à fermer grâce à l'utilisation de la réalité augmentée pour la construction et la consultation du SIG s'appuie sur des requêtes SIG effectuées via GDMS. Enfin les travaux d'analyse de morphologie urbaine comme urbSAT [100] ont constitué la première application réelle d'OrbisGIS et les projets en cours dans l'équipe «Données urbaines» (ANR AvuPur et EvalPDU notamment) s'appuient d'autant plus naturellement sur OrbisGIS que les participants à ces projets sont eux-mêmes les auteurs d'OrbisGIS.

3.4 Analyse

La capacité d'analyse d'un SIG couvre à l'origine l'ensemble des algorithmes de traitement des données disponibles afin de construire des connaissances plus synthétiques. Si nous couvrons actuellement quelque peu cette question en proposant une couche d'abstraction pour l'accès aux données, deux pistes fortes de recherche s'ouvrent maintenant à nous :

– d'une part, les travaux menés sur les modèles 3D+t une fois aboutis, se posera la question de l'accès à ces données et de la construction des traitements associés. D'un côté, nous devons définir un nouveau langage d'écriture de requêtes, les langages classiques type SQL spatial n'étant pas adaptés à ces contraintes ni en termes de

[2]http ://orbisgis.cerma.archi.fr

syntaxe ni de sémantique (comment prendre en compte la multi-représentation, la 3D voire l'incomplétude du SIG) ? Dans le cadre des travaux que nous entamons actuellement sur l'interface utilisateur du SIG et de la représentation cartographique, nous nous interrogeons également sur l'opportunité de définir une version gestuelle ou plus visuelle des langages de requêtes en nous appuyant sur les métaphores définies par Mainguenaud avec Cigales [30].

- d'autre part, afin de permettre l'acquisition et la mise à jour et le traitement des informations (cf. paragraphe suivant), de nouvelles méthodes d'analyse devront être développées : les différents projets auxquelles nous participons demandent le développement de méthodes spécifiques qui peuvent être liées au SIG ou aux images. Nous donnons par la suite quelques exemples tirés de projets actuels dans lesquels interviennent les chercheurs du PRF «Données urbaines» :
  - Calcul de bassins versants contraints par des obstacles liés à l'urbanisation : dans le projet Avupur[3] (E. Bocher, J.Y. Martin) qui vise à mesurer l'influence de l'étalement urbain sur l'hydrologie, l'objectif est de développer de nouveaux algorithmes de détermination de bassins versants (portion de territoire dont les eaux alimentent un exutoire commun : cours d'eau, lac, mer... selon [1]) qui prennent en compte de nouveaux objets créés par l'homme comme les routes et les réseaux d'eaux pluviales : ceux-ci modifient sensiblement le comportement des eaux ;
  - Évaluation des plans de déplacements urbains (PDU) : dans le projet EvalPDU[4] (E. Bocher, T. Leduc) qui se déroule en collaboration avec Nantes Métropole, l'objectif est l'évaluation des PDU en ce sens que le comportement des usagers n'est pas toujours celui prévu par les concepteurs du PDU. L'impact énergétique et écologique d'un nouveau PDU est difficile à mesurer parce l'état initial n'est pas connu et que les moyens de mesure terrain à disposition sont relativement faibles (compteurs de trafic ponctuels, enquêtes macroscopiques...). L'objectif du projet est d'établir la pertinence du remplacement de certaines chaînes lourdes de simulation physique par des quantifications approchées pour la construction d'indicateurs globaux : par exemple l'évaluation des incidences environnementales des changements de comportements liés à des actions phares du PDU. De par la globalité des phénomènes et des impacts en jeu dans le projet, la logique de mutualisation prônée par la plate-forme OrbisGIS (cf. paragraphe précédent) est une condition importante de réussite du projet.
  - Enfin, dans le cadre du projet urbSAT (N. Long, E. Bocher, F. Gonzaléz-Cortés, T. Leduc, G. Moreau) de constructions d'indicateurs morphologiques et aérodynamiques du tissu urbain (qui repose là aussi sur la plate-forme OrbisGIS), nous nous intéressons à l'influence des paramètres du modèle sur les indicateurs : urbSAT [100] est à la base la redéfinition du travail de N. Long [101]

[3] Assessing the Vulnerability of Peri-Urban Rivers

[4] Évaluation des impacts environnementaux d'un PDU et de leurs conséquences socio-économiques : développements méthodologiques et tests sur le PDU de Nantes Métropole

dans un SIG. Le module est construit comme une chaîne de traitements écrits en SQL spatial sous forme de plugin OrbisGIS. Afin de prendre en compte tous les types de données disponibles, l'utilisation d'un maillage est nécessaire. La suite du projet porte sur l'étude de l'influence des paramètres du maillage (maillage de type grille régulière nord-sud ou orientée selon des axes principaux, ou grille circulaire, taille des mailles...) comme le montrent les différents éléments de la figure 3.4.

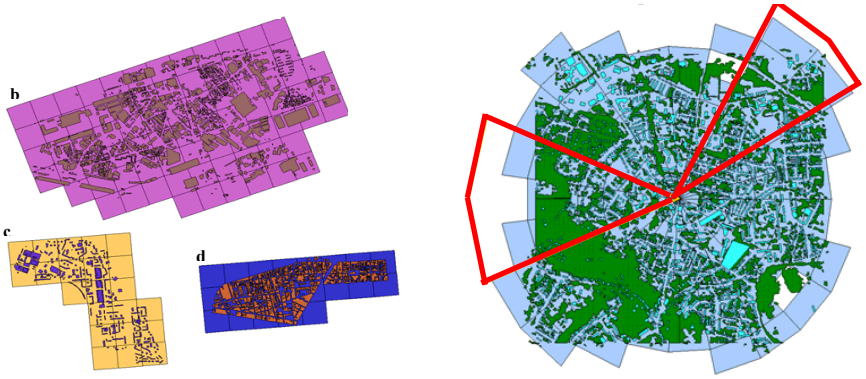


FIG. 3.4 – *Utilisation de différents types de maillages dans urbSAT*

3.5 ACQUISITION

Le titre de notre chapitre prend son sens à partir de ce paragraphe. L'acquisition d'informations, qu'elles soient de nature géométrique ou attributaire, passe par une étape obligée dans un SIG : le géo-référencement. Celui consiste à passer finalement de coordonnées locales, celle de l'objet à insérer/mettre à jour dans le SIG à des coordonnées relatives à un territoire, les coordonnées géographiques. C'est ainsi que les différents éléments de nos base de données pourront cohabiter lorsqu'ils font référence à une même portion de territoire.

Nous proposons ici de non seulement utiliser l'hypothèse classique qui consiste à utiliser un SIG 2D et une image géo-référencée pour construire un nouvel élément du SIG 3D, mais aussi d'inverser cette idée pour partir du problème dual : lors de l'acquisition d'une image sur site, un des problèmes à résoudre est de la géo-référencer. En milieu urbain, l'utilisation des GPS ne peut suffire à localiser avec suffisamment de précision l'image (et encore moins à l'orienter ce qui est incompatible avec l'utilisation dans un contexte réalité augmentée tel que nous souhaitons le proposer). Pour résoudre ce problème, nous avons proposé d'utiliser les informations présentes dans l'image et le SIG pour géo-référencer celle-ci. C'est l'objet du travail de thèse confié à N. Bioret qui donne aujourd'hui des premiers résultats intéressants.

3.5.1 Géo-référencement à partir d'images

L'état de l'art montre que le processus de géo-référencement est souvent découpé en deux phases : la première se réfère à une sorte de po-

sitionnement global ou initialisation généralement très peu abordé par les chercheurs. La seconde étape, largement couverte dans la littérature, est plus une étape de tracking, celui-ci pouvant de faire à partir de diverses sources d'informations (et de données capteurs qui peuvent être utilisées en plus des images). Le tracking est souvent adapté (qu'il utilise une base d'images classiques [77] ou sphériques [92] ou du suivi de caractéristiques «vectorielles» [13, 63]) à un faible nombre de bâtiments et ne sait donc généralement pas utiliser un SIG complet mais simplement un faible sous-ensemble de celui-ci. Dans le cas des bases de données d'images, le problème de taille devient rapidement critique pour envisager un passage à l'échelle urbaine. Seul [148] utilise aujourd'hui à la fois des images et des données SIG sur une échelle relativement large, au prix toutefois d'une construction d'un modèle 3D coûteuse.

Nous cherchons dans ce travail à adresser le problème du calcul de pose initial en extrayant certaines caractéristiques de l'image pour en déduire des requêtes SIG qui nous permettront un premier géo-référencement de l'image, fournissant ainsi un élément d'initialisation probant pour un calcul de pose plus précis basé sur les méthodes de l'état de l'art. Nous nous appuyons sur le fait que nous pouvons ainsi économiser le recours à un GPS et à son important temps d'initialisation. De plus cette technique pourra s'appliquer à des bases d'images non-géoréférencées existantes.

Les caractéristiques que nous extrayons des images sont volontairement simples afin de permettre une certaine robustesse. Dans un contexte urbain, la plupart des parois sont verticales, permettant ainsi l'extraction de points de fuite. Ayant eu l'occasion de travailler sur l'extraction automatique et robuste de points de fuite grâce à de simples résultats de géométrie projective [82, 83], nous nous appuyons sur la détection des points de fuite pour estimer :

- les ratios de longueurs r entre deux façades consécutives ;
- les angles ϕ entre deux façades consécutives (on notera en passant que les deux angles les plus courants sont les angles plats et les angles droits).

A cette estimation, nous associons une mesure de l'incertitude sur ces valeurs en nous appuyant sur les calculs d'incertitude des points de fuite présentés dans [83]. Nous pouvons ensuite effectuer une série de requêtes dans le SIG pour déterminer quelles peuvent être les couples (voire les n-uplets) de façades formant ces angles avec ces ratios de longueur. Nous utilisons pour cela dans un premier temps uniquement la couche bâti des données cadastrales. A ce jour, nous n'utilisons aucun élément métrique. A partir des résultats obtenus, nous pouvons effectuer certains post-traitements pour valider ou non une partie des résultats et effectuer un premier calcul de pose.

La figure 3.5 présente un premier résultat de calcul de poses possibles sur ces seuls critères dans une zone de rayon 1km contenant plus de 7000 bâtiments nantais. 54 poses sont proposées, dont la bonne. Si nous utilisons les trois façades visibles, ce nombre tombe à un. On notera que ce premier résultat porte sur un cas qu'on peut qualifier de facile en raison des angles particuliers entre les façades de ce bâtiment. On constate

FIG. 3.5 – *Premier résultat de calcul de poses possibles sur une paire de façades*

néanmoins que le nombre de poses possibles croit linéairement avec la tolérance utilisée dans les requêtes [15].

Sur un cas plus difficile comme celui de la figure 3.6, on obtient plus de 7000 poses qu'on peut réduire assez rapidement : à 77 si on utilise la troisième façade visible puis en vérifiant des contraintes sur les poses obtenues : celles-ci ne peuvent se trouver dans un bâtiment (contrainte donnée à la prise de vue) et aucun bâtiment ne doit venir effectuer une occlusion. Le nombre de poses est alors réduit à 18.

FIG. 3.6 – *Second cas plus standard*

Plusieurs pistes sont en cours d'étude :

- jusqu'ici aucune information métrique n'a été utilisée : or il est simple de donner des intervalles de valeurs pour ces informations. Ces intervalles pourraient par exemple se construire de façon relative par rapport à d'autres éléments présents dans l'image (piétons, véhicules, mobilier urbain, signalisation) dont la taille varie peu. On peut montrer que le nombre de poses du cas précédent, pourtant extrêmement générique, se réduit alors à 6 avec une tolérance de 50cm ;
- l'utilisation d'une information d'orientation absolue, même très grossière, permettrait de diminuer de façon très importante le nombre de poses (13/18 avec une tolérance de 20° dans l'exemple précédent). Cette mesure peut se faire avec un simple compas USB

et même potentiellement se mesurer sur l'image (position du soleil, orientation des ombres...) ;
- Dans l'état actuel du SIG, il est difficile de compter sur l'utilisation d'information 3D, plus difficilement détectable dans l'image et présente de façon très parcellaire dans le SIG. Bien sûr, plus le SIG 3D se construira, plus cette information pourra être utilisée. On peut y voir à termes, une forme d'apprentissage par renforcement ;
- La piste la plus intéressante pour l'instant consiste à utiliser non plus une image mais une paire ou une séquence d'images prises aux alentours du même endroit. Plusieurs stratégies sont possibles : des corrélations image à image ou entre deux applications de la méthode évoquée ci-dessus sur des prises de vue qui peuvent être proches en position ou en orientation. Ceci constitue une de nos pistes d'étude prioritaire.

Il est intéressant de noter que ce type de technique permet d'aider le géo-référencement d'images sur lesquelles nous n'avons aucune connaissance a priori. Cela inclut notamment les bases de données d'images comme celles disponibles sur Internet (Flickr ou bien les images disponibles sur Google Earth). Dans le paragraphe suivant, nous proposons de nous intéresser à la géolocalisation à partir d'images en utilisant les bases de données existantes comme les panoramas de Google Street View.

3.5.2 Approche alternative pour le géo-référencement : utilisation de bases d'images existantes

Comme nous l'avons mentionné en début de paragraphe précédent, l'utilisation de bases d'images soulève des questions de taille des bases de données et du nombre d'images à acquérir. Toutefois, il ne faut pas négliger le fait qu'il existe aujourd'hui des bases de données importantes d'images disponibles en ligne. Celles-ci sont déjà en cours d'exploitation pour la construction de modèles de bâtiments : Moslah [122] utilise les images du site `http://www.flickr.com` afin de reconstruire un modèle de la cathédrale Notre-Dame de Paris alors que celles-ci ne sont pas géo-référencées.

Une autre piste que nous avons très récemment commencée à explorer est l'utilisation des bases de données de panoramas comme Google Street View. En collaboration avec le Hideo Saito Lab de l'Université de Keio, nous cherchons une correspondance entre une image d'entrée quelconque et une base de données de panoramas. Les premiers résultats montrent qu'en utilisant des panoramas situés aux intersections et à condition que l'image d'entrée soit prise dans un rayon relativement proche, le système permet de restituer position et orientation (azimut seulement) de la prise de vue de façon relativement robuste, même avec des caméras de paramètres différents. En revanche, la robustesse aux objets mobiles et aux changement de conditions de prise de vue (jour/nuit, point de vue éloigné) reste à améliorer.

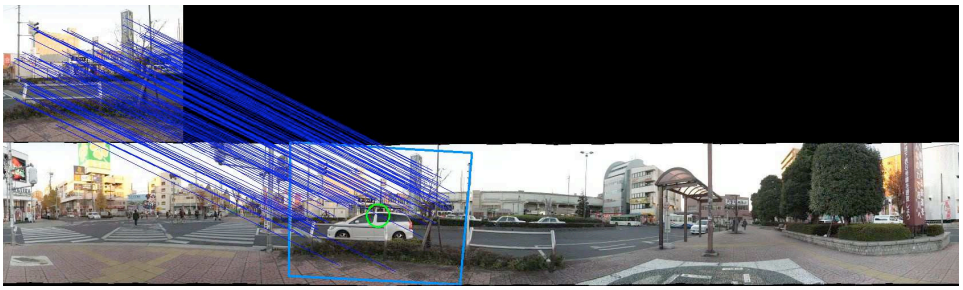


FIG. 3.7 – *Détermination de la position et de l'orientation d'une image dans une base de données de panoramas*

3.5.3 Vers la construction de SIG3D en réalité augmentée

Une fois l'étape de calcul de pose effectuée, il devient possible de superposer l'image prise par une caméra aux informations disponibles dans le SIG. Rappelons que la calcul de pose peut se décomposer en deux phases : la première consiste en une initialisation (phase de géo-référencement vue plus haut) qui n'est pas forcément possible en temps réel, la seconde est une phase de *tracking* : les techniques actuelles permettent le suivi d'un certain nombre de caractéristiques dans une suite d'images et d'en déduire la pose de la caméra plusieurs dizaines de fois par secondes. Nous ne souhaitons pas nous intéresser à cette partie du problème qui nous semble bien couverte par l'état de l'art. En revanche, pour la création/mise à jour du SIG sous-jacent se posent deux grands nouveaux types de problème que nous allons aborder à l'avenir : l'affichage d'informations et l'interaction «en écriture» avec le SIG. La problématique liée à l'affichage sera abordée de façon plus détaillée au paragraphe 3.6, nous nous intéressons ici au problème de l'insertion d'informations dans le SIG dans un contexte de réalité augmentée.

Nous distinguons différents cas en fonction de la complexité de la tâche : dans un premier temps, nous nous intéressons au problème d'insertion d'éléments attributaires (la construction de l'information à insérer/modifier est effectuée par l'observateur), puis aux informations de nature géométrique qui font appel à des algorithmes plus automatisés.

Annotation / insertion d'informations attributaires

Il s'agit de l'ajout ou de la mise à jour d'information relative à un élément géométrique du SIG comme le montre l'exemple factice de la figure 3.8. Le problème de l'annotation dans les environnements virtuels dispose d'un état de l'art conséquent qui tente de résoudre les problèmes suivants :

- la désignation de l'objet d'intérêt en utilisant généralement une interface tangible et un lancer de rayon dans la scène virtuelle ;
- l'insertion de l'information dans une base de données ;
- l'affichage des annotations de façon cohérente afin d'en permettre le partage depuis plusieurs points de vue.

Peu de travaux sont effectués dans un contexte de réalité augmentée en milieu extérieur. Les différences portent notamment sur la difficulté de désignation au sens du SIG : dans l'exemple factice donné figure 3.8, les

travaux mentionnés doivent être liés à la rue (en l'occurrence à la paire de rues), mais aussi au réseau souterrain qui est forcément impacté et qui n'est pas visible sur l'image.

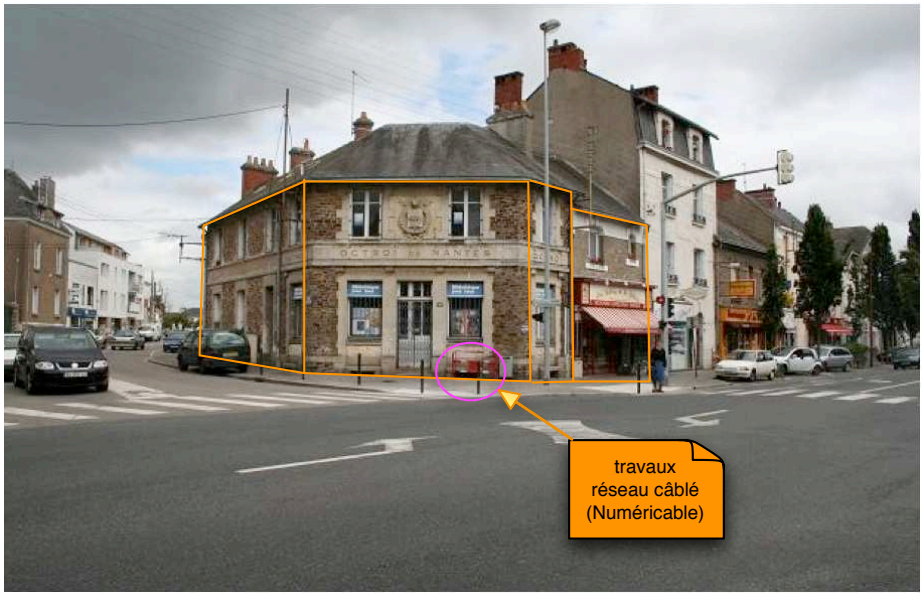


FIG. 3.8 – *Exemple factice d'annotation en réalité augmentée*

Au-delà du processus de désignation d'un objet à annoter qui peut être différent dans un contexte de RA sur site, il sera intéressant de mener une étude sémantique des annotations pour déterminer les objets réellement impactés. La forme des annotations pourra aussi poser question puisque la disponibilité d'outils classiques d'annotation (clavier notamment) en environnement extérieur pose là aussi problème. Une fois ces questions résolues, ces annotations seront insérées dans le cadre du SIG spatio-temporel évoqué au paragraphe 3.2. L'affichage sera évoqué au paragraphe 3.6.

Ajout et modification d'informations géométriques

Ce second cas est plus complexe en raison des problématiques d'interfaçage : dans un contexte de réalité augmentée sur site modifier les informations géométriques disponibles ou en ajouter de nouvelles. Nous proposons ici d'utiliser les possibilités des Tablets PC puis d'utiliser deux grands types de techniques :

- les techniques de vision par ordinateur : la construction de géométrie nécessite de pointer directement des points. En faisant l'hypothèse raisonnable que ces points correspondent à quelque chose de visible, nous proposons d'utiliser d'utiliser le pointage effectué par l'utilisateur comme l'initialisation d'un détecteur de coins (par exemple [67]). Ceci permettrait d'améliorer le processus du pointage, le risque d'imprécision étant élevé dans le contexte RA. Une fois ce pointage affinée, un lancer de rayon combiné à une requête SIG permettra de déterminer s'il existe des éléments géométriques

proches dans le modèle. Cette approche peut se généraliser à la détection d'autres types d'éléments géométriques que des points (lignes, primitives tridimensionnelles) ;
– la modélisation à partir d'esquisse : par le passé, certains travaux effectués dans des laboratoires nantais relatifs à la modélisation dite déclarative ont permis la reconstruction de modèle de bâtiments à partir d'esquisses [157].

3.6 Affichage

Cette dernière partie trace quelques perspectives quant aux problèmes d'affichage de l'information contenue dans les SIG 3D+t. Nous distinguons deux classes d'information à inclure dans les images : les informations relatives au contexte, qui permettent à l'observateur de se repérer (faire le lien entre l'image et un emplacement du territoire) et les autres informations qui permettent de se construire une connaissance supplémentaire. Pour le premier type, nous généralisons la notion de *fond de carte* en distinguant 4 cas présentés figure 3.9 qui seront abordés dans les paragraphes qui suivent : d'abord les environnements purement synthétiques utilisant des informations provenant uniquement du SIG, puis des environnements mixtes utilisant des images du monde réel auxquelles sont ajoutées des informations synthétiques.

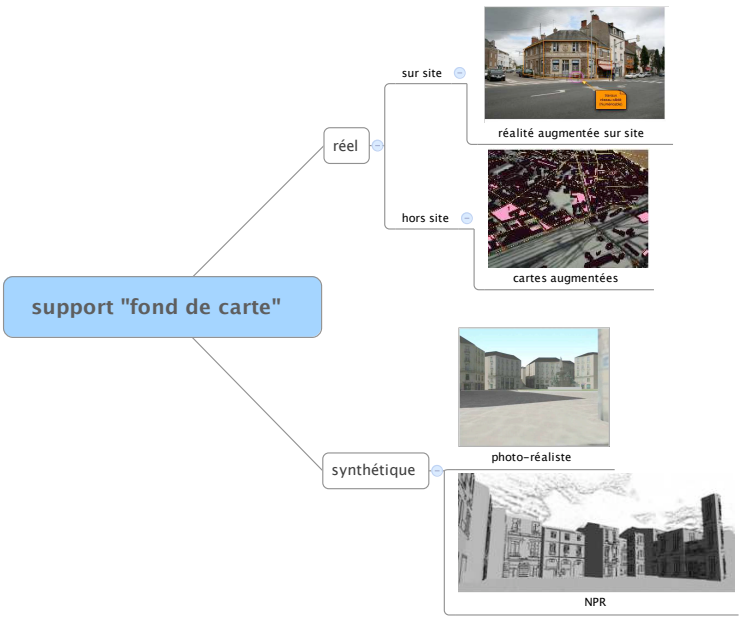


Fig. 3.9 – Fond de carte *généralisé*

3.6.1 Affichage d'informations 3D+t

Dans cette section, nous nous intéressons à l'affichage d'informations contenues dans un SIG3D+t, indépendamment du dispositif de rendu et

du *fond de carte* : Ceux-ci peuvent être un écran standard ou une installation de réalité virtuelle, voire de réalité augmentée utilisant n'importe quel algorithme de rendu. Nous sommes conscients que la sémiologie utilisée dépendra forcément à terme du support sur lequel elle est mise en œuvre mais nous préférons séparer ces questions dans un premier temps. Si l'on se réfère à la classification de la complexité de la perception des mondes virtuels effectuée figure 3, page 6, l'affichage d'informations se situe plutôt au niveau fonctionnelle tandis que la reconnaissance du lieu interpelle le niveau cognitif mais peut aussi faire appel à l'expérience sensible.

Nous distinguerons l'affichage d'informations 3D et celui d'informations temporelles.

Affichage d'informations temporelles

L'affichage d'informations variant dans le temps dans un SIG peut paraître facile dans un contexte de synthèse d'images temps-réel comme la réalité virtuelle, mais il doit prendre en compte le fait que le temps du SIG n'est pas forcément le temps utilisateur et utiliser des facteurs d'échelle, pas forcément linéaires. De plus, la lisibilité de certaines modifications géométriques n'est pas garantie. Le problème est abordé dans la littérature à travers des projets comme TimeMap [78] pour ce qui concerne les SIG et de façon plus générale pour ce qui concerne la visualisation scientifique : on peut citer la visualisation d'écoulements fluides comme exemple structurant.

A notre sens, ceci pose le problème plus général de la construction d'une sémiologie de la visualisation d'informations géographiques à caractère 3D+t.

Vers la définition d'une sémiologie 3D+t

La norme SLD (Style Layer Descriptors[5]) a introduit en deux dimensions la séparation des informations et de l'affichage à travers la notion de feuille de style pour l'information géographique. La feuille de style SLD est un simple fichier XML qui définit la façon dont seront rendues les différentes couches du SIG. Reste que cette norme est encore aujourd'hui incomplète en ce sens qu'elle ne prend pas en compte les informations thématiques, ni la 3D. Nous proposons d'avancer dans ce domaine en nous intéressant à la création de feuilles de style 3D+t permettant le rendu des informations stockées dans le SIG3D+t. Nous nous appuierons pour cela sur notre collaboration avec l'équipe GeoSysin de l'IICT qui développe TLD (Thematic Layer Description) et sur les travaux relatifs au SLD3D [128]. Nous devrons prendre en compte le caractère local ou global de ce rendu, lié à l'échelle : en effet, en réalité augmentée sur site, le type d'information qui pourra être affiché sera différent des indicateurs globaux qu'on peut avoir à l'échelle d'une ville. A l'instar de la cartographie classique, on peut envisager ici la notion de *généralisation thématique*.

La construction de feuilles de style est une étape dans le processus, mais de la même façon qu'une feuille de style inappropriée ne permet pas

[5]définie `http://www.opengeospatial.org/standards/sld`

de retirer toute l'information d'un document textuel, un travail d'évaluation de l'information acquise par l'observateur par rapport à l'information représentée est nécessaire. Ce évaluation au niveau fonctionnel sera nécessairement interdisciplinaire et devra déboucher des propositions d'automatisation et de partage contextualisés de l'utilisation des feuilles de style. Par exemple, si cette évaluation permet de mettre en évidence une utilisation particulière d'une feuille de style pour la visualisation de trajectoires de piétons, nous proposons d'affecter automatiquement cette feuille de style à des éléments similaires, suivant ainsi le principe des Systèmes d'Information Logique déjà évoqués en 3.2. C'est une des composantes du projet *GeoCognition* visant à partager les données mais aussi les connaissances et les méthodes de traitement.

Si ces travaux permettront de définir la manière d'afficher des informations de nature mesurée, interprétée ou dérivée, le problème du support d'affichage reste posé. En deux dimensions, ce support se résume à un *fond de carte* déterminé par l'utilisateur. En 3D, il devient nécessaire de s'intéresser à ce que doit contenir l'environnement virtuel urbain, puis de s'intéresser à la fidélité perceptive de cet environnement.

3.6.2 Environnements virtuels urbains : le problème du fond de carte

On peut imaginer créer tout ou partie d'un environnement virtuel urbain à partir d'un SIG 3D+t et d'un outil comme le SLD3D. Mais, comme nous l'avons vu au chapitre 1 sur la perception des environnements virtuels, cela ne suffira pas forcément à construire un environnement virtuel reconnaissable. En effet, si la fidélité géométrique de l'environnement virtuel est envisageable, sous réserve de la précision du SIG, à l'aune des travaux que nous avons menés précédemment, il est important d'évaluer dans quelle mesure le lieu reconstitué peut être fidèle, y compris en termes d'expérience sensible à son équivalent réel. A notre avis, la reconnaissance par l'observateur du lieu représenté est un pré-requis fondamental à cette fidélité. Cela pose la question de la génération automatisée de ces environnements virtuels urbains, contrairement à nos expériences précédentes où ceux-ci étaient construit manuellement. Qui plus est, nous pensons que les informations supplémentaires affichées peuvent constituer des difficultés supplémentaires en ce sens qu'elles peuvent perturber la reconnaissance du lieu représenté. En contrepartie, nous sommes à même ici d'afficher des éléments non présents dans l'environnement réel afin de faciliter son appropriation.

Pour faciliter la reconnaissance du lieu et améliorer sa fidélité, nous proposons d'aborder le problème du fond de carte sous plusieurs aspects complémentaires :

- par génération automatique à partir d'une feuille de style comme évoqué au paragraphe précédent : cette démarche présente l'avantage de l'automatisation mais n'engage que peu ou pas l'expérience sensible si l'on utilise les algorithmes classiques de rendu temps-réel ;
- en utilisant cette même approche à l'aide d'algorithmes de rendu «alternatifs», notamment les classes d'algorithmes de rendu expressif qui font appel à l'expérience sensible ;

– en faisant abstraction du fond de carte par l'utilisation d'images réelles acquises en direct sur site, via la réalité augmentée ; le système de réalité augmentée incluant directement ce fond de carte n'aura plus qu'à afficher l'information supplémentaire souhaitée.

3.6.3 Rendu expressif

Au lieu d'aller vers des environnements de plus en plus réalistes, nous proposons de nous intéresser aux algorithmes de rendu eux-mêmes. En effet, nous pensons que ces certains algorithmes de rendu non-photoréaliste (NPR) peuvent apporter une meilleure reconnaissance et appropriation de l'environnement représenté : un rendu de type silhouette présente intuitivement une forme du rendu plus caractéristique de la forme urbaine et de par sa «sobriété» un environnement laissant plus de place à l'affichage d'informations souhaitée qu'un environnement virtuel texturé classique de type photoréaliste. L'équipe iParla du LabRI a d'ailleurs défini un style de rendu expressif [144] qui présente en outre l'avantage d'être extrêmement léger au rendu et permet donc ainsi l'affichage de très grandes bases de données urbaines. Un exemple de rendu de façades provenant de cette méthode est proposé figure 3.10.

Cela dit, la capacité des méthodes de rendu expressif à provoquer des émotions peut également provoquer des biais dans l'appréhension de l'environnement virtuel ainsi produit. Ce biais peut présenter d'autres avantages en terme de rendu des ambiances et d'émergence d'effets (cf. 1.3, page 28).

FIG. 3.10 – *Rendu expressif d'environnements virtuels urbains (©iParla [144])*

Les travaux dans ce domaine porteront sur la validation de cette hypothèse, l'affichage d'informations (en surimpression ou autre) et des études sur la prise de repères en environnement virtuels.

3.6.4 Réalité augmentée

Sur site

L'utilisation de la réalité augmentée sur site (en environnement extérieur) semble naturellement la façon la plus intuitive d'afficher l'information contenue dans le SIG 3D+t. Outre l'affichage, elle devrait permettre à terme de modifier ces informations. Néanmoins, elle suppose résolus toutes les autres problématiques abordées jusqu'ici : calcul de pose en temps-réel (géo-référencement et tracking), affichage d'informations 3D+t,

interaction (pour l'édition)... C'est pour cette raison que ce paragraphe ne constitue aujourd'hui qu'une perspective de nos travaux.

En attendant la construction des différents composants de cette application de réalité augmentée, nous nous sommes intéressés, par le biais de notre collaboration avec le Hideo Saito Laboratory de l'Université de Keio, à l'utilisation de la réalité augmentée «hors-site» en construisant des cartes augmentées.

Cartes augmentées

En attendant d'arriver à mettre en place une application de réalité augmentée sur site, nous nous sommes intéressés à son utilisation hors-site sur une table de bureau en faisant une hypothèse très différente de celles effectuées jusque là : pour compléter les travaux sur l'appariement entre le lieu réel et le fond de carte, nous nous appuyions sur un principe de fidélité visuelle que nous déclinions en plusieurs algorithmes de rendu (réaliste, expressif ou purement absent). Ici, nous proposons de nous appuyer sur une carte réelle. Celle-ci, bien que représentant des informations purement symboliques, suit un principe de représentation parfaitement connu et constitue une interface tangible (au sens de Ishii [179]) parfaitement admise. C'est cette facilité de manipulation (contrairement à celles des environnements virtuels urbains à échelle 1) qui nous fait envisager cette hypothèse.

Le travail mené dans cette dernière partie consiste à afficher des informations supplémentaires (éventuellement à caractère 3D+t) sur des cartes réelles. Ces travaux font l'objet de la thèse de H. Uchiyama depuis fin 2007. Cette thèse s'intéresse au lien entre une carte filmée par une caméra vidéo et le SIG, sans utiliser de marqueurs. Dans un premier temps, nous avons proposé un algorithme qui localise une image de carte dans le SIG à partir de la distribution des intersections routières [178] : ainsi nous sommes capables de retrouver à la fois la partie de carte filmée dans le SIG et de connaître la pose de la caméra par rapport à cette carte réelle. Une fois ceci effectué, il ne reste plus qu'à projeter les informations du SIG sur la carte physique comme le montre la figure 3.11. Cette localisation est par ailleurs robuste (la main peut être posée sur la carte sans gêner le tracking), même si pour l'instant, nous utilisons une carte synthétique où les intersections sont mises en évidence plutôt qu'un algorithme de détection des intersections sur une carte véritable. Les travaux à venir portent sur l'utilisation de la distribution des tracés routiers plutôt que celle des intersections et la prise en compte de la généralisation cartographique afin de pouvoir utiliser des cartes d'échelles différentes.

Nous pensons – hypothèse qui reste à vérifier – que les cartes augmentées permettront à des personnes non familières des environnements virtuels urbains ou des SIG de manipuler de nouveaux types d'information sans la complexité inhérente aux SIG et sans le problème d'appropriation des EV urbains. De plus, contrairement à la réalité augmentée sur site, elle permet une vision globale plutôt que locale. La carte physique naturelle dépassera la simple interface tangible classique. De plus, ces travaux permettront de :

– commencer à réfléchir en avance de phase sur l'affichage d'informa-

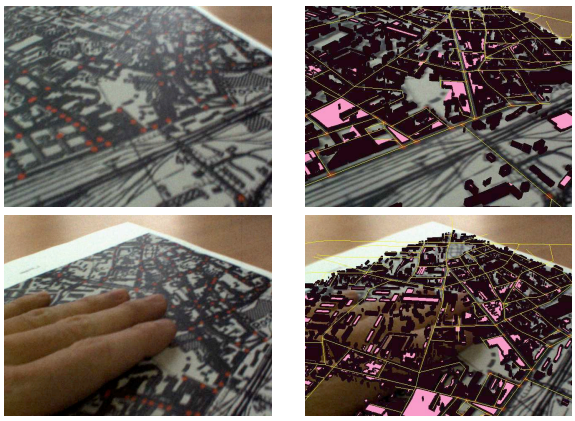


FIG. 3.11 – *Cartes augmentées : localisation robuste à partir de la distribution des intersections*

tions supplémentaires en réalité augmentée et de définir une sémiologie 3D+t adaptée à la réalité augmentée ;
- lancer les travaux sur l'interaction avec un SIG : aujourd'hui le modèle classique en couches des SIG est limité par la complexité de l'interface du SIG. En proposant une nouvelle forme d'interface [38], la réalité augmentée permettra peut-être d'étendre l'usage des SIG à un nouveau type de public. Ces travaux doivent aller de pair avec la réflexion sur les langages de requêtes spatio-temporels abordés en 3.4.

3.7 Vers l'interaction avec le SIG3D+t à l'aide d'images

Dans ce dernier chapitre, nous avons évoqué bon nombre de travaux en cours ou à démarrer prochainement sur le lien entre images et SIG3D+t ainsi que sur les pré-requis en matière de modélisation de l'information géographique et de suppport logiciel. Il n'est à ce jour que très faiblement question d'interaction avec le SIG3D+t. Or, nous avons vu dans les chapitres précédents que l'interaction et l'immersion constituent les deux piliers de la réalité virtuelle et qu'une bonne interaction améliore l'immersion dans les mondes virtuels.

Dans le cas des SIG, l'interaction peut se penser en deux versions :
- l'interaction 2D : il y a aujourd'hui des travaux visant à dépasser le modèle en couches traditionnellement utilisé dans les SIG, mais aussi des questionnements sur la représentation de l'information (requêtes versus navigation) sur les langages d'interrogation ;
- l'interaction 3D : la littérature actuelle sur l'interaction 3D est extrêmement riche afin de dépasser le paradigme WIMP difficilement accessible dans des installations de RV et RA. Ces travaux restent majoritairement dans le contexte réalité virtuelle où l'on trouve par exemple de nombreuses utilisations des gestes (pour rester dans le domaine de la coopération analyse-synthèse). Les travaux en réalité augmentée sont plus rares et font généralement appel à des interfaces tangibles ou à des éléments d'interface supplémentaires

comme des écrans tactiles. Le développement d'applications de RA sur téléphone mobile et/ou écran tactile est aujourd'hui dans une phase très active.

Fort de ces éléments, et compte tenu du contexte évoqué dans ces pages (société de l'image, société numérique, participation du public aux grandes décisions) nous pensons que les deux questions évoquées ci-dessus peuvent se fondre en une seule qu'il nous faudra aborder à terme : comment rendre simple l'interaction avec un système d'information géographique, qui plus est 3D+t afin que tous puissent en bénéficier ?

Bibliographie

[1] Directive-cadre sur l'eau,. Directive 2006/118/CE (JO L 372 du 27.12.2006, p. 19)., 2006. (Cité page 83.)

[2] Dominique Adde. Environnement graphique de spécification du comportement humain en animation. Mémoire de fin d'études d'ingénieur, Institut d'Informatique d'Entreprise, Evry, France, June 1998. (Cité page 71.)

[3] Omar Ahmad, James Cremer, Joseph Kearney, Peter Willemsen, and Stuart Hansen. Hierarchical, concurrent state machines for behavior modeling and scenario control. In *Conference on AI, Planning, and Simulation in High Autonomy Systems*, Gainesville, Florida, USA, 1994. (Cité page 62.)

[4] B. Arnaldi and G. Dumont. Vehicle simulation versus vehicle animation. In *3rd Eurographics Workshop on Animation and Simulation*, Cambridge, Royaume-Uni, 1992. (Cité page 62.)

[5] Bruno Arnaldi, Rémi Cozot, Stéphane Donikian, and Michel Parent. Simulation models for the french praxitèle project. In *Transportation Research Board Annual Meeting*, Washington DC, USA, January 1996. (Cité page 72.)

[6] Julian Aubourg, Guillaume Moreau, and Philippe Fuchs. Génération assistée de numérisation et de texturation d'objets muséographiques. Rapport de fin de projet GANTOM, Ecole des mines de Paris / programme PRIAMM CNC, 2002. (Cité page 55.)

[7] Julian Aubourg, Guillaume Moreau, and Philippe Fuchs. Three-dimensional reconstruction and texturation of museographic objects using multiple images and stereoscopic depth map fusion. In Brian D. Corner, Peng Li, and Roy P. Pargas, editors, *Electronic Imaging Three Dimensional Image Capture and Applications VI*, volume 5302, pages 136–147, San Jose, CA, janvier 2004. SPIE & IST. (Cité page 54.)

[8] Jean-François Augoyard and Henri Torgue. *Sonic Experience : A guide to everyday sounds*. McGill-Queen's University Press, Montreal, 2005. (Cité page 31.)

[9] Norman I. Badler, Barry D. Reich, and Bonnie L. Webber. Towards personalities for animated agents with reactive and planning behaviors. *Lecture Notes in Artificial Intelligence, Creating Personalities for synthetic actors*, (1195) :43–57, 1997. (Cité page 62.)

[10] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool. Surf : Speeded up robust features. In *Proceedings of the ninth European Conference on Computer Vision*, May 2006. (Cité page 18.)

[11] Bayarri, Fernandez, Perez, and Rodriguez. A database model for urban and motorway driving simulation. In *Driving simulation conference*, September 1996. (Cité page 73.)

[12] Olivier Bedel, Sébastien Ferré, Olivier Ridoux, and Erwan Quesseveur. GEOLIS : A logical information system for geographical data. *Revue Internationale de Géomatique*, 17(3-4) :371–390, 2008. (Cité page 81.)

[13] Reinhold Behringer, Jun Park, and Venkataraman Sundareswaran. Model-based visual tracking for outdoor augmented reality applications. In *ISMAR, Darmstadt, Germany, September 30-October 1*, pages 277–278, 2002. (Cité page 85.)

[14] G. Berry and G. Gonthier. The esterel synchronous programming language : design, semantics, implementation. Rapport de recherche RR-842, INRIA, 1988. (Cité page 63.)

[15] Nicolas Bioret, Myriam Servières, and Guillaume Moreau. Outdoor localization based on image/gis correspondence using a simple 2d building layer. In *2nd International Workshop on Mobile Geospatial Augmented Reality*, Lecture notes in Geoinformation and Cartography, Québec, Canada, aug 2008. Springer-Verlag. (Cité page 86.)

[16] Ian Bishop and David Miller. Visual assessment of off-shore wind turbines : the influence of distance, contrast, movement and social variables. *Renewable Energy*, 32(5) :814–831, apr 2007. (Cité page 38.)

[17] F. Blaise. *Simulation du parallélisme dans la croissance des plantes et application*. Thèse de doctorat, Université Louis Pasteur, Strasbourg, 1991. (Cité page 58.)

[18] W. Blohm, I.P. Beldie, K. Schenke, K. Fazel, and S.Pastoor. Stereoscopic image representation with synhtetic depth of field. *Journal of the Society for Information Display*, 5(3) :307–313, 1997. (Cité page 16.)

[19] Erwan Bocher, Thomas Leduc, Guillaume Moreau, and Fernando González Cortès. Gdms : An abstraction layer to enhance spatial data infrastructures usability. In *Agile 2008*, 2008. (Cité page 81.)

[20] C. Bonnet. *Manuel pratique de psychophysique*. Armand Collin, 1986. (Cité page 19.)

[21] M. Booth, Jim Cremer, and Joseph Kearney. Scenario control for real-time driving simulation. In *Fourth Eurographics Workshop on Animation and Simulation*, pages 103–119, Barcelone, Espagne, 1993. (Cité pages 61 et 62.)

[22] Janos Boudet. Appariement stéréoscopique utilisant la transformée en ondelettes multi-échelles. Technical report, Ecole Nationale Supérieure des Mines de Paris, Juin 2001. (Cité page 17.)

[23] Ronan Boulic, Nadia Magnenat-Thalmann, and Daniel Thalmann. A global human walking model with real-time kinematic personification. *Visual Computer*, 6 (6) :344–358, 1990. (Cité page 59.)

[24] Ronan Boulic and R. Mas. Hierarchical kinematic behaviors for complex articulated figures. In Daniel Thalmann and Nadia Magnenat-Thalmann, editors, *Advanced Interactive Animation*, chapter 3. Prentice-Hall Europe, 1996. (Cité page 59.)

[25] Patricia Bournai, B. Chéron, Thierry Gautier, Bernard Houssais, and Paul Le Guernic. Signal manual. Publication Interne 745, IRISA, Rennes, 1993. (Cité page 63.)

[26] Thierry Bouron. *Structures de communication et d'organisation pour la coopération dans un univers multi-agents*. Thèse de doctorat, Université de Paris VI, Paris, 1993. (Cité pages 60 et 64.)

[27] T.L. Brooks. Telerobotic response requirements. In *IEEE International Conference on systems, Man and Cybernetics*, pages 113–120, Los Angeles, USA, 1990. (Cité page 25.)

[28] Christopher Buehler, Wojciech Matusik, Steven Gortler, and Leonard McMillan. Creating and rendering image-based visual hulls. Technical Report MIT-LCS-TR-780, MIT Laboratory for Computer Science, Cambridge, MA USA, juin 1999. (Cité pages x et 52.)

[29] G.J. Buhyoff, P.A. Miller, R.B. Hull, and D.H. Schlager. Another look at expert visual assessment : validity and reliability. *AI Applications*, 9 :112–120, 1995. (Cité page 40.)

[30] D. Calcinelli and M. Mainguenaud. Cigales, a visual query language for a geographical information system : the user interface. *Journal of Visual Languages and Computing*, 5 :113–132, 1994. (Cité page 83.)

[31] Y. Uni Cao, A. Fukunaga, A. Kahng, and F. Meng. Cooperative mobile robotics : Antecedents and directions. In *International Conference on Intelligent Robots Systems*, volume 1, pages 226–233, 1995. (Cité page 62.)

[32] C. Chien and J. Aggarwal. Identification of 3d objects from multiple silhouettes using quadtrees/octrees. *Computer Vision, Graphics and Image Processing*, 36 :256–273, 1986. (Cité page 53.)

[33] J. F. Coeterier. Dominant attributes in the perception and evaluation of the dutch landscape. *Landscape and Urban Planning*, 34(1) :27–44, January 1996. (Cité page 38.)

[34] P.A. Costello and P.A. Howarth. The visual effects of immersion in four virtual environments. Rapport de recherche 9604, VISERG, Loughborough University, septembre 1996. (Cité page 16.)

[35] Rémi Cozot. From multibody systems modelling to distributed real-time simulation. In ACM, editor, *American Simulation Symposium*, New Orleans, USA, 1996. (Cité page 71.)

[36] James Cremer, Joseph Kearney, and Peter Willemsen. A directable vehicle behavior model for virtual driving environments. In *Conference on AI, Simulation and Planning in High Autonomy Systems*, La Jolla, USA, 1996. (Cité page 62.)

[37] Gordon Cullen. *Townscape*. The Architectural Press, London, 1961. (Cité pages 30 et 31.)

[38] Gilberto Câmara, Ricardo Cartaxo, Modesto De Souza, Antônio Miguel, Vieira Monteiro, Juan Carlos, Pinto De Garrido, Divisão Processamento, and Imagens Dpi. Handling complexity in gis interface design. In *In : Proceedings of the I Brazilian Workshop on GeoInformatics, Campinas, São*, 1999. (Cité page 95.)

[39] Jean Denègre and François Salgé. *Les systèmes d'information géographiques*. Presses Universitaires de France, 2ème edition, 2004. (Cité pages ix, xi, 3, 4 et 80.)

[40] François Deschênes and Djemel Ziou. Depth from defocus estimation in spatial domain. *Computer Vision and Image Understanding*, 81(2) :143–165, février 2001. (Cité page 53.)

[41] F. Devillers. *Langage de scénario pour des acteurs semi-autonomes*. Thèse de doctorat, University of Rennes I, September 2001. (Cité page 75.)

[42] Stéphane Donikian. Les modèles comportementaux pour la génération du mouvements d'objets dans une scène. *Revue Internationale de CFAO et d'Infographie*, 9(6) :847–871, 1994. (Cité page 60.)

[43] Stéphane Donikian. Multilevel modeling of virtual urban environments for behavioral animation. In *Computer Animation*, pages 127–133, Genève, Suisse, 1997. (Cité page 71.)

[44] Stéphane Donikian. Vuems : A virtual urban environment modeling system. In *Computer Graphics International*, pages 84–92, Hasselt, Belgique, 1997. (Cité page 71.)

[45] Stéphane Donikian and Alain Chauffaut. Gasp : a general animation and simulation platform. In J. Halttunen and R. Tuokko, editors, *ISMCR'97 Topical Workshop on Virtual Reality and Advanced Man-Machine Interfaces*, volume IXB, pages 158–63, Tampere, Finlande, 1997. (Cité pages 70 et 71.)

[46] Stéphane Donikian, Alain Chauffaut, Thierry Duval, and Richard Kulpa. Gasp : from modular programming to distributed execution. In *Computer Animation'98*, Philadelphia, USA, June 1998. IEEE Computer Society Press. (Cité pages 70 et 71.)

[47] Stéphane Donikian, Frédéric Devillers, and Guillaume Moreau. The kernel of a scenario language for animation and simulation. In Nadia Magnenat-Thalmann and Daniel Thalmann, editors, *Eurographics Workshop for Computer Animation and Simulation*, Milan, Italie, septembre 1999. Eurographics. ISBN 3-211-83392-7. (Cité page 75.)

[48] Stéphane Donikian and Eric Rutten. Reactivity, concurrency, dataflow and hierarchical preemption for behavioural animation. In *Fifth Eurographics Workshop on Programming Paradigms in Graphics*, Maastricht, Pays Bas, 1995. (Cité page 70.)

[49] Nicolas Dubreuil and Dominique Bechmann. Facial animation. In *Computer Animation'97*, pages 98–109, Genève, Suisse, 1997. (Cité page 59.)

[50] G. Egnal and R. P. Wildes. Detecting binocular half-occlusions : empirical comparisons of five approaches. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(8) :1127–1133, 2002. (Cité page 49.)

[51] Peter Eisert, Eckehard Steinbach, and Bernd Girod. Automatic reconstruction of 3-d stationary objects from multiple views. *IEEE Transactions on Circuits and Systems for Video Technology*, 10(2) :261–277, mars 2000. (Cité page 53.)

[52] Irfan Essa, Sumit Bassu, Trevor Darrell, and Alex Pentland. Modeling, tracking and interactive animation of faces and heads using input from video. In *Computer Animation'97*, pages 68–79, Genève, Suisse, 1997. (Cité page 59.)

[53] Olivier Faugeras. *Three-Dimensional Computer Vision. A Geometric Viewpoint*. MIT Press, 1996. (Cité page 49.)

[54] P. Fua, A. Gruen, R. Plaenkers, N. D'apuzzo, and Daniel Thalmann. Human body modeling and motion analysis from video sequences. In *International Symposium on Real Time Imaging and Dynamic Analysis*, Hakodate, Japon, 1998. (Cité page 59.)

[55] P. Fua and C. Miccio. Fitting sophisticated facial animation models to image data. In *Optical 3–D Measurement Techniques Conference*, Zurich, Suisse, 1997. (Cité page 59.)

[56] Philippe Fuchs and Guillaume Moreau, editors. *Le Traité de la Réalité Virtuelle*. Les Presses de l'Ecole des mines, Paris, 3 edition, mars 2005. (Cité pages 3, 16 et 33.)

[57] George A. Gescheider. *Psychophysics : method, theory and application*. Lawrence Erlbaum Associates, New Jersey, USA, 2ème edition, 1985. (Cité pages 19 et 24.)

[58] James J. Gibson. *The ecological approach to visual perception*. Lawrence Erlblaum Associates, Hillsdale, USA, 1986. (Cité pages 28 et 38.)

[59] S. Giroux. *Agents et systèmes, une nécessaire unité*. Thèse de doctorat, Faculté des arts et des sciences, Université de Montréal, Montréal, Canada, 1993. (Cité page 60.)

[60] Stefan Gottschalk, Ming Lin, and Dinesh Manocha. OBBTree : A hierarchical structure for rapid interference detection. In *Proceedings of ACM SIGGRAPH*, pages 171–180, 1996. (Cité page 25.)

[61] J. P. Granieri, Welton Beckett, Barry D. Reich, J. Crabtree, and Norman I. Badler. Behavioral control for real-time simulated agents. In *Symposium on interactive 3D Graphics*, pages 173–180, Monterey, USA, 1995. ACM Computer Press. (Cité page 60.)

[62] P. Guillaume. *La psychologie de la forme*. Flammation, 1979. (Cité page 28.)

[63] Norbert Haala and Jan Böhm. A multi-sensor system for positioning in urban environments. *ISPRS Journal of Photogrammetry & Remote Sensing*, 58 :31–42, January 2003. (Cité page 85.)

[64] Antoine Hahusseau. *Spécification d'automates parallèles hiérarchisés à l'aide du langage réactif Statemate pour la modélisation comportementale*. Rapport de d.e.a., Université de Rennes I, Rennes, 1997. (Cité page 70.)

[65] Stuart Hansen, Joseph Kearney, and James Cremer. Motion control through communicating, hierarchical state machines. In G. Hégron and O. Fahlander, editors, *5th Eurographics Workshop on Animation and Simulation*, Oslo, Norvège, 1994. (Cité page 62.)

[66] David Harel. Statecharts : a visual formalism for complex systems. *Science of Computer Programming*, 8(3) :231–275, 1987. (Cité page 70.)

[67] C. Harris and M.J. Stephens. A combined corner and edge detector. In *Alvey Vision Conference*, pages 147–152, 1988. (Cité page 89.)

[68] Fabrice Harrouët, Eric Cazeaux, and Thomas Jourdan. AReVi. In Philippe Fuchs and Guillaume Moreau, editors, *Le Traité de la Réalité Virtuelle*, volume 3, chapter 13, pages 369–392. Les Presses de l'Ecole des mines, Paris, 3ème edition, 2006. (Cité page 41.)

[69] Gérard Hégron and Bruno Arnaldi. *Computer Animation : Motion and Deformation Control*. Eurographics'92 Tutorial Notes, Eurographics Technical Series, Cambridge, Grande-Bretagne, September 1992. (Cité page 58.)

[70] C.H. Ho, C. Basdogan, and M.A. Srinivasan. Efficient point-based rendering techniques for haptic display of virtual objects. *Presence*, 8(5) :477–491, 1999. (Cité page 25.)

[71] Jessica Hodgins. Three-dimensional human running. In *IEEE Conference on Robotics and Automation ICRA'96*, 1996. (Cité page 59.)

[72] Ian Howard and B. Rogers. *Binocular vision and stereopsis*. Oxford Osychology Series, 1995. (Cité page 15.)

[73] M. Jaeger and Ph. de Reffye. Basic concepts of computer simulation of plant growth. *J. Biosci*, 17 (3) :275–291, September 1992. (Cité page 58.)

[74] Jihen Jallouli and Guillaume Moreau. An immersive path-based study of wind turbines landscape : a french case in plouguin. *Renewable Energy*, 34(3) :597–607, mar 2009. (Cité pages 40 et 41.)

[75] Jihen Jallouli, Guillaume Moreau, and Ronan Querrec. Wind turbines landscapes : using virtual reality for the assessment of multisensory perception in motion. In *ACM Virtual Reality Software and Technology (poster)*, Bordeaux, France, oct 2008. ACM. (Cité page 41.)

[76] Jihen Jallouli, Guillaume Moreau, and Davy Salerne. Free versus constrained motion for assessing wind turbines' impacts on landscape in virtual environments. In *International Conference on Artificial Reality and Telexistence (poster)*, pages 261–254, Yokohama, Japan, december 2008. The Virtual Reality Society of Japan. (Cité page 41.)

[77] Björn Johansson and Roberto Cipolla. A system for automatic pose-estimation from a single image in a city scene. In *Signal Processing, Pattern Recognition, and Applications, Crete, Greece, June 25-28*, 2002. (Cité page 85.)

[78] Ian Johnson. Putting time on the map : using TimeMap for map animatin and delivery. *GeoInformatics*, 2004-8, août 2004. (Cité page 91.)

[79] Thierry Joliveau. *Géomatique et gestion environnementale du paysage : recherche sur un usage géographique des SIG*. Mémoire d'habilitation à diriger des recherches, Université Jean Monnet, Saint-Etienne, 2004. (Cité page 37.)

[80] L.A. Jones and I.W. Hunter. Human operator perception of mechanical variables and their effects on tracking performances. In *ASME Winter Annual Meeting : Advances in Robotics*, volume 42, pages 49–53, 1992. (Cité page 24.)

[81] B. Julesz. *Fundation of cyclopean perception*. University of Chicago Press, 1971. (Cité page 16.)

[82] Mahzad Kalantari, Jean-Pierre Guédon, Franck Jung, and Guillaume Moreau. Détection entièrement automatique de points de fuite dans des scènes architecturales urbaines. In *CORESA*, 2006. (Cité page 85.)

[83] Mahzad Kalantari, Franck Jung, Jean-Pierre Guédon, and Nicolas Paparoditis. Détection automatique des points de fuite et calcul de leur incertitude à l'aide de la géométrie projective. In *RFIA*, Amiens, France, 2008. (Cité page 85.)

[84] Mohamed Benali Khoudja. Facteurs humais et études d'ergonomie pour les planches de bord automobiles. Rapport de dea réalité virtuelle et maîtrise des systèmes complexes, Ecole des mines de Paris, Paris, juin 2001. (Cité page 24.)

[85] Mohamed Benali Khoudja, Guillaume Moreau, Philippe Fuchs, and Panagiotis Stergiopoulos. Human factors and ergonomic studies for the car manufacturing industry. In *ICEE 2002, International Conference on Engineering Education*, Manchester, UK, août 2002. (Cité page 24.)

[86] Kiriakos Kutulakos and Steven Seitz. A theory of shape by space carving. *International Journal of Computer Vision*, 38(3) :199–218, 2000. (Cité page 54.)

[87] Eckart. Lange. The limits of realism : perceptions of virtual landscapes. *Landscape and Urban Planning*, 54 :163–182, 2001. (Cité page 38.)

[88] Aldo Laurentini. The visual hull concept for silhouette-based image understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(2) :150–162, février 1994. (Cité page 51.)

[89] Aldo Laurentini. How many 2d silhouettes does it take to reconstruct a 3d object ? *Computer Vision and Image Understanding*, 67(1) :81–87, juillet 1997. (Cité page 52.)

[90] Christophe Lecerf. *Contrôle du mouvement de systèmes mécaniques en animation*. Thèse de doctorat, Université de Rennes I, 1994. (Cité page 62.)

[91] Thomas Leduc, Erwan Bocher, Fernando Gonzaléz-Cortés, and Guillaume Moreau. Gdms-r : A mixed sql to manage raster and vector data. In *GIS 2009*, Ostrava, République Tchèque, jan 2009. (Cité page 81.)

[92] Jong Weon Lee, Suya You, and Ulrich Neumann. Tracking with omni-directional vision for outdoor ar systems. In *ISMAR, Darmstadt, Germany, September 30-October 1*, pages 47–58, 2002. (Cité page 85.)

[93] P. Lee, C. Phillips, E. Otani, and Norman I. Badler. The jack interactive human model. In *First Annual Symposium on Mechanical Design in a Concurrent Engineering*, volume 1, pages 179–198, Iowa City, USA, October 1989. (Cité page 59.)

[94] Yuencheng Lee, Demetri Terzopoulos, and Keith Waters. Realistic modeling for facial animation. In Robert Cook, editor, *SIGGRAPH'95*, pages 55–62, Los Angeles, USA, 1995. ACM Computer Press. (Cité page 59.)

[95] Nicolas Lemmer. Etude de la stéréovision : aspects ergonomiques. Rapport de DEA RV-MSC, Université d'Evry-Val d'Essonne, Paris, 2002. (Cité page 17.)

[96] Nicolas Lemmer, Guillaume Moreau, and Philippe Fuchs. Enhancement of stereoscopic comfort by fast control of frequency content with wavelet transform. In Andrew Woods, Mark Bolas, John Merritt, and Stephen Benton, editors, *Stereoscopic Displays and Virtual Reality Systems X*, volume 5006 of *Proceedings of Electronic Imaging*, pages 283–291, Santa Clara, CA, janvier 2003. SPIE. (Cité page 17.)

[97] Laure Leroy, Philippe Fuchs, Alexis Paljic, and Guillaume Moreau. Study of stereoscopic shape perception for virtual design. In *International Conference on Artificial Reality and Telexistence*, pages 144–151, Yokohama, Japon, décembre 2008. The Virtual Reality Society of Japan. (Cité page 18.)

[98] Ming Lin and Stefan Gottschalk. Collision detection between geometric models : A survey. In *Proceedings of IMA Conference on Mathematics of Surfaces*, 1998. (Cité page 25.)

[99] Ming Lin, Dinesh Manocha, and Jon Cohen ans Stefan Gottschalk. Collision detection : algorithms and applications. In *2nd Workshop on Algorithmic Foundations of Roboticscs Motion and Manipulation*, 1996. (Cité page 25.)

[100] Nathalie Long, Erwan Bocher, Thomas Leduc, and Guillaume Moreau. Sensitivity of spatial indicators for urban terrain characterization. In *IEEE International Geoscience and Remote Sensing Symposium*, Boston, MA, juillet 2008. IEEE. (Cité pages 82 et 83.)

[101] Nathalie Long, Patrice Mestayer, and Christophe Kergomard. Development of a software to describe the city morphology and to compute aerodynamic parameters from an urban database. In *Proc. of the 4th Symposium on the Urban Environment*, Norfolk, USA, 2002. AMS. (Cité page 83.)

[102] R. G. Lord and P. E. Levy. Moving from cognition to action : A control theory perspective. *Applied Psychology : an international review*, 43 (3) :335–398, 1994. (Cité page 61.)

[103] William E. Lorensen and Harvey E. Cline. Marching cubes : A high resolution 3d surface construction algorithm. *Computer Graphics (Proceedings of SIGGRAPH 87)*, 21(4) :163–169, July 1987. Held in Anaheim, California. (Cité page 53.)

[104] Kevin Lynch. *The image of the city*. MIT Press, 1960. (Cité pages ix et 30.)

[105] Anatole Lécuyer, Sabine Coquillart, Abderhramane Kheddar, Paul Richard, and Philippe Coiffet. Pseudo-haptice feedback : can isometric input devices simulate force feedback ? In *IEEE Virtual Reality*, pages 83–90, New Brunswick, USA, 2000. (Cité page 24.)

[106] I. Scott MacKenzie and Colin Ware. Lag as a determinant of human performance in interactive systems. In *Proceedings of the ACM Conference on Human Factors in Computing Systems - INTERCHI '93*, pages 488–493, New York, 1993. ACM. (Cité page 14.)

[107] Nadia Magnenat-Thalmann and Daniel Thalmann, editors. *Artificial Life and Virtual Reality*. John Wiley, Chichester, 1984. (Cité page 62.)

[108] Stéphane Mallat. Zero-crossings of a wavelet transform. *IEEE Transactions on Information Theory*, 37 :1019–1033, 1991. (Cité page 46.)

[109] D. Marr and T. Poggio. Cooperative computation of stereo disparity. *Science*, (194) :283–287, 1976. (Cité pages 45 et 49.)

[110] Y. Matsumoto, K. Fujimura, and T. Kitamura. Shape-from-silhouette/stereo and its application to 3-d digitizer. In G. Bertrand, M. Couprie, and L. Perroton, editors, *Proceedings of the International Conference on Discrete Geometry for Computer Imagery*, number 1568 in Lecture Notes in Computer Science, pages 177–188, Marne-la-Vallée, France, mars 1999. Springer Verlag. (Cité page 53.)

[111] R. M. Michaels and J. Fazio. Driver behavior model of merging. *Transportation Research Record*, 1213 :4–10, 1989. (Cité page 76.)

[112] A. Moles. *Information Theory and Aesthetic Perception*. University of Illinois Press, Illinois, 1966. (Cité page 38.)

[113] A. Moles and E. Rohmer. *Space psychosociology (in French)*. L'Harmattan, Paris, 1998. (Cité page 29.)

[114] Guillaume Moreau. *Modélisation du comportement pour la simulation interactive : application au trafic routier multimodal*. Thèse de doctorat, Université de Rennes I, Rennes, novembre 1998. (Cité pages 64, 65, 69, 71 et 73.)

[115] Guillaume Moreau. Techniques de modélisation 3d à partir d'images photographiques. Rapport de Recherche (contrat PRIAMM-CNC Gantom) CAOR/2001-06-1, Ecole des Mines de Paris, Paris, juin 2001. (Cité page 51.)

[116] Guillaume Moreau and Stéphane Donikian. From psychological and real-time interaction requirements to behavioural simulation. In Bruno Arnaldi and Gérard Hégron, editors, *Eurographics Workshop on Computer Animation and Simulation*, pages 29–44, Lisbonne, Portugal, septembre 1998. Springer. (Cité page 71.)

[117] Guillaume Moreau and Philippe Fuchs. Mise au point d'un outil de réalite virtuelle en vue d'études de visibilité. Rapport de fin de contrat Armines - PSA, Centre de Robotique, Ecole des Mines de Paris, 2000. (Cité page 21.)

[118] Guillaume Moreau and Philippe Fuchs. IMAVE : outil de réalité virtuelle pour la revue de projet. Rapport de fin de contrat Armines - Sommer Allibert Automotive Industries PPROB.0780093, Centre de Robotique, Ecole des Mines de Paris, Paris, France, 2001. (Cité page 23.)

[119] Guillaume Moreau and Philippe Fuchs. Virtual reality in the design process : from design review to ergnomic studies. In N. Giambiasi and C. Frydman, editors, *ESS'2001*, pages 123–130, Marseille, octobre 2001. SCS Europe. (Cité page 20.)

[120] Guillaume Moreau and Philippe Fuchs. Stereoscopic displays for virtual reality in the car manufacturing industry : application to desgin review and ergonomic studies. In Andrew Woods, John Merritt, Stephen Benton, and Mark Bolas, editors, *Proceedings of SPIE "Stereoscopic Displays and Virtual Reality Systems IX"*, volume 4660, pages 504–513, San Jose, Californie, USA, janvier 2002. IST & SPIE. (Cité pages 16 et 20.)

[121] Guillaume Moreau, Philippe Fuchs, Andrei Doncescu, and Sébastien Régis. Dense stereo matching method using a quarter of wavelet transform. In *Proceedings of the International Conference on Image Processing (ICIP'02)*, volume 1, pages 261–264, Rochester, NY, Sep. 2002. IEEE. (Cité page 47.)

[122] Oussama Moslah, Mathieu Klee, Antoine Grolleau, Vincent Guitteny, Serge Couvet, and Sylvie Philipp-Foliguet. Urban models texturing from un-calibrated photographs. In *23rd International Conference Image and Vision Computing New Zealand*, 2008. (Cité page 87.)

[123] Richard Mouli. *Un modèle d'animation comportementale basé sur le concept de personnage*. Thèse de doctorat, Université Paul Sabatier, Toulouse, 1994. (Cité page 61.)

[124] Richard Mouli, Yves Duthen, and René Caubet. Un modèle de simulation comportementale orienté objet pour l'animation intelligente. In *IMRV92 Interface des Mondes Réels et Virtuels*, pages 111–121, Montpellier, 1992. (Cité page 61.)

[125] Franck Multon. Animation d'humanoïdes synthétiques par un modèle biomécanique. In *4èmes Journées de l'Association Française Graphique*, Dijon, November 1996. (Cité page 59.)

[126] Franck Multon, Laure France, Marie-Paule Cani-Gascuel, and Giles Debunne. Computer animation of human walking : a survey. *The Journal of Visualization and Computer Animation*, 10(1) :39–54, January–March 1999. (Cité page 59.)

[127] D.D. Nelson, D.E. Johnson, and E. Cohen. Haptic rendering of surface-to-surface sculpted model interaction. In *Proceedings of the 8th Annual Symposium on Haptic User Interfaces for Virtual Environments and Teleoperator Systems*, Nashville, USA, novembre 1999. ASME. (Cité page 26.)

[128] S. Neubauer and A. Zipf. Suggestions for extending the ogc styled layer descriptor (sld) specification into 3d - towards visualization rules for 3d city models. In *Urban Data Management Society Symposium*, Stuttgart, Germany, 2007. (Cité page 91.)

[129] Allen Newell. *Unified Theories of Cognition*. Harvard University Press, 1990. (Cité page 64.)

[130] J. Thomas Ngo and Joe Marks. Spacetime constraints revisited. In James T. Kajiya, editor, *SIGGRAPH'93*, pages 343–350, Anaheim, USA, 1993. ACM Computer Press. (Cité page 60.)

[131] Wolfgang Niem. Robust and fast modelling of 3d natural objetcs from multiple views. In *Proceedings Image and Video Processing II*, volume 2182, pages 388–397, San Jose, USA, février 1994. SPIE. (Cité page 53.)

[132] Wolfgang Niem. Error analysis for silhouette-based 3d shape estimation from multiple views. In *International Workshop on Synthetic - Natural Hybrid Coding and Three Dimensional Imaging (IWSNHC3DI'97)*, Rhodes, Grèce, septembre 1997. (Cité page 52.)

[133] H.K. Nishishara. Practical real-time imaging stereo matcher. In M. Fichler and O. Firschein, editors, *Readings in Computer Vision*. Kauffman, Los Altos, 1987. (Cité page 45.)

[134] Hansrudi Noser. *A behavioral animation system based on L-systems and Synthetic Sensors for Actors*. Thèse de doctorat, École Polytechnique Fédérale de Lausanne, Lausanne, Suisse, 1997. (Cité page 62.)

[135] University of Newcastle. Visual assessment of windfarms : Best practice. Commissioned report fo1aa303a, Scottish Natural Heritage, 2002. (Cité page 38.)

[136] Y.E. Papelis and S. Bahauddin. Logical modeling of roadway environment to support real-time simulation of autonomous traffic. In *SIVE95 : the First Workshop on Simulation and Interaction in Virtual Environments*, pages 62–71, University of Iowa, Iowa City, U.S.A., July 1995. (Cité page 73.)

[137] Eja Pedersen and Kerstin Persson Waye. Exploring perception and annoyance due to wind turbine noise in dissimilar living environments. In Acoustic Society of Finland EAA and VTT, editors, *Euronoise 2006*, page 6, Tampere, Finlande, 30 Mai - 1 juin 2006. Colloque international Bordeaux, EAA, Acoustic Society of Finland and VTT. (Cité page 40.)

[138] Jérôme Perrin, Bruno Torrésani, and Philippe Fuchs. Une fonction de corrélation localisée pour la mise en correspondance des images stéréoscopiques. *Traitement du signal*, 16(1), 1999. (Cité pages 46 et 50.)

[139] Jérôme Perrin. *Profondeur et binocularité : algorithmie, étude psychophysique et intérêt pour l'ergonomie des interfaces stéréoscopiques*. Thèse de doctorat, Ecole des Mines de Paris, Paris, septembre 1998. (Cité pages ix, 14, 15, 16, 17, 46 et 50.)

[140] Jérôme Perrin, Philippe Fuchs, Corinne Roumes, and F. Perret. Improvement of stereoscopic comfort through control of disparity and spatial frequency content. In *Visual Information Processing VII*, volume 3387-14 of *Proceedings of SPIE*, Apr. 1998. (Cité page 16.)

[141] C. Phillips and Norman I. Badler. Jack : a toolkit for manipulating articulated figures. In *ACM/SIGGRAPH Symposium on user interface software*, Banff, Canada, October 1988. (Cité page 59.)

[142] V. Popescu, , Grigore Burdea, and M. Bouzit. Virtual reality simulation modelling for a haptic glove. In *Computer Animation'99*, pages 195–200, Genève, Suisse, 1999. (Cité page 26.)

[143] Michael Potmesil. Generating octree models of 3d objects from their silhouettes in a sequence of images. *Computer Vision, Graphics and Image Processing*, 40 :1–29, 1987. (Cité page 53.)

[144] Jean-Charles Quillet, Gwenola Thomas, Xavier Granier, Pascal Guitton, and Jean-Eudes Marvie. Using expressive rendering for remote

visualization of large city models. In *Web3D '06 : Proceedings of the eleventh international conference on 3D web technology*, pages 27–35. ACM, apr 2006. (Cité pages xi et 93.)

[145] Serge Rainjonneau. *Un modeèle orienté-objet pour la simulation comportementale*. Thèse de doctorat, Université Paul Sabatier, Toulouse, 1992. (Cité page 61.)

[146] Sébastien Régis, Andrei Doncescu, Guillaume Moreau, and Philippe Fuchs. Dense stereo correspondence using quarters of wavelet transform. *International Journal of Pattern Recognition and Artificial Intelligence*, 19(1) :21–43, janvier 2005. (Cité pages 48 et 50.)

[147] Barry D. Reich. *An architecture for behavioral locomotion*. Thèse de doctorat, Université de Pennsylvannie, Philadelphie, USA, 1997. (Cité page 62.)

[148] Gehrard Reitmayr and Tom W. Drummond. Initialisation for visual tracking in urban environments. In *ISMAR, Nara, Japan, November 13-16*, pages 1–9, 2007. (Cité page 85.)

[149] Craig W. Reynolds. Flocks, herds, and schools : A distributed behavioral model. In *SIGGRAPH'87*, volume 21, pages 25–34, Anaheim, 1987. (Cité page 60.)

[150] W. D. Reynolds and R. V. Kenyon. The wavelet transform and the suppression theory of binocular vision for stereo image compression. In *IEEE International Conference on Image Processing*, volume 2, pages 557–560, Lausanne, Suisse, 1996. (Cité page 46.)

[151] Corinne Roumes, Justin Plantier, and J.P. Menu. Local contrast : a quantification tool for stereo image. In *Proceedings of the Annuel International Conference of the IEEE*, volume 5, pages 2098–2099. IEEE, 1992. (Cité page 14.)

[152] David Roussel. *Reconstruction de courbes et de surfaces 3D en stéréo-acquisition*. Thèse de doctorat, Université de Paris XI, Orsay, janvier 1999. (Cité page 53.)

[153] Davy Salerne. Realité virtuelle et augmentée pour la gestion urbaine et paysagère. Rapport de master mnrv, ENSAM Angers, Laval, septembre 2008. (Cité pages 41 et 42.)

[154] C. Sénat, S. Garrigues, and Gamba R. Calculate noise of wind farms. In *First international conference on wind turbines noise : perspectives for Control proceedings*, Berlin, 2005. INCE/Europe. (Cité page 40.)

[155] Camillo Sitte. *Art of building cities : city building according to its artistic fundamentals*. Hyperion Press Inc., Westport, CT, 1979. (Cité pages 30 et 31.)

[156] Gregory Slabaugh, Thomas Malzbender, Bruce Culbertson, and Ronald Schafer. Improved voxel coloring via volumetric optimization. Center for Signal and Image Processing Tech Report 3, Georgia Institute of Technology, Atlanta, GA, 2000. (Cité page 52.)

[157] Alex Sosnov, Pierre Macé, and Gérard Hégron. Semi-metric formal 3d reconstruction from perspective sketches. In *International Workshop on Computer Graphics and Geometric Modeling*, number 2330 in LNCS, pages 285–294, Berlin-Heidelberg, 2002. Springer-Verlag. (Cité page 90.)

[158] Eckehard Steinbach, Peter Eisert, A. Betz, and Bernd Girod. 3-d object reconstruction using spatially extended voxels and multi-hypothesis voxel coloring. In *International Conference on Pattern Recognition*, volume 1, pages 774–777, Barcelone, Espagne, septembre 2000. (Cité page 53.)

[159] L. Stelmach, W.J. Tam, and D. Meegan. Perceptual basis of stereoscopic video. In *Proceedings of SPIE*, volume 3639, pages 260–265, 1999. (Cité page 12.)

[160] L. Stelmach, W.J. Tam, D. Meegan, A. Vincent, and P. Corriveau. Human perception of mismatched stereoscopic 3D inputs. In *International Conference on Image Processing*, volume 1, pages 5–8. IEEE, septembre 2000. (Cité page 12.)

[161] Panagiotis Stergiopoulos, Guillaume Moreau, Mehdi Ammi, and Philippe Fuchs. A framework for the haptic rendering of the human hand. In *Virtual Reality HAPTICS 2003*, pages 340–347, Los Angeles, CA, 2003. IEEE. (Cité page 26.)

[162] Panagiotis Stergiopoulos, Guillaume Moreau, Domitile Lourdeaux, and Philippe Fuchs. Fast and accurate collision detection for haptic interaction with solid models and virtual prototypes. In *VRIC*, Laval, France, juin 2002. (Cité page 25.)

[163] Richard Szeliski and Ramin Zabih. An experimental comparison of stereo algorithms. In *International Workshop on Vision Algorithms*, pages 1–19, Kerkyra, Greece, Sept. 1999. Springer Verlag. (Cité page 50.)

[164] Souha Tahrani. Les indicateurs physico-sensibles pour une exploitation dans les ambiances lumineuses. Rapport de DEA ambiances architecturales et urbaines, Ecole Nationale Supérieure d'Architecture de Nantes, Nantes, 2002. (Cité page 33.)

[165] Souha Tahrani. *Vers un outil de réalité virtuelle pour l'analyse et la communication des ambiances lumineuses dans le projet urbain*. Thèse de doctorat, Université de Nantes, Nantes, Nov. 2006. (Cité pages ix, 29, 32 et 35.)

[166] Souha Tahrani, Jihen Jallouli, Guillaume Moreau, and Philippe Woloszyn. Towards a virtual reality tool for lighting communication and analysis in urban environments. In Bob Martens and André Brown, editors, *CAADS Future 2005*, pages 115–124, Vienne, Autriche, juin 2005. Springer Verlag. (Cité page 33.)

[167] Souha Tahrani and Guillaume Moreau. Analyzing urban daylighting ambiences by walking in a virtual city. In William Sherman, Ming Lin, and Anthony Steed, editors, *Virtual Reality 2007*, pages 227–230, Charlotte, USA, mars 2007. IEEE. (Cité page 33.)

[168] Souha Tahrani and Guillaume Moreau. Integration of immersive walking to analyze urban daylighting ambiences. *Journal of Urban Design*, 13(1) :99–123, février 2008. (Cité pages 32 et 35.)

[169] W.J. Tam, L.B. Stelmach, and P. Corriveau. Psychovisual aspects of viewing stereoscopic video sequences. In *Proceedings of SPIE*, volume 3295, pages 226–235, avril 1998. (Cité page 12.)

[170] H.Z. Tan, X.D. Pang, and N.I. Durlach. Manual resolution of lenght, force and comppliance. In *ASME Winter Annual Meeting : Advances in Robotics*, volume 42, pages 13–18, 1992. (Cité page 24.)

[171] Daniel Thalmann. Interactive perceptive virtual humans. In *Computer Animation'98*, Philadelphie, USA, 1998. (Cité page 62.)

[172] Jean-Paul Thibaud. Frames of visibility in public places. *Places : A forum of environment design*, 14(1) :42–47, 2001. (Cité pages 31 et 40.)

[173] Gwenola Thomas and Stéphane Donikian. Virtual humans animation in informed urban environments. In *IEEE Computer Animation*, pages 129–136, Philadelphia, PA, 2000. IEEE Computer Society Press. (Cité page 75.)

[174] D. J. Tolhurst and Y. Tadmor. Band-limited contrast in natural images explains the detectability of changes in the amplitude spectra. *Vision Research*, 37(23) :3203–3215, 1997. (Cité page 46.)

[175] M.R. Tremblay, C. Ullrich, D.H. Gomez, R. Carmel, and J. Tian. Whole-hand interaction with 3d environments. In *Workshop on perceptual user interfaces*, San Francisco, USA, 1998. (Cité page 26.)

[176] Roger Tsai. An efficient and accurate camera calibration technique for 3d machine vision. In *IEEE Conference on Computer Vision and Pattern Recognition*, pages 364–374, 1986. (Cité pages 52 et 56.)

[177] Constantinos Tzafestas. *Synthèse de retour kinésthésique et perception haptique lors de tâches de manipulation*. Thèse de doctorat, Université de Paris VI, Paris, 1998. (Cité page 24.)

[178] Hideaki Uchiyama, Hideo Saito, Vivien Nivesse, Myriam Servières, and Guillaume Moreau. AR representation system for 3D GIS based on camera pose estimation using distribution of intersections. In *International Conference on Artificial Reality and Telexistence*, pages 218–225, Yokohama, Japon, december 2008. The Virtual Reality Society of Japan. (Cité page 94.)

[179] B. Ullmer and H. Ishii. The metaDESK : Models and prototypes for tangible user interfaces. In *Proceedings of User Interface Software Technology (UIST '97)*, pages 223–232, Banff, octobre 1997. ACM, ACM. (Cité page 94.)

[180] Raymond Unwin. *Town planning in practice : an introduction to the art of designing cities ans suburbs*. Classic Reprint. Princeton Architectural Press, New York, 1994. (Cité pages 30 et 31.)

[181] Martin Usoh, Kevin Arthur, Mary C. Whitton, Rui Bastos, Anthony Steed, Mel Slater, and Frederick P. Brooks Jr. Walking > walking in place > flying in virtual environments. In *SIGGRAPH'99*, pages 359–364. ACM, 1999. (Cité page 33.)

[182] N.A. Valyus. *Sterescopy*. Focal Press, 1962. (Cité pages 14 et 22.)

[183] Michiel van de Panne and Eugene Fiume. Sensor-actuator networks. In James T. Kajiya, editor, *Computer Graphics (SIGGRAPH'93 Proceedings)*, pages 335–342, Anaheim, USA, 1993. ACM Computer Press. (Cité page 60.)

[184] J.P. Wann, S. Rushton, and M. Mon-Williams. Natural problems for stereoscopic depth perception in virtual environments. *Vision Res.*, 35 :2731–2736, 1995. (Cité page 12.)

[185] Bonnie Webber and Norman Badler. Animation through reactions, transition nets and plans. In *Internation Workshop on Human Interface Technology*, Aizu, Japon, 1995. (Cité page 62.)

[186] Toto Widyanto, Andrew Marriot, and Malcom West. Applying a virtual perception model to behavioral animation system. In *Eurographics Workshop on Animation and Simulation*, pages 89–98, Vienne, Autriche, 1991. (Cité page 60.)

[187] Jane Wilhelms and Robert Skinner. a notion for interactive behavioral animation control. *IEEE Computer Graphics and Applications*, 10 (3) :14–22, May 1990. (Cité page 60.)

[188] M. Wöpking. Viewing comfort with stereoscopic. *SID*, 3 :101–103, 1995. (Cité pages ix, 14 et 15.)

[189] S. Yano, S. Ide, T. Mituhashi, and H. Twaites. A study of visual fatigue and visual comfort for 3D HDTV/HDTV images. *Display*, 213 :191–201, 2002. (Cité page 12.)

[190] Chamseddine Zaki, Myriam Myriam Servières, and Guillaume Moreau. An ontology-based model for representing evolution of both data and semantic in gis. In Edzer Pebesma, Mohamed Bishr, and Thomas Bartoschek, editors, *GI Days*, pages 157–164, Münster, Allemagne, juin 2008. IfGI. (Cité page 81.)

[191] David Zeltzer. Task-level graphical simulation : abstraction, representation and control. In N. I. Badler, B. A. Barsky, and D. Zeltzer, editors, *Making them move : Mechanics, control and animation of articulated figures*, pages 3–33. Morgan Kaufmann, 1991. (Cité page 60.)

[192] Ruo Zhang, Ping-Sing Tsai, James Edwin Cryer, and Mubarak Shah. Shape from shading : A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(08) :690–706, août 1999. (Cité page 53.)

[193] X. Zhao, D. Tolani, B.J. Ting, and Norman I. Badler. Simulating human movements using optimal control. In *Computer Animation and Simulation'96*, pages 109–120, 1996. (Cité page 59.)

Titre Représentation et construction des systèmes d'information par l'image

Résumé Nous nous intéressons au lien entre l'image numérique et un système d'information dans un cadre de conception de produit centré sur l'usager.
Dans une première partie, nous nous intéressons à la perception par l'usager des environnements virtuels représentant un système d'information. La fidélité perceptive de l'environnement virtuel est étudié à plusieurs niveaux : psychophysique, fonctionnel et enfin sensible.
La seconde partie est consacrée à la construction d'environnements virtuels fidèles au monde réel : nous abordons d'abord leur fidélité géométrique via la vision par ordinateur, puis la fidélité comportementale en reproduisant le comportement humain dans les environnements virtuels.
Enfin, dans la dernière partie, nous proposons de considérer l'interaction entre le monde réel et le système d'information comme un système fonctionnant en boucle fermée : le système peut être représenté par des images mais aussi construit et mis à jour directement avec des images. Le domaine d'application visé est le système d'information géographique.

Mots-clés perception humaine et synthétique, réalité virtuelle, réalité augmentée, système d'information, synthèse d'images

Title Representation and construction of information system with images

Abstract Our work focuses on the link between digital images and information systems in the used-centered design domain.
In the first part, we deal with user perception of virtual environments representing an information system. Perceptual fidelity is assessed on several levels : psychophysical, functional and sensitive.
The second part is dedicated to the construction of virtual environments faithful to the real world : we first focus on the geometrical fidelity through computer vision then on behavioral fidelity by reproducing human behavior in virtual environments.
At last, we propose to consider interaction between the real world and information systems as a closed-loop system : the system can be represented by images but can also be constructed and updated by images. The application domain is geographical information systems.

Keywords human and synthetic perception, virtual reality, augmented reality, information systems, computer graphics